Supplement of

Evaluating the contribution of the unexplored photochemistry of aldehydes on the tropospheric levels of molecular hydrogen (H₂)

Maria Paula Pérez-Peña et al.

Correspondence to: Jenny A. Fisher (jennyf@uow.edu.au) and Maria Paula Pérez-Peña (m.perez pena@unsw.edu.au)

The copyright of individual parts of the supplement might differ from the article licence.
1 Mixing ratios results from the box modelling

The mixing ratios in the baseline simulations provide context for our box modelling results. Select modelled species from the baseline simulations for London, Cape Verde and Borneo are shown in Figures S1-S3 respectively. The baseline mixing ratio of H$_2$ increased continuously from its initial value of 530 ppbv to 639.97 ppbv at London and 533.96 ppbv at Borneo. At Cape Verde, the mixing ratio of H$_2$ decreased during the simulated period to 517.23 ppbv. These changes are equivalent to an increase in H$_2$ at London and 533.96 ppbv at Borneo. At Cape Verde, the mixing ratio of H$_2$ decreased during the simulated period to 517.23 ppbv. These changes are equivalent to an increase in H$_2$ at London and 533.96 ppbv at Borneo. For the London and Borneo simulations, the H$_2$ increase over time in the baseline run was effectively caused by the photolysis of formaldehyde and glyoxal (the only H$_2$ sources in this simulation). The decrease in modelled H$_2$ at Cape Verde was a result of imbalances in the chemical sources and sinks in this regime. For the Cape Verde simulation, neither formaldehyde nor glyoxal had available measurement constraints (see Table Table 1 in the manuscript); however, both were produced chemically, with formaldehyde production from the degradation of methane, ethene, propene, toluene and benzene and glyoxal production from ethene and toluene precursors (Stavrakou et al., 2009). These five precursor species were all constrained in the Cape Verde simulations (see Table Table 1, footnote b). Figure S2 shows the time series of selected species modelled for Cape Verde, including formaldehyde, which had an average modelled value of ~800 pptv (~2×10$^{19}$ molecules cm$^{-3}$). For comparison, Whalley et al., 2010 reported an average noon value of 328 pptv for their MCM simulations at Cape Verde during May-June 2007 (note that they do not report values for glyoxal). Considering that our modelled formaldehyde mixing ratios were higher than those reported previously by Whalley et al., 2010, and because of the formaldehyde and glyoxal lifetimes of a few hours, we conclude that our simulations included sufficient precursor concentrations and therefore that the decrease in H$_2$ in the baseline Cape Verde simulation implies that the available H$_2$ was consumed by OH more rapidly than it could be produced by formaldehyde and glyoxal, yielding an effective loss over the 12 days modelled.

2 Dry deposition velocity calculation in Yashiro et al., 2011

The dry deposition velocity on the surface of the inactive layer was calculated by Yashiro et al., 2011 as:

$$ V_d = \frac{1}{\frac{\delta}{D_s \theta_a} + \sqrt{\frac{\delta}{D_s \theta_a}}} $$

(1)

where $\theta_a$ is volume of gas per unit volume of soil (air filling porosity) \([m^3 \ m^{-3}]\), $D_s$ is the diffusivity in soil, $k$ is the biological uptake rate \([s^{-1}]\) dependent on soil temperature and soil moisture (see Yashiro et al., 2011 equations 13 to 16). For the diffusivity

$$ D_s = D_a \frac{\theta_a^{3.1}}{\theta_{sat}^{2}} $$

(2)

where $D_a$ is the molecular diffusion coefficient with a value for H$_2$ of 0.611 cm$^2$ s$^{-1}$ and $\theta_{sat}$ is the maximum aerial or liquid water volume per unit volume of soil (total porosity).

The parameterisation to derive the dry deposition velocity for H$_2$ used by Yashiro et al., 2011 implements the same variables as that used by Ehhalt and Rohrer, 2013. The Ehhalt and Rohrer, 2013 parameterisation, applied recently by Paulot et al., 2021, differs from the one by Yashiro et al., 2011 in that the latter considers the diffusivity in the soil to be uniform from the soil surface to a sufficient depth (because the diffusivity is within the first layer of the parent land model), while Ehhalt and Rohrer, 2013 use two different soil diffusivities. Further, the biological activity (uptake rate $k$), soil moisture and soil temperature dependencies are also different between Yashiro et al., 2011 and Ehhalt and Rohrer, 2013. Yashiro et al., 2011 follow the variation of the biological activity from Smith-Downey et al., 2006, while Ehhalt and Rohrer, 2013 rely on the dependencies from the reanalysis performed by Ehhalt and Rohrer, 2011. Also, the thickness of the inactive layer used by Yashiro et al., 2011 is considered to be uniform (with a value of 0.7 cm) while Ehhalt and Rohrer, 2013 provide values that are a function of the average volumetric soil water content. Even though both models start from equation 3 (where $F_s$ is the flux and $C$ is the mass concentration of H$_2$) to derive the dry deposition velocity as a function of the H$_2$ flux and diffusivities, the two models differ when the thickness of the inactive layer $\delta = 0$ as a consequence of the difference in the definition of the flux.

