Article:
Dalsøren, Stig B., Myhre, Gunnar, Hodnebrog, Øivind et al. (14 more authors) (2018) Discrepancy between simulated and observed ethane and propane levels explained by underestimated fossil emissions. Nature Geoscience. pp. 178-184. ISSN 1752-0908

https://doi.org/10.1038/s41561-018-0073-0

Discrepancy between simulated and observed ethane and propane levels explained by underestimated fossil emissions

Stig B. Dalsøren[1,*], Gunnar Myhre[2], Øivind Hodnebrog[2], Cathrine Lund Myhre[3], Andreas Stohl[3], Ignacio Pisso[3], Stefan Schwietzke[4,5], Lena Höglund-Isaksson[6], Detlev Helmig[7], Stefan Reimann[8], Stéphane Sauvage[9], Norbert Schmidbauer[3], Katie A. Read[10], Lucy J. Carpenter[10], Alastair C. Lewis[10], Shalini Punjabi[10] and Markus Wallasch[11]

[1]CICERO-Center for International Climate and Environmental Research Oslo, 0318, Oslo, Norway. Now at Institute for Marine Research, Flødevigen, 4817 His, Norway.

[2]CICERO-Center for International Climate and Environmental Research Oslo, 0318, Oslo, Norway.

[3]NILU-Norwegian Institute for Air Research, 2027 Kjeller, Norway.

[4]CIRES-Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado 80309, USA.

[5]NOAA Earth System Research Laboratory, Global Monitoring Division, Boulder, Colorado 80305-3337, USA.

[6]International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria.

[7]Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado 80305, USA.

[8]Empa, Laboratory for Air Pollution/Environmental Technology, Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland.

[9]IMT Lille Douai, Univ. Lille, SAGE - Département Sciences de l'Atmosphère et Génie de l'Environnement, 59000 Lille, France.

[10]Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom.

[11]Umweltbundesamt, Messnetzzentrale Langen, D-63225 Langen, Germany.

*e-mail: stig.dalsoeren@hi.no

Ethane and propane are the most abundant non-methane hydrocarbons in the atmosphere. However, their emissions, distribution in the atmosphere, and trends in their atmospheric concentrations are insufficiently understood. Atmospheric model simulations using standard community emission inventories do not reproduce available measurements in the Northern Hemisphere. Here, we show that observations of preindustrial and present-day ethane and propane can be reproduced in simulations with a detailed atmospheric-chemistry transport model, provided natural geologic emissions are taken into account and anthropogenic fossil fuel emissions are assumed to be two to three times higher than indicated in current inventories. Accounting for these enhanced ethane and propane emissions results in simulated surface ozone concentrations that are 5 to 13 % higher than previously assumed in some polluted regions in Asia. The improved correspondence with observed ethane and propane in model simulations with greater emissions suggests that the level of fossil (geologic + fossil fuel) methane emissions in current inventories may need re-evaluation.

Direct emissions at the surface are the only sources of ethane and propane to the atmosphere[1,2], and several studies suggest that they are underestimated in global inventories[1,3-7]. A major source of uncertainty is that these inventories first calculate total non-methane hydrocarbon (NMHC) emissions and then disaggregate them into individual species (ethane, propane, etc.) based on limited amounts of data[8-11]. Over the past decade the inventories do not fully account for an abrupt increase in the exploitation of unconventional natural gas in the United States[3,4] and therefore likely underestimate present-day emissions[3,4,12-23]. Recent atmospheric model simulations applying current global emission inventories tend to underestimate observed ethane and propane concentrations in wintertime in the Northern Hemisphere[3-5,7,12,24-29].

Anthropogenic fossil fuel (conventional and unconventional) emissions are presently the largest emission source of ethane and propane in most global inventories. Fugitive emission is the main fossil fuel NMHC source and includes venting and flaring, evaporative losses, and equipment leaks but not fuel combustion. Changes in these particular emissions are regarded as the main cause of observed ethane trends[3,4,26,30-34]. Recent studies[9,35] calculated fugitive fossil fuel emissions from oil, natural gas, and coal systems for ethane based on a joint inventory and

