Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/4/4/eaar2547/DC1

Supplementary Materials for

Isoprene photo-oxidation products quantify the effect of pollution on hydroxyl radicals over Amazonia

Yingjun Liu, Roger Seco, Saewung Kim, Alex B. Guenther, Allen H. Goldstein, Frank N. Keutsch, Stephen R. Springston, Thomas B. Watson, Paulo Artaxo, Rodrigo A. F. Souza, Karena A. McKinney, Scot T. Martin

Published 11 April 2018, *Sci. Adv.* **4**, eaar2547 (2018)
DOI: 10.1126/sciadv.aar2547

This PDF file includes:

- table S2. Uncertainty estimates for inferred $C_{OH,noon}$ via error propagation.
- table S3. Summary of inferred OH concentrations over tropical forests in South America.
- References (*41–50*)

section S1. Use of observed C_{NO_y} to represent the average NO_x exposure

The observed NO_y concentration C_{NO_y} was used to quantify NO_x exposure of an air parcel passing over T3 site, which had affected the observed C_{PROD}/C_{ISOP}. The major NO_x sources for air masses passing over T3 site included background emissions from forests, regional biomass burning, and anthropogenic emissions in the city of Manaus. Emissions from forests and regional biomass burning were diffusive NO_x sources. If these emissions dominated, the use of observed C_{NO_y} to represent the average NO_x exposure of the air parcel is a reasonable assumption. Manaus, located 4 to 6 h upwind of T3 site, was a point source of NO_x. Observed C_{NO_y} at T3 site can indicate NO_x exposure of the air parcel after passing over Manaus, but may not reflect the NO_x exposure before that. Nevertheless, the 4 to 6 h from Manaus in mid-morning to T3 site in mid-afternoon was the most intense period of isoprene photooxidation during a day, i.e., high isoprene concentration (Fig. 2a) and high OH concentration (fig. S2). That is, NO_x concentrations during that period had a dominant influence on the ratio C_{PROD}/C_{ISOP} observed at T3. The use of C_{NO_y} to represent NO_x exposure can be problematic in extreme cases when observed NO_y was primarily emitted in the vicinity of the observation site instead of far upwind, for instance from nearby biomass burning. This scenario occurred occasionally, indicative by elevated ratio of C_{NO}/C_{NO_y} or C_{NO_x}/C_{NO_y}. For the data presented in Fig. 3c, no nearby biomass burning event was, however, clearly identified by the ratios.

section S2. Determining relation $y_i^*\left(C_{NO_y}\right)$

Relation $y_i^*\left(C_{NO_y}\right)$ was determined based on observations and analysis presented in ref (*29*) for the same observation site. Ref (*29*) showed that the concentration ratio $C_{ISOPOOH}/C_{MVK+MACR}$ correlated tightly with C_{NO_y} and that the concentration ratio can be transformed to the effective production ratio $y^*_{ISOPOOH}/y^*_{MVK+MACR}$ by combining the measurements with kinetic modeling. In Ref (*29*), only $C_{ISOPOOH}/C_{MVK+MACR}$ under background condition was transformed to $y_{ISOPOOH}/y_{MVK+MACR}$. In the current study, the same transformation was applied to all the data points in the plot of $C_{ISOPOOH}/C_{MVK+MACR}$ versus C_{NO_y} in Ref (*29*), so that a plot of $y^*_{ISOPOOH}/y^*_{MVK+MACR}$ versus C_{NO_y} can be obtained. A fit of the plot led to the following empirical relation of $y^*_{ISOPOOH}/y^*_{MVK+MACR}$ and C_{NO_y}

$$y^*_{ISOPOOH}/y^*_{MVK+MACR} = 3.8\ \mathrm{Exp}(-3.4\ C_{NO_y}) \qquad \text{(S1)}$$

which is presented in fig. S3a. The empirical relation $y^*_{ISOPOOH}/y^*_{MVK+MACR}(C_{NO_y})$ in Equation (S1) was then used to derive $y_i^*\left(C_{NO_y}\right)$ based on two approximations. First, $y_{ISOPOOH} + y_{MVK+MACR}$ was 0.7 across the observed NO_y range (*29*, *30*). Secondly, $y_{(1,2)\text{-}ISOPOOH}/y_{(4.3)\text{-}ISOPOOH}$ had the same value as y_{MVK}/y_{MACR} determined in laboratory experiments (*17*). The obtained relations $y_i^*\left(C_{NO_y}\right)$, where *i* is one of MVK, MACR, (1,2)-ISOPOOH, and (4,3)-ISOPOOH, were shown in fig. S3b.

section S3. Approximation of C_{O_3}

The observed diel variation of ozone concentrations C_{O_3} (fig. S4a) and correlation of afternoon values of C_{O_3} and C_{NO_y} (fig. S4b) was used to constrain C_{O_3} as a function of t and C_{NO_y}. As shown in fig. S4a, C_{O_3} started to increase after sunrise and peaked in midafternoon. The observed time dependence before 16:00 local time was represented empirically by the following equation

$$C_{O_3}(t) = C^*_{O_3}\left(1 - \mathrm{Exp}[-\frac{t}{12000}]\right) \quad \text{(S2)}$$

where $C^*_{O_3}$ was a daily scaling factor representing the variability in ozone concentrations for each day, and t was the time since sunrise in seconds. Values of $C^*_{O_3}$ of 15 and 40 ppb in Equation S2 were used to approximate the time dependence of median C_{O_3} in the wet and dry seasons, respectively (fig. S4a). As shown in fig. S4b, C_{O_3} and C_{NO_y} correlated in afternoon hours, as often observed for photochemically aged air (*41*). Considering an average reaction time of 8.5 h in the afternoon, the linear fit of C_{O_3} and C_{NO_y} in fig. S4b allowed for approximating the dependence of $C^*_{O_3}$ on C_{NO_y}

$$C^*_{O_3} = \frac{C_{O_3,afternoon}}{1 - \mathrm{Exp}[-\frac{8.5\times 3600}{12000}]} = \frac{6.4 + 18C_{NO_y}}{0.92} \quad \text{(S3)}$$