$$ V_d = \frac{F_s}{\rho C} $$

(3)

Ehhalt and Rohrer, 2013 used equation 4, where $M_s$ corresponds to the H$_2$ mixing ratio in the soil air.
\[ F_s = -D_s \rho \frac{\delta M_s}{\delta z} \] (4)

On the other hand, Yashiro et al., 2011 used equation 5

\[ F_s = -D_s \rho \theta a \frac{\delta M_s}{\delta z} \] (5)

This equation originated from equations (2), (3) and (4) in Yonemura et al., 2000. The version used by Yashiro et al., 2011 (equation 5) means that the gradient of the gas concentration between the two layers is determined not only by the mixing ratio, but also by the air-filled porosity.

References


Figure S1: Time series for selected modelled species in molecules cm$^{-3}$ from the baseline box model simulation in London starting at noon on the 22$^{nd}$ of July 2012.
Figure S2: Time series for selected modelled species in molecules cm$^{-3}$ from the baseline box model simulation in Cape Verde starting at midnight on the 2nd of January 2015.
Figure S3: Time series for selected modelled species in molecules cm$^{-3}$ from the baseline box model simulation in Borneo starting at 18:00 on the 11th of July 2008.
Figure S4: Percentage change in the total rate of production of $H_2$ from aldehydes for varying quantum yields calculated relative to the baseline simulation at the three tested sites for 1 modelled day (local time).

Figure S5: Percentage change in the total rate of production of $H_2$ from aldehydes for varying quantum yields at the three tested sites for 5 modelled days (local time).
Figure S6: Average modelled mixing rations of $H_2$ per season estimated with GEOS-Chem at 500 hPa for 2015 and 2016.

Figure S7: Seasonal comparison between average modelled (triangles) and average observed (circles) $H_2$ estimated ratios at different heights with respect to the surface layer. Modelled averaged values are from 2015 and 2016. Observations averaged values are those from the Aircraft (AIA) flask sampling data from Krummel et al., 2021.


Figure S9: Seasonal cycle comparisons of CO at five sites from the CSIRO dataset (Krummel et al., 2021a, 2021b, 2021c, 2021d, 2021e, 2021f, 2021g, 2021h, 2021i) reported Flask Data for 2015 and 2016, the dash line with circle markers corresponds to observed values and the continuous line with triangle markers corresponds to modelled values in GEOS-Chem.
Supplementary material

Figure S10: Percentage difference on the chemical production of H₂ between the 1% production of H₂ scenario and the baseline for a year a) At the surface level, b) At 500 hPa, c) Longitude section at 2 N and d) Latitude section at -70 W

Figure S11: Percentage difference on the chemical production of H₂ between the 1% production of H₂ scenario and the baseline for a year at the surface layer from the photolysis of a) methylglyoxal b) acetaldehyde, HPALD, glycolaldehyde, methacrolein, lumped aldehydes with more than 3 carbon atoms RCHO (integrated)

Figure S12: Simulated mixing ratios of methylglyoxal at the surface layer for January and July 2015. The modelled magnitudes of methylglyoxal in our GEOS-Chem simulation are comparable to those shown by Fu et al., 2008 in their Figure 2b.
Figure S13: Percentage difference on the modelled mixing ratios of H$_2$ between the 1% production of H$_2$ scenario and the baseline for a year a) At the surface level, b) At 500 hPa, c) Longitude section at 2deg N and d) Latitude section at -70deg W


Table S1: List of aldehydes in the MCM, along with photolysis products and rates ($J \,[\text{s}^{-1}]$) for the $H_2$ channel estimated with 1%, 2%, 5% and 10% quantum yields applied across wavelengths.

