Satellite evidence for a large source of formic acid from boreal and tropical forests

Trissevgeni Stavrakou, J.-F. Müller, J. Peeters, A. Razavi, L. Clarisse, Cathy Clerbaux, Pierre-François Coheur, D. Hurtmans, M. de Mazière, C. Vigouroux, et al.

Satellite Evidence for a Large Source of Formic Acid From Boreal and Tropical Forests

T. Stavrakou[1⋆], J.-F. Müller[1], J. Peeters[2], A. Razavi[3], L. Clarisse[3], C. Clerbaux[4,3], P.-F. Coheur[3], D. Hurtmans[3], M. De Mazière[1], C. Vigouroux[1], N. M. Deutscher[5,6], D. W. T. Griffith[5], N. Jones[5], C. Paton-Walsh[5]

[1] Belgian Institute for Space Aeronomy, Avenue Circulaire 3, 1180, Brussels, Belgium

[2] Department of Chemistry, University of Leuven, B-3001, Heverlee, Belgium

[3] Spectroscopie de l'Atmosphère, Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, Belgium

[4] UPMC Univ. Paris 6; Université Versailles St.-Quentin; CNRS/INSU, LATMOS-IPSL, Paris, France

[5] School of Chemistry, University of Wollongong, Wollongong, Australia

[6] Institute of Environmental Physics, University of Bremen, Germany

⋆ E-mail : Trissevgeni.Stavrakou@aeronomie.be

Formic acid (HCOOH) has been identified as a major contributor to acidic rain in remote environments[1, 2] but its atmospheric cycle is far from being understood. Along with direct emissions by human activities, vegetation fires, and green plants, its major and most uncertain source is photochemical, and predominantly biogenic[3, 4]. Severe underpredictions of observed formic acid concentrations by large scale models in earlier studies[5, 6, 7] pointed to the existence of missing sources. Here we use spaceborne formic acid column measurements[8] as input to an advanced source inversion algorithm coupled with a global atmospheric model to obtain an improved formic acid global budget. We deduce an annual formic acid source of 100-120 Tg, i.e. 2-3 times higher than estimated from known sources, with a biogenic contribution of about 90%, mostly from tropical and boreal forests. Model comparisons with independent formic acid observations strengthen our conclusions and provide indirect validation for the satellite measurements. The implications of the larger formic acid source on precipitation acidity are substantial, especially over boreal forests in summertime, where the extra biogenic source results in a pH decrease of 0.25-0.5.

Known sources of formic acid in the atmosphere include fossil fuel and biofuel combustion[9], biomass burning[10], plants[11], dry savanna soils[12], formicine ants[13], cloud processing[2], abiological formation on rock surfaces[14],

and photochemical oxidation of volatile organic precursors[15]. Among these sources, the contributions of savanna soils, ants, rocks, and in-cloud formation are very uncertain, but most probably minor.

Based on current inventories, primary formic acid emissions amount to ca. 10 Tg annually on the global scale (Table 1). The largest contribution to the global formic acid budget is due to the photo-oxidation of non-methane hydrocarbons, representing more than 80% of the secondary source (Table 1). In the base version of the IMAGESv2 global chemistry-transport model[16, 17] used in this work to simulate the formic acid budget, the annual production from biogenic precursors is estimated at ca. 20 Tg, about twice lower than in a recent modelling study[7]. The major part of this secondary flux is due to isoprene oxidation by OH (8.9 Tg) and by ozone (3.9 Tg), followed by monoterpene oxidation (3 Tg). The contribution of isoprene and monoterpenes is, however, particularly uncertain due to the scarcity of laboratory experiments of HCOOH formation under atmospheric conditions. For example, the estimated impact of isoprene is largely dependent on the uncertain fate of key intermediates produced at high yields according to recent studies, like dihydroxy epoxides from HO_2-reactions of the isoprene peroxy radicals[18] and hydroperoxy-enones from peroxy isomerisations[19]. Formic acid is removed from the atmosphere through oxidation by OH, accounting for 27% of the global sink, and dry and wet deposition (cf. Supplementary Section 3), resulting in a global lifetime of 3-4 days.

Global models substantially underpredict the observed HCOOH abun-

dances from available ground-based and aircraft measurements[5, 6, 7], pointing to the existence of missing sources. Their quantitative estimation is, however, very difficult due to the scarcity and limited representativity of these measurements. Recently acquired vertical profiles from two satellite sensors (ACE-FTS and MIPAS)[20, 21] are of limited usefulness for probing the emissions, as they sample only the upper troposphere and lower stratosphere.

The new generation IASI/MetOp satellite sensor, launched in 2006, measures in the thermal infrared and has two major advantages compared to its predecessors: high spatial resolution and twice daily global coverage[22, 23]. Recently global day-time measurements of formic acid have been obtained and are discussed in detail in Ref.[8]. By limiting the retrieval to clear sky scenes with a large thermal contrast (>5K), the theoretical error on the total column does not exceed 60%, but leads to fewer high-latitude data and to exclusion of oceanic observations. Figure 1 and Supplementary Figures S1, S2 illustrate the IASI-retrieved HCOOH monthly column abundances in 2009. The enhanced values observed above the mid and high-latitudes of the Northern Hemisphere during the growing season testify to the existence of a strong source, most likely of biogenic origin, since the biomass burning emission patterns do not generally coincide with enhanced columns (Supplementary Figure S3). Elevated values are also observed in the tropics, above densely vegetated areas. Comparison of the IASI columns with the IMAGESv2 model predictions (Fig. 1) corroborate earlier studies reporting large model underestimations[5, 6, 7].

To help interpret this discrepancy, we test the impact of two possible additional sources: (a) the heterogenous oxidation of organic aerosols by OH[7], assumed to form one HCOOH molecule per OH lost, and (b) the generation of 0.5 HCOOH in the photolysis of hydroperoxy-enones from isoprene[19]. These hypotheses acknowledge the poor characterisation of organic aerosol aging and the large uncertainty associated with the isoprene oxidation mechanism. The global modelled annual organic aerosol source ($\sim$100 Tg) accounts for direct emissions and secondary organic aerosol formation (Supplementary Section 4) and results in an extra global annual HCOOH flux of 27 Tg, whereas a larger source (ca. 40 Tg) is issued by the hypothesised HCOOH production through the hydroperoxy-enones (Supplementary Section 2). Although both scenarios lead to significant enhancements in the HCOOH columns allowing for some improvement in the model predictions, especially in tropical regions like Indonesia and Amazonia, they prove inadequate to reconcile the model with the high observed columns in mid and high-latitude areas (Fig. 1, Supplementary Figure S4), underscoring the need for an even larger HCOOH source.

To properly quantify the formic acid source required to reproduce the space-based constraints we use the adjoint source inversion method to infer "top-down" emissions at the resolution of the global model[24] (cf. Methods). Two optimisation experiments are designed and performed, both constrained by monthly IASI columns. Along with the vegetation fire source, we optimise either a direct HCOOH emission from vegetation (Opt1), or, a secondary

HCOOH source from the OH-oxidation of an as-yet-unidentified biogenic precursor, with a global lifetime taken equal to about one day (Opt2, Table 1). These simple settings are meant to represent situations in which the missing HCOOH source is due to direct emission or photochemical formation on very short time frames, or through a collection of compounds leading to HCOOH formation after a number of unspecified intermediate steps.

Both optimisations predict the existence of a substantial biogenic source of HCOOH estimated at about 90-110 Tg annually, i.e. 3.5 times larger than in the a priori budget, as illustrated in Table 1. This brings the contribution of biogenic sources to 90% of the global HCOOH budget, the remainder being due to fires and human activities. The optimisation improves significantly the agreement between the model and IASI columns in terms of both column amplitude and seasonality, also at high northern latitudes where the differences were more pronounced, as shown by comparisons in Fig. 1, Supplementary Figures S1-S2 and Table S1. The extra emission over boreal forests is presumably largely due to oxidation of biogenic volatile compounds from coniferous trees, for which HCOOH formation pathways remain so far unexplored. A large contribution of primary biogenic HCOOH emissions cannot be excluded, but appears less likely, since high emissions of HCOOH from plants are not corroborated by reported flux measurements[25, 26].

The extra secondary HCOOH source inferred from the Opt2 inversion is estimated at 65 Tg annually, of which tropical ecosystems (30 S-30 N) and extratropical latitudes (30-90 N) contribute 40 Tg and 24 Tg annually, re-

spectively. The strong contribution of boreal forests seen in its geographical distribution (Fig. 3) clearly suggests that oxidation of terpenoids emitted by these forests generates substantial amounts of HCOOH. The seasonal variation of the extra source displays a summertime maximum at high latitudes, which is however less pronounced than in isoprene and terpene emission inventories used in global models (Supplementary Figure S5). Interestingly, relatively high emissions are derived also during spring at these latitudes, possibly reflecting higher base emission rates in spring (compared to summer and autumn), as found in recent field studies of monoterpene emissions[27]. In the tropics the HCOOH precursor emission is largest at the end of the dry season, but is weakly correlated with the vegetation fire source, which is however of low magnitude. Note that, should the entire missing source of HCOOH be attributed to monoterpenes only, a total molar yield of 200% HCOOH in the oxidation of monoterpenes would be required; however, the contribution of terpenoid compounds other than isoprene and monoterpenes might be substantial, since field measurements over boreal forests have suggested the existence of large emissions of undetected short-lived organic compounds[28].

Both the magnitude and distribution of the source inferred from IASI are found to be only weakly sensitive to model uncertainties as shown by the results of sensitivity studies detailed in Supplementary Section 8. In particular, we investigated the influence of uncertainties in the wet and dry deposition parameterisations, the chemical scheme, and the assumed errors on the spaceborne data. In most cases, the deduced annual global biogenic

source is found to differ by less than 10% from the reference case, confirming the robustness of the inferred estimates.

In an attempt to evaluate the IASI-derived source, we have conducted extensive comparisons of the model with independent HCOOH measurements (Supplementary Section 5). Comparisons with infrared column measurements (FTIR) at Wollongong and Reunion Island show substantial improvements after optimisation, as seen from the average bias reduction by a factor of 3 or more at both locations (Fig. 2). At ground level, the a priori model underprediction of HCOOH concentration measurements in air and precipitation by large factors gives way to an a posteriori mean underestimation by a factor of about two or less in all regions (Supplementary Figure S6-S7). More precisely, the average gas-phase (liquid) concentration is increased from 0.29 ppbv (1.32 μmol/L) in the a priori simulation to 0.78-1.17 ppbv (4.0-4.2 μmol/L) in the optimisation runs, to be compared to the observed 1.33 ppbv (6.7 μmol/L). Consistent with these results, the optimisation brings also the model significantly closer to airborne concentration measurements above North America and the Pacific (Supplementary Figure S8). Overall, these comparisons lend confidence to the IASI dataset, and therefore to the strong biogenic source inferred from the optimisation. The comparisons, however, also point to a limited ability of the model to capture the observed variability and vertical profile. Although this is largely due to the limited representativity of local measurements, it might also reflect the existence of important shortcomings in the representation of the formic acid sources and

sinks.

We quantify the global impact of the IASI-constrained HCOOH source on precipitation acidity using the calculated wet deposition fluxes of nitrate, sulphate, ammonium, formate and acetate ions (Supplementary Section 7). The inferred decrease in pH due to the additional HCOOH source is estimated at 0.25-0.5 over boreal forests in summertime, and 0.15-0.4 above tropical vegetated areas throughout the year (Supplementary Figure S9). Our model simulations predict that formic acid alone accounts for as much as 60-80% of the rainwater acidity over Amazonia, in accordance with in situ measurements[29], but also over boreal forests during summertime. Its contribution is also substantial at mid-latitudes, in particular over much of the US, where it reaches 30-50% during the summer (Supplementary Figure S10). Given the remaining underestimation of a posteriori modelled HCOOH concentrations against independent measurements (Supplementary Table S5), the large calculated impacts on precipitation acidity are probably conservative estimates. They underscore the importance of HCOOH in acid deposition in different environments, although its consequences for ecosystems are likely limited due to the assimilation of simple carboxylic acids by the biota[30].

In this letter we have revisited the formic acid global distribution and budget, using source inversion constrained by space observations. We found a biogenic source of formic acid of the order of 100 Tg annually, much larger than the current state-of-the-art estimates. A large fraction of this source

originates in boreal and tropical forests, and although it remains mostly unidentified, it is likely to be of secondary origin. These findings suggest that formic acid is a high-yield product in the oxidation of organic compounds emitted abundantly by plant ecosystems, such as monoterpenes and other terpenoids, and underscore the need for further research on the emissions and chemistry of biogenic precursors.

Methods

Modelling

The simulations are performed with the IMAGESv2 global chemistry transport model for the year 2009, after a four-month spin-up time. The model resolution is $4° \times 5°$ with 40 vertical levels. The accompanying Supplementary material includes : details on HCOOH sources and sinks used in the base model (Sections 1-3), a description of the organic aerosol module (Section 4), a description of ground-based and aircraft measurements used for validation purposes and comparisons with the model predictions (Section 5), discussion on the uncertainties of the IASI columns (Section 6), on the impact of formic acid sources on precipitation acidity (Section 7), and on the conducted sensitivity inversions (Section 8). Results are illustrated in Supplementary Figures S1-S12 and Tables S1-S7.

Method for source inversion

The source inversion is realised through minimisation of the cost function

J, which measures the misfit between the model and the observations. By using the adjoint model method, the derivatives of the cost function J are computed with respect to a number of control variables $\mathbf{f}$ (emission parameters)

$$J(\mathbf{f}) = \frac{1}{2}[(H(\mathbf{f}) - \mathbf{y})^T \mathbf{E}^{-1}(H(\mathbf{f}) - \mathbf{y}) + \mathbf{f}^T \mathbf{B}^{-1}\mathbf{f}], \quad (1)$$

where $H(\mathbf{f})$ is the model operator acting on the control variables, $\mathbf{y}$ is the observation vector, $\mathbf{E}$, $\mathbf{B}$ are the covariance matrices of the errors on the observations and the emission parameters $\mathbf{f}$, respectively, and ^{T} is the transpose. The method makes use of a priori distributions for biogenic and pyrogenic emissions, obtained from available inventories (Supplementary Section 1). The adjoint method enables handling of problems with large numbers of control variables, e.g. emission inversions at the model resolution. The cost function is minimised through an iterative descent algorithm which uses the forward and the adjoint model of IMAGESv2. About 50 iterations are usually needed to reach the minimum and derive the "top-down" emission estimates. The norm of the gradient of the cost function is reduced by a factor ≥ 1000 after optimisation. In both optimisations, the pyrogenic and biogenic emission source strengths are updated (ca. 12,000 emission parameters). The errors on the emission parameters, i.e. the square roots of the diagonal elements of $\mathbf{B}$, are assumed to be a factor of two for biogenic emissions, and a factor of 2.5 for biomass burning, whereas spatiotemporal correlations are introduced through the off-diagonal elements[24]. The matrix $\mathbf{E}$ is assumed

diagonal. The errors on IASI columns are taken as the quadratic sum of a 30% relative error and a 4×10^{15} molec.cm^{-2} absolute error. This estimate does not account for a possible bias in the IASI measured columns. The latter could not be quantified due to the lack of correlative measurements. At Wollongong, however, we find IASI columns to be biased 30% low as compared to FTIR ground-based measurements in 2008-2009, with a high degree of correlation (0.84, Supplementary Section 6). Note also that the model inversions do not account for the heterogeneous vertical sensitivity of IASI. The application of a vertical smoothing to the model columns could potentially decrease the global HCOOH source by ca. 10%.

References

[1] Galloway, J. N., Likens, G. E., Keene, W. C. & Miller, J. M. The composition of precipitation in remote areas of the world. *J. Geophys. Res.* **87**, 8771–8786 (1982).

[2] Chameides, W. L. & Davis, D. D. Aqueous-phase source of formic acid in clouds. *Nature* **304**, 427–429 (1983).

[3] Kavouras, I. G., Mihalopoulos, N. & Stephanou, E. G. Formation of atmospheric particles from organic acids produced by forests. *Nature* **395**, 683–686 (1998).

[4] Glasius, M. Sources to formic acid studied by carbon isotopic analysis and air mass characterization. *Atmos. Environ.* **34**, 2471-2479 (2000).

[5] von Kuhlmann, R., Lawrence, M. G., Crutzen, P. J. & Rasch, P. J. A model for studies of tropospheric ozone and non-methane hydrocarbons: Model evaluation of ozone related species. *J. Geophys. Res.* **108**(D23), 4729, doi:10.1029/2002JD003348 (2003).

[6] Ito, A., Sillman, S. & Penner, J. E. Effects of additional nonmethane volatile organic compounds, organic nitrates, and direct emissions of oxygenated organic species on global tropospheric chemistry. *J. Geophys. Res.* **112**, D06309, doi:10.1029/2005JD006556 (2007).

[7] Paulot, F. *et al.* Importance of secondary sources in the atmospheric budgets of formic and acetic acids. *Atmos. Chem. Phys.* **11**, 1989–2013 (2011).

[8] Razavi, A. *et al.* Global distributions of methanol and formic acid retrieved for the first time from the IASI/MetOp thermal infrared sounder. *Atmos. Chem. Phys.* **11**, 857–872 (2011).

[9] Kawamura, K., Ng., L. L. & Kaplan, I. R. Determination of organic acids (C_1-C_{10}) in the atmosphere, motor exhausts and engine oils. *Environ. Sci. Tech.* **19**, 1082–1086 (1985).