Combining Eqns. (S2) and (S3), the daily course of ozone concentration can be approximated as a function of afternoon concentration C_{NO_y} as follow

$$C_{O_3}(C_{NO_y}, t) = \frac{6.4 + 18C_{NO_y}}{0.92}\left(1 - \mathrm{Exp}[-\frac{t}{12000}]\right) \quad \text{(S4)}$$

section S4. Derivation of Eq. 5 in the main text

For products species *i* of isoprene oxidation, where *i* is one of (1,2)-ISOPOOH, (4,3)-ISOPOOH, MVK, or MACR, the time course of product concentrations $C_i(t)$ in an air mass is governed by the following family of equations (i.e., Equation (1) in the main text)

$$\frac{dC_i}{dt} = y_i(\mathrm{NO})\, k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH}}(t)\, C_{\mathrm{ISOP}}(t) + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}} C_{\mathrm{O_3}}(t)\, C_{\mathrm{ISOP}}(t) - k_i(t)\, C_i(t) \qquad \text{(S5)}$$

where y_i is the production yield of product species *i* from the reaction between hydroxy radical and isoprene, and it varies according to the fate of ISOPOO, which is largely controlled by the NO concentration and hence susceptible to pollution; $C_{\mathrm{OH}}(t)$,$C_{\mathrm{ISOP}}(t)$, and $C_{\mathrm{O3}}(t)$ are the concentrations of hydroxyl radical, isoprene, and ozone, respectively, which vary strongly with time of day; $k_{\mathrm{ISOP,OH}}$ and $k_{\mathrm{ISOP,O3}}$ are the reaction rate constant of isoprene with hydroxyl radical and ozone, respectively; $y_{\mathrm{i,O_3}}$ is the production yield of product speicies *i* from the reaction between ozone and isoprene; and k_i is a composite, pseudo-first order loss coefficient of species *i*, given by $k_i(t) = k_{i,\mathrm{OH}}\, C_{\mathrm{OH}}(t) + k_{i,\mathrm{O_3}}\, C_{\mathrm{O_3}}(t) + k_{i,en} + k_{i,de}$ for bimolecular reaction between species *i* and the hydroxyl radical (OH), atmospheric entrainment (*en*), and surface deposition (*de*). Table S1 present typical values of $y_{\mathrm{i,O_3}}$, $k_{i,\mathrm{OH}}$, $k_{i,\mathrm{O_3}}$, $k_{i,en}$, and $k_{i,de}$ for central Amazonia, partly reproduced from Liu *et al.* (*29*).

The time course of ratio $\varsigma_i(t)$ of concentration of products species *i* to isoprene concentration is governed by the following equation

$$\frac{d\varsigma_i}{dt} = \frac{d(C_i / C_{\mathrm{ISOP}})}{dt} = \frac{1}{C_{\mathrm{ISOP}}(t)}\frac{dC_i}{dt} - \frac{C_i(t)}{C^2_{\mathrm{ISOP}}(t)}\frac{dC_{\mathrm{ISOP}}}{dt} \tag{S6}$$

Taking Equation (S5) into Equation (S6), we have

$$\begin{aligned}\frac{d\varsigma_i}{dt} &= y_i(\mathrm{NO})\, k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH}}(t) + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}}\, C_{\mathrm{O_3}}(t) - k_i(t)\frac{C_i(t)}{C_{\mathrm{ISOP}}(t)} - \frac{C_i(t)}{C^2_{\mathrm{ISOP}}(t)}\frac{dC_{\mathrm{ISOP}}}{dt} \\ &= y_i(\mathrm{NO})\, k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH}}(t) + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}}\, C_{\mathrm{O_3}}(t) - (k_i(t) + \frac{1}{C_{\mathrm{ISOP}}(t)}\frac{dC_{\mathrm{ISOP}}}{dt})\varsigma_i(t)\end{aligned} \tag{S7}$$

The term $C_{\mathrm{ISOP}}(t)$ is empirically represented by the following linear equation (i.e., Equation (2) in the main text)

$$C_{\mathrm{ISOP}}(t) = \xi\, C_{\mathrm{ISOP},0}\left(1 + t / t^*_{\mathrm{ISOP}}\right) \tag{S8}$$

where ξ was a daily scaling factor representing the variability in isoprene concentrations for each day, $C_{\mathrm{ISOP},0}$ was the typical isoprene concentration at time zero (sunrise), and t^*_{ISOP} was the typical characteristic time for C_{ISOP} to double its initial value. Taking Equation (S8) into Equation (S7), we have

$$\frac{d\varsigma_i}{dt} = y_i(\mathrm{NO})\, k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH}}(t) + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}}\, C_{\mathrm{O_3}}(t) - (k_i(t) + \frac{1}{\xi\, C_{\mathrm{ISOP,0}} \left(1 + t / t^*_{\mathrm{ISOP}}\right)} \frac{\xi\, C_{\mathrm{ISOP,0}}}{t^*_{\mathrm{ISOP}}})\varsigma_i(t) \qquad \text{(S9)}$$

$$= y_i(\mathrm{NO})\, k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH}}(t) + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}}\, C_{\mathrm{O_3}}(t) - (k_i(t) + \frac{1}{t + t^*_{\mathrm{ISOP}}})\varsigma_i(t)$$

where quantities ξ and $C_{\mathrm{ISOP,0}}$ are dropped off in the analysis. Further replacing $y_{\mathrm{i}}(\mathrm{NO})$ in Equation (S9) with $y_i^*\left(C_{\mathrm{NO}_y}\right)$, replacing $C_{\mathrm{O_3}}(t)$ with $C_{\mathrm{O_3}}(C_{\mathrm{NO}_y}, t)$ (Equation (S4)), and assuming $C_{\mathrm{OH}}(t) = C_{\mathrm{OH,noon}} \frac{J_{\mathrm{O_3}}(t)}{J_{\mathrm{O_3,\,noon}}}$ (Equation (3) in the main text), we have

$$\frac{d\varsigma_i}{dt} = y_i^*\left(C_{\mathrm{NO}_y}\right) k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH,noon}} \frac{J_{\mathrm{O_3}}(t)}{J_{\mathrm{O_3,noon}}} + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}} C_{\mathrm{O_3}}(C_{\mathrm{NO}_y}, t) - \left(k_i(t) + \frac{1}{t^*_{\mathrm{ISOP}} + t}\right)\varsigma_i(t) \qquad \text{(S10)}$$

which constitutes Equation (5) in the main text.