<table>
<thead>
<tr>
<th>Aldehyde</th>
<th>MCM ID</th>
<th>Photolysis products$^a$</th>
<th>JID 1% qy</th>
<th>J 2% qy</th>
<th>J 5% qy</th>
<th>J 10% qy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formaldehyde (Methanal)</td>
<td>HCHO$^b$</td>
<td>$H_2 + CO$</td>
<td>J12</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Glyoxal (oxalde-</td>
<td>GLYOX$^b$</td>
<td>$CO + CO + _H_2$</td>
<td>J31</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>CH3CHO</td>
<td>$CH_2CO + _H_2$</td>
<td>J62</td>
<td>4.37E-07</td>
<td>2.18E-06</td>
<td>4.37E-06</td>
</tr>
<tr>
<td>Propanal</td>
<td>C2H5CHO</td>
<td>$C_2H_4 + H_2$</td>
<td>J63</td>
<td>5.68E-07</td>
<td>2.84E-06</td>
<td>5.68E-06</td>
</tr>
<tr>
<td>Butanal</td>
<td>C3H7CHO</td>
<td>$C_3H_6 + H_2 + CO$</td>
<td>J64</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>Isobutylaldehyde</td>
<td>IPRCHO</td>
<td>$C_3H_6 + H_2 + CO$</td>
<td>J65</td>
<td>5.49E-07</td>
<td>2.74E-06</td>
<td>5.49E-06</td>
</tr>
<tr>
<td>Acrolein (2-</td>
<td>ACR</td>
<td>$C_2H_2 + H_2 + CO$</td>
<td>J66</td>
<td>2.44E-06</td>
<td>1.22E-05</td>
<td>2.44E-05</td>
</tr>
<tr>
<td>propanal)</td>
<td>MACR</td>
<td>$C_3H_4 + H_2 + CO$</td>
<td>J67</td>
<td>3.05E-07</td>
<td>1.52E-06</td>
<td>3.05E-06</td>
</tr>
<tr>
<td>Methylglyoxal (2-</td>
<td>MGLYOX</td>
<td>$CH_2CO + _H_2$</td>
<td>J68</td>
<td>4.42E-07</td>
<td>2.21E-06</td>
<td>4.42E-06</td>
</tr>
<tr>
<td>(oxopropanal)</td>
<td>H2OCH2OOH</td>
<td>$H_2 + Other$</td>
<td>J69</td>
<td>4.65E-08</td>
<td>2.32E-07</td>
<td>4.65E-07</td>
</tr>
<tr>
<td>Glycolaldehyde (2-</td>
<td>HOCH2CHO</td>
<td>$HCHO + H_2 + CO$</td>
<td>J70</td>
<td>1.07E-07</td>
<td>5.37E-07</td>
<td>1.07E-06</td>
</tr>
<tr>
<td>(hydroxyacetaldheyde)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Hydroxy-2-Methylpro-</td>
<td>HOIPRCHO</td>
<td>$H_2 + CO$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>panal</td>
<td>CH3CHOHCHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>3-Oxopentanal</td>
<td>CO3C4CHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>3-Hydroxypentanal</td>
<td>HO3C4CHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>3-Hydroxypropanal</td>
<td>HOCC2H4CHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>2-Hydroxybutanal</td>
<td>HO3C3CHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
<tr>
<td>4-Hydroxybutanal</td>
<td>HO3C3H6CHO</td>
<td>$H_2 + Other$</td>
<td>J71$^d$</td>
<td>6.76E-07</td>
<td>3.38E-06</td>
<td>6.76E-06</td>
</tr>
</tbody>
</table>

$^a$ Expressed with the identification used in the MCM.

$^b$ Species in MCM that already include the photolytic generation of $H_2$. The photolysis rates were kept the same for all tests.

$^c$ The ketene $CH_2CO$ generated in the photolysis channel was replaced with glycolaldehyde as described in the text.

$^d$ In absence of aldehyde-specific cross section measurements, the photolysis rate of butanal was used as a surrogate in each case.

Names in parenthesis are IUPAC names.
Table S2: Calculated metrics at each measuring site from the Krummel et al., 2021j, 2021k, 2021l, 2021m, 2021n, 2021o, 2021p, 2021q dataset

