Seasonal changes in the tropospheric carbon monoxide profile over the remote Southern Hemisphere evaluated using multi-model simulations and aircraft observations

J. A. Fisher¹, S. R. Wilson¹, G. Zeng², J. E. Williams³, L. K. Emmons⁴, R. L. Langenfelds⁵, P. B. Krummel⁵, and L. P. Steele⁵

¹University of Wollongong, Wollongong, New South Wales, Australia
²National Institute of Water and Atmospheric Research, Lauder, New Zealand
³Royal Netherlands Meteorological Institute, De Bilt, the Netherlands
⁴National Center for Atmospheric Research, Boulder, Colorado, USA
⁵CSIRO Oceans and Atmosphere Flagship, Aspendale, Victoria, Australia

Correspondence to: J. A. Fisher (jennyf@uow.edu.au)

Received: 7 October 2014 – Published in Atmos. Chem. Phys. Discuss.: 3 November 2014
Revised: 7 February 2015 – Accepted: 24 February 2015 – Published: 23 March 2015

Abstract. The combination of low anthropogenic emissions and large biogenic sources that characterizes the Southern Hemisphere (SH) leads to significant differences in atmospheric composition relative to the better studied Northern Hemisphere. This unique balance of sources poses significant challenges for global models. Carbon monoxide (CO) in particular is difficult to simulate in the SH due to the increased importance of secondary chemical production associated with the much more limited primary emissions. Here, we use aircraft observations from the 1991–2000 Cape Grim Overflight Program (CGOP) and the 2009–2011 HIAPER (High-performance Instrumented Airborne Platform for Environmental Research) Pole-to-Pole Observations (HIPPO), together with model output from the SH Model Intercomparison Project, to elucidate the drivers of CO vertical structure in the remote SH. Observed CO vertical profiles from Cape Grim are remarkably consistent with those observed over the southern mid-latitudes Pacific 10–20 years later, despite major differences in time periods, flight locations, and sampling strategies between the two data sets. These similarities suggest the processes driving observed vertical gradients are coherent across much of the remote SH and have not changed significantly over the past 2 decades. Model ability to simulate CO profiles reflects the interplay between biogenic emission sources, the chemical mechanisms that drive CO production from these sources, and the transport that redistributes this CO throughout the SH. The four chemistry-climate and chemical transport models included in the intercomparison show large variability in their abilities to reproduce the observed CO profiles. In particular, two of the four models significantly underestimate vertical gradients in austral summer and autumn, which we find are driven by long-range transport of CO produced from oxidation of biogenic compounds. Comparisons between the models show that more complex chemical mechanisms do not necessarily provide more accurate simulation of CO vertical gradients due to the convolved impacts of emissions, chemistry, and transport. Our results imply a large sensitivity of the remote SH troposphere to biogenic emissions and chemistry, both of which remain key uncertainties in global modeling. We suggest that the CO vertical gradient can be used as a metric for future model evaluation as it provides a sensitive test of the processes that define the chemical state of the background atmosphere.

1 Introduction

Carbon monoxide (CO) plays multiple fundamental roles in tropospheric chemistry, in particular serving as a major reactant of the hydroxyl radical OH (Logan et al., 1981) and...
as an indirect greenhouse gas (Myhre et al., 2013). A product of incomplete combustion, CO has primary sources from fossil fuel and biomass burning (BB) as well as secondary sources from oxidation of methane (CH₄) and non-methane volatile organic compounds (NMVOCs), with a typical tropospheric lifetime of 1–2 months. In the Southern Hemisphere (SH), the distribution of CO is strongly impacted by emissions from BB (Edwards et al., 2006; Gloudemans et al., 2006) and biogenic sources (Williams et al., 2013), while anthropogenic emissions play only a minor role due to an inter-hemispheric transport barrier caused by the Inter-Tropical Convergence Zone (Hamilton et al., 2008). Much of the SH is characterized by very low CO emissions, and in these remote regions CO is largely controlled by the balance between long-range transport, production from methane oxidation, and chemical removal via reaction with OH. Seasonal variability in CO sources, transport pathways, and loss processes leads to a complex seasonal cycle that is different in the free troposphere than at the surface (Pak et al., 2003). The ability of large-scale global atmospheric models to represent the processes driving this seasonality has been difficult to evaluate due to a paucity of measurements in the SH free troposphere. Particularly rare are observations of the CO vertical profile in the SH, despite the importance of such measurements for testing model processes including source attribution and vertical transport (Liu et al., 2013, 2010). Here, we use simulations from four global chemical transport and chemistry-climate models conducted for the Southern Hemisphere Model Intercomparison Project (SHMIP) to interpret a unique 9-year record of airborne CO vertical profiles in the remote SH from the Cape Grim Overflight Program (CGOP) (Langenfelds et al., 1996).

Evaluation of CO distributions in atmospheric models has largely focused on the Northern Hemisphere where observations are more widely available, with some limited evaluation in the SH as part of global comparisons (e.g., Shindell et al., 2006). The SHMIP was devised to provide a more focused evaluation of current large-scale atmospheric chemistry models in the SH. A central goal of SHMIP is to quantify model ability to represent the seasonal and spatial distributions of trace gases including CO. An overview of SH CO distributions in the four SHMIP models is provided by Zeng et al. (2015), who compare simulated CO to observations from surface in situ and ground-based total column measurements at selected SH sites. They show that using different biogenic emission inventories leads to marked differences in modeled CO at these sites and that accurate representation of biogenic emissions is critical to reproducing observed SH background CO. They also find that the underlying chemical and transport characteristics of each model greatly impact model ability to reproduce background SH CO. In some cases, the inter-model differences are larger than those associated with uncertainties in biogenic emissions, especially for locations further from tropical biogenic and BB sources. Detailed analyses of these uncertainties are addressed by Zeng et al. (2015) using column and surface observations; here we expand on this analysis using in situ observations from the remote free troposphere.

As in the SHMIP model evaluation of Zeng et al. (2015), previous model comparison to observations in the SH has generally been limited to in situ surface data (e.g., Duncan et al., 2007; Wai et al., 2014) and ground- or satellite-based remotely sensed total column data (e.g., De Laat et al., 2007; Edwards et al., 2006; Gloudemans et al., 2006; Kopacz et al., 2010; Morgenstern et al., 2012; Shindell et al., 2006; Zeng et al., 2012). Total column comparisons provide an advantage over in situ surface comparisons for model validation in the free troposphere (Deutscher et al., 2010). However, neither surface nor total column data are able to constrain the vertical structure of CO, which is still poorly understood in the SH mid-latitudes. For example, Shindell et al. (2006) showed that a 26-model ensemble mean was able to reproduce mid-tropospheric CO measurements from the MOPITT satellite instrument in the extra-tropical SH, but the same models uniformly overestimated the upper-to-lower troposphere ratio seen by the satellite (as well as the seasonal cycle of the ratio). This comparison relied on qualitative differences between MOPITT upper and lower tropospheric retrievals (Shindell et al., 2006), as MOPITT sensitivity was different at these two altitudes in the version 3 data used (the newer versions 5 and 6 provide more sensitivity to the lower troposphere). More generally, remote-sensing instruments typically display different sensitivities at different altitudes, making it difficult to use these data to study vertical structure. For quantitative evaluation of vertical gradients, independent in situ data from the free troposphere are essential.

To date, in situ observations of CO in the SH remote free troposphere are sparse. Aircraft campaigns carried out in the SH over the last 2 decades have largely taken place near major emission sources (e.g., BARCA and GABRIEL in South America, SAFARI in southern Africa, ACTIVE/SCOUT in northern Australia) or their outflow regions (e.g., TRACE-A in the South Atlantic). Ongoing programs such as IAGOS/MOZAIC that conduct measurements from aboard commercial aircraft have been limited in the SH, with most concentrated over the African outflow region of the equatorial Atlantic. Neither of these programs included flights over the Pacific or Indian oceans; however, IAGOS flights to Australia began in late 2013 and will likely provide a valuable additional SH data set in the future. More extensive remote sampling of SH CO occurred over the South Pacific during NASA’s PEM-Tropics A (1996) and B (1999) campaigns (Chattfield et al., 2002; Staudt et al., 2001). These campaigns provided detailed characterization of free tropospheric distributions during austral spring and autumn but were temporally limited and unable to capture a full annual cycle. More recently, the HIAPER (High-performance Instrumented Airborne Platform for Environmental Research) Pole-to-Pole Observations (HIPPO) traversed the South Pacific during multiple seasons over the period 2009–2011 (Wofsy, 2011),
offering a previously inaccessible view of seasonal variability in the remote SH free troposphere. However, with only one set of flights in each season (including 4–6 individual flights in the SH), it remains difficult to quantify the seasonal and interannual representativeness of these data, complicating their interpretation.