Equation (5) is a first-order linear differential equation, and its general analytical solution is given by

$$\varsigma_i(t) = \frac{1}{e^{\int P(t)dt}} \int e^{\int P(t)dt} Q(t)dt \qquad \text{(S11)}$$

where

$$P(t) = k_i(t) + \frac{1}{t^*_{\mathrm{ISOP}} + t}$$

$$Q(t) = y_i^*\left(C_{\mathrm{NO}_y}\right) k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH,noon}} \frac{J_{\mathrm{O_3}}(t)}{J_{\mathrm{O_3,noon}}} + y_{i,\mathrm{O_3}} k_{\mathrm{ISOP,O_3}} C_{\mathrm{O_3}}(C_{\mathrm{NO}_y}, t)$$

Equation (S11) contains complex integral terms, which cannot be done analytically. Equation (5) was thus integrated numerically in the current analysis.

section S5. Additional note of Fig. 4 in the main text

The simulated increase of C_{PROD}/C_{ISOP} with increasing C_{NO_y} at zero ozone (dashed lines in Fig. 4) arises from the different lifetimes of MVK, MACR, (1,2)-ISOPOOH, and (4,3)-ISOPOOH. For reference, these four products have respective lifetimes against OH reaction of 14 h, 10 h, 4 h, and 2 h for $C_{OH} = 1 \times 10^6$ cm^{-3} (cf. table S1 and references therein). The lifetimes of the ISOPOOH isomers, which are the dominant products for low NO concentrations, are about 70% shorter than those of MVK and MACR. As a consequence, the concentration ratio of C_{PROD}/C_{ISOP} is lower at lower C_{NO_y}, all other factors being equal. The ratio levels off at high C_{NO_y} once MVK and MACR become the dominant products.

section S6. Sensitivity tests regarding entrainment process

Entrainment of air masses aloft associated with the growth of the convective boundary layer can influence C_{PROD}/C_{ISOP} and C_{NO_y}. Treatment of entrainment process in the model thus affects the inferred C_{OH}. In our base-case model, the entrainment process is considered but in a simplified way (cf. note in table S1). This simplification is evaluated through a series of sensitivity tests using a mixed boundary-layer model (MXL), which couples chemistry with boundary layer dynamics and allows for explicit modeling of the entrainment processes (*42*, *43*).

In the MXL approach, the evolution of the boundary layer height h is explicitly simulated, driven by the surface heat fluxes that are prescribed to the model. The entrainment velocity v_{en} is taken as the growth rate of the boundary layer height (dh/dt). The concentration discontinuity between the mixed layer and the inversion layer/free troposphere is represented by a zero-order jump. The approach assumes that under convective conditions strong turbulent mixing causes perfect mixing of quantities over the entire depth of the boundary layer. Arellano *et al.* (*43*) showed that the concentration profiles of isoprene simulated using MXL for Amazon condition were comparable to the profiles simulated using the Large-Eddy Simulation technique coupled to a chemistry module. For the MXL model, the counterpart of Equation (5) in the main text is given by

$$\frac{d\varsigma_i}{dt} = y_i^*\left(C_{\mathrm{NO}_y}\right) k_{\mathrm{ISOP,OH}}\, C_{\mathrm{OH,noon}} \frac{J_{\mathrm{O}_3}(t)}{J_{\mathrm{O}_3,\mathrm{noon}}} + y_{i,\mathrm{O}_3} k_{\mathrm{ISOP,O}_3} C_{\mathrm{O}_3}(C_{\mathrm{NO}_y}, t) - \left(k_{\mathrm{i,OH}}\, C_{\mathrm{OH,noon}} \frac{J_{\mathrm{O}_3}(t)}{J_{\mathrm{O}_3,\mathrm{noon}}} + k_{\mathrm{i,O}_3} C_{\mathrm{O}_3}(C_{\mathrm{NO}_y}, t) + \frac{v_{de,i}}{h(t)} + \frac{1}{h(t)} \frac{dh}{dt} + \frac{1}{t^*_{\mathrm{ISOP}} + t} \right) \varsigma_i(t) + \frac{1}{h(t)} \frac{dh}{dt} \frac{C_{i,upper}(h,t)}{C_{\mathrm{ISOP}}(t)} \quad \text{(S12)}$$

To use Equation (S12), we have to know the isoprene concentration $C_{\mathrm{ISOP}}(t)$ in the mixed layer as well as the product concentration $C_{i,\mathrm{upper}}(h, t)$ above the boundary layer. In the real atmosphere, $C_{i,\mathrm{upper}}$ is expected to be a function of time and height, yet there is little observational constraint of the dependence. In the sensitivity analysis below, $C_{i,\mathrm{upper}}$ is assumed across a range of values, bounded by an upper limit equivalent to typical product concentration around sunset when the mixed layer collapses (i.e., no nocturnal dilution of residual products in the inversion layer) and a lower limit of 0 ppb (i.e., residual products from previous day fully removed by dilution or wet deposition).

Figure S5 presents an example of simulation results using this approach. The model was run under typical conditions in the wet season. The simulated evolution of boundary layer height is in good agreement with retrievals from the discrete radiosonde measurements at the site (fig. S5a). The simulated evolution of $C_{\mathrm{PROD}}/C_{\mathrm{ISOP}}$ is shown in fig. S5b for $C_{\mathrm{PROD,upper}}$ ranging from 0 to 0.6 ppb. The value of 0.6 ppb is a typical product concentration in the boundary layer before sunset during the wet season (cf. Fig. 2 in the main text). As shown in fig. S5b, $C_{\mathrm{PROD}}/C_{\mathrm{ISOP}}$ is higher for higher $C_{\mathrm{PROD,upper}}$, yet the difference decreases with time of the day. In the morning when the boundary layer is shallow and fast-growing, entrainment from air aloft has a large impact on $C_{\mathrm{PROD}}/C_{\mathrm{ISOP}}$. The value of $C_{\mathrm{PROD}}/C_{\mathrm{ISOP}}$ simulated using the upper limit of $C_{\mathrm{PROD,upper}}$ is up to 5 times higher than that simulated using the lower limit. Later in the day, growth of the boundary layer slows and photochemistry speeds up, each of which lessens the impact of

entrainment on C_{PROD}/C_{ISOP}. The difference in simulated C_{PROD}/C_{ISOP} of the two limiting cases diminishes to 30% in afternoon hours. In the current study, the OH retrieval is based on C_{PROD}/C_{ISOP} in the afternoon hours, when C_{PROD}/C_{ISOP} was less sensitive to the entrainment processes.