[10] Andreae, M. O. & Merlet, P. Emission of trace gases and aerosols from biomass burning. *Global Biogeochem. Cy.* **15**, 955–966 (2001).

[11] Gabriel R., Schäfer, L., Gerlach, C., Rausch, T. & Kesselmeier, J. Factors controlling the emissions of volatile organic acids from leaves of Quercus ilex L. (Holm oak). *Atmos. Environ.* **33**, 1347–1355 (1999).

[12] Sanhueza, E. & Andreae, M. O. Emissions of formic and acetic acids from tropical savanna soils. *Geophys. Res. Lett.* **18(9)**, 1707–1710 (1991).

[13] Graedel, T. E. & Eisner, T. Atmospheric formic acid from formicine ants: a preliminary assessment. *Tellus B* **40**, 335–339 (1988).

[14] Ohta, K., Ogawa, H. & Mizuno, T. Abiological formation of formic acid on rocks in nature. *Appl. Geochem.* **15**, 91–95 (2000).

[15] Neeb, P., Sauer, F., Horie, O. & Moortgat, G. R. Formation of hydroxymethyl hydroperoxide and formic acid in alkene ozonolysis in the presence of water vapor. *Atmos. Environ.* **31**, 1417–1423 (1997).

[16] Müller, J.-F. & Stavrakou, T. Inversion of CO and NOx emissions using the adjoint of the IMAGES model. *Atmos. Chem. Phys.* **5**, 1157–1186 (2005).

[17] Stavrakou, T. *et al.* Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns. *Atmos. Chem. Phys.* **9**, 1037–1060 (2009).

[18] Paulot, F. *et al.* Unexpected epoxide formation in the gas-phase photooxidation of isoprene. *Science* **325**, 730–733 (2009).

[19] Peeters, J. & Müller, J.-F. HOx radical regeneration in isoprene oxidation via peroxy radical isomerisations, II: Experimental evidence and global impact. *Phys. Chem. Chem. Phys.* **12(42)**, 14,227–14,235 (2010).

[20] González Abad, G. *et al.* Global distribution of upper tropospheric formic acid from the ACE-FTS. *Atmos. Chem. Phys.* **9**, 8039–8047 (2009).

[21] Grutter, M. *et al.* Global distribution and variability of formic acid as observed by MIPAS-ENVISAT. *J. Geophys. Res.* **115**, D10303, doi:10.1029/2009JD012980 (2010).

[22] Clerbaux, C. *et al.* Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmos. Chem. Phys.* **9**, 6041–6054 (2009).

[23] Clarisse, L., Clerbaux, C. Dentener, F. Hurtmans, D. & Coheur, P. F. Global ammonia distribution derived from infrared satellite observations. *Nature Geoscience* **2**, 479–483 (2009).

[24] Stavrakou, T. *et al.* First space-based derivation of the global atmospheric methanol emission fluxes. *Atmos. Chem. Phys.* **11**, 4873–4898 (2011).

[25] Kuhn, U. *et al.* Exchange of short-chain monocarboxylic acids by vegetation at a remote tropical forest site in Amazonia. *J. Geophys. Res.* **107**, D20, doi:10.1029/2000JD000303 (2002).

[26] Kesselmeier, J., Bode, K., Gerlach, C. & Jork, E. M. Exchange of atmospheric formic and acetic acids with trees and crop plants under controlled chamber and purified air conditions. *Atmos. Environ.* **32**, 1765–1775 (1998).

[27] Holzinger, R., Lee, A., McKay, M. & Goldstein, A. H. Seasonal variability of monoterpene emission factors for a Ponderosa pine plantation in California. *Atmos. Chem. Phys.* **6**, 1267–1274 (2006).

[28] Di Carlo, P. *et al.* Missing OH reactivity in a forest: evidence for unknown reactive biogenic VOCs. *Science* **304**, 722–725 (2004).

[29] Andreae, M. O., Talbot, R. W., Andreae, T. W. & Harriss, R. C. Formic and Acetic Acid Over the Central Amazon Region, Brazil 1. Dry Season. *J. Geophys. Res.* **93(D2)**, 1616–1624 (1988).

[30] Keene, W. C. & Galloway, J. N. Organic acidity in precipitation of North America. *Atmos. Environ.* **18**, 2491–2497 (1984).

Additional information

Correspondence and requests for materials should be addressed to T. S.

Acknowledgements

This study has been supported by the projects PRODEX A3C of the European Space Agency funded by the Belgian Science Policy Office, and the IBOOT, BIOSOA, AGACC and AGACC-II projects within the "Science for a Sustainable Development" research programme funded by the Belgian Science Policy Office. Financial support by the "Actions de Recherche Concertées" (Communauté Française de Belgique) is also acknowledged. IASI has been developed and built under the responsibility of the Centre National d'Etudes Spatiales (CNES, France). It is flown onboard the Metop satellites as part of the EUMETSAT Polar System. The IASI L1 data are received through the EUMETCast near real time data distribution service. L.C. and P.-F.C. are respectively Postdoctoral Researcher and Research Associate with F.R.S.-FNRS. The Australian Research Council (Grant DP110101948) is gratefully acknowledged for their funding of the Wollongong HCOOH measurements. C.V. and M.D.M. are grateful to the BIRA and LACY team members who support the FTIR observations at Reunion Island.

Author contributions

T.S. and J.-F.M. obtained the results, drafted the manuscript and prepared the figures. J.P. developed the isoprene degradation mechanism used to estimate the photochemical source of formic acid in the model. A.R.,

L.C., P.-F.C., D.H. and C.C. carried out the first retrievals of formic acid observations from space. C.C. contributed also actively in the development of the IASI sensor. C.V. and M.D.M. retrieved the FTIR data at Reunion Island. N.D., D.G., N.J. and C.P.-W. retrieved the FTIR observations at Wollongong.

Competing Financial Interests Statement

The authors declare no competing financial interests as defined by the Nature Publishing Group or other interests that might be perceived to influence the results and discussion reported in this paper.

Figure legends

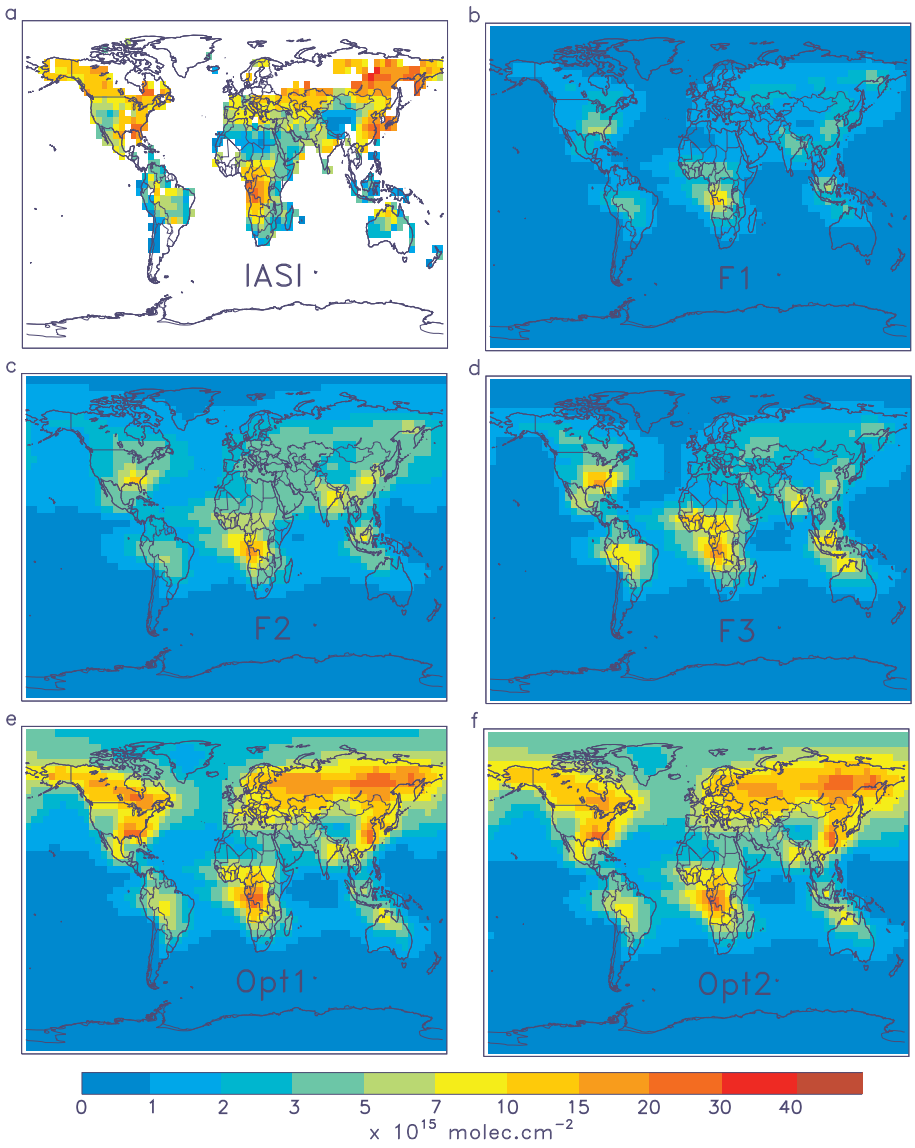

Figure 1: **Monthly averaged HCOOH columns in June 2009 (expressed in 10^{15} molec.cm^{-2}).** **a,** Observed by IASI. **b,** Simulated by the standard model (F1, Table 1). **c,** Simulated assuming that HCOOH is produced in the heterogeneous oxidation of organic aerosols by OH (F2). **d,** Simulated accounting for a production of HCOOH in the photolysis of hydroperoxy-enones from isoprene (F3). **e-f,** Inferred from source inversion assuming either a primary (Opt1) or a secondary (Opt2) biogenic HCOOH source (Table 1).

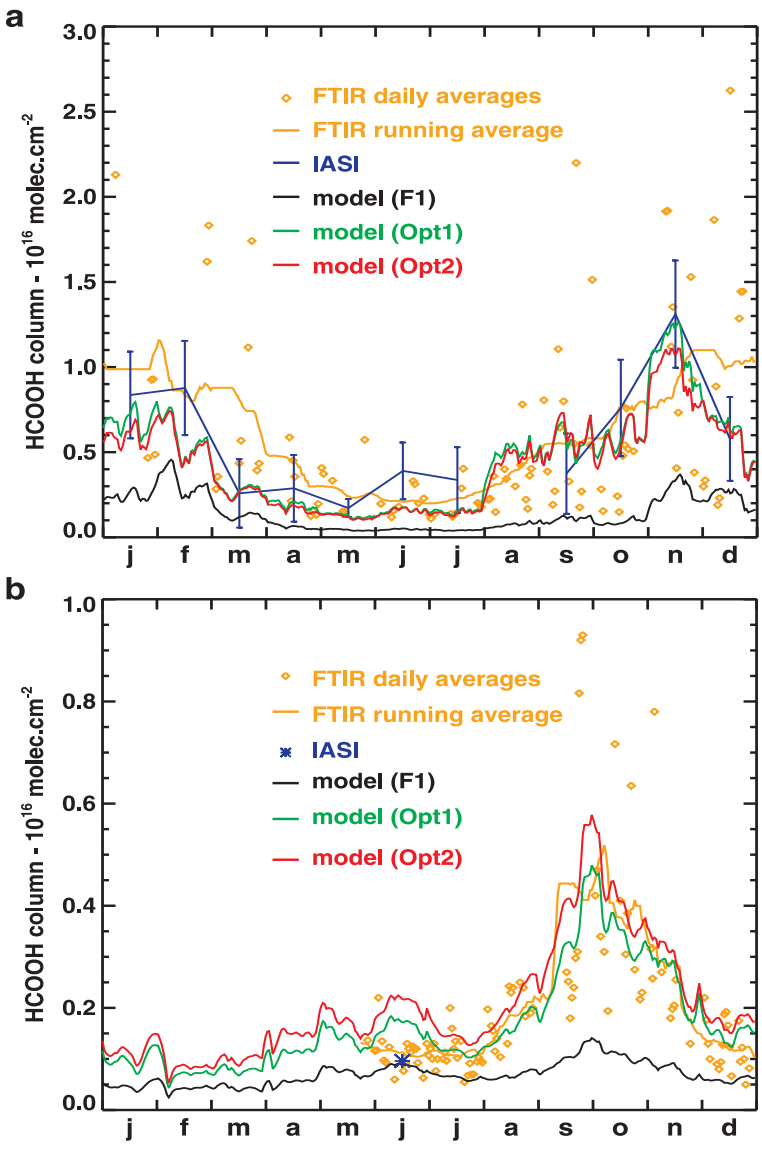

Figure 2: **Comparisons between FTIR, IASI and modelled HCOOH columns in 2009. a,** Daily and 30-day running FTIR column averages at Wollongong (34.41 S, 150.88 E) (orange) and model results from the standard simulation F1 (black) and the two inversions Opt1 (green) and Opt2 (red). Average IASI columns within a radius of 4 degrees around the site and their standard deviations are shown in blue. **b,** Same for Reunion Island (21 S, 55 E). Except for June, IASI columns are not available at this site due to the thermal contrast requirement for their retrieval.

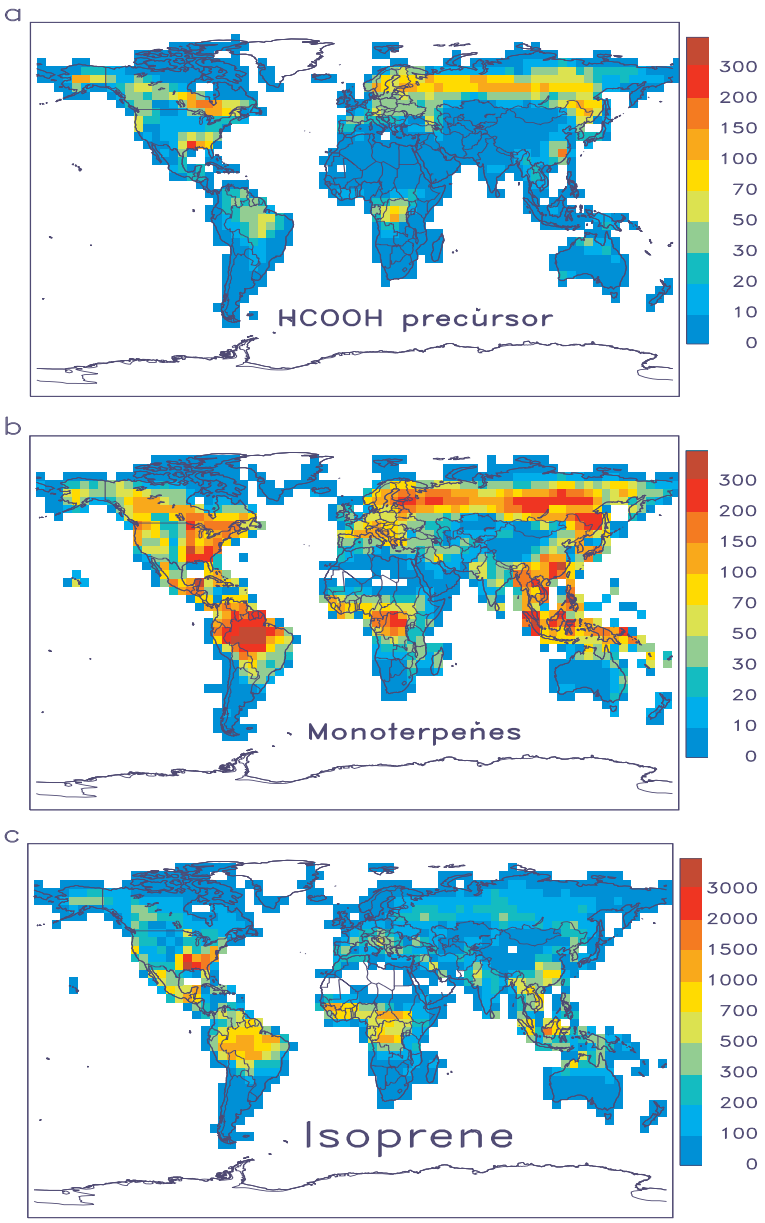

Figure 3: **Global distribution of biogenic emissions in μg C m^{-2} s^{-1} in July 2009. a,** HCOOH precursor as deduced from the Opt2 inversion (Table 1). **b-c,** Monoterpene and isoprene emissions from MEGAN inventory, cf. Supplementary Section 1.

Table 1: Description of the simulations and inversions performed in this study, and global HCOOH budget calculated with IMAGESv2 and comparison with previous work. Notes : (a) includes biofuel burning and cattle emissions, (b) includes a soil source of 1.8 Tg yr^{-1}, (c) of which 19.3 Tg yr^{-1} is due to known sources and 64.7 Tg yr^{-1} to unidentified precursors, (d) A priori budget, but an estimated extra source of 2 Tmol yr^{-1} is invoked to reconcile the model with HCOOH observations, (e) includes a dust sink of 1.38 Tg yr^{-1}.