The 9-year record of aircraft data from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) CGOP (Langenfelds et al., 1996) provides a unique data set to quantify seasonal variability at altitudes from the surface to 8 km in the remote SH. With monthly flights over the Southern Ocean during clean air conditions, this record contains significant information on the seasonal and vertical structure of CO in the SH free tropospheric background. We use this record to develop a climatological picture of CO seasonal cycles and vertical gradients in the remote SH that can be used to test both the temporal representativeness of other data sets (e.g., HIPPO) and the capabilities of models in these data-poor environments. We first describe both the models and the observations used for constructing the SH CO climatology (Sect. 2) and examine the ability of the models to match observed CO vertical gradients across different seasons (Sect. 3). We then use sensitivity studies to quantify the roles of emissions, transport, and chemistry in driving inter-model variability and examine the sensitivity of the simulations to the various uncertainties introduced (Sect. 4). Finally, we evaluate model differences in chemical mechanisms and vertical transport in terms of their impacts on model ability to match observed CO vertical gradients in the remote SH (Sect. 5). A summary and conclusions are presented in Sect. 6.

2 Observations and models

2.1 Cape Grim Overflight Program

Australia’s CSIRO has had long involvement in aircraft-based sampling of atmospheric composition above the southeastern Australian region (Francey et al., 1999; Langenfelds et al., 1996). Between 1972 and 1991, multiple sampling programs were maintained at different times, involving various sampling strategies and locations. From August 1991, upgraded analytical equipment and techniques allowed improved sampling relative to earlier flights, focused on obtaining regular (approximately monthly) vertical profiles of the clean marine troposphere. The CGOP ran from August 1991 through December 1999, with additional sampling taking place during August–September 2000. Flights were conducted out of Melbourne, Victoria, flying southward over the Bass Strait towards Cape Grim, Tasmania, with spatial coverage spanning between 38.6°–41.5° S and 142.1°–146.0° E. Approximately 85 flights were carried out over the life of the program, with sampling locations shown in Fig. S1 in the Supplement. The program was designed to measure background concentrations of CO and greenhouse gases in conditions representative of the remote SH. Flights were therefore conducted only during anticipated clean air conditions, typically characterized by southwestward surface winds (Pak et al., 2003). Vertical profiles from 0 to 8 km were measured due west of Cape Grim on most flights (centered around 40.5° S, 144.3° E; see Fig. 2) but with some variation in the exact location to avoid sampling outflow from Tasmania. Air was collected in glass flasks, with on average 17–20 samples per flight, and subsequently analyzed in the CSIRO Global Atmospheric Sampling Laboratory (GASLAB). Measurements are reported in units of nanomoles of CO per mole of dry air, which we refer to here using the shorthand ppbv. CO was measured using a gas chromatograph with a precision of ±1% over the calibrated range of 20–400 ppbv (Pak et al., 2003).

2.2 HIAPER Pole-to-Pole Observations

CGOP provides multi-year temporal coverage but limited spatial coverage. We supplement this record using observations from the HIPPO aircraft campaign, allowing us to test the representativeness of both airborne data sets. HIPPO consisted of five deployments across different seasons from 2009 to 2011 and took place primarily over the western Pacific. Flights involved repeated vertical profiles from the surface to 8 km with 4–6 flights in the SH during each deployment. CO was measured during HIPPO using five instruments: the Quantum Cascade Laser System (QCLS), the GV AeroLaser VUV CO Sensor, the Unmanned Aircraft Systems Chromatograph for Atmospheric Trace Species (UCATS), the PAN and other Trace Hydrohalocarbon ExpeRiment-Electron Capture Detectors (PANTHER-ECD), and the NOAA Whole Air Sampler – Measurement of Atmospheric Gases that Influence Climate Change (NWAS-MAGICC). From these, a 10 s merged data set for CO based on best available data (CO.X) was constructed (Wofsy et al., 2012). Here we use CO.X from the most recent available revision (R_20121129). We select HIPPO data representative of clean SH extra-tropical air, defined here as all observations over the South Pacific mid-latitudes (20–50° S, 160° E–90° W) except those recorded at low altitudes near cities (mainly Christchurch). Flight dates with available CO data meeting these criteria included 18, 20, 23, 26, 28 January 2009 (HIPPO-1); 7, 9, 11, 14 November 2009 (HIPPO-2); 2, 5, 8, 10 April 2010 (HIPPO-3); 22, 25, 28 June 2011 (HIPPO-4); and 24, 27, 29 August and 1, 3 September 2011 (HIPPO-5). Sampling locations meeting these criteria are shown in Fig. S1 in the Supplement.

2.3 Southern Hemisphere Model Intercomparison Project

We compare the Cape Grim and HIPPO aircraft observations to output from a suite of model runs conducted for
Table 1. Details of model simulations used in SHMIP.a

<table>
<thead>
<tr>
<th>Model type</th>
<th>GEOS-Chemb</th>
<th>NIWA-UKCAt</th>
<th>TMSd</th>
<th>CAM-cheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal resolution (long × lat)</td>
<td>2.5° × 2°</td>
<td>3.75° × 2.5°</td>
<td>3° × 2°</td>
<td>3° × 1.9°</td>
</tr>
<tr>
<td>Vertical levels: total (&lt; 8 km)</td>
<td>47 (26)</td>
<td>60 (20)</td>
<td>34 (16)</td>
<td>56 (27)</td>
</tr>
<tr>
<td>Meteorology</td>
<td>GEOS-5, GEOS-4</td>
<td>Forced by SSTs &amp; sea ice</td>
<td>ERA-interim</td>
<td>MERRA</td>
</tr>
<tr>
<td>Emission injection heights</td>
<td>Surface</td>
<td>Surface</td>
<td>Surface</td>
<td>Surface</td>
</tr>
<tr>
<td>Methane mixing ratios</td>
<td>Prescribed from observations</td>
<td>Prescribed from hemispheric gradient</td>
<td>Simulated</td>
<td>Prescribed from observations</td>
</tr>
<tr>
<td>Global mean tropospheric OHf (molec cm(^{-3}))</td>
<td>10.5 × 10^5</td>
<td>10.6 × 10^5</td>
<td>8.6 × 10^5</td>
<td>9.1 × 10^5</td>
</tr>
<tr>
<td>Cape Grim background regiong</td>
<td>138.75–143.75° E, 41.0–45.0° S</td>
<td>136.875–144.375° E, 41.24–46.25° S</td>
<td>138.0–144.0° E, 42.0–46.0° S</td>
<td>138.75–143.75° E, 41.7–45.5° S</td>
</tr>
</tbody>
</table>

a A full description of the models including references is provided in Zeng et al. (2015).

b GEOS-Chem (www.geos-chem.org) was modified from the standard version 9-01-03 with Caltech isoprene mechanism (Paulot et al., 2009b, b) to include HO\(_2\) uptake by aerosols with \(\gamma = 0.2\) (Mao et al., 2013b), add methanol as an interactive tracer based on the offline simulation of Millet et al. (2008), and use pre-computed biogenic emissions with imposed diurnal variability tied to solar zenith angle.

c NIWA-UKCA comprises a coupled stratosphere–troposphere chemistry scheme (Morgenstern et al., 2013). The background climate model for NIWA-UKCA is HadGEM3-A (Hewitt et al., 2011). The updated version used here includes C\(_2\)H\(_4\), C\(_3\)H\(_8\), CH\(_3\)OH, isoprene, and monoterpene in addition to those described in Morgenstern et al. (2013). The isoprene oxidation is based on Puxbaum et al. (2000) and the monoterpene oxidation is as described in Brasseur et al. (1998).

d The TMS version used here employs the modified CB05 mechanism (Williams et al., 2013) using the configuration outlined in Williams et al. (2014). The isoprene and monoterpene oxidation schemes are taken from Yarwood et al. (2005) subsequently modified according to Archibald et al. (2010).

e CAM-chem is described in Lamarque et al. (2012). Community Earth System Model (CESM)-1.1.1 is used here, with tropospheric (MOZART-4) chemistry.