Figure S6 displays retrieved $C_{OH,noon}$ for given C_{PROD}/C_{ISOP} at 18.5 UTC (14.5 local time), determined using the base-case model and the MXL model, respectively. Model parameters were kept the same as the simulation in fig. S5. The MXL simulations with $C_{PROD,upper}$ of 0 and 0.6 ppb provides upper and lower bounds for retrieved $C_{OH,noon}$, respectively, and the band width is up to 0.7×10^6 cm^{-3}. The $C_{OH,noon}$ retrieved using the base-case model generally falls into the range suggested by the MXL model, close to the lower limit at upper C_{PROD}/C_{ISOP} and vice versa. For high C_{PROD}/C_{ISOP}, $C_{OH,noon}$ retrieved using the base-case model can be an underestimate by up to 20%. The level of agreement between the MXL and base-case model simulation indicates that parameterization of entrainment in the base-case model is a reasonable simplification. The potential bias associated with the simplification (i.e., underestimating $C_{OH,noon}$ at higher C_{PROD}/C_{ISOP} and overestimating at lower C_{PROD}/C_{ISOP}) indicates the line drawn in Fig. 5 ($C_{OH,noon}$ versus C_{NO_y}) in the main text can be an underestimate, given that higher C_{PROD}/C_{ISOP} was observed at higher C_{NO_y} (Fig. 3c in the main text).

section S7. Error analysis

Uncertainties in inferred $C_{OH,noon}$ can arise from uncertainties in measurements, uncertainties in model parameterization, and assumptions in the inference (forward) model. The uncertainties associated with measurements and model parameterization are evaluated using a numeric error propagation approach described below.

The current work tackles an inverse problem, which is to infer $C_{OH,noon}$ for given C_{NO_y} and C_{PROD}/C_{ISOP} at a specific time point t, based on a forward model

$$C_{PROD}/C_{ISOP}(t) = F(C_{NO_y}, C_{OH,noon}, t, p_1, p_2...p_n) \quad \text{(S13)}$$

where function F can be obtained from Equation (S11) and $p_1, p_2...p_n$ represent various modeling parameters. Assuming that G is the inverse of function F, defined as

$$C_{OH,noon} = G(C_{PROD}/C_{ISOP}, C_{NO_y}, t, p_1, p_2...p_n) \quad \text{(S14)}$$

the standard deviation $S_{C_{OH,noon}}$ of retrieved $C_{OH,noon}$ can be estimated by error propagation via function G as follows

$$S_{C_{OH,noon}} = \sqrt{\sum \left(\frac{\partial G}{\partial x}\right)^2 S_x^2}\ ;\ x \in \{C_{Prod}/C_{ISOP}, C_{NO_y}, p_1, p_1 \cdots p_n\} \quad \text{(S15)}$$

where S_x is the standard deviation of observable/parameter x. Equation (S15) is based on linear characteristics of the gradient of G. Specifically, a linear approximation of G has to be valid inside a neighborhood of radius S_x. Equation (S15) also neglects the correlation of variables.

Based on Eq. (S15), the key of uncertainty estimation through error propagation is to determine $\partial G/\partial x$ for any observable/parameter x. Since Equation (S11) and hence the function F contain complex integral terms, the inverse function G cannot be obtained analytically. Instead we took an alternative numeric approach to derive $\partial G/\partial x$ for any given $\{C_{PROD}/C_{ISOP}', C_{NO_y}', t'\}$. It involves four steps:

(1) numerically solving Eq. (S11) at $\{C_{NO_y}', t'\}$ and a range of prescribed $C_{OH,noon}$ values to get a list of $\{C_{PROD}/C_{ISOP,} C_{OH,noon}\}$ pairs;

(2) interpolating the $\{C_{PROD}/C_{ISOP,} C_{OH,noon}\}$ pairs to determine the corresponding $C_{OH,noon}'$ for given C_{PROD}/C_{ISOP}';

(3) varying parameter x in the forward model around the central value x_0 and redoing steps (1-2) to generate a list of $\{x, C_{OH,noon}'\}$ pairs; and

(4) interpolating the $\{x, C_{OH,noon}'\}$ pairs and taking the derivative at $x = x_0$, which is the value of $\partial G/\partial x(x_0)$.

The numerical approach outlined above determines the value of $\partial G/\partial x$ locally at a specific point $\{C_{PROD}/C_{ISOP}', C_{NO_y}', t'\}$ and thereby estimates the local uncertainty of $C_{OH,noon}$. In addition, the numerical simulation (x versus $C_{OH,noon}'$) also helps to validate the major underlying assumption of Eq. (S15), i.e., a linear approximation of G inside a neighborhood of radius S_x.

Table S2 presents the results of uncertainty estimates for three cases. The uncertainty analysis considers random errors of observables C_{PROD}/C_{ISOP} and C_{NO_y} as well as systematic errors of three modeling parameters t^*_{ISOP}, $p_{NO_y,yield}$ and $p_{OH,t}$. They are respectively associated to three approximations quantitatively most important in the retrieval of $C_{OH,noon}$, including $C_{ISOP}(t)$, $y_i^*\left(C_{NO_y}\right)$, and $C_{OH}(t)$. Lognormal error distribution is assumed for each of the three modeling parameters. Table S2 presents the best-estimate standard error S associated to each observable or parameter. The resultant variation of $C_{ISOP}(t)$, $y_i^*\left(C_{NO_y}\right)$, and $C_{OH}(t)$ due to standard variation of t^*_{ISOP} , $p_{NO_y,yield}$ and $p_{OH,t}$ is presented in fig. S7 (a-c), respectively.