F1	standard simulation					
F2	including a production of HCOOH in the heterogeneous oxidation of organic aerosols (OA) by OH					
F3	including a production of HCOOH in the photolysis of hydroperoxy-enones from isoprene					
Inversion - Opt1	derive a primary HCOOH biogenic source using IASI					
Inversion - Opt2	derive a secondary HCOOH biogenic source using IASI					
Sources (Tg/yr)	F1	F2	F3	Opt1	Opt2	Paulot et al. 2011
Anthropogenic	4	4	4	4	4	2.28[a]
Pyrogenic	2.9	2.9	2.9	4	4	1.5
Biogenic primary	5.6	5.6	5.6	88	5.6	4.37[b]
Photochemical						
Biogenic	19.3	19.3	58.3	19.3	84[c]	42.2
Anthrop./pyrog.	4.2	4.2	4.2	4.6	4.6	6.35
OA+OH	0	27	0	0	0	(15)
Total	**36**	**63**	**75**	**120**	**102**	**56.7**[d]
Sinks (Tg/yr)						
OH oxidation	9.6	19	20	27	28.4	10.6
Dry deposition	12.7	19.9	26.8	49.5	33.6	26.0[e]
Wet deposition	13.7	24.1	28.2	43.4	40	20.1
Lifetime (days)	4	4.5	3.8	3.5	4.3	3.2

Supplementary Information for

Satellite Evidence for a Large Source of Formic Acid From Boreal and Tropical Forests

T. Stavrakou[1⋆], J.-F. Müller[1], J. Peeters[2], A. Razavi[3], L. Clarisse[3], C. Clerbaux[4,3], P.-F. Coheur[3], D. Hurtmans[3], M. De Mazière[1], C. Vigouroux[1], N. M. Deutscher[5,6], D. W. T. Griffith[5], N. Jones[5], C. Paton-Walsh[5]

[1] Belgian Institute for Space Aeronomy, Avenue Circulaire 3, 1180, Brussels, Belgium
[2] Department of Chemistry, University of Leuven, B-3001, Heverlee, Belgium
[3] Spectroscopie de l'Atmosphère, Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, Belgium
[4] UPMC Univ. Paris 6; Université Versailles St.-Quentin; CNRS/INSU, LATMOS-IPSL, Paris, France
[5] School of Chemistry, University of Wollongong, Wollongong, Australia
[6] Institute of Environmental Physics, University of Bremen, Germany
⋆ E-mail : Trissevgeni.Stavrakou@aeronomie.be

Contents

1 Emissions of Formic Acid in the IMAGESv2 global CTM

The IMAGESv2 global chemistry-transport model is run at $4° \times 5°$ resolution and is resolved at 40 vertical levels from the surface up to the lower stratosphere. The model meteorology is provided by ERA-Interim analyses of the ECMWF data center. The model has been extensively described in previous work [S1, S2, S3, S4], except for the model updates detailed in this Supplement. The model simulations are performed for year 2009, after a four-month spin-up time starting on September 1, 2008.

Global fossil fuel and biofuel NMVOC emissions are obtained from the RETRO[S5] database for the year 2000, which is overwritten by the 2009 REAS[S6] inventory over Asia. The global annual anthropogenic source of formic acid (including biofuel use) amounts to 2.2 Tg yr^{-1}, when the contribution of HCOOH to the carboxylic acids emissions in these inventories is assumed to be 25%. Formic acid emissions from cattle are estimated at 1.8 Tg yr^{-1}[S7]. The emissions of the main biogenic precursors of formic acid, isoprene and the monoterpenes, are both obtained from inventories based on the MEGAN model[S8]. The global annual sources of isoprene and monoterpenes in 2009 amount to 424 Tg and 97 Tg, respectively.

Vegetation fire emissions are obtained from the latest version of the Global Fire Emission Database[S9], by applying updated (in 2007) emission factors[S10]. For HCOOH, the emission factors are 1.13, 2.43 and 0.63 g HCOOH/kg of dry matter for tropical, extratropical, and savanna burning, respectively, and the global biomass burning flux is estimated at 2.9 Tg for the year 2009. Monthly HCOOH biomass burning emissions between May and August 2009 are shown in Fig. S3. Monthly emissions of HCOOH from leaves are the average of the fluxes estimated between 1983 and 1995 by the ORCHIDEE global dynamic vegetation model[S11], and amount at 5.6 Tg yr^{-1}.

2 Photochemical Sources of Formic Acid

This section deals with the photochemical HCOOH formation in the standard (F1, Table 1) version of IMAGESv2.

2.1 Isoprene oxidation by OH

The OH-initiated isoprene oxidation mechanism in IMAGESv2 is based on the new theoretically-based LIM0 mechanism[S12, S13], which has been implemented in IMAGESv2 and evaluated against aircraft observations[S4]. This scheme puts forward that isomerisation reactions of specific isomer/conformer peroxy radicals from isoprene lead to the formation of HOx radicals and photolabile hydroperoxy-enones, most likely hydroperoxy methylbutenals (HPALDs) at rates faster than those of traditional reactions of the isoprene peroxy radicals

with NO and HO_2 in most remote atmospheric regimes, offering thereby a likely explanation for the high abundances of HOx radicals observed in isoprene-rich areas[S14, S15]. Unfortunately, the rates of key isomerisation reactions in LIM0 remain uncertain, in particular the 1,6-H shift of the δ-hydroxy peroxy radicals leading to the formation of the HPALDs. Based on the results of isoprene oxidation experiments conducted at low NOx levels[S16], the theoretical best estimate for these rates appeared to be likely overestimated by a factor of 2 or 3[S13, S17]. However, even such a rate reduction has only limited consequences on the yields of the primary products, due to the strongly non-linear dependence of the yields on the 1,6-H shift rates[S13, S4].

In IMAGESv2 the reaction of isoprene with OH is written as

$$\mathrm{ISOP + OH \rightarrow 0.6\ ISOPBO2 + 0.3\ ISOPDO2 + 0.1\ ISOPEO2,} \qquad \text{(S1)}$$

where ISOPBO2, ISOPDO2 are the peroxys formed through OH addition to the terminal carbons, and ISOPEO2 represent the peroxys formed through addition to the inner carbons. The δ-hydroxy peroxy radicals are not explicitly included, since their traditional reaction products have negligible yields according to LIM0 in most atmospheric conditions. The isomerisation of the isoprene hydroxyperoxys is accounted through the reactions :

$$\begin{array}{lll}
\mathrm{ISOPBO2} & \rightarrow \mathrm{HPALD1 + HO_2} & k_1 = 1.203 \cdot 10^9 \exp(-7127/T) \\
\mathrm{ISOPBO2} & \rightarrow \mathrm{MVK + HCHO + OH} & k_2 = 2.04 \cdot 10^{11} \exp(-8987/T) \\
\mathrm{ISOPDO2} & \rightarrow \mathrm{HPALD2 + HO_2} & k_3 = 1.571 \cdot 10^9 \exp(-7098/T) \\
\mathrm{ISOPDO2} & \rightarrow \mathrm{MACR + HCHO + OH} & k_4 = 2.04 \cdot 10^{11} \exp(-8987/T),
\end{array}$$

where T is the temperature (in K) and MVK, MACR denote methylvinylketone and methacrolein, respectively. The adopted rates k_1–k_4 (in s^{-1}) account for a factor of 2 reduction of the theoretical best estimates for the 1,6-H shift rates of the δ-hydroxy peroxy radicals[S4]. Photolysis is believed to be the main sink reaction of the HPALDs accounting for about two thirds of the global sink[S4], and leading to the regeneration of one or several OH radicals. However, the further oxidation of the photolysis products remains to be elucidated. In the current version of the mechanism[S4], the photolysis of the HPALDs is represented as

$$\mathrm{HPALD1 + h\nu \rightarrow 3\,OH + HO_2 + 0.5(\,HYAC + MGLY + GLYALD) + HCHO} \qquad \text{(S2)}$$

$$\mathrm{HPALD2 + h\nu \rightarrow 3\,OH + HO_2 + 0.5(\,HYAC + MGLY + GLY) + HCHO,} \qquad \text{(S3)}$$

where GLY, MGLY, HYAC, and GLYALD are abbreviations for glyoxal, methylglyoxal, hydroxyacetone and glycolaldehyde, respectively. The possibility that HPALD oxidation leads to additional HCOOH formation is explored in the F3 simulation (Table 1).

A yield of 6% of isoprene nitrates (ISOPN) is assumed in the reaction of isoprene peroxy radicals with NO. Their oxidation by OH leads to HCOOH formation[S18], as summarised below :

$$\mathrm{ISOPN + OH} \rightarrow 0.6\,(\mathrm{GLYALD + HYAC + NO_2}) + 0.4\,(\mathrm{HCHO + HO_2}) + 0.14\,\mathrm{MVKN} + 0.26\,\mathrm{MACRN} \quad \text{(S4)}$$

$$\mathrm{MVKN + OH} \rightarrow 0.65\,(\mathrm{MGLY + HCOOH}) + 0.35\,(\mathrm{HCHO + CH_3COCOOH}) + \mathrm{NO_3} \quad \text{(S5)}$$

$$\mathrm{MACRN + OH} \rightarrow 0.85\,(\mathrm{HYAC + NO_2}) + 0.08\,(\mathrm{CH_3COOH + HCHO}) + 0.07\,(\mathrm{HCOOH + MGLY}) + 0.15\,\mathrm{NO_3}, \quad \text{(S6)}$$

where MVKN and MACRN are nitrates produced in the oxidation of MVK and MACR by OH :

$$\mathrm{MVK + OH} \rightarrow \mathrm{MVKO2} \quad \text{(S7)}$$
$$\mathrm{MVKO2 + NO} \rightarrow 0.11\,\mathrm{MVKN} + 0.623\,\mathrm{GLYALD} + \mathrm{products} \quad \text{(S8)}$$
$$\mathrm{MACR + OH} \rightarrow 0.57\,\mathrm{MCO3} + 0.43\,\mathrm{MACRO2} \quad \text{(S9)}$$
$$\mathrm{MACRO2 + NO} \rightarrow 0.15\,\mathrm{MACRN} + \mathrm{products}. \quad \text{(S10)}$$

The oxidation by OH of isoprene hydroxy-hydroperoxides issued by the reaction of isoprene peroxy radicals with HO_2 leads to the formation of epoxides[S16] (IEPOX) :

$$\mathrm{ISOPBOOH + OH} \rightarrow \mathrm{IEPOX + OH} \quad \text{(S11)}$$
$$\mathrm{ISOPDOOH + OH} \rightarrow \mathrm{IEPOX + OH}. \quad \text{(S12)}$$

Their further oxidation by OH is assumed[S16] to lead to (among others) HCOOH and GLYALD :

$$\mathrm{IEPOX + OH} \rightarrow \mathrm{IEPOXO2} \quad \text{(S13)}$$

$$\mathrm{IEPOXO2 + NO} \rightarrow 0.074\,\mathrm{HCOOH} + 0.275\,\mathrm{GLYALD} + \mathrm{products} \quad \text{(S14)}$$
$$\mathrm{IEPOXO2 + HO_2} \rightarrow 0.074\,\mathrm{HCOOH} + 0.275\,\mathrm{GLYALD} + \mathrm{products}. \quad \text{(S15)}$$

Note however that HCOOH is not produced from IEPOX according to the protocol rules of the Master Chemical Mechanism[S17].

According to a laboratory study of the Orléans group[S19], the oxidation of GLYALD is also a source of HCOOH :

$$\mathrm{GLYALD + OH} \rightarrow 0.17\,(\mathrm{GLY + HCOOH}) + 0.67\,\mathrm{HCHO} + 0.75\,\mathrm{HO_2} + 0.25\,\mathrm{OH} + 0.5\,\mathrm{CO} + 0.34\,\mathrm{CO_2}. \quad \text{(S16)}$$

Glycolaldehyde is mainly produced from the oxidation of MVK, ISOPN, IEPOX and the HPALDs (see above). It is also a major product in the oxidation of ethylene by OH[S20]. Another laboratory study by the Orléans group[S21] proposed that HCOOH is formed in the OH-oxidation of hydroxyacetone, a major secondary product in the oxidation of isoprene; however, the proposed formation mechanism appears rather unusual. More recent experiments[S22] give evidence that this production does not occur at atmospheric conditions, and hence is not accounted in this study.

2.2 Ozonolysis of alkenes

The ozonolysis of alkenes proceeds via the addition of O_3 to the C=C bond to form a primary ozonide which rapidly decomposes to one or two sets of Criegee intermediate (CI) plus a carbonyl[S23, S24]. In the case of terminal alkenes ($H_2C{=}R_1R_2$), ozonolysis forms the smallest CI, $[H_2COO]^*$:

$$H_2C{=}CR_1R_2 + O_3 \rightarrow [H_2C(OOO)CR_1R_2]^* \quad \text{(S17)}$$

$$[H_2C(OOO)CR_1R_2]^* \rightarrow [H_2COO]^* + R_1C(O)R_2 \quad \text{(S18)}$$

$$[H_2C(OOO)CR_1R_2]^* \rightarrow [R_1R_2COO]^* + HCHO. \quad \text{(S19)}$$

Due to the high exothermicity of the overall reaction (S17-S19), both products are chemically activated and only a fraction of the energy-rich CIs becomes collisionally stabilised, while the rest undergoes various "prompt" unimolecular reactions. This stabilised fraction is close to 0.37 (average of 5 experimental studies)[S25] for $[H_2COO]^*$ originating from C_2H_4 + O_3. Larger values are expected in the case of higher alkenes, for which a larger fraction of the energy will be partitioned to the larger fragment $R_1C(O)R_2$. Nevertheless, we adopt here the value of 0.37 for the cases of C_3H_6, MACR and MVK.

In tropospheric conditions, reaction of the stabilised Criegee intermediates (SCIs) with H_2O is generally by far the dominant sink of the SCIs. Laboratory experiments[S24, S26, S27, S28, S29] of alkene ozonolysis indicate that the reaction of H_2COO with H_2O forms formic acid and hydroxymethyl hydroperoxide (HMHP). It has been shown[S24, S28], that the formation of formic acid in the laboratory was partly or entirely due to the decomposition of HMHP to HCOOH + H_2O. The precise extent to which this formation occurred by heterogeneous reactions on the reactor walls is uncertain. A theoretical investigation[S30] confirmed that the reaction of H_2COO with H_2O leads to the formation of HMHP. The predominant pathway in atmospheric conditions was found to be the reaction with the H_2O dimer: the rate constants for the reaction of H_2COO with the H_2O monomer and the water dimer were calculated to be $8.2\cdot 10^{-18}$ and $1.46\cdot 10^{-12}$ molec.$^{-1}$cm^{3}s^{-1}, respectively.

Assuming that HMHP is the only product in the reaction of H_2COO with H_2O, HMHP yields in the ozonolysis of C_2H_4 (0.37), C_3H_6 (0.16), MVK (0.32) and MACR (0.22) are adopted, based on the $[H_2COO]^*$ stabilisation fraction

observed for C_2H_4 + O_3 (0.37) and on the IUPAC recommendation[S25] for the $[H_2COO]^*$ branching ratio in the alkene + ozone reaction. For isoprene, we adopt the average HMHP yield from two studies (0.23)[S24, S31]. The case of monoterpenes is discussed in the subsection 2.3.

HMHP has been commonly observed at levels comparable to those of methyl hydroperoxide in the boundary layer over isoprene-rich areas[S32, S33, S34, S35], supporting the view that isoprene and other biogenic alkenes are an important source of HMHP. These observations also show that fast decomposition of HMHP is unlikely at atmospheric conditions. We therefore neglect HMHP decomposition in the model. Should some decomposition (to HCOOH + H_2O) occur, our assumption would not lead to a large underestimation of HCOOH formation, because the atmospheric degradation of HMHP leads primarily to HCOOH anyway, as detailed below.

HMHP photolysis proceeds by cleavage of the peroxidic bond[S25]:

$$HOCH_2OOH + h\nu \rightarrow HOCH_2O + OH, \tag{S20}$$

which is immediately followed by

$$HOCH_2O + O_2 \rightarrow HCOOH + HO_2. \tag{S21}$$

The HMHP absorption cross sections are slightly lower than those of methyl hydroperoxide[S36], which makes photolysis a very minor sink. Reaction with OH proceeds by H-abstraction from the carbon, from the hydroperoxide or from the hydroxy group[S37]:

$$\begin{aligned} HOCH_2OOH + OH &\rightarrow HOCHOOH + H_2O && \text{(S22)} \\ &\rightarrow HOCH_2O_2 + H_2O && \text{(S23)} \\ &\rightarrow OCH_2OOH + H_2O. && \text{(S24)} \end{aligned}$$

In lower tropospheric conditions, $HOCH_2O_2$ decomposes thermally to HCHO and HO_2[S38, S39]. The radicals formed in Reactions (S22) and (S24) decompose rapidly to HCOOH + OH[S40] and HC(O)OOH + H, respectively[S37]. The first channel (Reaction S22) is by far dominant[S37], as confirmed by the experimental HCOOH yield[S36] of 95% in the reaction of HMHP with OH. As the stronger hydroxy O–H bond should keep the third channel negligible, we write Reactions (S22)-(S24) as

$$HOCH_2OOH + OH \rightarrow 0.95\,HCOOH + 0.95\,OH + 0.05\,HCHO + 0.05\,HO_2 \tag{S25}$$

with an adopted reaction rate of $3 \cdot 10^{-11}$ molec.$^{-1}$ cm^3 s^{-1}.

HMHP also undergoes dry and wet deposition. Dry deposition is calculated using the deposition velocity calculated for H_2O_2. The parameterisation of wet deposition accounts for the high Henry's law's constant of HMHP[S41], $1.24{\cdot}10^{-8}$ exp$(9700/T)$ M atm^{-1}.

An additional loss process for HMHP is provided by its dissolution in liquid cloud droplets followed by (water-assisted) decomposition to HCHO + H_2O_2[S41, S42]. Whereas the decomposition lifetime is less than an hour at neutral pH conditions, it is relatively slow at typical pH values found in clouds (4-5), with a lifetime ranging between several hours and several days[S42]. It is therefore neglected in the model.