f Chemical transport model (CTM) or chemistry-climate model (CCM).

g Altitudes are approximated from model pressure levels. The number of levels below 8 km is for the Cape Grim background region, with bounds given below.

h Goddard Earth Observing System (GEOS) fields are from the NASA Global Monitoring and Assimilation Office (GMAO). GEOS-5 was used for the base simulations, and GEOS-4 was used for a 1-year sensitivity study. NIWA-UKCA sea surface temperatures (SSTs) are from the Program for Climate Model Diagnostic and Intercomparison (PCDMI).

i ERA-interim fields are from the European Centre for Medium-Range Weather Forecasts (ECMWF). Modern Era Retrospective-analysis for Research and Applications (MERRA) fields are from the NASA GMAO.

j With the exception of the aircraft source, emissions are generally injected at the surface or in the first few model layers. Aircraft emissions are introduced throughout the troposphere depending on airport location and flight paths. In TM5, isoprene emissions between 20° S and 20° N are introduced into the first two layers of the model to represent canopy height. Also in TM5, fire emissions are distributed over different altitude regimes depending on fire type following Dentener et al. (2006), except in the tropics where injection heights are increased from 1 to 2 km based on recent satellite observations (Labonne et al., 2007). Surface observations from the NASA Global Monitoring Division (GMD) are used to prescribe methane mixing ratios in GEOS-Chem (all altitudes) and CAM-chem (surface only).


l Multi-year mean air density-weighted OH below the climatological tropopause defined as \(p = 300 × 215\cos(lat)\) hPa (Lawrence et al., 2001).

m Cape Grim background region is the region in each model used for comparison with clean air observations from the Cape Grim Overflight Program (CGOP), as described in the text and shown in Fig. 2.
Table 2. Global and southern hemispheric CO budgets (Tg yr$^{-1}$) for 2004 in the SHMIP simulations.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Global</th>
<th>Southern Hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary emission$^a$</td>
<td>1010</td>
<td>306</td>
</tr>
<tr>
<td>Anthropogenic</td>
<td>579</td>
<td>70</td>
</tr>
<tr>
<td>Biomass burning</td>
<td>334</td>
<td>189</td>
</tr>
<tr>
<td>Biogenic</td>
<td>77</td>
<td>35</td>
</tr>
<tr>
<td>Ocean</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>Chemical production$^b$</td>
<td>1340–1920$^c$</td>
<td>578–821$^d$</td>
</tr>
<tr>
<td>Methane$^e$</td>
<td>1050–1070</td>
<td>431–445</td>
</tr>
<tr>
<td>NMVOC$^e$</td>
<td>643–855</td>
<td>303–390</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sinks</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical loss$^b$</td>
<td>2200–2790$^f$</td>
<td>848–1120$^g$</td>
</tr>
<tr>
<td>Dry deposition$^h$</td>
<td>101–115</td>
<td>27–36</td>
</tr>
</tbody>
</table>

$^a$ Primary emissions are from REASv2.1 nested in MACCity (anthropogenic), GFEDv3 (biomass burning), and MEGAN2.1 calculated using CLM (biogenic), as described in the text. Direct ocean emissions of CO are from POET (Granier et al., 2005).

$^b$ Tropospheric chemical production and loss terms are expressed as the range over the SHMIP models. Values for the individual models are given in the footnotes.

$^c$ 1340 (CAM-chem), 1590 (TM5), 1690 (GEOS-Chem), 1920 (NIWA-UKCA).

$^d$ 578 (CAM-chem), 748 (GEOS-Chem), 744 (TM5), 821 (NIWA-UKCA).

$^e$ Production of CO from methane is estimated assuming a 100% yield of CO from methane oxidation. Production from NMVOCs is then estimated as the difference between total production and production from methane. These assumptions are made for diagnostic purposes only, and are not assumed in the chemical mechanisms. These values are only available from GEOS-Chem and NIWA-UKCA, and the value shown is their range.

$^f$ 2200 (CAM-chem), 2520 (TM5), 2770 (GEOS-Chem), 2790 (NIWA-UKCA).

$^g$ 848 (CAM-chem), 1020 (TM5), 1050 (NIWA-UKCA), 1120 (GEOS-Chem).

$^h$ Range from NIWA-UKCA (lower limit) and TM5 (upper limit). Loss via dry deposition was not included in GEOS-Chem and was not archived in CAM-chem.

Zeng et al. (2015) provide a detailed analysis of SH CO distributions simulated by the four SHMIP models as well as the models’ varied abilities to reproduce surface and total column CO observations from selected SH sites. Here, we provide an additional test of the models’ abilities to represent vertical structure in the SH free troposphere (and the associated inter-model differences) using observed vertical profiles representative of SH mid-latitudes clean background air. For comparison with observations from CGOP, which measured only clean background air, we sample each model over the Southern Ocean southwest of Tasmania. We reduce the influence of model spatial variability on the comparisons by averaging each model over four representative grid squares in this region (referred to hereafter as the Cape Grim background region). These grid squares, shown in Fig. 2, were chosen to minimize the influence of outflow from the Australian continent (which we cannot filter directly as only monthly mean model output was archived and radon was not simulated as part of SHMIP). We tested the influence of our choice of sampling region by also performing our analyses using either the grid square containing the CGOP profiles or the nearest ocean-only grid square (as done for TRANSCOM, e.g., Law et al., 2002; Loh et al., 2015). We found that changing the sampling region did not significantly impact the shape of the model profiles or the relative differences between the models, suggesting our results are robust to this choice. Coordinates of the grid squares in each model that define the Cape Grim background region are given in Table 1, with minor differences stemming from model resolution and grid spacing as shown in Fig. 2.

Because of the temporal offset between CGOP (1990s), HIPPO (2009–2011), and the SHMIP simulations (2004–2008), we do not compare individual flights or profiles but instead focus on average behavior seen across multiple years in the observations and models. Multiple studies have shown that trends in SH CO over similar time periods are either small (Zeng et al., 2012; Worden et al., 2013) or insignificant (Warner et al., 2013; Yoon and Pozzer, 2014), depending on the period and region analyzed, especially when El Niño years are neglected. We evaluated long-term CO trends specific to the Cape Grim region over the 1991–2008, 2004–2008, and 1991–2011 time periods relevant to this work using CSIRO flask samples collected in surface air at the Cape Grim Baseline Air Pollution Station. Results of this analysis, shown in Table S1 in the Supplement, indicate that CO trends at Cape Grim over these periods were not statistically
significant on an annual basis or for any individual season, justifying our use of long-term temporal averages.

Figure 3 shows the median observed seasonal cycle of CO at Cape Grim averaged over 0–2, 2–5, and 5–8 km altitude bins (black line). The observations show increasing CO mixing ratios with altitude in all months, as previously reported by Francey et al. (1999) in an analysis of 5 years of the same data set. Peak mixing ratios were observed in austral spring during the tropical BB season. At altitudes below 2 km, the seasonal maximum occurred in October, as seen also in flask samples collected in surface air. This October peak in the boundary layer appears to represent a 1 month offset from higher altitudes, where peak CO was observed in September. However, the September maximum above 2 km is not statistically significant and is skewed by a large number of samples from September 2000 collected as part of the SAFARI aircraft campaign (Pak et al., 2003). As no measurements were made in other months in 2000, the SAFARI data cannot be considered indicative of slow mixing of transported BB plumes into the boundary layer.

The colored lines in Fig. 3 show the simulated seasonal cycles in the Cape Grim background region for the four individual SHMIP models. Despite large differences in absolute mixing ratios (discussed below), the models are generally able to reproduce the shape of the observed seasonal cycle especially above 2 km, as expected from previous studies (e.g., Shindell et al., 2006). In the 5-year mean, the models show peak mixing ratios in September rather than October at all altitudes, but this timing varies from year to year. A particularly strong September peak is simulated by all models.
for 2005, reflecting significantly enhanced BB emissions in South America and southern Africa in the GFEDv3 inventory for this year (and leading to an outsized influence on the 5-year mean). None of the models capture the delay in peak mixing ratios in the boundary layer, suggesting errors in model representation of vertical mixing and/or boundary layer heights, both known issues in atmospheric transport models (e.g., Gerbig et al., 2008; Locatelli et al., 2013). Model ability to match other aspects of the seasonal changes in the relationship between different altitudes is varied and is the subject of further discussion in Sect. 3.