The relative standard error $S_{C_{OH,noon}}$ of inferred $C_{OH,noon}$ ranged from 23% to 30% for the three cases in table S2, which represent data points at the three corners of the triangle-shape scatter plot of C_{PROD}/C_{ISOP} versus C_{NO_y} (Fig. 3c in the main text). This level of uncertainty in inferred $C_{OH,noon}$ values would not affect the derived dependence of $C_{OH,noon}$ on C_{NO_y}. The largest contribution of $S_{C_{OH,noon}}$ is from errors of observable C_{PROD}/C_{ISOP} , followed by systematic errors of parameters $p_{OH,t}$ (i.e., time dependence of C_{OH}) and t^*_{ISOP}. Errors in NO_y related terms, including C_{NO_y} or $p_{NO_y,yield}$, have smaller contribution, in particular at higher C_{NO_y}.

section S8. Comparison of $C_{OH,noon}$ obtained in this study with other OH studies

The OH values obtained in the current study can be compared with several other studies.

(1) LIF measurement over rainforest in South America (*5*)

Aircraft measurements over rainforests on the east coast of South America reported daytime OH concentration of (5.6 ± 1.9) × 10^6 cm^{-3} for background condition (*5*). Laser-Induced-Florescence (LIF) technique was used for the OH measurement. By comparison, the equivalent noontime OH concentration for background central Amazon conditions derived in the current study is (6.8 ± 2.1) × 10^5 cm^{-3} (Fig. 5), which is less than 15% of daytime OH value reported in Ref (*5*). The study in Ref (*5*) was, however, 1500 km northeast of the present study in a coastal region, so factors other than possible measurement discrepancies could lead to a difference in OH concentrations in the two regions.

(2) Concurrent CIMS measurement at the same observation site

During GoAmazon 2014/5 campaign, OH concentrations were monitored at the same site of current study using Chemical Ionization Mass Spectrometry (CIMS) (*44*). The CIMS measurements suggested an average midday maximum OH concentration of 1.2× 10^6 cm^{-3} in the wet season (sparse data in the dry season). By comparison, the inter-quartile range of $C_{OH,noon}$ derived using the current method was (0.5-1.2) × 10^6 cm^{-3} for the wet season, with a standard error of ±30%. The statistics of inferred $C_{OH,noon}$ characterize the regional oxidation capacity over the Amazon basin upwind of the observation site, since the air parcels arriving at the site can have varied backward trajectories. By comparison, the OH concentrations measured by CIMS were instantaneous values specific to the observation site. Keeping these differences in the mind, the inferred OH values are consistent with the CIMS observations.

(3) Inferred OH concentration over rainforests in South America

Table S3 summarizes OH concentrations inferred from VOC observations over rainforests in South America. The earlier studies were based on airborne measurements in the wet seasons (*45–47*). The OH concentrations were simulated using observation-constrained box model with simplified chemical mechanism. Simulated daytime maximum OH concentrations were in the range of (0.6-1.5) × 10^6 cm^{-3}, which is consistent with the inferred $C_{OH,noon}$ in this study for the wet season, with an inter-quartile range of (0.5-1.2) × 10^6 cm^{-3}. The more recent OH estimates were based on airborne measurements in the dry seasons (*22*, *23*). The OH concentration were inferred using two methods: (1) vertical gradient of concentration ratio C_{PROD}/C_{ISOP} in the boundary layer and (2) budget of biogenic VOCs in the boundary layer. The estimated daytime OH concentration ranged from (1.3 ± 0.5) × 10^6 cm^{-3} to (3-8) × 10^6 cm^{-3}. In this study, the interquartile range of inferred $C_{OH,noon}$ for the dry season was (0.8-1.5) × 10^6 , which fell into the lower range of previous estimates.

Supplementary Figures

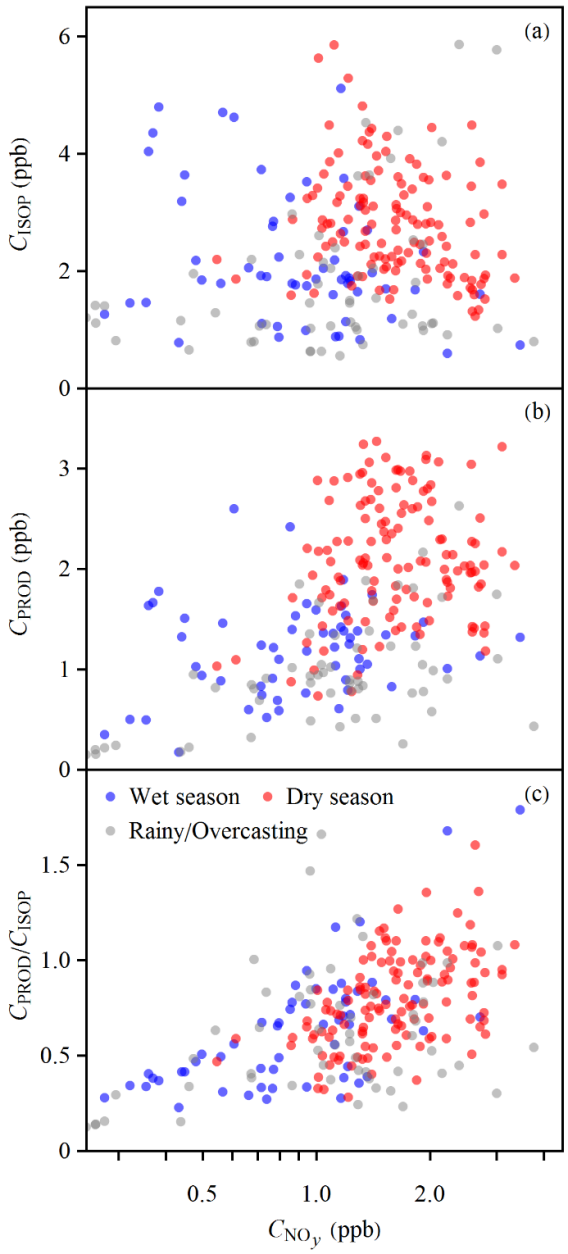

fig. S1. Scatterplots of VOC concentrations with NO_y concentration for all-weather condition. (**a**) Isoprene concentration C_{ISOP}. (**b**) Sum concentration C_{PROD} of isoprene oxidation products. (**c**) Concentration ratio C_{PROD}/C_{ISOP}. Data points represent hourly averages from 17:00 - 20:00 UTC (13:00 - 16:00 local time). Data are shown for the wet and dry seasons in blue and red colors, respectively, for fair-weather conditions. There are 57 data points for the wet season and 128 for the dry season. The gray points, representing data recorded in either season at times of heavy rainfall or overcast conditions, are not included in the analysis.