2.3 Oxidation of monoterpenes

HMHP can also be produced in the ozonolysis of exo-cyclic (or epi-cyclic) monoterpenes (e.g. β-pinene and sabinene[S25], but not α-pinene). For such large alkenes, decomposition of the primary ozonide to HCHO and a large Criegee intermediate (reaction S19) is largely dominant (> 80%)[S43]. In this case, even though the energy channelled to $[H_2COO]^*$ in the minor ozonolysis route (reaction S18) is at most a few kcal/mol, such that all will be stabilized, the yield of HMHP remains small, of the order of 20% or less. This formation is neglected in the model. Instead, a direct HCOOH formation is assumed, with a yield of 10% in the reactions of monoterpenes with O_3 and OH, to account for the possible but very uncertain formation of HCOOH, mostly from secondary chemistry, in the still largely unexplored chemical mechanism of monoterpenes. First-generation molar yields of formic acid were found to range between 2 and 11% in laboratory studies of monoterpene oxidation by OH[S44, S45] and O_3[S46]. Evidence for higher yields due to the further degradation of the primary products is provided by the OH-initiated monoterpene oxidation experiments of Larsen *et al.*[S47]. Yields ranging between 28% (for α-pinene) and 54% (for limonene) were obtained after 20 minutes of intense photooxidation. Since the total yield of identified small organic compounds (HCOOH, HCHO, CO, CO_2 and acetone) represented between 28% and 46% of the reacted terpene on a carbon basis in these experiments, the further oxidation of the remaining products (i.e. 54-72% of the carbon flux) might bring the overall molar yield of HCOOH in these laboratory conditions to much larger values, possibly of the order of 100%. Caution is required, however, since an unrealistic radiation source (Hg lamp peaking at 254 nm) was used in these experiments, which probably caused a strong photolysis of aldehydic compounds. Still, these results clearly suggest that HCOOH is a likely end-product in a large number of terpene degradation pathways.

3 Sinks of Formic Acid

The only gas-phase sink of formic acid is reaction with OH, represented in the model as

$$HCOOH + OH \rightarrow CO_2 + H_2O + HO_2. \quad \text{(S26)}$$

Its rate expression, k=2.27·10^{-14}· exp(786/T) + 9.85·10^{-13}· exp(-1036/T) implies higher rate values in the cold upper troposphere, compared to boundary

layer conditions[S7]. The use of a temperature-independent rate[S25] will be tested in a sensitivity inversion, as described in Section 8.

The dry deposition velocity over land is parameterised as a function of the leaf area index (LAI) obtained from MODIS[S48]. It increases linearly from 0 to 1.2 cm s^{-1} for an LAI of 6 m^2m^{-2}. Wet scavenging is calculated based on ECMWF precipitation, cloud fraction and convective updraft fields[S49]. The effective Henry's law constant (in M atm^{-1}) is calculated with[S50, S51]

$$K_{H,\text{eff}} = [2.5 \cdot 10^{-5} \cdot \exp(5700/T)] \cdot (1 + 1.8 \cdot 10^{-3}/[\text{H}^+]) \tag{S27}$$

and assuming a pH of 4.5.

Finally, the effect of reversible uptake by ice clouds is also considered. Partitioning to ice particles is calculated with

$$[X_s]/[X_g] = S_{\text{ice}} \cdot K_{\text{linC}} \tag{S28}$$

with $[X_s]$ and $[X_g]$ the ice surface and gas phase concentrations, respectively (in molec.cm^{-3}), S_{ice} the surface area density of ice (cm^2cm^{-3}) and K_{linC} (in cm) the partitioning coefficient in the linear regime of the adsorption isotherm[S52]

$$K_{\text{linC}} = 4 \cdot 10^{-12} \cdot \exp(7000/T), \tag{S29}$$

where it has been taken into account that the fractional surface coverage ($\theta = K_{\text{linC}} \cdot [X_g]/\text{N}_{\max}$) is much less than unity in atmospheric conditions. The ice surface area density is parameterised based on the ice water content (IWC, in g m^{-3}) assuming a scaling ratio of 10 between the cross-sectional area and the surface area density[S53, S54]:

$$S_{\text{ice}} = 10^{-3} \cdot \text{IWC}^{0.9} \tag{S30}$$

The effect of cirrus gravitational settling on the vertical distribution of HCOOH is estimated using Eq. (S28) and the parameterisation of the ECMWF IFS model for the ice settling velocity[S49]. This effect, however, is found to be very small.

4 Organic Aerosol Modeling in IMAGESv2

This section presents the organic aerosol module in the model. The heterogeneous oxidation of organic aerosols by OH is assumed to generate HCOOH in the F2 run (Table 1). Three sources of organic aerosol (OA) are considered :

- Direct emission of Primary Organic Aerosol (POA) due to anthropogenic sources[S55] (13.5 TgC/year or 23.6 Tg/year) and biomass burning[S9] (12.2 TgC/year or 21.3 Tg/year). POA has hydrophilic and hydrophobic components. Conversion of hydrophobic to hydrophilic POA due to aging is represented by an e-folding time of 1 day[S56].

- Reversible Secondary Organic Aerosol (SOA) formation due to the partitioning of semi-volatile organic compounds. We adopt the approach of the two-product model[S57, S58] for estimating this production of SOA from isoprene, terpenes and aromatics. The product yields and partitioning parameters of the condensable compounds from aromatics[S57], isoprene[S59] and sesquiterpenes[S58] are based on laboratory data. For monoterpenes, we adopt a ten-product model[S60], with parameters derived from simulations using a detailed α-pinene oxidation and SOA formation model (BOREAM)[S61, S62]. The model accounts for a dependence of the yields on the NOx regime and on the nature of the primary oxidant (O_3, OH or NO_3). The partitioning coefficients (not only for α-pinene SOA, also for other condensable compounds) are corrected to account for water uptake, based on BOREAM model simulations[S60].

- Irreversible SOA formation due to the production of essentially non-volatile compounds (e.g. oligomers and polymers). In IMAGESv2, the formation of SOA from glyoxal and methylglyoxal is included as an irreversible process[S63, S49]. It is assumed to be very fast in liquid clouds. On aqueous aerosols, a constant reactive uptake coefficient ($2.9{\cdot}10^{-3}$) is assumed.

The model accounts for wet and dry deposition of aerosols and aerosol precursors. The wet removal scheme[S49] is based on ECMWF cloud and precipitation data. The semi-volatile gaseous SOA precursors are assumed to be as water-soluble as glyoxal, given the high solubility of typical multi-functional SOA components.

The global SOA source calculated by IMAGESv2, ca. 100 Tg/year, compares well with the recent estimate of 140 $\pm$ 90 Tg/year derived from aerosol mass spectrometer (AMS) measurements[S64]. Note however that our modelled SOA is overwhelmingly ($> 90\%$) biogenic, whereas AMS measurements suggest a large (100 Tg/year) anthropogenically-controlled contribution. Comparisons with surface OC concentrations measurements of the IMPROVE (Interagency Monitoring of Protected Visual Environments)[S65] and EMEP (European Monitoring and Evaluation Programme)[S66] network suggests that our modelled SOA source is overestimated by almost a factor of two over the Eastern US during summertime and that it is largely underestimated over Europe.

5 Ground-based and Aircraft Measurements

To evaluate the model performance, we use Fourier Transform Infrared (FTIR) column measurements of HCOOH at Wollongong in southeast Australia (34.41 S, 150.88 E), and Reunion Island in the southern Indian Ocean (21 S, 55 E)[S7]. The retrievals adopted the improved spectroscopic parameters for the HCOOH ν_6 absorption band near 1105.4 cm^{-1} (HITRAN 2008)[S67, S68].

Further, we use a compilation of in situ concentration measurements in surface air and in precipitation water (Tables S2, S3 and S4). The observed concentrations are compared with the modelled concentrations in 2009, interpolated at the measurement locations and averaged over the same months (Figs. S6, S7). The significant dispersion before and after optimisation most likely reflects the large spatiotemporal variability of HCOOH and the limited representativity of local measurements from many different years in comparison with model values for 2009 at $4^\circ \times 5^\circ$ resolution. When considering all measurements, the correlation coefficient is equal to 0.5 for surface air concentrations and 0.3 for precipitation water concentrations in the a priori simulations. After optimisation, the correlation coefficient for both air and water concentrations is equal to ca. 0.45.

Aircraft campaign measurements of HCOOH from three expeditions are also used for comparisons (Fig. S8) :

1. The PEM-Tropics A (Pacific Exploratory Mission-Tropics A) mission conducted in August-October 1996 as part of NASA's Global Tropospheric Experiment (GTE) with focus on the remote tropical regions of the Pacific ocean. The data, compiled onto a 5×5 degrees grid with a vertical resolution of 1 km[S69], can be accessed through the data composites web page (http://acd.ucar.edu/ ˜emmons/DATACOMP/camp_table.htm).

2. The SONEX (Subsonic assessment, Ozone and Nitrogen oxide EXperiment) airborne field campaign conducted between October and November 1997 aboard a DC-8 NASA aircraft in the vicinity of the North Atlantic flight corridor[S70].

3. The National Aeronautical Space Administration (NASA) Intercontinental Chemical Transport Experiment, Phase B (INTEX-B) aircraft mission conducted in spring 2006 (1 March to 15 May) over the region of Mexico in March (first phase) and over the Pacific in April and May (second phase)[S71].

A comparison between the observed and the modelled HCOOH concentrations averaged over large regions before and after optimisation is provided in Table S5.

6 Uncertainties on IASI HCOOH columns

The error of 60% on the IASI HCOOH columns is a theoretical error based only on forward simulations for an individual measurement. Other sources of error include propagating errors from level 2 data (e.g. total water column, thermal contrast) and possible bias errors, such as uncertainty in the assumed profile of formic acid or on the spectroscopy. The total error can only be rigorously quantified with an extensive validation of the IASI columns but unfortunately

there are currently few correlative measurements of the formic acid total columns for such an analysis.

An explicit comparison with FTIR ground-based measurements at Wollongong and Reunion Island site (Figure 2 of the letter and Fig. S11) shows a high degree of correlation ($R^2 = 0.84$). Data from other FTIR sites, like Thule and Bremen have been obtained from Figs. 3 and 4 of Ref. [S7]. However, at these sites spatially co-located IASI measurements are not available for comparison, unless a very large $8°\times8°$ area is considered around each station, making the comparison of very little value for validation. From these restricted comparisons there is thus no indication for a significant IASI overall bias.

The error adopted in the inversions consists in an absolute error of 4×10^{15} molec.cm^{-2} and a 30% relative error. This was chosen in order to avoid overly optimistic error bars for the low concentrations. We acknowledge the crudeness of this error estimate, and therefore, sensitivity inversions using either doubled or halved errors were performed. These tests are presented and discussed in Section 8.

7 Impact on Precipitation Acidity

The pH of precipitation is estimated in the model based on the wet deposition fluxes of HNO_3/NO_3^-, SO_4^{2-}, NH_3/NH_4^+, CO_2/HCO_3^- and the carboxylic acids HCOOH and CH_3COOH, accounting for their respective dissociation constants, $1.8{\cdot}10^{-4}$ and $1.8{\cdot}10^{-5}$ mol/L[S51]. Dust aerosols, and therefore, alkaline components such as Ca^{2+}, Na^+ and K^+, are neglected in the model, leading to a likely underestimation of precipitation pH.

8 Sensitivity Inversions

This section explores the impact of possible errors in key model parameters on the inversion results. In each sensitivity inversion (Table S6), a parameter or input dataset of the Opt2 standard optimisation is varied, and new emissions are obtained. The resulting global emissions and sinks are summarised in Table S7.

Acknowledging the broad IASI column error estimate used in the reference optimisations (see "Methods" section), additional inversions were performed using either doubled (Opt2-Errx2) or halved (Opt2-Err:2) IASI errors. Changing these errors affects the balance between the first term of the cost function (cf. Eq. (1), "Methods" section) representing the model-data bias and the second term which measures how far the optimised emissions are from the a priori. Halving the errors results in an almost negligible increase (6% globally). Doubling the IASI column errors leads to a decrease (13%) in the optimised emissions, as the optimisation is more constrained by the a priori. The latitudinal distribution of the emissions (Fig. S12) remains unaffected, except in the summertime Southern hemisphere in the 20°-35° S band, where the optimisation results appear less robust. However, the comparison with ground-based

FTIR data at Wollongong (Fig. 2) indicates that IASI errors are not likely to be larger than our reference estimates.

Dry deposition accounts for about one third of the global HCOOH sink in IMAGESv2, based on a simple parameterisation. The contribution of dry deposition could be higher, however, as reactive uptake by vegetation has been recently shown to enhance the deposition of many oxygenated VOCs[S72]. We therefore conducted a source inversion with HCOOH deposition velocities over land increased by 50% (Opt2-Dry). In spite of this strong increase, the inferred global HCOOH emission is increased by only 6% (Table S7), because dry deposition over land represents only a relatively small fraction of the global sink, and because the stronger surface sink reduces the vertical tropospheric gradient and therefore the surface HCOOH concentrations.

Uncertainties related to wet scavenging (about 38% of the global HCOOH sink in IMAGESv2) include uncertainties in the meteorological fields and in the gas/liquid partitioning ratio for HCOOH, calculated from the effective Henry's law constant for HCOOH, itself dependent on pH. Using an alternative dataset for convective and stratiform precipitation (the NCAR/NCEP Reanalysis[S73]) (http://www.esrl.noaa.gov/psd/) in inversion Opt2-NCEP results in a longer HCOOH lifetime with respect to wet deposition (10.5 vs. 12.2 days), primarily because stratiform precipitation in NCEP is much less widespread at tropical latitudes, compared to ERA-Interim fields. This change has however only very little impact on the optimised HCOOH emissions, cf. Table S7.

Including an on-line dynamical calculation of cloud pH in the model leads to a negligibly (3%) longer HCOOH lifetime, compared to the reference model calculations assuming pH=4.5. However, the pH might be underestimated, given that alkaline components from e.g. dust are neglected. The results of optimisation Opt2-pH5, assuming pH=5, indicate a low sensitivity of the overall HCOOH sink to the cloud pH: the global lifetime is decreased by about 5%, and the global HCOOH source needed to match the IASI column measurements is increased by about 4 Tg yr^{-1}.

The rate used in the model for HCOOH+OH, and obtained from a theoretical study[S74], implies a negative temperature dependence in the atmospheric range, leading to higher rates in the upper troposphere. However, laboratory measurements conducted at and above 297 K (IUPAC recommendation[S25]) indicate a negligible temperature dependence. The sensitivity inversion Opt2-kOH conducted with the IUPAC recommendation ($4.5{\cdot}10^{-13}$ molec.$^{-1}$cm^3 s^{-1}, independent of temperature) leads to a 7% increase of the global HCOOH lifetime, in spite of the higher value of the rate in the lower troposphere. The global biogenic source of HCOOH is decreased by almost 4 Tg yr^{-1} (i.e. by 4%) compared to the reference optimisation.

The isoprene chemical mechanism used in the model includes peroxy radical isomerisations leading to compounds (hydroperoxy-enones) for which the subsequent chemistry remains largely unexplored; furthermore, the isomerisation rates are themselves uncertain. In optimisation Opt2-Isom, these isomerisations are simply ignored. This mechanistic change increases the production

of HCOOH due to isoprene oxidation by 5.7 Tg yr^{-1}, mostly due to enhanced HCOOH formation through (1) the epoxide channel of the isoprene peroxys, and (2) HMHP formation due to isoprene ozonolysis. The contribution of ozonolysis to the isoprene sink is increased from 8% to 14.5% when the 1,6 H-shifts are suppressed, due to lower OH abundances over forested areas. The global production of isoprene epoxides is increased from 40 to 106 Tg annually. Isoprene ozonolysis and the epoxide oxidation by OH are both significant pathways to HCOOH formation in our mechanism. An additional effect of the mechanistic change is the increased overall lifetime of HCOOH, from 4.3 to 4.6 days, also due to the lower OH levels calculated when the isomerisation reactions of the isoprene peroxys are ignored. Note that the higher OH levels calculated with the LIM0 mechanism are supported by comparisons with observations over Amazonia and over the US[S4]. As a consequence of the higher formation rate and overall lifetime of HCOOH in this sensitivity inversion, the annual global source needed to match the IASI measurements is slightly decreased from 102 to 98 Tg.

The optimisation results for the biogenic HCOOH source are also dependent on assumptions for the pyrogenic source, also retrieved in the optimisations. E.g. increasing the pyrogenic a priori emissions, or increasing the fraction of biomass burning emissions released at high altitudes, where HCOOH is longer-lived, could in principle increase the contribution of fires to the total column, thereby reducing the need for large biogenic emissions. The optimisation Opt2-Inj uses a fire injection profile[S75] with substantially higher fractions injected in the middle (30%) and upper (30%) troposphere, compared to the profile[S76] used in IMAGESv2. The results (Table S7) show that the vertical distribution of pyrogenic emissions has only a negligible influence on the derived budget and lifetime. The optimised global emissions due to vegetation fires is decreased (from 4.0 to 3.5 Tg yr^{-1}) when adopting the alternative injection height profile, as a consequence of the longer lifetime of formic acid at higher altitudes. This confirms the minor role played by vegetation fires in determining the HCOOH column distribution.

In conclusion, the optimisation results appear to be robust, since the annual global biogenic source is found to range between 77 and 96 Tg, when the additional HCOOH production is assumed to be secondary. The lowest emissions (77 Tg yr^{-1}) are obtained when assuming unrealistically large errors on the IASI column data. Still, it must be acknowledged that the real uncertainty might be larger, since the above sensitivity experiments cannot possibly cover the full range of possible uncertainties.