The inter-model and model-observation differences near Cape Grim shown in Figs. 2 and 3 are sizeable, consistent with the detailed analysis of the simulations by Zeng et al. (2015). Annual mean mixing ratios in surface air in this region range from less than 50 ppbv in CAM-chem to nearly 65 ppbv in TM5 (compared to 53 ppbv observed; Fig. 2). GEOS-Chem CO is artificially enhanced as the model does not include a CO sink from dry deposition. A sensitivity test including dry deposition over all vegetated surfaces led to a 1–2 ppbv decrease in GEOS-Chem CO at all altitudes (equivalent to ∼50 Tg yr\(^{-1}\) or 2% of the total global CO sink) but did not substantially change the vertical, horizontal, or seasonal distributions. The TM5 overestimate is consistent with the high bias in surface CO identified previously using monthly mean surface CO measurements at Cape Grim from the year 2000 (Williams et al., 2013). The CO differences between models persist with similar magnitude at all altitudes (Fig. 3). These differences in background CO are influenced by a number of factors including grid resolution, meteorological drivers, and chemical mechanisms as discussed in detail by Zeng et al. (2015). In particular, they find that consistent inter-model differences in the SH CO background are largely driven by differences in CO production efficiency, with an additional contribution from differences in oxidizing capacity (especially for TM5, which has the lowest OH of the four models as shown in Table 1). As our focus here is on relative rather than absolute vertical and seasonal gradients, we remove the influence of consistent differences in the CO background from our comparisons by showing CO mixing ratios expressed as ΔCO, the deviation (in ppbv) from a specific baseline value, as done previously for CGOP data by Francey et al. (1999). We use as baseline value the median mixing ratio in surface air (below 1 km) in a given season, computed separately for each model and each set of observations. Expressing the vertical gradients as deviations rather than absolute values also allows us to compare the CGOP and HIPPO observations, which are on different absolute scales due to different sampling locations and strategies.

3 Observed and simulated vertical gradients

3.1 Cape Grim and HIPPO observations

The median climatological vertical gradients of CO from CGOP are shown as black lines in Fig. 4. For each season, medians were computed after binning observations from all years into 1 km altitude ranges. Observed variability in each 1 km altitude bin was estimated using the median absolute deviation (MAD) statistic for all observations in the bin, shown as the thin horizontal lines. Profiles are expressed as ΔCO, the deviation in ppbv from the observed median value in surface air (0–1 km) for each season, as described in Sect. 2. Observations were grouped seasonally to increase the number of data points used to construct each profile, with seasonal groupings selected based on inspection of seasonal cycles in the data. In particular, observed behavior in June showed more similarity to that in the preceding months than in July–August in terms of both magnitude and interannual variability, especially at higher altitudes (Fig. 3). This reflects variability in the onset of the SH BB season, which typically occurs sometime in July or August (Edwards et al., 2006). We therefore grouped June data with austral autumn (MAMJ) rather than austral winter (JA) and retained austral spring (SON) as a definitive season.

Figure 4 shows moderate seasonal variability in observed CO vertical gradients of a few ppbv km\(^{-1}\), with larger gradients during the winter–spring burning seasons (JA–SON) than during the rest of the year. As reported previously by Pak et al. (1996) using the first 4 years of this data set, the observations show an increase in gradient above 2 km, with the suppressed gradient at lower altitudes indicative of mixing throughout the local boundary layer (Pak et al., 1996). We quantify the observed CO vertical gradients using a linear
Table 3. Observed and simulated 0–8 km CO vertical gradients near Cape Grim, in ppbv km$^{-1}$ for each season.

<table>
<thead>
<tr>
<th></th>
<th>Summer (DJF)</th>
<th>Autumn (MAMJ)</th>
<th>Winter (JA)</th>
<th>Spring (SON)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGOP Observations</td>
<td>1.95 ± 0.36</td>
<td>1.58 ± 0.28</td>
<td>1.90 ± 1.05</td>
<td>2.22 ± 0.47</td>
</tr>
<tr>
<td>GEOS-Chem</td>
<td>1.08</td>
<td>0.65</td>
<td>1.57</td>
<td>2.42</td>
</tr>
<tr>
<td>NIWA-UKCA</td>
<td>1.38</td>
<td>1.50</td>
<td>1.95</td>
<td>2.51</td>
</tr>
<tr>
<td>TM5</td>
<td>2.27</td>
<td>1.77</td>
<td>1.95</td>
<td>3.46</td>
</tr>
<tr>
<td>CAM-chem</td>
<td>1.13</td>
<td>1.12</td>
<td>1.57</td>
<td>2.13</td>
</tr>
</tbody>
</table>

* Vertical gradients were calculated using a linear regression of the median 0–8 km observed and simulated profiles, binned in 1 km altitude bins. Simulated gradients are for the Cape Grim background region (see text). Errors on the observed gradients show the 95% confidence intervals calculated using the bootstrap method with 10,000 random samplings from the original data points. Bold values indicate the simulated vertical gradient is within the 95% confidence interval of the observed slope from CGOP.

fit to the median profiles for each season. Calculated gradients are given in Table 3 and show a minimum in autumn (1.6 ppbv km$^{-1}$) and maximum in spring (2.2 ppbv km$^{-1}$) that are significantly different from one another.

We evaluate the large-scale spatial representativeness of the CGOP data using independent HIPPO observations from the South Pacific. Seasonal profiles are shown in gray in Fig. 4 and were constructed from one HIPPO deployment each, with the exception of MAMJ which includes both HIPPO-3 and HIPPO-4 flights. HIPPO-5 profiles for JA also include data from the two flights in early September to increase the data available in that season and to keep all flights from each deployment together. The figure shows that although the relative variability in CO (thin lines) differs somewhat between HIPPO and CGOP, there is generally overlap in the observed ΔCO from each data set (thick lines). Small differences between the two are likely driven by (1) BB plumes from Africa and South America experiencing more dilution during transport to the Pacific than to Cape Grim, and (2) sampling of Australian BB outflow during HIPPO but not CGOP. Both of these factors should be most influential in austral winter–spring, when burning in the SH is at its peak (also the period when the data sets show the most variability).

The most notable difference between data sets (although still within the observed variability of both campaigns) is seen above 4 km in JA. We examined this difference using regional BB tracers in the four models (tracers are described below) and found that the offset between CGOP and HIPPO in JA is consistent with differences in transport from southern African BB sources to the two different sampling locations. Outflow from Africa is frequently southeastward at this time of year, passing directly over Cape Grim. BB plumes are not well mixed by the time they arrive at Cape Grim, resulting in large and distinct peaks in observed CO anywhere from 4 to 8 km (Francey et al., 1999; Pak et al., 2003). In contrast, simulated transport to the southwest Pacific is both less frequent and less direct in JA, leading to more diffuse BB plumes and lower CO mixing ratios. Simulated CO profiles over the Pacific (not shown) display a broad peak of moderately enhanced CO from 2 to 8 km, consistent in shape with airborne observations of BB-influenced air from PEM-Tropics A (Staudt et al., 2002).

With the exception of the mid-troposphere in JA, the observed CO vertical gradients are very similar between CGOP and HIPPO, despite major differences in flight locations (Southern Ocean vs. Pacific), observation years (1990s vs. late 2000s), and sampling strategies (number of profiles, frequency of flights). This remarkable correspondence lends confidence to our use of vertical gradients derived from the CGOP data as being representative of the remote SH (except perhaps in regions of continental outflow). It also suggests the HIPPO CO observations are representative of long-term seasonal patterns, facilitating future interpretation of these data.

3.2 SHMIP simulations in the Cape Grim background region

Figure 5 and Table 3 compare the observed vertical gradients from CGOP to the SHMIP simulations in the Cape Grim background region. Simulated vertical gradients for each model are derived from monthly mean output (and therefore not specifically selected for baseline conditions). Modeled monthly means for each year in 2004–2008 were averaged over the four grid squares shown in Fig. 2. From these spatial means, a seasonal median model profile was derived by calculating the median model value for each 1 km altitude bin (median over all model levels in the altitude bin and all months/years in the season). As for the observations, the model profiles are expressed as ΔCO, the deviation from the median model value at 0–1 km in each season. Simulated vertical gradients in Table 3 were calculated from a linear fit to the median simulated profiles.