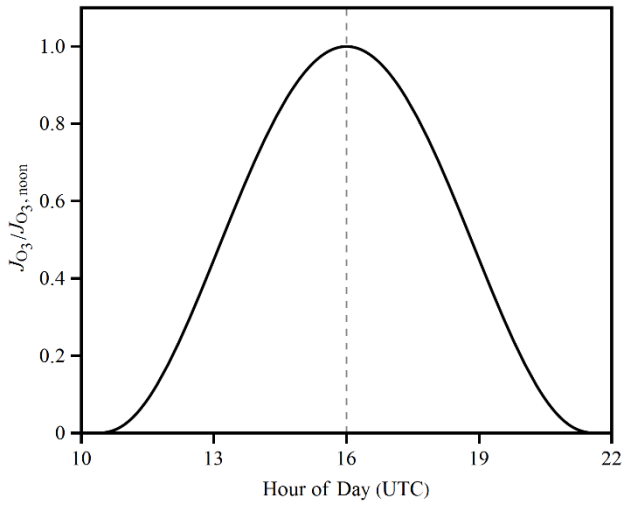

fig. S2. Simulated daily variation of photolysis frequency of ozone J_{O_3}, normalized to the noontime value $J_{O3,noon}$, based on Master Chemical Mechanism. The dashed line represents local noon.

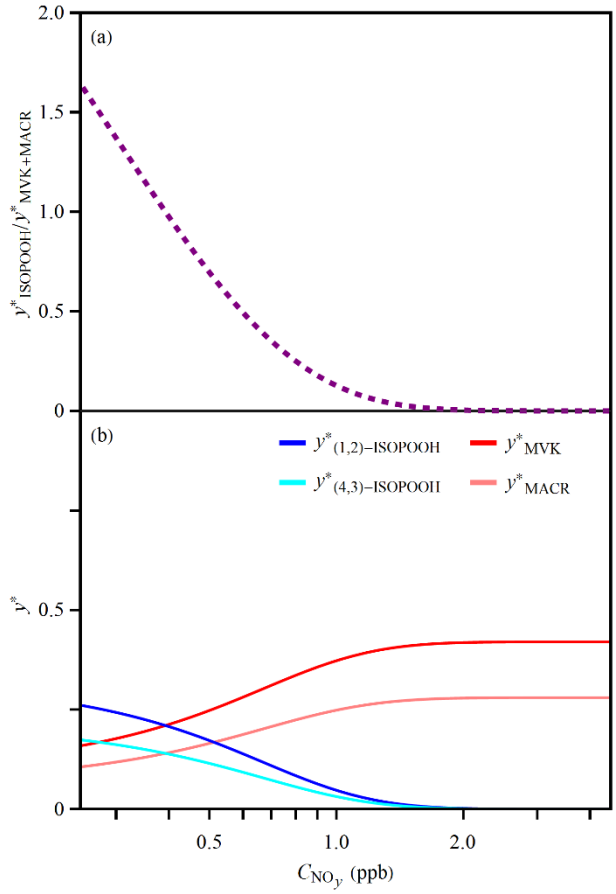

fig. S3. Simulated NO_y dependence of effective production yields. Dependence on NO_y concentration of (**a**) ratio $y^*_{ISOPOOH}/y^*_{MVK+MACR}$ of product yield and (**b**) product yield y^*. The presented result is based on observations and analysis presented in ref (*29*). See also section S2 in the Supplementary Materials.

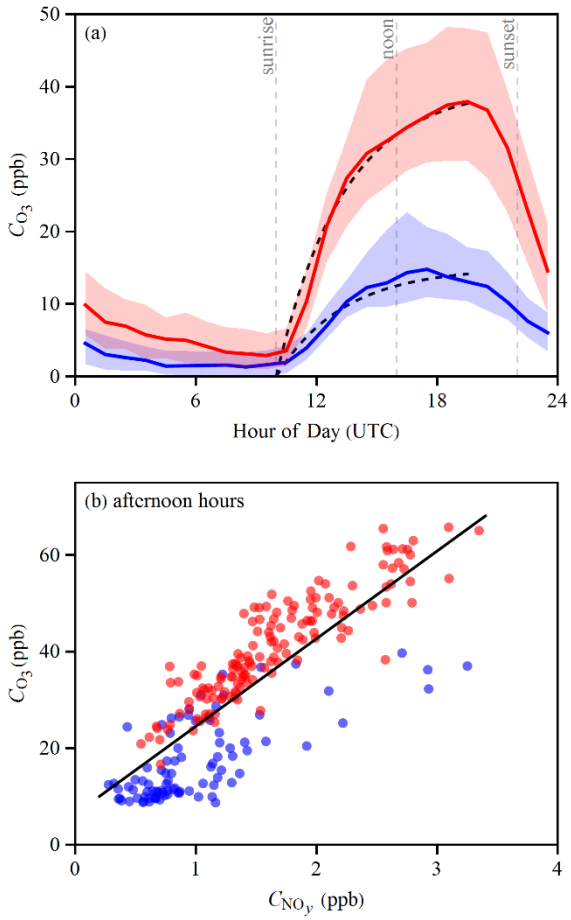

fig. S4. Observation and simulation of ozone concentration C_{O_3}. (**a**) Hourly variation and (**b**) correlation with NO_y concentration in afternoon hours. Data are shown in the wet and dry seasons in blue and red colors, respectively. For panel (a), the solid line and shaded regions respectively represent the median and interquartile ranges of the data sets for each hour of the day. The two black dashed lines show the simulated increase of C_{O_3} from sunrise to midafternoon using Equation (S2) for $C^*_{O_3}$ values of 15 and 40 ppb. The vertical dashed gray lines demarcate local sunrise, noon, and sunset (UTC less 4 h). For panel (b): hourly concentrations within a time window of 13:00-16:00 (local time) were shown, for fair-weather conditions (the same as Fig. 3 in the main text). The black line represent linear fit of observed C_{O_3} and C_{NO_y}.