9 Tables

Table S1: Mean percentage biases (model-observations) and correlation coefficients calculated for the standard model simulation F1 and the Opt2 inversion constrained by IASI data. Results from the Opt1 inversion are very similar to those from Opt2 and are not shown here. N is the number of observational elements used in the inversion, i.e. monthly averaged IASI HCOOH columns at the resolution of the model.

Region	*N*	*Mean bias*		*Correlation coefficient*	
		F1	Opt2	F1	Opt2
N.America	475	-80.8	-11.0	0.43	0.72
30-70 N, 60-170 W					
Tropical America	489	-72.3	-12.5	0.84	0.91
30 S-20 N, 30-80 W					
Africa	814	-71.7	-10.8	0.75	0.93
40 S-30 N, 20 W-45 E					
Asia	856	-73.2	-14.2	0.34	0.80
10 S-55 N, 60-160 E					
Oceania	402	-72.9	-14.1	0.44	0.84
40 S-10 N, 110-170 E					
Boreal (55-88 N)	419	-85.8	-13.4	0.33	0.71
Tropics (25 S-25 N)	1869	-70.5	-12.2	0.70	0.91
Globe	3768	-75.4	-13.6	0.59	0.87

Table S2: In situ measurements of HCOOH mixing ratio expressed in pptv.

Location	Coordinates	*Period*	HCOOH	*Ref.*
North America				
Mt. Lemmon, Tucson	32.4N 249.2E	Feb	130	[S77]
Socorro, New Mexico	36N 254E	Jun-Aug	700	[S78]
Langmuir, New Mexico	33.9N 253E	Jun-Aug	600	id.
id.	id.	Jun-Aug	1050	id.
Pagosa Springs, Colorado	37.2N 253E	Jun	1310	[S77]
Wolf Creek, Colorado	37.5N 253.3E	Jun	1190	id.
NW Dakota	45.5-49N			
	256-260E	Jun	2640	id.
id.	id.	Jul	1810	id.
id.	id.	May	110	id.
Oregon, Cape Meares	45.5N 236E	Sep	180	id.
id.	id.	Oct	60	id.
Pensylvania	41N 283E	Jul-Aug	2500	[S79]
Charlottesville, Virginia	38N 281.7E	Jun	1820	[S80]
Shenandoah park, Virginia	38N 281.7E	Sep	5400	[S81]
Virginia	37N 283.7E	Mar-Sep	1890	[S82]
Virginia	id.	Oct-Feb	695	id.
South America				
Venezuelan savannah	8N 297E	Apr	450	[S83]
Venezuelan savannah	8.5N 298E	Sep	450	[S84]
N. Venezuela	10.5N 293E	Mar	1700	[S85]
id.	id.	Jun	710	id.
id.	id.	Oct	960	id.
id.	id.	Dec	610	id.
La Selva, Costa Rica	10.4N 276.1E	Apr-May	905	[S86]
Central Amazonia	4S 300E	Jul-Aug	1590	[S87]
Rondonia, Brazil	10S 298E	May	1600	[S88]
id.	id.	Sep-Nov	10500	id.
Ducke Reserve-Manaus	3S 300E	Jul-Aug	1880	[S80]
id.	id.	Apr-May	510	[S89]
Other regions				
Schoeneben, Austria	48.7N 13.9E	Mar-Apr	1066	[S90]
id.	id.	Sep	620	id.
N.Congo rainforest	2N 18E	Feb	500	[S91]
Mayombe, SW Congo	4.5N 12.5E	Jul-Aug	260	[S92]
Dayalbagh, India	28.5N 77.2E	Jan-Dec	1700	[S93]
id.	id.	May-Oct	1300	id.
id.	id.	Nov-Apr	1800	id.
Gopalpura, India	27N 78E	Jul-Sep	1700	[S94]

Table S3: In situ measurements of HCOOH concentration in precipitation (μmol/L, volume-weighted averages) in North and South America.

Location	Coordinates	*Period*	HCOOH	*Ref.*
North America				
Virginia	38.1N 281.2E	Apr-Sep	13.6	[S95]
Hampton,Virginia	38N 281.7E	Mar-Sep	6.8	[S82]
id.	id.	Oct-Feb	1.6	id.
Tennessee	35.5N 276.7E	Jan-Feb	1.0	[S96]
Tallahassee	30.5N 275.8E	Jun-Aug	18.2	[S80]
North Carolina	34.2N 282.1E	Jan-Dec	9.9	[S97]
North Carolina, 1987-89	id.	Apr-Sep	8.2	id.
North Carolina, 1987-89	id.	Oct-Mar	3.5	id.
North Carolina, 1996-98	id.	Apr-Sep	17.6	id.
North Carolina, 1996-98	id.	Oct-Mar	3.0	id.
Los Angeles, California	34.1N 241.6E	Jan-Dec	12.4	[S98]
South America				
Calabozo, Venezuela	8.9N 298.4E	Jan-Dec	6.5	[S99]
Parupa, Venezuela	5.7N 298.4E	Jan-Dec	6.3	[S100]
Luepa, Venezuela	5.9N 298.6E	Jan-Dec	4.2	id.
Kavanayen, Venezuela	5.6N 298.2E	Jan-Dec	2.1	id.
Yuruani, Venezuela	5N 298.8E	Jan-Dec	2.4	id.
Auyantepuy, Venezuela	5.9N 298E	Apr	4.2	[S101]
Canaima, Venezuela	6.3N 297.1	Apr	2.9	id.
J.del Tigre, Venezuela	8.7N 296.7E	Jan-Dec	8.2	[S84]
Central Amazonia	2.6S 300E	Jul-Aug	17.9	[S87]
Central Amazonia	2.6-3.1S 300E	Apr-May	3.2	[S102]
Balbina, Centr. Amazonia	1.9S 300.5E	Jan-Dec	0.5	[S103]
French Guyana	5N 306.9E	Jan-Dec	5.5	[S104]
Lake Calado, Brazil	3.2S 299.5E	Jan-Dec	2.9	[S105]
Acegua, Uruguay	32.1S 306E	Jan-Dec	7.6	[S106]
Melo, Uruguay	32.4S 306E	Jan-Dec	7.1	id.
Treinta y Tres, Uruguay	32.4S 306E	Jan-Dec	2.2	id.
Torres del Paine, Chile	51.1S 288E	Jan-Dec	5.5	[S107]

Table S4: As Table S3, for Africa, Europe and Asia-Australia.

Location	Coordinates	*Period*	HCOOH	*Ref.*
Africa				
Banizoumbou, Niger	13.5N 2.7E	Jun-Oct	4.7	[S108]
Amersfort, S. Africa	27.1S 30E	Jan-Dec	7.5	[S109]
L. Trichardt, S. Africa	23S 30E	Jan-Dec	12.9	id.
Lamto, Ivory Coast	6N 355E	Jan-Dec	11.1	[S104]
Zoetele, Cameroon	3.2N 12E	Jan-Dec	8.7	[S110]
Mayombe, Congo	4.5S 12.5E	Jun-Sep	10.0	[S92]
id.	id.	Jun-Oct	10.6	[S111]
id.	id.	Nov-May	6.3	id.
Europe				
Basque Coun., Spain	43N 357.4E	Mar-Sep	8.1	[S112]
id.	id.	Oct-Feb	3.8	id.
Faroe Isl., Denmark	62.1N 353E	summer	2.0	[S113]
Anholt Isl., Denmark	56.7N 11.5E	summer	15.0	id.
Galicia, Spain	42.5-43.2N			
	351.5-352.5E	Mar-Sep	12.7	[S114]
id.	id.	Oct-Feb	2.8	id.
Patras, Greece	38.2N 21.7E	Jan-Dec	3.8	[S115]
Asia-Australia				
Amsterdam Isl., Indian Oc.	37.8S 77.5E	Jan-Dec	3.2	[S116]
Dorrigo, SE Australia	30.3S 152.7E	Jan-Apr	4.6	[S117]
Bar.Tops, SE Australia	31.9S 151.6E	Jan-Apr	6.8	id.
Katherine, Australia	14.5S 132.3E	Sep-Apr	10.5	[S118]
Darwin Airport, Australia	12.5S 130.8E	Nov-Apr	4.9	[S119]
West Guizhou, China	26.3N 105.9E	Jun-Oct	4.6	[S120]
Lijiang, China	26.9N 100.2E	Jan-Dec	3.1	[S121]
Gopalpura, India	27N 78E	Jul-Sep	5.4	[S94]
Rampur, India, monsoon	27.2N 78.1E	Jul-Sep	5.1	[S122]

Table S5: Average HCOOH concentrations over large regions measured (a) at surface sites (Table S2) ; (b) in precipitation water (Table S3 and S4) ; (c) during aircraft campaigns (SONEX, INTEX-B and PEM-Tropics-A (PEM-TA), and (d) average HCOOH columns measured by FTIR. The data are compared with model results in simulations F1-F3 and in inversion experiments assuming either a biogenic source of primary (Opt1) or secondary (Opt2) origin (see Table 1).

	Obs.	F1	F2	F3	Opt1	Opt2
(a) Mixing ratio in surface air (pptv)						
North America	1150	220	340	400	840	680
South America	1820	290	350	820	1410	740
Other regions	1120	510	770	740	1790	1170
(b) Concentration in precipitation water ($\mu mol/L$)						
North America	8.7	1.3	1.9	2.3	4.7	4.5
Europe	6.9	0.7	1.2	0.9	3.2	3.0
South America	5.2	1.0	1.4	2.8	3.6	2.9
Africa	9.0	2.8	4.8	5.4	10.1	8.4
Asia-Australia	5.4	1.4	2.3	3.5	3.1	3.0
(c) Aircraft campaign mixing ratio (pptv)						
North America (SONEX)	94	31	49	39	110	130
Western US (INTEX-B)	490	66	100	88	360	230
Mexico (INTEX-B)	1550	180	300	320	580	440
North Pacific (INTEX-B)	140	28	58	32	94	120
N. Trop. Pacific (PEM-TA)	40	11	31	18	20	25
S. Trop. Pacific (PEM-TA)	89	19	44	36	46	57
(d) Vertical columns (10^{14} $molec.cm^{-2}$)						
Reunion Isl., Indian Ocean	20	7.8	18	14	20	24
Wollongong, Australia	59	12.5	20	36	47.5	45

Table S6: Description of the performed sensitivity inversions.

Description	Abbreviation
Standard inversion with secondary HCOOH source	Opt2
Errors on IASI HCOOH columns doubled	Opt2-Errx2
Errors on IASI HCOOH columns halved	Opt2-Err:2
HCOOH deposotion velocity increased by 50%	Opt2-Dry
Use precipitation fields from NCEP	Opt2-NCEP
Cloud water pH assumed equal to 5	Opt2-pH5
Use IUPAC recommendation for HCOOH+OH reaction rate	Opt2-kOH
Use alternative fire injection heights[S75]	Opt2-Inj
Suppress isomerisation of peroxy radicals from isoprene	Opt2-Isom

Table S7: Annual estimates of sources, sinks and global lifetime of HCOOH inferred by sensitivity inversions. Units are in Tg per year. The biogenic emission estimates represent the sum of the primary biogenic source and of the IASI-derived secondary biogenic source of HCOOH.

Abbreviation	Global emission	Biogenic emission	Dry dep.	Wet dep.	OH oxidation	Lifetime (days)
Opt2	102.4	90.0	33.6	40.0	28.4	4.30
Opt2-Errx2	89.3	76.7	29.2	35.5	24.6	4.26
Opt2-Err:2	108.0	95.0	35.5	42.5	30.0	4.28
Opt2-Dry	108.4	96.3	40.0	40.0	28.4	4.05
Opt2-NCEP	101.5	89.0	33.8	37.4	30.3	4.55
Opt2-pH5	106.4	94.2	33.1	45.2	28.1	4.08
Opt2-kOH	98.8	86.4	31.8	39.3	27.7	4.62
Opt2-Inj	102.2	90.3	32.5	39.7	29.5	4.32
Opt2-Isom	97.8	84.8	31.1	39.8	26.9	4.57

10 Figures

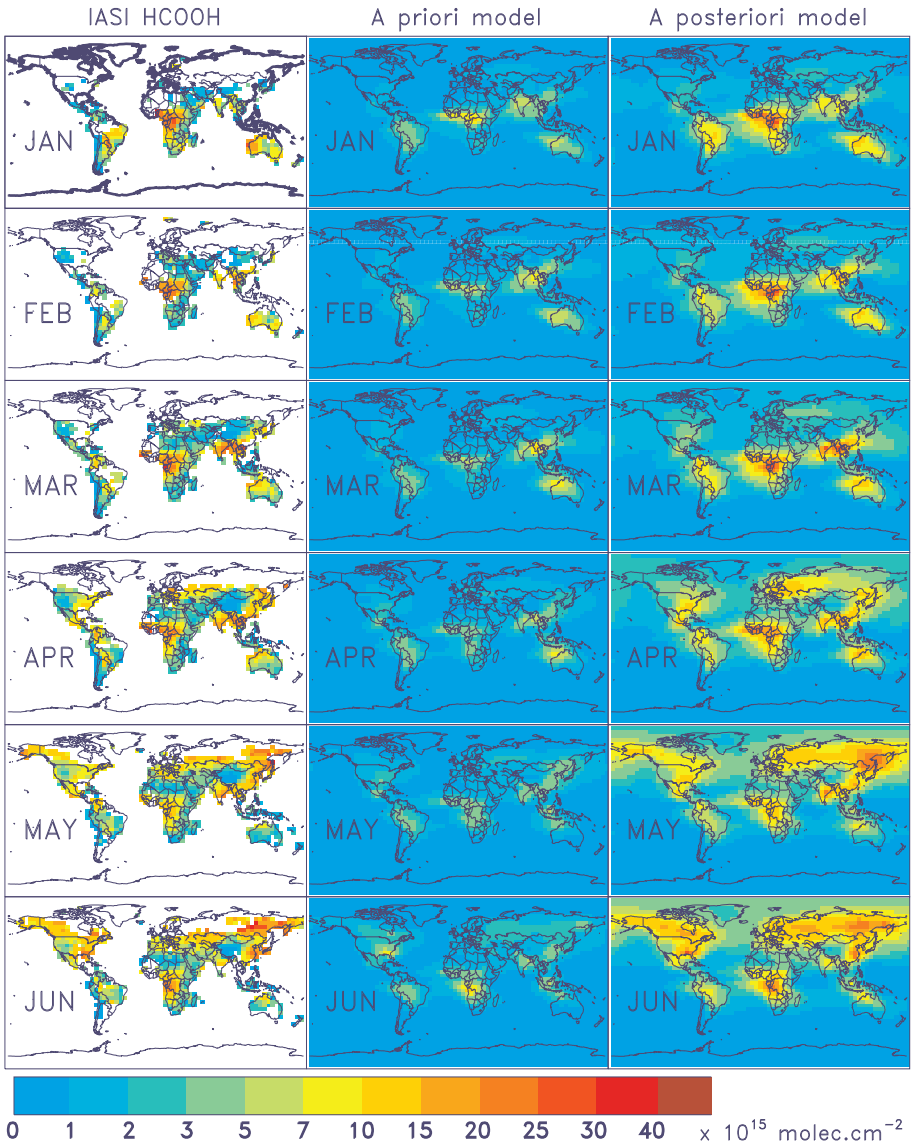

Figure S1: **Monthly averaged observed and modelled HCOOH columns between January and June 2009.** IASI columns (left column) are compared with a priori model columns (F1, middle) and optimised Opt2 HCOOH columns (right) inferred when a secondary HCOOH source of biogenic origin is assumed. Units are 10^{15} molec.cm^{-2}. Results from the Opt1 inversion are very similar to those from Opt2 and are not shown.