As seen in Fig. 5, the models generally provide a good simulation of 0–8 km CO vertical gradients in austral winter (JA) and spring (SON). With the exception of TM5 below 3 km in SON, simulated gradients are within the large variability of the observations in these seasons. At this time of year, the dominant influence on SH CO is the intense BB that takes place across the tropics and in the SH extra-tropics. Burning
peaks in August–September in southern Africa, September–October in South America, and October–November in Australia (Edwards et al., 2006; Gloudemans et al., 2006). Fire emissions have been shown to influence Australia and the Cape Grim region via long-range transport in the mid-upper troposphere (UT) (Bowman, 2006; Gloudemans et al., 2006; Pak et al., 2003), driving the enhanced gradient above the surface in these months. Simulated tracers of regional influence (CO2s, described in Sect. 4.1) show peak contributions from southern African BB at 4–7 km and from South American BB at 6–10 km.

The ability of the models to capture the observed BB enhancement indicates that the models (all using GFEDv3 emissions) are successfully capturing the long-range transport of BB sources. The main exception is the positive gradient simulated from 7 to 8 km in SON (versus the observed decrease over this altitude range in CGOP). The cause of the discrepancy is unclear. In the models the increase above 7 km reflects a larger contribution from South American than African BB at these altitudes, primarily in October. BB plumes are likely very dispersed at these altitudes following long-range transport, and this dispersion complicates simulation of the gradient. The otherwise good agreement between observed and simulated JA–SON gradients suggests that there has not been significant change in the major SH burning source regions that contribute to background CO in the Cape Grim region since the 1990s (when the observations were collected). This is consistent with a number of recent studies (and with our own analysis of Cape Grim surface flask data, Sect. 2.3) showing observed trends in SH CO are much smaller than interannual variability (Zeng et al., 2012; Wai et al., 2014; Warner et al., 2013; Worden et al., 2013; Yoon and Pozzer, 2014). Significant peaks in BB have been observed for individual years in both periods (in particular, the 1997 and 2006 El Niño years), and these are reflected in the large interannual variability shown for these seasons in Fig. 5 (horizontal lines).

Outside of the burning season, model ability to match observed vertical gradients deteriorates, as does inter-model agreement (Fig. 5). GEOS-Chem and CAM-chem in particular show a sharp drop in gradient from spring to summer/autumn that is unmatched by the observations; the change in NIWA-UKCA and TM5 is more gradual but still too large (Table 3). Across all models, the overall decrease in the vertical gradient from spring to autumn is between 1 and 2 ppbv km$^{-1}$, larger than the observed change from CGOP of $\sim 0.5$ ppbv km$^{-1}$. In the following section, we evaluate possible reasons for the model-CGOP and inter-model discrepancies in the summer–autumn CO vertical gradients.

4 Drivers of inter-model variability

Model-observation differences can result from model errors in emissions, chemistry, meteorology/transport, or a mix of these. As all SHMIP models used identical emissions (except for parameterized lightning, soil, and volcanic sources with limited impact on CO), the inter-model differences seen here should result primarily from differences in chemistry and meteorology/transport (resolution may also play a small role). Here, we investigate the role of model differences in transport and chemistry on differences in simulated vertical gradients using a set of sensitivity simulations, run for a 2-year period (2004–2005) to reduce the influence of interannual variability on the results. Figure 6a shows the simulated CO vertical gradients from the standard simulations in 2004–2005 as a point of reference for the sensitivity simulations. As seen in the figure, simulated profiles during the 2004–2005 test period are generally similar to those for the full SHMIP period (Fig. 5).
4.1 Transport

The first sensitivity simulation uses an idealized CO-like tracer \( \text{CO}_{25} \) designed to quantify the impact of model transport, independent of the influence of model chemistry. The \( \text{CO}_{25} \) tracer used the same emissions as CO globally with a fixed 25-day lifetime and was not subject to chemical production or chemical loss. In remote regions, the \( \text{CO}_{25} \) mixing ratio therefore represents the balance between primary emission and long-range transport, with differences between models caused exclusively by differences in transport over the 25-day tracer lifetime. Vertical gradients of the \( \text{CO}_{25} \) tracer are shown in Fig. 6b. In DJF and MAMJ, all models display a greatly diminished ability to match observed gradients when chemistry is neglected, indicating transported primary emissions play only a small role in driving CO vertical gradients at this time of year. In winter–spring, \( \text{CO}_{25} \) vertical gradients are only slightly shallower than those of total CO and are within the observed interannual variability, consistent with the gradients being driven by primary BB emissions that are well represented in all four models.

The differences in \( \text{CO}_{25} \) between models are much smaller than differences in total CO, especially in summer–autumn. This is consistent with results from Zeng et al. (2015), who examine \( \text{CO}_{25} \) columns over the entire SH and find both the magnitude and distribution to be similar across the four models. They also show that the small inter-model differences in \( \text{CO}_{25} \) columns are not reflected in the distributions of total CO columns, indicating a limited role for horizontal transport differences as a source of inter-model variability, at least over the 25-day lifetime of the tracer. As seen in Fig. 6b, all four models show similarly shallow \( \text{CO}_{25} \) vertical gradients in DJF and MAMJ. These similarities could reflect similar transport of primary emissions to the Cape Grim region and/or similarly rapid vertical mixing relative to the 25-day tracer lifetime, which would obscure the role of transport differences in driving inter-model variability. Given the lack of primary CO sources near the Cape Grim region, the latter is unlikely to have a major impact on primary CO gradients but may be important for inter-model differences in secondary CO. We explore this effect further in Sect. 5.

We further investigate the impacts of inter-model transport differences using regional \( \text{CO}_{25} \) tracers. Figure 7 shows the contribution from six regions (Australia, South America, southern Africa, Southeast Asia, East Asia, and all other sources) to total \( \text{CO}_{25} \) at Cape Grim in three altitude ranges. Total \( \text{CO}_{25} \) amounts are highest in GEOS-Chem, indicating more rapid transport to this region than in the other models, and are typically lowest in NIWA-UKCA. In summer, the relative contributions of different sources are largely consistent across the models, with a slight dominance from Australia below 2 km and a slight dominance from South America above. The other contribution shown in gray in Fig. 7 represents the difference between the global \( \text{CO}_{25} \) tracer and the sum of the regional \( \text{CO}_{25} \) tracers and mainly reflects the contribution from northern Africa. Inter-model differences are larger for this contribution (other) than for any of the regional \( \text{CO}_{25} \) tracers, with in particular NIWA-UKCA showing less influence than the other models. This contribution peaks in austral summer, likely driven by the seasonal source from NH African burning, which is at its annual maximum in DJF (Roberts et al., 2009). In summer–autumn, differences between models are largely constant with altitude and result in very similar vertical gradients, as seen previously in Fig. 6b.

During the tropical BB seasons (JA–SON), inter-model differences in the \( \text{CO}_{25} \) sources at Cape Grim are larger, as shown in Fig. 7. In JA, the contribution from southern Africa is dominant and also varies most, responsible for 50–60 % of total \( \text{CO}_{25} \) in NIWA-UKCA compared to only 20–30 % in CAM-chem, with the other models falling between these values. Absolute differences in this source of up to 7 ppbv at 8 km can explain much of the difference in the JA \( \text{CO}_{25} \) gradient shown in Fig. 6b, suggesting that long-range transport of African BB emissions contributes to inter-model variability during the early BB season. In SON, the South American contribution dominates, reflecting a 1-month offset in peak emissions from these regions in 2004–2005 in the GFEDv3 inventory. The southern African contribution is also more consistent across models in SON, with inter-model differences of similar magnitude to those from the South American source (2–3 ppbv).

4.2 Chemical loss

Fixed-lifetime tracers do not account for model differences in chemistry, which for CO include differences in both OH-driven loss and secondary chemical production. We isolate the impact of the former using a second set of idealized CO-like tracers (\( \text{CO}_{25} \)). In this case, the \( \text{CO}_{25} \) tracers again have the same primary emissions as CO but with tracer loss driven by each model’s OH fields and CO+OH rate constant. Differences in the rate constant at standard temperature and pressure are on the order of 10 % (e.g., between the IUPAC recommendation used in NIWA-UKCA and the JPL recommendation used in GEOS-Chem). Differences in OH mixing ratio are on the order of 5–20 % for the global tropospheric mean (Table 1) but can be much larger regionally. Like \( \text{CO}_{25} \), \( \text{CO}_{25} \) tracers are subject to differences in model transport, with differences between \( \text{CO}_{25} \) and \( \text{CO}_{25} \) indicative of the impacts of OH-driven chemical loss. The \( \text{CO}_{25} \) lifetime varies spatially and seasonally (due to OH variability), and in winter–spring can be significantly longer than 25 days. As described by Zeng et al. (2015), the \( \text{CO}_{25} \) mixing ratios therefore provide a more realistic metric than \( \text{CO}_{25} \) for evaluating the combined impacts of transport and loss of primary CO.