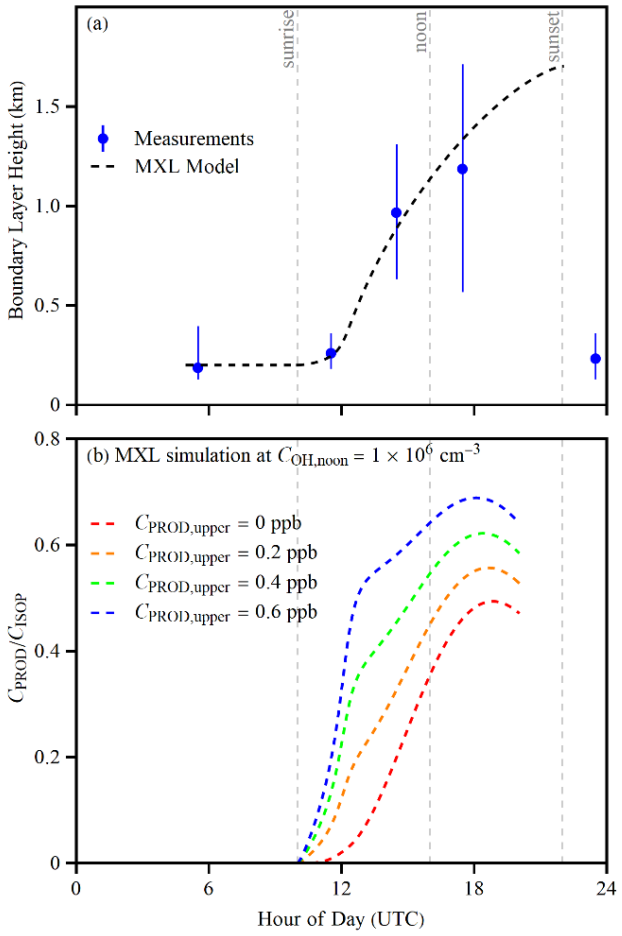

fig. S5. Simulation using a mixed boundary layer model. Simulated daytime evolution of (**a**) boundary layer height and (**b**) concentration ratio C_{PROD}/C_{ISOP} using a mixed boundary layer (MXL) model, which better accounts for entrainment process. Details of MXL model are presented in section S6 of the Supplementary Materials. The simulation is run under typical condition of the wet season ($y_{MVK+MACR}$ = $y_{ISOPOOHs}$ = 0.35; C_{O_3} = 12 ppb, $C_{ISOP,0}$ = 0.35 ppb). The noontime OH concentration $C_{OH,noon}$ is set as 1.0×10^6 cm^{-3}. In panel (a), blue dot and line represent median and inter-quartile range of boundary layer height retrieved from radiosonde measurements in the wet season. In panel (b), simulation results are shown for $C_{PROD,upper}$ of 0, 0.2, 0.4, and 0.6 ppb.

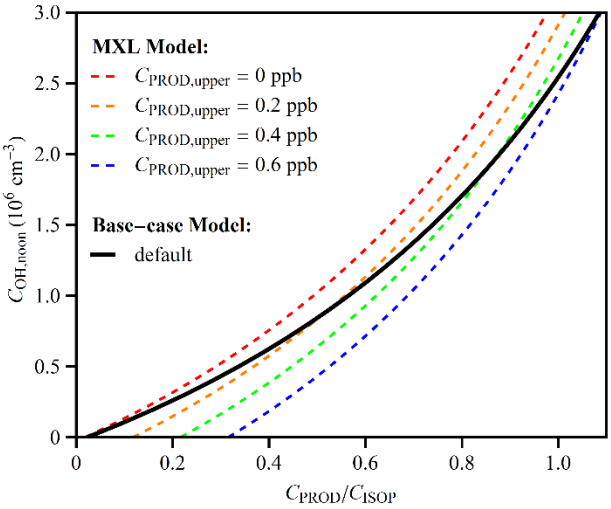

fig. S6. Simulated relationship of equivalent noontime OH concentration $C_{OH,noon}$ and concentration ratio C_{PROD}/C_{ISOP} using the base-case model and a mixed boundary-layer model. Details of the MXL model are presented in section S6 of the Supplementary Materials. The simulation is run under typical condition of the wet season ($y_{MVK+MACR} = y_{ISOPOOHs} = 0.35$; $C_{O_3} = 12$ ppb, $C_{ISOP,0} = 0.35$ ppb, $t = 8.5$ h). For MXL modeling, simulation results are shown for $C_{PROD,upper}$ of 0, 0.2, 0.4, and 0.6 ppb.

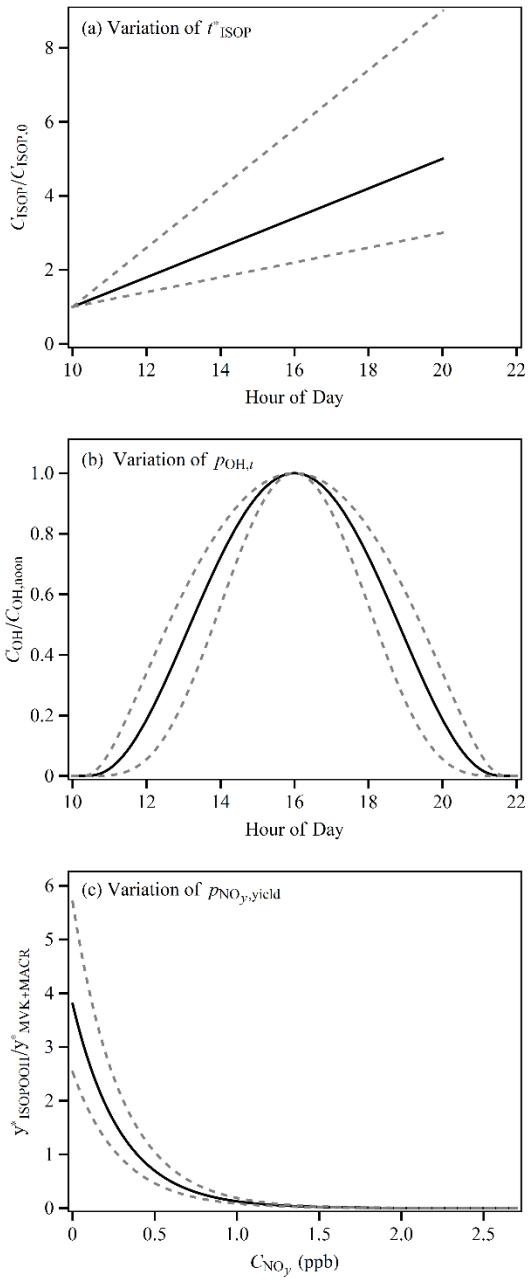

fig. S7. Variation in model parametrizations for error analysis. Panels (**a**), (**b**), (**c**) present parametrizations for $C_{ISOP}(t)$, $C_{OH}(t)$, and $y_i^*\left(C_{NO_y}\right)$, respectively. The black line represents best-estimate parameterization. The grey dashed lines represent parameterizations using best-estimate value +/- standard error of associated modeling parameters (cf. table S2).