<table>
<thead>
<tr>
<th>Site ID</th>
<th>Site name</th>
<th>Lat</th>
<th>Lon</th>
<th>Hemisphere</th>
<th>Lat. band</th>
<th>Gas</th>
<th>Start(^a)</th>
<th>End(^b)</th>
<th>N. obs(^c)</th>
<th>Mean Mod. [ppb]</th>
<th>Mean Obs. [ppb]</th>
<th>r</th>
<th>RMSE</th>
<th>MB</th>
<th>NMB</th>
<th>MFB</th>
<th>NME</th>
<th>FE</th>
<th>FB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHT</td>
<td>Alert</td>
<td>82.4991</td>
<td>-62.3415</td>
<td>NH</td>
<td>HNH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>494.86</td>
<td>489.93</td>
<td>0.93</td>
<td>9.30</td>
<td>4.93</td>
<td>1.01</td>
<td>0.01</td>
<td>1.62</td>
<td>1.63</td>
<td>1.00</td>
</tr>
<tr>
<td>MHD</td>
<td>Mace Head</td>
<td>53.33</td>
<td>-9.9</td>
<td>NH</td>
<td>HNH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>502.40</td>
<td>515.91</td>
<td>0.95</td>
<td>15.99</td>
<td>-13.51</td>
<td>-2.62</td>
<td>-0.03</td>
<td>2.02</td>
<td>2.70</td>
<td>-2.70</td>
</tr>
<tr>
<td>MLO</td>
<td>Mauna Loa</td>
<td>19.5362</td>
<td>-155.5762</td>
<td>NH</td>
<td>LNH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>530.12</td>
<td>536.51</td>
<td>0.54</td>
<td>11.53</td>
<td>-6.39</td>
<td>-1.19</td>
<td>-0.01</td>
<td>1.60</td>
<td>1.62</td>
<td>-1.21</td>
</tr>
<tr>
<td>CFA</td>
<td>Cape Ferguson</td>
<td>-19.2773</td>
<td>147.0584</td>
<td>SH</td>
<td>LSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>20</td>
<td>483.79</td>
<td>546.30</td>
<td>0.52</td>
<td>60.87</td>
<td>-60.26</td>
<td>-11.03</td>
<td>-0.10</td>
<td>11.03</td>
<td>9.74</td>
<td>-9.74</td>
</tr>
<tr>
<td>CGO</td>
<td>Cape Grim</td>
<td>-40.6822</td>
<td>144.6883</td>
<td>SH</td>
<td>LSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>508.42</td>
<td>547.37</td>
<td>0.74</td>
<td>39.44</td>
<td>-38.95</td>
<td>-7.12</td>
<td>-0.07</td>
<td>7.12</td>
<td>7.39</td>
<td>-7.39</td>
</tr>
<tr>
<td>MQA</td>
<td>Macquarie Island</td>
<td>-54.4985</td>
<td>158.9805</td>
<td>SH</td>
<td>LSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>513.94</td>
<td>547.78</td>
<td>0.84</td>
<td>34.04</td>
<td>-33.84</td>
<td>-6.18</td>
<td>-0.06</td>
<td>6.18</td>
<td>6.38</td>
<td>-6.38</td>
</tr>
<tr>
<td>CYA</td>
<td>Casey</td>
<td>-66.2853</td>
<td>110.5167</td>
<td>SH</td>
<td>HSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>516.11</td>
<td>546.93</td>
<td>0.67</td>
<td>30.98</td>
<td>-30.82</td>
<td>-5.64</td>
<td>-0.06</td>
<td>5.64</td>
<td>5.90</td>
<td>-5.90</td>
</tr>
<tr>
<td>MAA</td>
<td>Mawson</td>
<td>-67.6014</td>
<td>102.8706</td>
<td>SH</td>
<td>HSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>14</td>
<td>516.20</td>
<td>531.72</td>
<td>0.54</td>
<td>34.05</td>
<td>-33.90</td>
<td>-6.14</td>
<td>-0.04</td>
<td>6.14</td>
<td>3.70</td>
<td>-3.70</td>
</tr>
<tr>
<td>SPO</td>
<td>South Pole</td>
<td>-89.9969</td>
<td>-24.8</td>
<td>SH</td>
<td>HSH</td>
<td>H2</td>
<td>2015-01</td>
<td>2016-12</td>
<td>24</td>
<td>516.12</td>
<td>547.13</td>
<td>0.87</td>
<td>31.16</td>
<td>-31.01</td>
<td>-5.67</td>
<td>-0.06</td>
<td>5.67</td>
<td>5.83</td>
<td>-5.83</td>
</tr>
</tbody>
</table>

\(^a\)Corresponds to the first date available in the measurements from Krummel et al., (2021a-i) used in the calculation of metrics.

\(^b\)Corresponds to the last date available in the measurements from Krummel et al., (2021a-i) used in the calculation of metrics.

\(^c\)Number of measurements used to compare during the period covered by the start \(^a\) and end \(^b\) dates.