Figure 6c shows the vertical gradients of the global \( \text{CO}_{25} \) tracer in the Cape Grim region as simulated by GEOS-Chem, NIWA-UKCA, and TM5 (CAM-chem did not include a global \( \text{CO}_{25} \) tracer). Both the relative vertical gradients...
Figure 6. Median CO profiles from CGOP observations (black) compared to model simulations for 2004–2005 using (a) the standard simulation; (b) a global CO-like tracer with a 25-day lifetime (CO$_{25}$; see text); (c) a global CO-like tracer with OH-driven loss but no secondary production (CO$_{OH}$; see text); and (d) LPJ-GUESS isoprene and monoterpene emissions. Solid colored lines represent the standard simulations and dashed lines the sensitivity simulations for GEOS-Chem (red), INASA-UKCA (orange), TM5 (purple), and CAM-chem (blue; no OH-loss tracer). Profiles are shown as ΔCO, the deviation (in ppbv) from the observed or modeled value in surface air in each season, with the surface value calculated separately for each sensitivity test.

and the regional contributions are generally similar between the two idealized tracers (regional contributions are shown in Fig. S2 in the Supplement). In DJF, tropospheric OH production leads to a small decrease in mid-tropospheric CO$_{OH}$ relative to surface values in all three models. As for CO$_{25}$, CO$_{OH}$ gradients in DJF–MAMJ are greatly reduced relative to those of total CO (Fig. 6a), suggesting both transport and chemical loss of primary emissions are insufficient in these seasons to explain the large inter-model variability, which instead must be driven by secondary CO production.

4.3 Chemical production

The difference between CO$_{OH}$ and total CO for each model represents the contribution from in situ chemical production, estimated to account for roughly half of the total CO source.
Figure 7. Contributions to CO$_{25}$ in the Cape Grim background region from sources in Australia (111–156° E, 11–44° S), South America (34–84° W, 57° S–1° N), southern Africa (9–44° E, 5–37° S), Southeast Asia (94–156° E, 11° S–7° N), East Asia (91–144° E, 7–44° N), and elsewhere, as simulated by GEOS-Chem (G.C.), NIWA-UKCA (N.U.), CAM-chem (C.C.), and TM5 for different altitude bands and seasons in 2004–2005. Note the difference in scale between DJF–MAMJ and JA–SON.

Figure 8. Five-year mean DJF mixing ratios in near-surface air (< 1 km) of (a) CO, (b) isoprene, (c) CH$_2$O, (d) OH, and (e) HO$_2$ as simulated by GEOS-Chem (left) and NIWA-UKCA (right). (f) The bottom row shows the difference between rates of CO chemical production (P$_{CO}$) and CO chemical loss (L$_{CO}$) in molecules cm$^{-3}$ s$^{-1}$ for DJF 2004 only. Regions where production outweighs loss are shown in red and the inverse in blue.
globally (Jiang et al., 2011; Kopacz et al., 2010) and an even larger proportion in the SH (Pfister et al., 2008). Comparing Fig. 6a and c show that chemical production plays a dominant role in controlling the simulated CO vertical gradient in DJF and MAMJ but has much less influence during the tropical BB seasons when primary emissions dominate. Chemical production also appears to be the major source of inter-model variability in DJF and MAMJ, and uncertainties in this term may help explain the large underestimates of the observed summer gradient seen in particular by GEOS-Chem and CAM-chem (Fig. 5).

Chemical production of CO originates from oxidation of both methane and NMVOCs, and inter-model variability in the vertical gradients may reflect contributions from both. In remote regions, the methane source dominates the CO burden while the NMVOC source dominates the variability (Pfister et al., 2008). Differences in the methane mixing ratios in the four models (Table 1) are thus more likely to affect overall concentration differences (e.g., Fig. 3) than differences in the vertical gradient. However, the methane contribution cannot be quantified from the archived SHMIP output. Instead, we perform a final sensitivity test to evaluate the role of the NMVOC source in driving the simulated CO vertical gradients. Figure 6d shows the result of replacing MEGAN-CLM biogenic emissions with LPJ-GUESS for isoprene and monoterpenes. Methane, OH, and other emissions remain unchanged from the standard simulation. Since emissions are the same across models, they cannot explain inter-model variability; however, they can help attribute sources of model-observation bias as well as provide insight into the dependence of the simulated vertical gradients on biogenic NMVOC sources. The figure shows that relative to the standard simulation, the LPJ-GUESS emissions reduce the simulated CO vertical gradient in summer–autumn in all models. In winter–spring, the differences are negligible. The small increases in gradient from Fig. 6c to d reflect both methane and NMVOC contributions (which are smaller but still significant in LPJ-GUESS). These results present a picture consistent with the previous sensitivity tests; namely, that observed vertical gradients are driven in winter–spring by primary BB emissions and in summer–autumn by secondary CO, largely of biogenic NMVOC origin.

Biogenic source regions are located far upwind of Cape Grim, so model error in the Cape Grim background region can result from errors in both model chemistry and the transport of secondary CO. Distinguishing between these factors is not straightforward. Using GEOS-Chem, we performed an additional 1-year sensitivity test for 2004 designed to partially discriminate between these terms by replacing the standard GEOS-5 meteorology with GEOS-4. The latter has been shown to have more rapid vertical uplift over tropical source regions (Liu et al., 2013, 2010), where biogenic emissions are also large (Guenther et al., 2006). The same chemical mechanism was used in both simulations, and the CO+OH reaction rate changed by less than 2% from differences in temperature and pressure, so simulated differences at Cape Grim can be attributed to model transport. Results from this sensitivity simulation (not shown) indicated virtually no impact on the CO vertical gradient in summer–autumn, implying a dominant influence from the chemistry controlling secondary CO production.

5 Chemistry and transport of biogenic-sourced secondary CO

In preceding sections, we have shown that inter-model differences in the vertical distribution of CO in the remote SH are largest in austral summer–autumn, and that these differences cannot be explained by the transport or chemical loss of primary emitted CO; instead, they are clearly driven by differences in CO produced chemically from biogenic NMVOC emissions. Here we evaluate model differences in the chemistry and transport of secondary CO from biogenic source regions in the context of their impacts on SH background CO in summer (DJF), when inter-model variability is largest. We focus our analysis in this section on GEOS-Chem and NIWA-UKCA, the two models that best reproduce absolute CO mixing ratios at Cape Grim (Fig. 3) but with significant differences in the simulated vertical gradient (Fig. 5).

Chemical mechanisms differ substantially across the SHMIP models (Zeng et al., 2015), and differences are difficult to interpret due to varying levels of complexity, especially for NMVOC speciation and oxidation. Of particular importance here are differences in the oxidation of isoprene, summarized for all models in Table 4, and monoterpenes. The GEOS-Chem SHMIP simulations use the Caltech isoprene mechanism as implemented in v9-01-03 (http://wiki.seas.harvard.edu/geos-chem/index.php/New_isoprene_scheme_prelim), which includes formation of first and second generation isoprene nitrates under high-NOx conditions (Paulot et al., 2009a) and formation of isoprene hydroperoxides and subsequently epoxidiols under low-NOx conditions (Paulot et al., 2009b). Isoprene oxidation in NIWA-UKCA is from the original Mainz Isoprene Mechanism (MIM; Pöschl et al., 2000) but with updated rate coefficients for reactions between OH and isoprene nitrates and between NO and isoprene peroxyl radicals from Paulot et al. (2009a, b); still, the NIWA-UKCA mechanism contains a limited number of species and predates many recent advances in isoprene chemistry available in newer mechanisms like the Caltech scheme or MIM2 (Taraborrelli et al., 2009).

Monoterpenes oxidation is not included explicitly in GEOS-Chem v9-01-03 as used here; instead, monoterpene emissions produce CO with an assumed 20% molar yield (Duncan et al., 2007). NIWA-UKCA includes simple monoterpene oxidation reactions based on Brasseur et al. (1998). Oxidation products of isoprene and monoterpenes are similar,
and we do not distinguish between these two sources in either model.