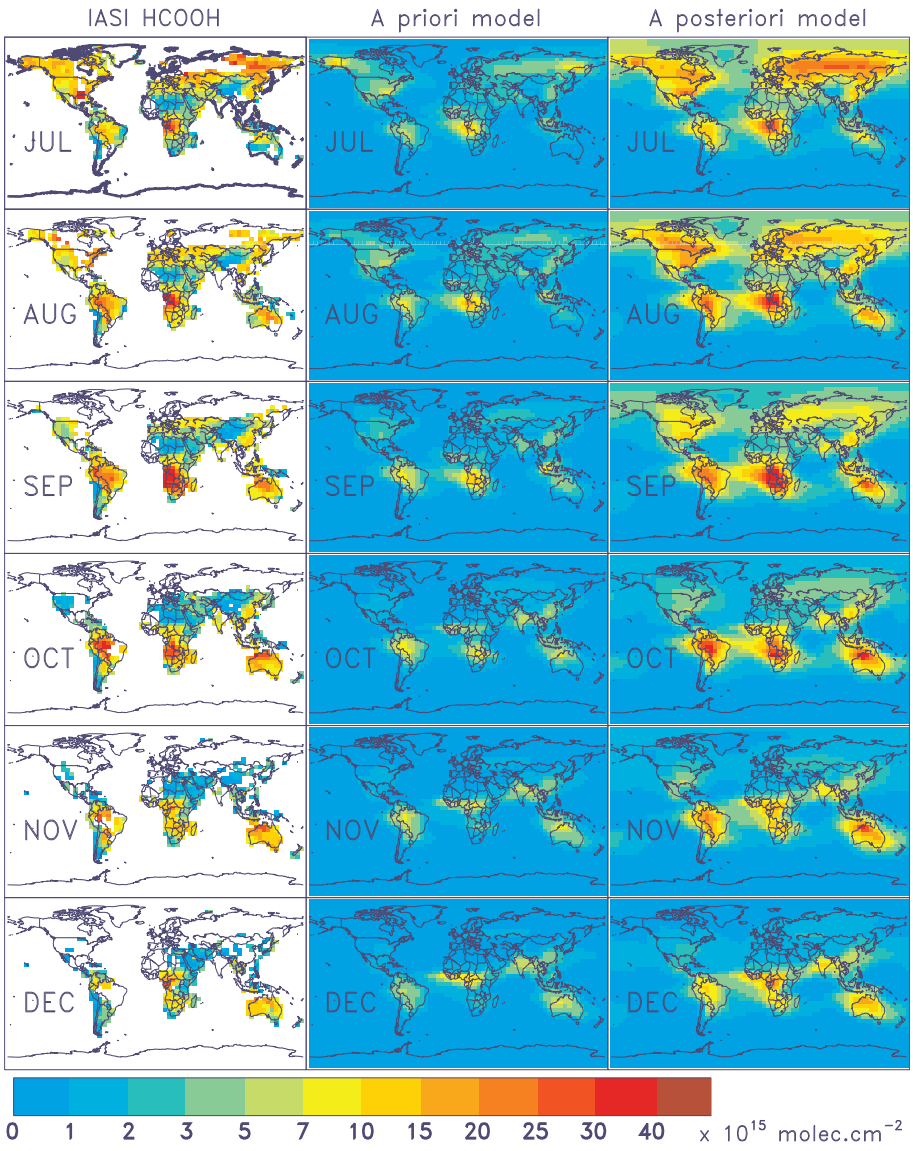

Figure S2: **Same as Fig. S1, for July–December 2009.**

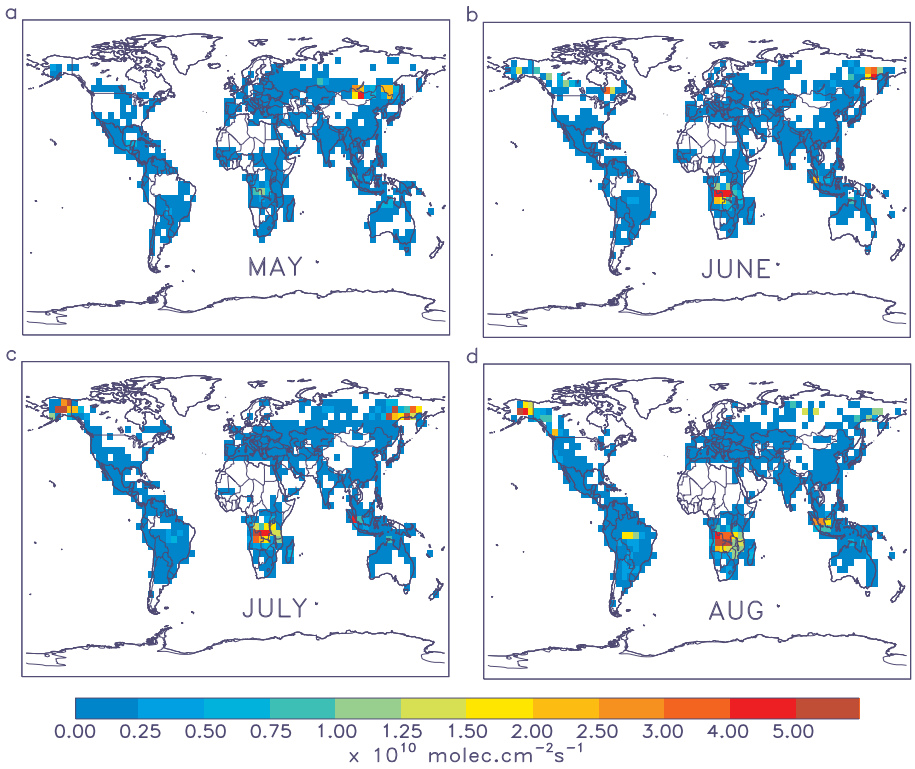

Figure S3: **Monthly averaged emissions of HCOOH by biomass burning (10^{10} molec.cm^{-2} s^{-1}) from the GFEDv3 inventory[S9]. a,** May. **b,** June. **c,** July. **d,** August. The fire locations on this figure generally do not coincide with enhanced HCOOH columns observed by IASI (see Fig. S1-S2), except to some extent in June in boreal regions.

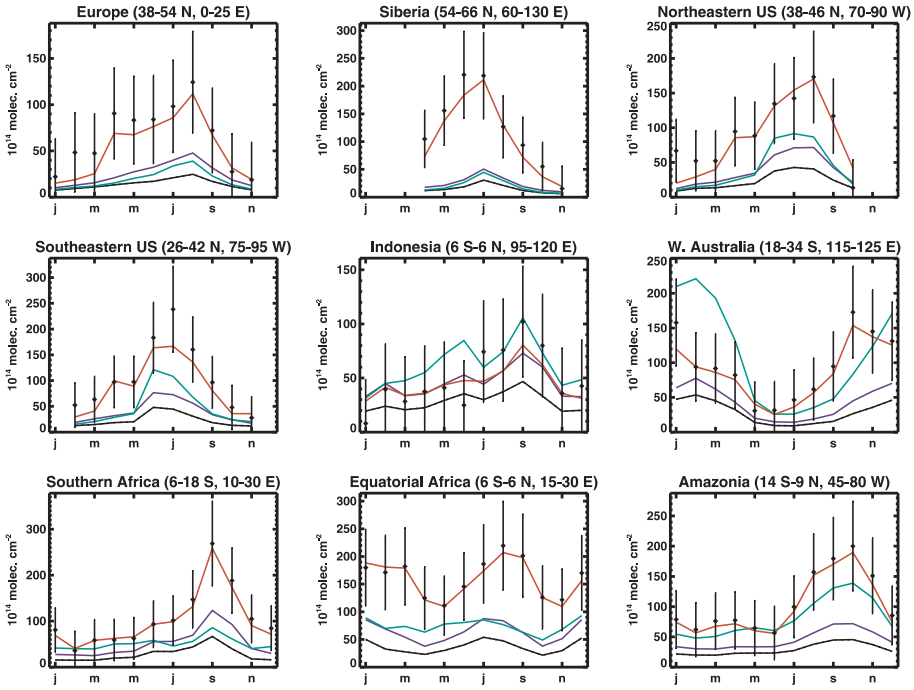

Figure S4: **Seasonal evolution over selected regions of monthly mean observed and modelled HCOOH columns.** IASI HCOOH columns (black diamonds) are compared with modelled columns of the F1 (black), F2 (purple), F3 (green), and Opt2 (red) simulations of Table 1. The error bars represent the assumed errors on IASI columns adopted in the source inversions. Results from Opt1 simulation are very close to Opt2 and are therefore not shown.

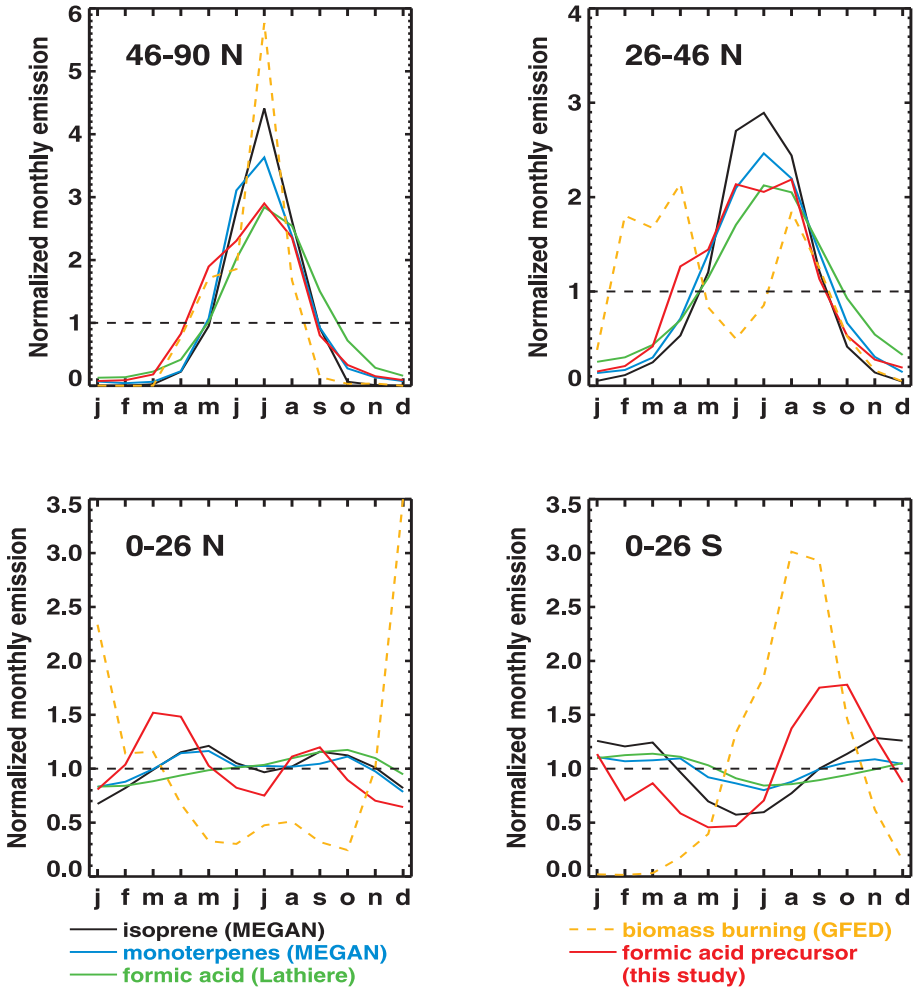

Figure S5: **Monthly emissions (normalised by their annual average) for four latitude bands.** Curves correspond to isoprene (in black), monoterpenes (blue), primary biogenic HCOOH (green), biomass burning HCOOH (orange), and the biogenic HCOOH precursor optimised in inversion Opt2 (red).

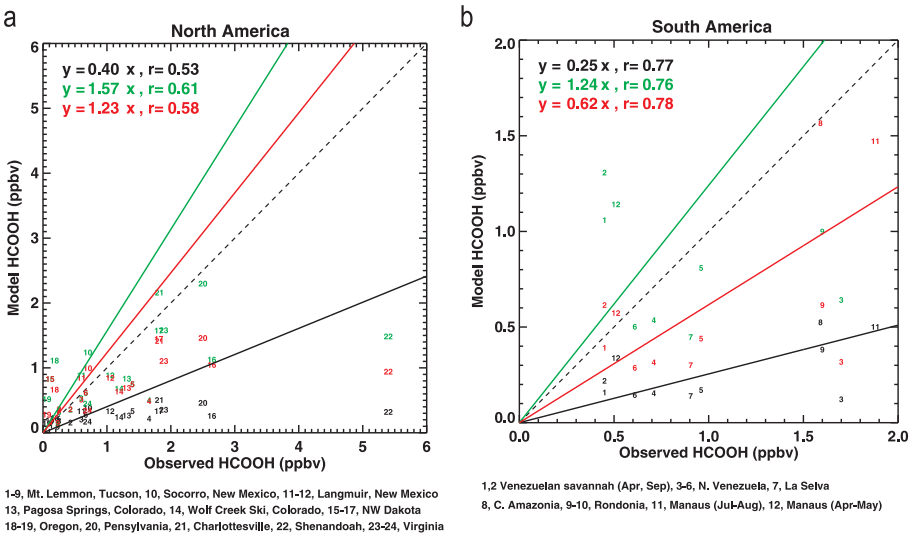

Figure S6: **Ground-based HCOOH mixing ratios at North and South American sites (Table S1) against model values.** Model concentrations are obtained at the same locations using the standard model (in black), and after optimisation assuming either a primary (in green) or a secondary (in red) biogenic source of HCOOH. The slopes of the regression lines indicated in the upper left corner are calculated as the arithmetic mean ratios of modelled to observed values. The correlation coefficient r for each region is given inset.

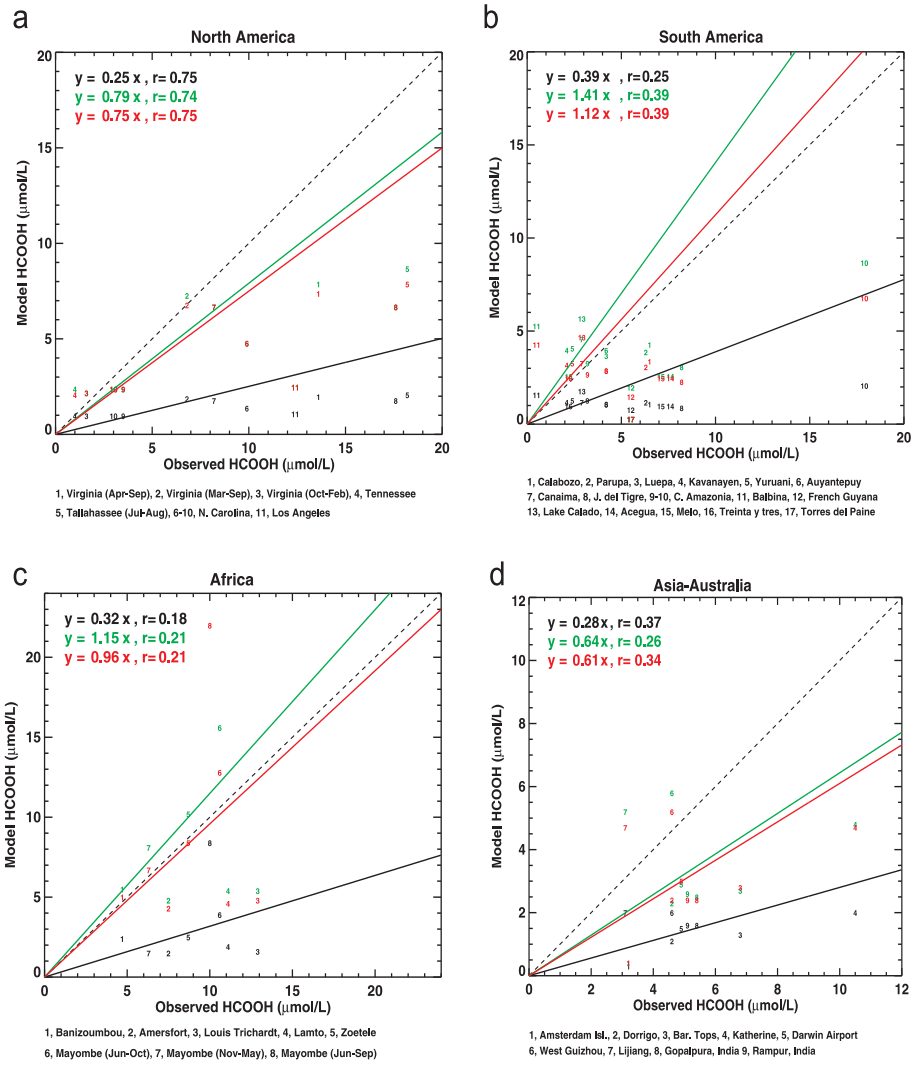

Figure S7: **Observed HCOOH concentrations in precipitation water against model values (Table S2 and S3) in different continents.** Color code is as in Fig. S6. The slopes of the regression lines indicated in the upper left corner are calculated as the arithmetic mean ratios of modelled to observed values. The correlation coefficient r for each region is given inset.

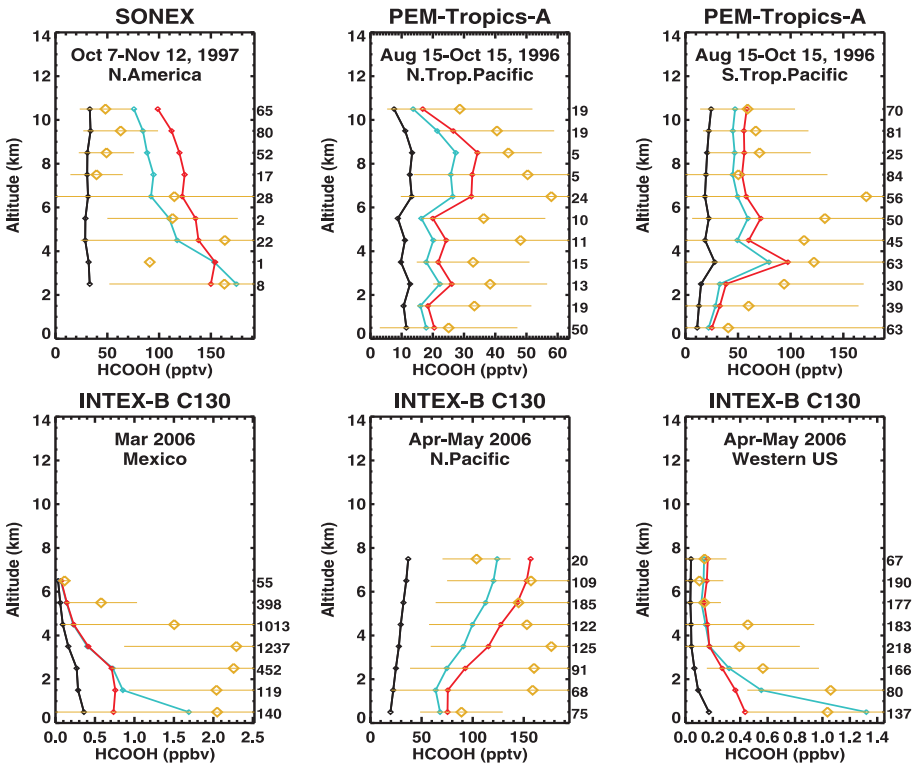

Figure S8: **Observed and modelled vertical distributions of HCOOH concentrations.** SONEX, PEM-Tropics-A and INTEX-B aircraft data over North America (25-55 N, 45-130 W), Mexico (0-25 N, 85-110 W), Western US (35-45 N, 105-125 W), North Pacific (25-55 N, 160-230 E), North Tropical Pacific (0-25 N, 180-275 W), and South Tropical Pacific (0-25 S, 180-275 W) are compared with model output in 2009 from the standard simulation F1 (black) and from Opt1 (green) and Opt2 (red). The number of observations at each altitude bin is given on the right end of each plot. Error bars are standard deviations.