Supplementary Tables

table S1. Production yields and loss rate coefficients for isoprene oxidation products used in the model. This table is partly reproduced from Liu et al. (*29*). A 1000-m deep, well mixed planetary boundary layer (PBL) is assumed. A temperature of 298 K is used.

Species,*i*	y_{i,O_3}*	$k_{i,OH}$† (10^{-11} cm^3 s^{-1})	k_{i,O_3}† (10^{-18} cm^3 s^{-1})	$k_{i,de}$‡ (10^{-5} s^{-1})	$k_{i,en}$§ (10^{-5} s^{-1})
(1,2)-ISOPOOH	0	7.4	0	2.0	1.0
(4,3)-ISOPOOH	0	11.8	0	2.0	1.0
MVK	0.2	2.0	5.2	0.2	1.0
MACR	0.3	2.8	1.2	0.2	1.0

* Yields of the four products via isoprene ozonolysis were obtained from Master Chemical Mechanism v3.3.1 (*30*).

† For MVK and MACR, the reactions rate coefficients with OH or ozone were taken from IUPAC recommendation (*48*). For (1,2)-ISOPOOH and (4,3)-ISOPOOH, the reaction rate coefficients with OH were taken from St. Clair et al. (*49*), the reaction rate coefficients with ozone have not been reported yet and are hence taken as zero.

‡ A deposition velocity of 2.0 cm s^{-1} was used for ISOPOOH isomers, as recommended recently by Nguyen et al. based on measurements over a temperate forest (*50*). For MVK and MACR, the deposition rate was assumed to be one magnitude lower (i.e., 0.2 cm s^{-1}) considering its lower water solubility and reactivity.

§ An entrainment rate coefficient was determined by considering the entrainment velocity and concentration gradient between PBL and cloud layer. An entrainment velocity of 2.0 cm s^{-1} was taken based on the simulated evolution of PBL height over a tropical forest during daytime (10:00-17:00 LT) (*43*). Based on reported vertical profiles of isoprene concentration and the ratio of the concentration of (MVK+MACR+ISOPOOH) to that of isoprene over central Amazonia (*22*, *23*), the concentration jump of MVK, MACR, and ISOPOOH between the PBL and the cloud layer is approximated by half of PBL concentration. Compared with a concentration jump of isoprene of one order of magnitude, the concentration jump of MVK, MACR, and ISOPOOH is smaller because of enhanced oxidation in the cloud layer (*22*, *23*).

table S2. Uncertainty estimates for inferred $C_{OH,noon}$ via error propagation.

			Results at{C_{PROD}/C_{ISOP}, C_{NO_y} (ppb), t (h)}		
			{0.4, 0.4, 7.5}	{0.7, 2.5, 8.5}	{1.2, 2.5, 9.5}
Inferred value $C_{OH,noon}$ (10^5 cm^{-3})			6.7	9.1	20
Variable-specific error $\left\|\frac{\partial G}{\partial x}\right\| S_x$ (10^5 cm^{-3})	x	S_x			
	C_{PROD}/C_{ISOP}	15%	1.3	1.9	4.8
	C_{NO_y}	20%	0.24	0.20	0.18
	Log(t^*_{ISOP}) *	Log(2)	0.39	0.62	1.5
	Log($p_{NO_y,yield}$)*	Log(1.5)	0.30	0.001	0.005
	Log($p_{OH,t}$) *	Log(2)	0.60	1.0	3.2
Absolute standard error $S_{C_{OH,noon}}$ † (10^5 cm^{-3})			1.5	2.2	6.0
Relative standard error $S_{C_{OH,noon}}/C_{OH,noon}$			23%	25%	30%

* Error of respective variable follows lognormal distribution. The resultant variation of $C_{ISOP}(t)$, $y_i^*\left(C_{NO_y}\right)$, and $C_{OH}(t)$ due to standard variation of t^*_{ISOP}, $p_{NO_y,yield}$ and $p_{OH,t}$ is presented in fig. S7 (a-c), respectively.

† Standard error $S_{C_{OH,noon}}$ is estimated as square root of sum of squares of the variable-specific errors (cf. Eq. (S15)).

table S3. Summary of inferred OH concentrations over tropical forests in South America.

Location & Time	Measurement	Methods	Inferred OH (cm^{-3})	Ref.
Amazonia; wet season (1987)	Airborne (ABLE 2B)	Obs-constrained forward model	1.5×10^6 (daymax)	(*45*)
Suriname; wet season (1998)	Airborne (LBA-CLAIRE)	Obs-constrained forward model	$(1\text{-}3) \times 10^5$ (24-h avg)	(*46*)
		Obs-constrained forward model	$(0.6\text{-}1.1) \times 10^6$ (daymax)	(*47*)
Amazonia; dry season (2001)	Airborne (LBA-CLAIRE)	Isoprene budget analysis	$(1\text{-}4) \times 10^6$ (daytime)	(*22*)
		Vertical gradient of C_{PROD}/C_{ISOP}	$(3\text{-}8) \times 10^6$ (daytime)	(*22*)
Amazonia; dry season (2004)	Airborne (TROFFEE)	Isoprene budget analysis	$(1.3 \pm 0.5) \times 10^6$ (daytime)	(*23*)
		Vertical gradient of C_{PROD}/C_{ISOP}	$(4.3 \pm 2.4) \times 10^6$ (daytime)	(*23*)
Amazonia; wet season (2014)	Ground (GoAmazon 2014/5)	C_{PROD}/C_{ISOP}	$(0.5\text{-}1.2) \times 10^6$ (daymax; IQR*)	This study
Amazonia; dry season (2014)	Ground (GoAmazon 2014/5)	C_{PROD}/C_{ISOP}	$(0.8\text{-}1.5) \times 10^6$ (daymax; IQR*)	This study

* IQR: inter-quartile range