Figure 8 shows mean summertime mixing ratios of CO and key related species (isoprene, formaldehyde, OH, and HO$_2$) in near-surface air (<1 km) as simulated by GEOS-Chem and NIWA-UKCA. Similar maps for TM5 and CAM-chem can be found in Fig. S5 in the Supplement. At the surface, CO hotspots across the tropics show similar magnitudes in GEOS-Chem and NIWA-UKCA, especially in Africa and Southeast Asia where primary emissions dominate (Fig. 8a). Surface isoprene – indicative of biogenic source regions – is also similar across models (Fig. 8b), with maximum values of more than 10 ppbv over South America. Comparison to observations from the October 2005 GABRIEL campaign over the northeast Amazon shows a 40–70 % high isoprene bias in the boundary layer (modeled means of 2.9–3.4 ppbv in the 3–6°N, 50–60°W flight region vs. observed mean of 2.00±0.76 ppbv from Lelieveld et al., 2008). In the free troposphere, mean simulated isoprene ranges from 0.04 to 0.2 ppbv across models, generally within the variability of the GABRIEL observations (0.07±0.12 ppbv; Lelieveld et al., 2008). The inter-model consistency of the surface overestimate points to a high bias in the MEGAN-CLM emissions, which are common to all SHMIP models.

The models show large discrepancies in surface distributions of formaldehyde (CH$_2$O), with much higher surface CH$_2$O in NIWA-UKCA than GEOS-Chem (Fig. 8c). In non-urban continental boundary layers, the dominant source of CH$_2$O is atmospheric oxidation of NMVOCs, and in particular isoprene (Palmer et al., 2003). The higher mixing ratios simulated by NIWA-UKCA are therefore indicative of more rapid chemical processing following isoprene oxidation. As inter-model differences are small for isoprene mixing ratios (Fig. 8b), OH mixing ratios (Fig. 8d), and the rate of the initial isoprene+OH oxidation reaction (within ~1 % at standard temperature and pressure), the differences in surface CH$_2$O shown in Fig. 8 are likely driven by the chemistry (including photolysis) of second and later generation isoprene oxidation products. CH$_2$O oxidation provides a source of CO over short timescales, and the faster production of CH$_2$O therefore also results in more rapid production of CO in NIWA-UKCA. This is seen in Fig. 8f, which shows that the net balance between CO chemical production (P$_{CO}$) and CO chemical loss (L$_{CO}$) is more strongly weighted towards production in NIWA-UKCA, leading to slight enhancements in boundary layer CO over biogenic source regions (e.g., South America, Fig. 8a). While differences in CO loss rates are likely partially responsible, we expect that CO production contributes more to the P$_{CO}$−L$_{CO}$ differences given the similarity of surface OH between models, particularly over South America where all models show OH titration (Fig. 8d). The near-source surface differences between the two models are consistent with the whole troposphere budgets for the SH given in Table 2, which show total CO production is about 10 % higher in NIWA-UKCA than GEOS-Chem, while total loss is about 5 % lower.

The implications of these chemistry differences for the broader vertical and horizontal distributions of CO depend on subsequent transport and chemical processing. Figure 9 shows mean summertime longitude–altitude cross sections (averaged over 15–45°S) for isoprene, OH, CH$_2$O, and CO (see Fig. S4 for TM5 and CAM-chem). The isoprene cross sections (Fig. 9a) show key differences in vertical transport between models. Relative to GEOS-Chem, NIWA-UKCA shows less deep convective injection of isoprene to the UT over Africa and Australia but more over South America, where isoprene mixing ratios are at their maximum. As a result, NIWA-UKCA displays an enhancement of isoprene mixing ratios at roughly 12 km over South America while isoprene is largely depleted at these altitudes in GEOS-Chem. The effects of the enhanced isoprene uplift in NIWA-UKCA are compounded by lower OH in the UT in this region (Fig. 9b). The net result for both CH$_2$O (Fig. 9c) and CO (Fig. 9d) is more UT production, less UT destruction, and therefore higher UT mixing ratios in NIWA-UKCA than GEOS-Chem. Subsequent zonal transport distributes this additional CO across the SH mid-latitudes UT. Because isoprene emissions are much higher in South America than other SH source regions (Fig. 8), the differences in vertical transport over Africa and Australia play a much more minor role in defining SH UT CO distributions.

The mean location and vertical extent of the profiles from CGOP are shown as the blue lines in Fig. 9d. The figure shows that the inter-model differences in CO vertical gradient seen in Fig. 5 are consistent with the combined effects of differences in chemistry and transport. Slower near-source oxidation of isoprene products in GEOS-Chem leads
to a horizontal smearing effect in the lower mid-troposphere, resulting in relatively more CH$_2$O and CO (largely of Australian biogenic origin) reaching Cape Grim below $\sim$3 km in GEOS-Chem compared to NIWA-UKCA. Meanwhile, NIWA-UKCA’s rapid isoprene uplift and subsequent CO production and UT transport combined with reduced UT loss result in relatively more CO (largely of South American biogenic origin) reaching Cape Grim above $\sim$6 km in NIWA-UKCA than in GEOS-Chem. Combined, these two factors drive a stronger vertical gradient in NIWA-UKCA in the Cape Grim region. Impacts are similar over the western Pacific region sampled by HIPPO.

In austral autumn (MAMJ), inter-model differences in surface mixing ratios and vertical uplift are similar to those shown in Figs. 8 and 9 for austral summer. We have shown previously that biogenic-derived secondary sources continue to drive simulated CO gradients in this season (Fig. 6). Combined, these results suggest that the inter-model variability in autumn is caused by the same differences in model chemistry and transport as seen for summer.

6 Summary and conclusions

We have used a 9-year data set of monthly airborne observations of CO from the Cape Grim Overflight Program (CGOP) to evaluate CO distributions in the remote southern hemispheric free troposphere as simulated by four global 3-D atmospheric chemistry models using identical emissions. Observations above the surface in this region are rare and are typically limited to a single year and/or season, so interpretation of the Cape Grim data provides a unique picture of climatological CO seasonal cycles and vertical gradients in the remote SH. Our analysis focused on the models’ relative abilities to reproduce observed vertical gradients of CO from the surface to 8 km in different seasons. Through model sensitivity analysis and comparison of simulated spatial distributions, we evaluated the importance of primary vs. secondary sources on CO vertical gradients and diagnosed the causes of inter-model divergence.

Observations from both CGOP near Tasmania (1991–2000) and the recent HIPPO campaigns over the SH Pacific (2009–2011) showed similar seasonality, with larger gradients during the austral winter–spring burning seasons (JA–SON) than during the rest of the year. The close correspondence between these two data sets despite differences in location, time period, and sampling strategies suggests the processes driving observed vertical gradients are coherent across much of the remote SH and have not changed significantly over the past 2 decades. The consistency between the two data sets further suggests that quantitative metrics derived from the CGOP observations can be used to diagnose model performance, both for the SHMIP models used here and more generally for future revisions of these and other models. Tables 5 and 6 provide tabulated observation-based metrics for the two salient features of the CGOP data: the seasonal cycle at different altitudes (represented in Table 5 by a harmonic fit), and the vertical profile in different seasons (represented in Table 6 by a polynomial fit). Tables S2 and S3 in the Supplement provide the equivalent parameters for the SHMIP models as a baseline against which to test future improvements to these models. The fitting methodologies are described in detail in the Supplement and can be easily applied to any atmospheric chemistry model for quick-look diagnosis of the ability to represent the SH free tropospheric CO background.