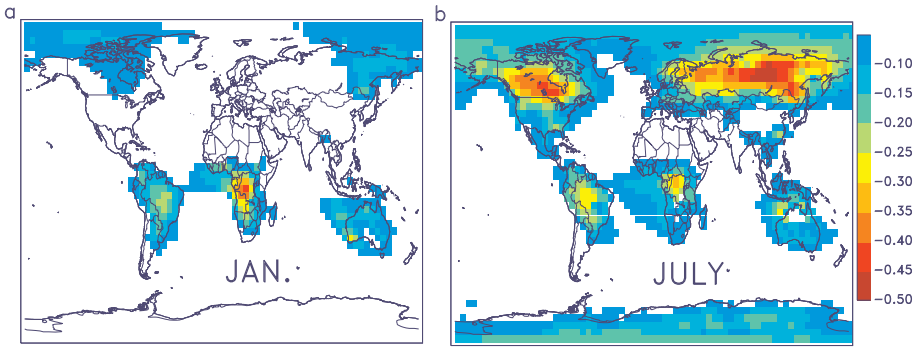

Figure S9: **Calculated change in precipitation pH due to the additional secondary biogenic source of HCOOH inferred from IASI data (Opt2). a,** In January. **b,** In July.

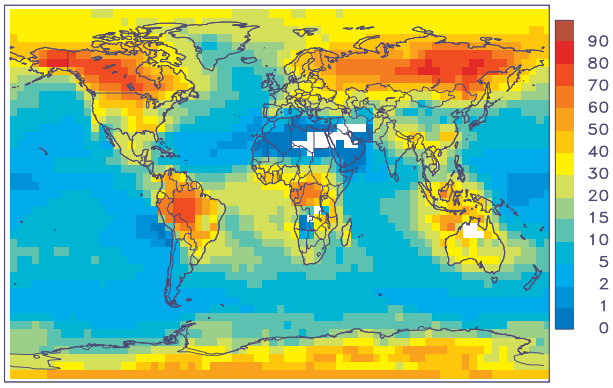

Figure S10: **Calculated contribution (%) of HCOOH to the total concentration of hydrogen cations [H^+].** Results are from the Opt2 source inversion in July.

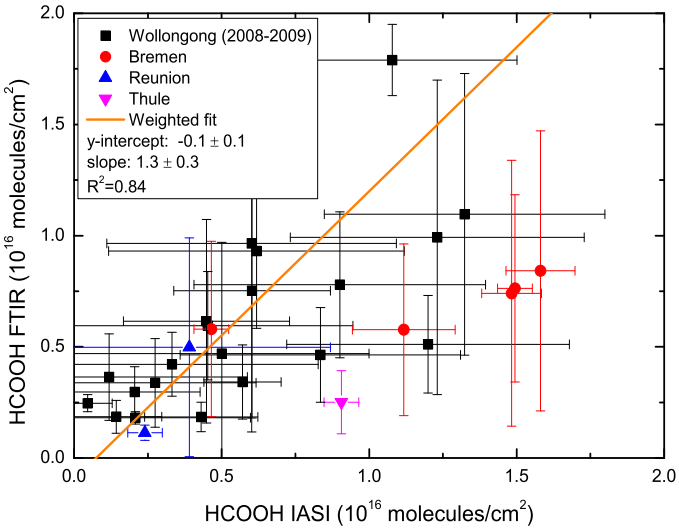

Figure S11: **Comparison between IASI columns and FTIR measurements at different sites.** FTIR data at Wollongong and Reunion Island are presented in this study, at Thule (2004-2008) and Bremen (2002-2010) are obtained from Ref. [S7]. The regression line applies to Wollongong and Reunion Island data.

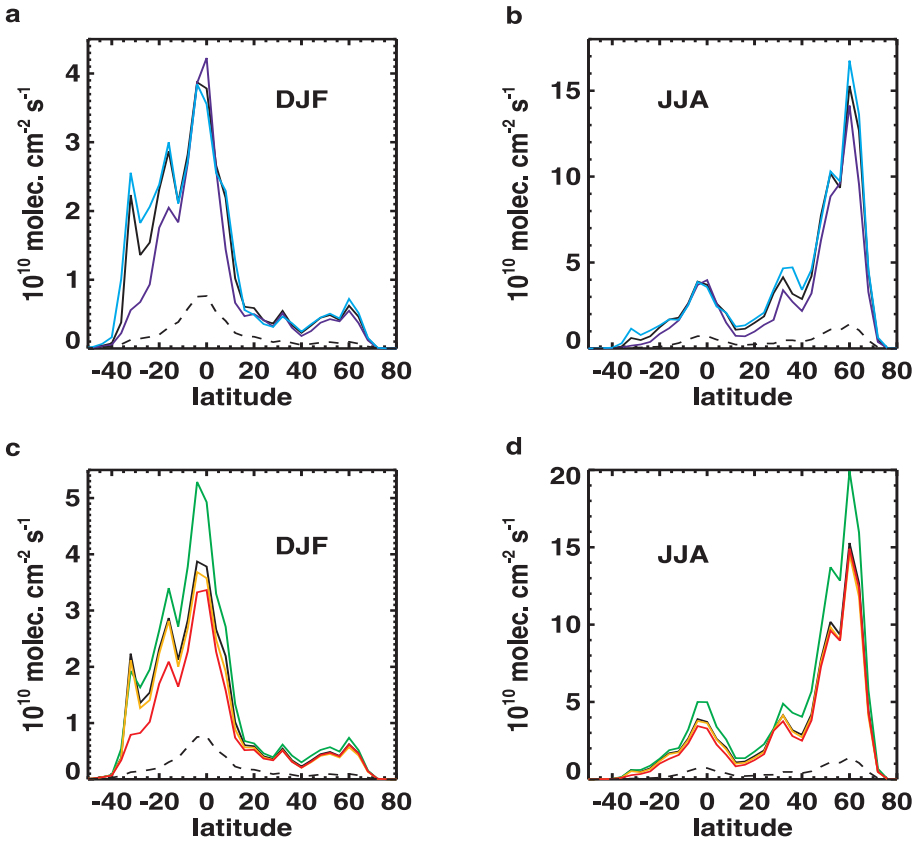

Figure S12: **Latitudinal profiles of "top-down" biogenic emissions of the HCOOH precursor averaged over December-January-February (DJF) and June-July-August (JJA). a-b,** Opt2 in black, Opt2-Err:2 in light blue, Opt2-Errx2 in dark blue. **c-d,** Opt2 in black, Opt1 in green, Opt2-kOH in orange, Opt2-Isom in red. The dashed lines represent the a priori distribution of the HCOOH precursor used in the optimisations. The other sensitivity inversions lie very close to the standard Opt2 results and are not illustrated here.

References

[S1] Müller, J.-F. & Stavrakou, T. Inversion of CO and NO_x emissions using the adjoint of the IMAGES model. *Atmos. Chem. Phys.* **5**, 1157–1186 (2005).

[S2] Stavrakou, T., Müller, J.-F., Boersma, F., De Smedt, I. & van der A, R. Assessing the distribution and growth rates of NOx emission sources by inverting a 10-year record of NO_2 satellite columns. *Geophys. Res. Lett.* **35**, L10810, doi:10.1029/2008GL033521 (2008).

[S3] Stavrakou, T. *et al.* Evaluating the performance of pyrogenic and biogenic emission inventories against one decade of space-based formaldehyde columns. *Atmos. Chem. Phys.* **9**, 1037–1060 (2009).

[S4] Stavrakou, T., Peeters, J. & Müller, J.-F. Improved global modelling of HOx recycling in isoprene oxidation : evaluation against the GABRIEL and INTEX-A aircraft campaign measurements. *Atmos. Chem. Phys.* **10**, 9863–9878 (2010).

[S5] Schultz, M. G. *et al.* Global wildland fire emissions from 1960 to 2000. *Global Biogeochem. Cy.* **22**, GB2002, doi:10.1029/2007GB003031 (2008).

[S6] Ohara, T. *et al.* An Asian emission inventory of anthropogenic emission sources for the period 1980–2020. *Atmos. Chem. Phys.* **7**, 4419–4444 (2007).

[S7] Paulot, F. *et al.* Importance of secondary sources in the atmospheric budgets of formic and acetic acids. *Atmos. Chem. Phys.* **11**, 1989–2013 (2011).

[S8] Guenther, A. *et al.* Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys.* **6**, 3181–3210 (2006).

[S9] van der Werf, G. R. *et al.* Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009). *Atmos. Chem. Phys.* **10**, 11,707–11,735 (2010).

[S10] Andreae, M. O. & Merlet, P. Emission of trace gases and aerosols from biomass burning. *Global Biogeochem. Cy.* **15**, 955–966 (2001).

[S11] Lathière, J. *et al.* Impact of climate variability and land use changes on global biogenic volatile organic compound emissions. *Atmos. Chem. Phys.* **6**, 2129–2146 (2006).

[S12] Peeters, J., Nguyen, T. L. & Vereecken, L. HOx radical regeneration in the oxidation of isoprene. *Phys. Chem. Chem. Phys.* **11**, 5935–5939 (2009).

[S13] Peeters, J. & Müller, J.-F. HOx radical regeneration in isoprene oxidation via peroxy radical isomerisations, II: Experimental evidence and global impact. *Phys. Chem. Chem. Phys.* **12**, 14,227–14235 (2010).

[S14] Lelieveld, J. *et al.* Atmospheric oxidation capacity sustained by a tropical forest. *Nature* **452**, 737–740 (2008).

[S15] Martinez, M. *et al.* Hydroxyl radicals in the tropical troposphere over the Suriname rainforest: airborne measurements. *Atmos. Chem. Phys.* **10**, 3759–3773 (2010).

[S16] Paulot, F. *et al.* Unexpected epoxide formation in the gas-phase photooxidation of isoprene. *Science* **325**, 730–733 (2009).

[S17] Archibald, A. T. *et al.* Impacts of mechanistic changes on HOx formation and recycling in the oxidation of isoprene. *Atmos. Chem. Phys.* **10**, 8097–8118 (2010).

[S18] Paulot, F. *et al.* Isoprene photooxidation: new insights into the production of acids and organic nitrates. *Atmos. Chem. Phys.* **9**, 1479–1501 (2009).

[S19] Butkovskaya, N. I., Pouvesle, N., Kukui, A. & Le Bras, G. Mechanism of the OH-initiated oxidation of glycolaldehyde over the temperature range 233-296 K. *J. Phys. Chem. A* **110**, 13,492–13,499 (2006).

[S20] Saunders, S.M., Jenkin, M. E., Derwent, R. G. & Pilling, M. J. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): tropospheric degradation of non-aromatic volatile organic compounds. *Atmos. Chem. Phys.* **3**, 161–180 (2003).

[S21] Butkovskaya, N. I., Pouvesle, N., Kukui, A., Mu, Y. & Le Bras, G. Mechanism of the OH-initiated oxidation of hydroxyacetone over the temperature range 236-298 K. *J. Phys. Chem. A.* **110**, 6833–6843 (2006).

[S22] Orlando, J. J. & Tyndall, G. S. Mechanism for the oxidation of hydroxyacetone under atmospheric conditions. *AGU Fall Meeting, Abstract A11F-0115* (2010).

[S23] Atkinson, R. Gas-phase tropospheric chemistry of volatile organic compounds: 1. Alkanes and alkenes. *J. Phys. Chem. Ref. Data* **26**, 215–290 (1997).

[S24] Neeb, P., Sauer, F., Horie, O. & Moortgat, G. K. Formation of hydroxymethyl hydroperoxide and formic acid in alkene ozonolysis in the presence of water vapour. *Atmos. Environ.* **31**(10), 1417–1423 (1997).

[S25] Atkinson, R. *et al.* Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II - gas phase reactions of organic species. *Atmos. Chem. Phys.* **6**, 3625–4055 (2006).

[S26] Hatakeyama, S. & Akimoto, H. Reactions of Criegee intermediates in the gas phase. *Res. Chem. Intermed.* **20**, 503–524 (1994).

[S27] Horie, O., Neeb, P., Limbach S. & Moortgat G. K. Formation of formic acid and organic peroxides in the ozonolysis of ethene with added water vapour. *Geophys. Res. Lett.* **21**, 1523–1526 (1994).

[S28] Hasson, A. S., Orzechowska, G. E. & Paulson, S. E. Production of stabilized Criegee intermediates and peroxides in the gas phase ozonolysis of alkenes 1. Ethene, trans-2-butene, and 2,3-dimethyl-2-butene. *J. Geophys. Res.* **106**, 34,131–34,142 (2001).

[S29] Leather, K. E. *et al.* Acid-yield measurements of the gas-phase ozonolysis of ethene as a function of humidity using Chemical Ionisation Mass Spectrometry (CIMS). *Atmos. Chem. Phys. Discuss.* **11**, 25173–25204, 2011.

[S30] Ryzhkov, A. B. & Ariya, P. A. A theoretical study of the reactions of parent and substituted Criegee intermediates with water and the water dimer. *Phys. Chem. Chem. Phys.* **6**, 5042–5050 (2004).

[S31] Hasson, A. S., Ho, A. W., Kuwata, K. T. & Paulson, S. E. Production of stabilized Criegee intermediates and peroxides in the gas phase ozonolysis of alkenes 2. Asymmetric and biogenic alkenes. *J. Geophys. Res.* **106**, 34,143–34,153 (2001).

[S32] Lee, J. H., Leahy, D. F., Tang, I. N. & Newman, L. Measurement and speciation of gas phase peroxides in the atmosphere. *J. Geophys. Res.* **98**, 2911-2915 (1993).

[S33] Weinstein-Lloyd, J. B. *et al.* Measurements of peroxides and related species during the 1995 summer intensive of the Southern Oxidants Study in Nashville, Tennessee. *J. Geophys. Res.* **103**, 22,361–22,373 (1998).

[S34] Sauer, F., Beck, J., Schuster, G., & Moortgat, G. K. Hydrogen peroxide, organic peroxides and organic acids in a forested area during FIELDVOC'94. *Chemosphere - Global Change Science* **3**, 309–326 (2001).

[S35] Valverde-Canossa, J. *et al.* First measurements of H_2O_2 and organic peroxides surface fluxes by the relaxed eddy-accumulation technique. *Atmos. Environ.* **40**, S55–S67 (2006).

[S36] Bauerle, S. & Moortgat, G. K. Absorption cross-sections of $HOCH_2OOH$ vapor between 205 and 360 nm at 298 K. *Chem. Phys. Lett.* **309**, 43–48 (1999).

[S37] Francisco, J. S. & Eisfeld, W. Atmospheric oxidation mechanism of hydroxymethyl hydroperoxide. *J. Phys. Chem. A* **113**, 7593–7600 (2009).

[S38] Veyret, B. *et al.* Kinetics and mechanism of the photooxidation of formaldehyde. 1. Flash photolysis study. *J. Phys. Chem.* **93**, 2368–2374 (1989).

[S39] Hermans, I., Müller, J.-F., Nguyen, T. L., Jacobs, P. A. & Peeters, J. Kinetics of α-hydroxy-alkylperoxyl radicals in oxidation processes. HO_2-initiated oxidation of ketones/aldehydes near the tropopause. *J. Phys. Chem. A* **109**, 4303-4311 (2005).

[S40] Vereecken, L., Nguyen, T. L., & Peeters, J. Computational study of the stability of α-hydroperoxyl- or α-alkylperoxyl substituted alkyl radicals. *Chem. Phys. Lett.* **393**, 432–436, 2004.

[S41] O'Sullivan, D. W., Lee, M., Noone, B. C. & Heikes, B. G. Henry's law constant determinations for hydrogen peroxide, methyl hydroperoxide, hydroxymethyl hydroperoxide, ethyl hydroperoxide, and peroxyacetic acid. *J. Phys. Chem.* **100**, 3241–3247 (1996).

[S42] Chen, Z. M. *et al.* Aqueous-phase ozonolysis of methacrolein and methyl vinyl ketone: a potentially important source of atmospheric aqueous oxidants. *Atmos. Chem. Phys.* **8**, 2255–2265 (2008).

[S43] Nguyen, T. L., Peeters, J. & Vereecken, L. Theoretical study of the gas-phase ozonolysis of β-pinene ($C_{10}H_{16}$). *Phys. Chem. Chem. Phys.* **11**, 5643–5656 (2009).

[S44] Orlando, J. J. *et al.* Product studies of the OH- and ozone-initiated oxidation of some monoterpenes. *J. Geophys. Res.* **105** (D9), 11,561–11,572, doi: 10.1029/2000JD900005 (2000).

[S45] Fantechi, G. *Atmospheric oxidation reactions of selected biogenic volatile organic compounds (BIOVOCs): A smog chamber study.* (Ph.D. Thesis, University of Leuven, 1999).