The four SHMIP models (GEOS-Chem, NIWA-UKCA, TM5, and CAM-chem) were all able to reproduce observed vertical gradients during winter–spring, but observed gradients were underestimated in austral summer (DJF) and autumn (MAMJ) by GEOS-Chem and CAM-chem. All models overestimated the seasonal cycle of the vertical gradient to some degree. Sensitivity analysis showed that transport of primary BB CO is the main driver of the observed gradients in winter–spring, when models and observations agree. Regional tracers with CO-like primary emissions and either fixed (CO$_{25}$) or OH-driven (CO$_{OH}$) lifetimes suggested

**Figure 9.** Five-year mean DJF longitude–altitude cross sections averaged over 15–45° S of (a) isoprene, (b) OH, (c) CH$_2$O, and (d) CO as simulated by GEOS-Chem (left) and NIWA-UKCA (right). Numbers in (a) correspond to locations of continental source regions: 1 = South America, 2 = Africa, 3 = Australia. The blue lines in (d) show the location and vertical extent of the CGOP aircraft profiles.
a dominant influence in winter–spring from southern African BB in JA and South American BB in SON, with the seasonal offset due to the timing of peak emissions from these two regions. Inter-model variability was relatively small in both seasons and could generally be attributed to variability in the influence of the southern African source. In summer–autumn, model ability to match observed gradients was significantly diminished when secondary CO sources were not included. Inter-model differences in both CO$_{25}$ and CO$_{OH}$ tracers were much smaller than differences in total CO during non-BB seasons, suggesting that neither transport nor loss of primary CO are sufficient to explain inter-model variability at this time of year. Instead, simulated gradients and inter-model variability in these gradients are driven by secondary CO of biogenic origin, implying a strong sensitivity of tropospheric composition in the remote SH to long-range transport of biogenic emissions and their oxidation products.

We compared simulated austral summer (DJF) horizontal and vertical distributions of CO and related species between NIWA-UKCA and GEOS-Chem (the models with the most realistic CO mixing ratios at Cape Grim) and found significant differences driven by chemical processing and vertical transport. While OH-driven oxidation of isoprene is similar between the models, the ensuing chemistry of isoprene oxidation products appears to proceed faster in NIWA-UKCA than in GEOS-Chem, leading to more rapid production of formaldehyde and CO. The slower chemistry in GEOS-Chem leads to a smearing effect, with CO produced further downwind from source regions, and this effect is particularly pronounced in the lower mid-troposphere near biogenic sources. Inter-model chemistry differences are compounded by differences in vertical transport. More rapid uplift over South America in NIWA-UKCA leads to a secondary isoprene maximum at roughly 12 km that is not seen in GEOS-Chem. Subsequent oxidation produces additional CO in the UT near biogenic source regions, and zonal transport distributes this CO across the SH mid-latitudes. The net effect of the differences in chemistry and vertical transport is less CO at the surface and more at altitude in NIWA-UKCA than GEOS-Chem, resulting in a stronger gradient that is more consistent with CGOP observations.

It is important to note that the simulated summer–autumn CO vertical gradients shown in Fig. 5 reflect the convolved effects of biogenic emissions, model chemistry, and model transport, and the ability to match the observed gradients cannot unambiguously test whether any of these are correct (e.g., the emissions sensitivity test in Fig. 6d). NIWA-UKCA’s superior ability to match the observed DJF gradient relative to GEOS-Chem or CAM-chem is achieved despite the fact that its isoprene oxidation scheme (MIM with some updates) is relatively simple and has known deficiencies (Butler et al., 2008; Taraborrelli et al., 2009). Many recent advances in
Understanding the clean background atmosphere is essential for accurately attributing the impacts of ongoing anthropogenic and natural global change. With relatively few primary source locations, the remote SH serves as a large-scale test bed for quantifying background processes. Although much of the previous work on SH atmospheric composition has focused on the impacts of tropical burning, we have shown here that the non-BB seasons (austral summer and autumn) provide a more nuanced and critical test of the chemistry of the background atmosphere. We have also shown that the vertical gradient of CO is a particularly sensitive test of this chemistry as it is driven by chemical production in summer and autumn. Regular measurements of CO vertical profiles in the remote SH, such as those conducted during the 1990s under the Cape Grim Overflight Program, would thus provide an extremely valuable data set for probing the state of the background atmosphere and its response to ongoing change. Current models display varying degrees of fidelity in reproducing observed CO gradients in a way that is consistent with a state-of-the-science understanding of isoprene chemistry, and increasing the complexity of the chemical mechanisms does not necessarily improve simulation of CO gradients. Disentangling the impacts on model biases of uncertainties in emissions from those in chemistry and transport will necessitate broader in situ sampling during non-burning seasons of multiple species with different chemical lifetimes (including CO, NMVOCs, and HOX), at altitudes throughout the tropospheric column, and in a range of SH environments including near-source, direct outflow, and remote downwind regions.
Appendix A: Abbreviated Campaign and Model Names

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGOP</td>
<td>Cape Grim Overflight Program</td>
</tr>
<tr>
<td>HIAPER</td>
<td>High-performance Instrumented Airborne Platform for Environmental Research</td>
</tr>
<tr>
<td>HIPPO</td>
<td>Pole-to-Pole Observations</td>
</tr>
<tr>
<td>SHMIP</td>
<td>Southern Hemisphere Model Intercomparison Project</td>
</tr>
<tr>
<td>MOPITT</td>
<td>Measurements of Pollution in the Troposphere</td>
</tr>
<tr>
<td>BARCA</td>
<td>Regional Carbon Balance in Amazonia</td>
</tr>
<tr>
<td>GABRIEL</td>
<td>Guyanas Atmosphere-Biosphere exchange and Radicals Intensive Experiment with the Learjet</td>
</tr>
<tr>
<td>SAFARI</td>
<td>Southern African Regional Science Initiative</td>
</tr>
<tr>
<td>ACTIVE</td>
<td>Aerosol and Chemical Transport In tropical conVEction</td>
</tr>
<tr>
<td>SCOUT</td>
<td>Stratospheric-Climate Links with Emphasis on the Upper Troposphere and Lower Stratosphere</td>
</tr>
<tr>
<td>TRACE-A</td>
<td>Transport and Atmospheric Chemistry in the Atlantic</td>
</tr>
<tr>
<td>IAGOS</td>
<td>In-service Aircraft for a Global Observing System</td>
</tr>
<tr>
<td>MOZAIIC</td>
<td>Measurement of Ozone and Water Vapor by Airbus In-Service Aircraft</td>
</tr>
<tr>
<td>PEM-Tropics</td>
<td>Pacific Exploratory Mission – Tropics</td>
</tr>
<tr>
<td>REAS</td>
<td>Regional Emission inventory in ASia</td>
</tr>
<tr>
<td>GFED</td>
<td>Global Fire Emissions Database</td>
</tr>
<tr>
<td>MACC</td>
<td>Monitoring Atmospheric Composition and Climate</td>
</tr>
<tr>
<td>MACCity</td>
<td>MACC-CityZen</td>
</tr>
<tr>
<td>MEGAN</td>
<td>Model of Emissions of Gases and Aerosols from Nature</td>
</tr>
<tr>
<td>CLM</td>
<td>Community Land Model</td>
</tr>
<tr>
<td>LPJ-GUESS</td>
<td>Lund-Potsdam-Jena General Ecosystem Simulator</td>
</tr>
<tr>
<td>TRANSCOM</td>
<td>Atmospheric Tracer Transport Model Intercomparison Project</td>
</tr>
<tr>
<td>IUPAC</td>
<td>International Union of Pure and Applied Chemistry</td>
</tr>
<tr>
<td>JPL</td>
<td>Jet Propulsion Laboratory</td>
</tr>
<tr>
<td>MIM</td>
<td>Mainz Isoprene Mechanism</td>
</tr>
</tbody>
</table>
Acknowledgements. This work was funded by a University of Wollongong Vice Chancellor’s Postdoctoral Fellowship to J. A. Fisher, with the assistance of resources provided at the NCI National Facility systems at the Australian National University through the National Computational Merit Allocation Scheme supported by the Australian Government. We gratefully acknowledge the staff of CSIRO GASLAB involved in the collection and processing of these data and the Australian Bureau of Meteorology/Cape Grim Baseline Air Pollution Station for funding the Cape Grim Overflight Program for 9 years. We also thank the entire HIPPO team and in particular Steve Wofsy, Rodrigo Jimenez, Bruce Daube, Eric Kort, Jasna Pittman, Greg Santoni, and Teresa Campos for providing CO measurements during the campaign. G. Zeng acknowledges NeSI high-performance computing facilities for NIWA-UKCA simulations and UKMO for using the UM. NZ’s national facilities are provided by the NZ eScience Infrastructure and funded jointly by NeSI’s collaborator institutions and through the MBIE’s Research Infrastructure programme. The National Center for Atmospheric Research is operated by the University Corporation for Atmospheric Research with funding from the National Science Foundation. We thank Jingqiu Mao for help in implementing the passive CO-like tracers in GEOS-Chem; Dagmar Kubistin and Clare Paton-Walsh for helpful discussions; and D. Parrish, B. Pak, and an anonymous reviewer for insightful comments that improved the manuscript.

Edited by: A. Pozzer

References