[S46] Lee, A. *et al.* Gas-phase products and secondary aerosol yields from the ozonolysis of ten different terpenes. *J. Geophys. Res.* **111**, D07302, doi:10.1029/2005JD006437 (2006).

[S47] Larsen, B. R. *et al.* Gas-Phase OH Oxidation of Monoterpenes: Gaseous and Particulate Products. *J. Atmos. Chem.* **38**, 231–276 (2001).

[S48] Zhang, P., Anderson, M., Barlow, B., Tan, B., & Myneni, R. B. Climate-related vegetation characteristics derived from Moderate Resolution Imaging Spectroradiometer (MODIS) leaf area index and normalized difference vegetation index. *J. Geophys. Res.* **109**, D20105, doi:10.1029/2004JD004720 (2004).

[S49] Stavrakou, T. *et al.* The continental source of glyoxal estimated by the synergistic use of spaceborne measurements and inverse modelling. *Atmos. Chem. Phys.* **9**, 8431–8446 (2009b).

[S50] Sellegri, K. *et al.* Contribution of gaseous and particulate species to droplet solute composition at the Puy de Dôme, France. *Atmos. Chem. Phys.* **3**, 1509–1522 (2003).

[S51] Chameides, W. L. The photochemistry of a remote marine stratiform cloud. *J. Geophys. Res.* **89**, 4739–4755 (1984).

[S52] Crowley, J. N. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume V - heterogeneous reactions on solid substrates. *Atmos. Chem. Phys.* **10**, 9059–9223 (2010).

[S53] Heymsfield, A. J. & McFarquhar, G. M. High albedos of cirrus in the tropical pacific warm pool: Microphysical interpretations from CEPEX and from Kwajalein, Marshall Islands. *J. Atmos. Sci.* **53**, 2424–2451 (1996).

[S54] Schmitt, C. G. & Heymsfield, A. J. Total surface area estimates for individual ice particles and particle populations. *J. App. Met.* **44**, 467–474 (2005).

[S55] Bond, T. *et al.* A technology-based global inventory of black and organic carbon emissions from combustion. *J. Geophys. Res.* **109**, doi:10.1029/2003JD003697 (2004).

[S56] Martin, R. V. *et al.* Global and regional decreases in tropospheric oxidants from photochemical effects of aerosols. *J. Geophys. Res.* **108**, 4097, doi:10.1029/2002JD002622 (2003).

[S57] Henze, D. K. *et al.* Global modeling of secondary organic aerosol formation from aromatic hydrocarbons: high- vs. low-yield pathways. *Atmos. Chem. Phys.* **8**, 2405–2421 (2008).

[S58] Chang, E. I. & Pankow, J. F. Organic particulate matter formation at varying relative humidity using surrogate secondary and primary organic compounds with activity corrections in the condensed phase obtained using a method based on the Wilson equation. *Atmos. Chem. Phys.* **10**, 5475–5490 (2010).

[S59] Henze, D. K. & Seinfeld, J. H. Global secondary organic aerosol from isoprene oxidation. *Geophys. Res. Lett.* **33**, L09812, doi:10.1029/2006GL025976 (2006).

[S60] Ceulemans, K., Müller, J.-F. & Compernolle, S. Parameterising secondary organic aerosol from α-pinene using a detailed oxidation and aerosol formation model. *Atmos. Chem. Phys. Discuss.*, submitted (2011).

[S61] Capouet, M., Müller, J.-F., Ceulemans, K., Compernolle, S., Vereecken, L. & Peeters, J. Modeling aerosol formation in α-pinene photooxidation experiments. *J. Geophys. Res.* **113**, doi:10.1029/2007JD008995 (2008).

[S62] Ceulemans, K., Compernolle, S., Peeters, J., & Muller, J.-F.. Evaluation of a detailed model of secondary organic aerosol formation from α-pinene against dark ozonolysis experiments. *Atmos. Environ.* **40**, 5434–5442 (2010).

[S63] Fu, T.-M. *et al.* Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols. *J. Geophys. Res.* **113**, doi:10.1029/2007JD009505 (2008).

[S64] Spracklen, D. V. *et al.* Aerosol mass spectrometer constraint on the global secondary organic aerosol budget. *Atmos. Chem. Phys. Discuss.* **11**, 5699–5755 (2011).

[S65] Malm, W. C., Schichtel, B. A., Pitchford, M. L., Ashbaugh, L. L., & Eldred, R. A. Spatial and monthly trends in speciated fine particle concentration in the United States. *J. Geophys. Res.* **109**, D03306, doi:10.1029/2003JD003739 (2004).

[S66] Yttri, K. E. *et al.* Elemental and organic carbon in PM10: a one year measurement campaign within the European Monitoring and Evaluation Programme EMEP. *Atmos. Chem. Phys.* **7**, 5711–5725 (2007).

[S67] Perrin , A. & Vander Auwera, J. An improved database for the 9 micron region of the formic acid spectrum. *J.Quant. Spectrosc. Radiat. Transfer* **108**, 363–370 (2007).

[S68] Rothman, L. S. *et al.* The HITRAN 2008 molecular spectroscopy database. *J. Quant. Spectrosc. Radiat. Transfer* **110**, 533–572 (2009).

[S69] Emmons, L. K. *et al.* Data composites of airborne observations of tropospheric ozone and its precursors. *J. Geophys. Res.* **105**, 20,497–20,538 (2000).

[S70] Singh, H. B., Thompson, A. M. & Schlager, H. SONEX airborne mission and coordinated POLINAT-2 activity: Overview and accomplishments. *Geophys. Res. Lett.* **26**, 3053-3056 (1999).

[S71] Singh, H. B., Brune, W. H., Crawford, J. H., Flocke, F. & Jacob, D. J. Chemistry and transport of pollution over the Gulf of Mexico and the Pacific: spring 2006 INTEX-B campaign overview and first results. *Atmos. Chem. Phys.* **9**, 2301–2318 (2009).

[S72] Karl., T. *et al.* Efficient Atmospheric cleansing of oxidized organic trace gases by vegetation. *Science* **330**, 816–819, doi: 10.1126/science.1192534 (2010).

[S73] Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc.* **77**, 437–470 (1996).

[S74] Galano, A., Alvarez-Idaboy, J. R., Ruiz-Santoyo, M. E. & Vivier-Bunge, A. Rate coefficient and mechanism of the gas phase OH hydrogen abstraction reaction from formic acid: A quantum mechanical approach. *J. Phys. Chem. A* 106, 9520–9528 (2002).

[S75] Turquety, S. *et al.* Inventory of boreal fire emissions for North America in 2004: Importance of peat burning and pyroconvective injection, *J. Geophys. Res.* **112**, D12S03, doi:10.1029/2006JD007281 (2007).

683 [S76] Dentener, F. *et al.* Emissions of primary aerosol and precursor gases in
684 the years 2000 and 1750 prescribed data-sets for AeroCom. *Atmos. Chem.*
685 *Phys.* **6**, 4321–4344 (2006).

686 [S77] Johnson, B. J. & Dawson G. A. A preliminary study of the carbon-isotopic
687 content of ambient formic acid and two selected sources : automobile exhaust
688 and formicine ants. *J. Atmos. Chem.* **17**, 123–140 (1993).

689 [S78] Villanueva-Fierro, I., Popp, C. J. & Martin, R. S. Biogenic emissions and
690 ambient concentrations of hydrocarbons, carbonyl compounds and organic
691 acids from ponderosa pine and cottonwood trees at rural and forested sites
692 in Central New Mexico. *Atmos. Environ.* **38**, 249–260 (2004).

693 [S79] Martin, R. S. *et al.* Measurement of isoprene and its atmospheric oxidation
694 products in a Central Pensylvania desiduous forest. *J. Atmos. Chem.* **13**, 1–92
695 (1991).

696 [S80] Andreae, M. O., R. W. Talbot & Li, S.-M. Atmospheric Measurements of
697 Pyruvic and Formic Acid. *J. Geophys. Res.* **92(D6)**, 6635–6641 (1987).

698 [S81] Talbot, R. W. *et al.* Carboxylic acids in the rural continental atmosphere
699 over the eastern United States during the Shenandoah Cloud and Photochem-
700 istry Experiment. *J. Geophys. Res.* **100(D5)**, 9335–9343 (1995).

701 [S82] Talbot, R., Beecher, K., Harriss, R. & Cofer W. III : Atmospheric geo-
702 chemistry of formic and acetic acids at a mid-latitude temperate site. *J. Geo-*
703 *phys. Res.* **93(D2)**, 1638–1652 (1988).

704 [S83] Sanhueza, E. & Andreae, M. O. Emissions of formic and acetic acids from
705 tropical savanna soils. *Geophys. Res. Lett.* **18(9)**, 1707–1710 (1991).

706 [S84] Sanhueza, E., Figueroa, L. & Santana, M. Atmospheric formic and acetic
707 acids in Venezuela. *Atmos. Environ.* **30**, 1861–1873 (1996a).

708 [S85] Sanhueza, E., Santana, M. & Hermoso, M. Gas- and aqueous- phase formic
709 and acetic acids at a tropical cloud forest site. *Atmos. Environ.* **26**, 1421–1426
710 (1992).

711 [S86] Karl, T. *et al.* Exchange processes of volatile organic compounds above a
712 tropical rain forest: Implications for modelling troposheric chemistry above
713 dense vegetation. *J. Geophys. Res.* **109**, D18306, doi:10.1029/2004JD004738
714 (2004).

715 [S87] Andreae, M. O., Talbot, R. W., Andreae, T. W. & and Harriss, R. C.
716 Formic and Acetic Acid Over the Central Amazon Region, Brazil 1. Dry
717 Season. *J. Geophys. Res.* **93(D2)**, 1616–1624 (1988).

718 [S88] Kesselmeier, J. *et al.* Concentrations and species composition of atmo-
719 spheric volatile organic compounds (VOCs) as observed during the wet and
720 dry season in Rondônia (Amazonia). *J. Geophys. Res.* **107(D20)**, 8053,
721 doi:10.1029/2000JD000267 (2002).

[S89] Talbot, R. W., Andreae, M. O., Berresheim, H., Jacob, D. J. & Beecher, K. M. Sources and sinks of formic, acetic and pyruvic acids over Central Amazonia, 2. Wet season. *J. Geophys. Res.* **95(D10)**, 16,799–16,811 (1990).

[S90] Puxbaum, H. *et al.* Atmospheric concentrations of formic and acetic acid and related compounds in eastern and northern Austria. *Atmos. Environ.* **22(12)**, 2841–2850 (1988).

[S91] Helas, G., Bingemer, H. & Andreae, M. O. Organic acids over Equatorial Africa : Results from DECAFE 88. *J. Geophys. Res.* **97(D6)**, 6187–6193 (1992).

[S92] Servant, J., Kouadio, G., Cros, B. & Delmas, R. Carboxylic Monoacids in the air of Mayombe forest (Congo) : Role of the forest as a source or sink. *J. Atmos. Chem.* **12**, 367–380 (1991).

[S93] Kumar, N. *et al.* Measurements of formic and acetic acid levels in the vapour phase at Dayalbagh, Agra, India. *Atmos. Environ.* **20**, 3545–3550 (1996).

[S94] Khare, P., Sastangi, G. S., Kumar, N., Kumari, K. M. & Srivastava, S. S. HCHO, HCOOH and CH_3COOH in air and rain water at a rural tropical site in North Central India. *Atmos. Environ.* **31**, 3867–3875 (1997).

[S95] Keene, W. C. & Galloway, J. N. Organic acidity of precipitation of North America. *Atmos. Environ.* **18**, 2491–2497 (1984).

[S96] Schaefer, D. A., Lindberg, S. E. & Hoffman, W. A. Fluxes of undissociated acids to terrestrial ecosystems by atmospheric deposition. *Tellus* **41B**, 207–218 (1989).

[S97] Brooks Avery, G. Jr., Yang, Y., Kieber, R. J. & Willey, J. D. Impact of recent urbanization on formic and acetic acid concentrations in coastal North Carolina rainwater. *Atmos. Environ.* **35**, 3353–3359 (2001).

[S98] Sakugawa, H., Kaplan, I. R. & Shepard, L. S. Measurements of H_2O_2, aldehydes, and organic acids in Los Angeles rainwater: their sources and deposition rates. *Atmos. Environ.* **27B**, 203–219 (1993).

[S99] Sanhueza, E. *et al.* Field measurement evidence for an atmospheric chemical source of formic and acetic acids in the tropic. *Geophys. Res. Lett.* **23(9)**, 1045–1048 (1996b).

[S100] Sanhueza, E., Alfonzo, Y. & Santana, M. Compuestos orgánicos volatiles en la atmósfera de la Gran Sabana. II: HCHO, HCOOH y CH_3COOH en lluvias. *Interciencia* **28(1)**, 29–35 (2003).

[S101] Sanhueza, E., Donoso, L., Santana, M., Fernández, E. & Romero, J. Atmospheric chemistry over the Auyantepuy (5° 46′ N; 62° 30′ W; 2100 meters a.s.l.). *Interciencia* **24**, 372–380 (1999).

[S102] Andreae, M., Talbot, R., Berresheim, H. & Beecher, K. Precipitation Chemistry in Central Amazonia. *J. Geophys. Res.* **95(D10)**, 16,987–16,999 (1990).

[S103] Pauliquevis, T., Lara, L. L., Antunes, M. L. & Artaxo, P. Aerosol and precipitation chemistry in a remote site in Central Amazonia : The role of biogenic contribution. *Atmos. Chem. Phys. Discuss.* **7**, 11,465–11,509 (2007).

[S104] Yoboué, Y., Galy-Lacaux, C., Lacaux, J. P. & Silué, S. Rainwater chemistry and wet deposition over the wet savanna ecosystem of Lamto (Côte d'Ivoire). *J. Atmos. Chem.* **52**, 117–141 (2005).

[S105] Williams, R. M., Fisher, T. R. & Melack, J. M. Chemical composition and deposition of rain in the Central Amazon, Brazil. *Atmos. Environ.* **31**, 207–217 (1997).

[S106] Zunckel, M., Saizar, C. & Zarauz, J. Rainwater composition in northeast Uruguay. *Atmos. Environ.* **37**, 1601–1611 (2003).

[S107] Galloway, J. N., Keene, W. C. & Likens, G. E. Processes controlling the composition at a remote southern hemispheric location : Torres del Paine National Park, Chile. *J. Geophys. Res.* **101(D3)**, 6883–6897 (1996).

[S108] Galy-Lacaux, C. & Modi, A. I. Precipitation chemistry in the Sahelian savanna of Niger, Africa. *J. Atmos. Chem.* **30**, 319–343 (1998).

[S109] Mphepya, J. N., Pienaar, J. J., Galy-Lacaux, C., Held, G. & Turner, C. R. Precipitation chemistry in semi-arid areas of Southern Africa : a case study for a rural and an industrial site. *J. Atmos. Chem.* **47**, 1–24 (2004).

[S110] Sigha-Nkamdjou, L. *et al.* Rainwater chemistry and wet deposition over the Equatorial forested ecosystem of Zoétélé (Cameroon). *J. Atmos. Chem.* **46**, 173–198 (2003).

[S111] Lacaux, J. P., Delmas, R., Kouadio, G., Cros, B. & Andreae, M. O. Precipitation chemistry in the Mayombe Forest of Equatorial Africa. *J. Geophys. Res.* **97(D6)**, 6195–6206 (1992).

[S112] Durana, N. *et al.* Organic acids in precipitation in the Basque Country (North of Spain), *Atmos. Res.* **28**, 93–101 (1992).

[S113] Glasius, M. *et al.* Relative contribution of biogenic and anthropogenic sources to formic and acetic acids in the atmospheric boundary layer. *J. Geophys. Res.* **106(D7)**, 7415–7426 (2001).

[S114] Peña, R. M. *et al.* Organic acids and aldehydes in rainwater in a northwest region of Spain. *Atmos. Environ.* **36**, 5277–5288 (2002).

[S115] Moschonas, N. & Glavas, S. Weak organic acidity in a wet-only precipitation study at a Mediterranean coastal site, Patras, Greece. *Atmos. Res.* **63**, 147–157 (2002).

[S116] Galloway, J. N. & Gaudry, A. The composition of precipitation on Amsterdam Island, Indian Ocean. *Atmos. Environ.* **18**, 2649–2656 (1984).

[S117] Post, D., Bridgman, H. A. & Ayers, G. P. Fog and rainwater composition in rural SE Australia. *J. Atmos. Chem.* **13**, 83–95 (1991).

[S118] Likens, G. E., Keene, W. C., Miller, J. M. & Galloway, J. N. Chemistry of precipitation from a remote terrestrial site in Australia. *J. Geophys. Res.* **92(D11)**, 13,299–13,314 (1987).

[S119] Ayers, G. P., Gillett, R. W. & Selleck, P. A pilot study on rain-water composition at Darwin Airport. *Aust. Met. Mag.* **42**, 143–150 (1993).

[S120] Zhang, Y.-L., Lee, X.-Q., Huang, D.-K., Huang, R.-S. & Jiang, W. Low molecular weight carboxylic acids in precipitation during the rainy season in the rural area of Anshun, West Guizhou province. *Huanjing Kexue/Environmental Science* , **30(3)**, 644–649 (2009).

[S121] Keene, W. C. *et al.* Processes controlling variability in the chemical composition of precipitation at a remote, temperate site in south-central China, Fourth Scientific Advisory Council, International Geosphere-Biosphere Programme, Beijing (1995).

[S122] Singh, S. P. *et al.* Rainwater composition at a regional representative site of semi-arid region of India. *Water, Air, and Soil Pollution* **127**, 93–108 (2001).