atmospheric box-model approach. Their emission dataset[9] was recently updated with new data[36]. Another recent study[10] used a detailed inventory approach to identify cold venting of associated petroleum gas containing also methane, propane, and butane as significant emission sources potentially underestimated in existing emission inventories. These new studies combine field measurements and country-specific information from published sources along with observed flaring of associated gas from satellite images. In combination these take into account that the emission factors from venting and flaring of associated gas released during extraction vary considerably across different oil, coal, and gas fields around the world. Such considerations have not been made in most community emission datasets, which apply emission factors reported by countries to the United Nations Framework Convention on Climate Change, or from measurements representative for North America for those countries that are not reporting. Deficiencies in fugitive fossil fuel emission estimates in community emission datasets were also recently found for black carbon at high latitudes[37] and SO_2 at low latitudes[38] (the Middle East). Natural geologic emissions are another suggested fossil hydrocarbon source missing in inventories[39,40], receiving little attention in previous model studies. Major geologic sources include seepage from onshore and submarine petroleum basins, volcanoes, and degassing from geothermal manifestations (see Methods). Based on the few available estimates[39,40], geologic emissions may have been the largest preindustrial source of ethane to the atmosphere (Figure 1a).

Another cause of poor model performance could be inaccurate representation of atmospheric sinks. Oxidation by hydroxyl in the troposphere is the main sink for ethane and propane[1,2,41]. Estimates of global mean hydroxyl levels and hemispheric ratios differ substantially between observation-based estimates and results from model ensembles[42-44]. Studies also find large inter-model differences in the atmospheric distributions of oxidants[45].

In this study we first apply the OsloCTM3 model[46] to investigate the preindustrial atmospheric ethane budget. We compare model results from simulations with and without geological emissions to ice-core measurements. Thereafter, we model current conditions represented by the year 2011, which is the last year available in all applied fugitive fossil fuel emission datasets. In the year 2011 baseline simulation we use the state of the art global anthropogenic emission inventory CEDS CMIP6[47]. We also include natural emissions (treated as negligible in many

model studies) from oceanic, biogenic (vegetation) and geologic sources (see Methods). We compare with surface ethane and propane measurements from global and regional surface networks (see Methods) with a focus on observations in the mid- to high-latitude Northern Hemisphere where previous model studies underestimated observations. We then generate alternative gridded emissions by replacing the fugitive fossil fuel emissions in the CEDS CMIP6 inventory with new datasets that better represent fossil fuel activity and emissions[9,10,35,36] with corrections to avoid double counting from potential overlap with natural geologic emissions (see Methods). In the resulting alternative simulations (ALT1 and ALT2, Tables 1-2) the fossil fuel emissions are factors of about two (ethane, Figure 1b) and three (propane, Figure S1) higher than in the baseline simulation with very different geographical distributions (Figure S2). We also suggest further modifications of the ALT1 and ALT2 emissions based on studies with the Flexpart model[48]. Back-trajectories from Flexpart are used to identify source regions related to OsloCTM3 under- and overestimation of observed ethane concentrations (see Methods). Finally, we explore the uncertainty of the atmospheric sinks in a sensitivity study where we perturb the OsloCTM3 hydroxyl level within its uncertainty range (see Methods).

Preindustrial ethane budget

Ice core measurements[39] reveal a large, factor of about 4, north/south inter-polar ratio for ethane. Figure 2a shows that this can be reproduced by the model with a geologic source of 3 Tg/yr, constituting about 40% of the total preindustrial ethane emissions (Figure 1a). With geologic emissions included in the simulations the modeled abundance at Summit in Greenland agrees with observations, and the simulated inter-polar ratio improves substantially relative to observations. Without geologic sources the simulated abundance at Summit is 50 % too low. This is in agreement with the findings of ref. 39, where a simpler model without interactive oxidation chemistry was used.

Our baseline simulation slightly overestimates the observed Antarctic ethane concentration. Transport to high southern latitudes and biomass burning emissions have high inter-annual variability. However, sensitivity simulations with meteorological input data for a different year and an alternative inventory with different geographical distribution and emission totals for biomass burning emissions (see Figure 1a and Methods) resulted in minor changes (see

Supplementary). We therefore suggest that a small Antarctic overestimation relates to uncertainties in the magnitude and distribution of geologic and oceanic emissions in the Southern Hemisphere.

Current ethane and propane budgets

The baseline simulation for 2011 does not reproduce the observed inter-polar ethane ratio well, even if geologic emissions are included (Figure 2b). The modeled concentration at Summit in Greenland is only about 50 % of measured values (Figure 2b). Underestimations of ethane and propane concentrations at high northern latitudes, particularly during wintertime (Figure 3, Figures S3-S7), are similar to most other model studies[3-5,7,12,24-29] using standard emission inventories. As shown below, the likely cause is underestimated fossil fuel emissions in the standard community emission data set CEDS CMIP6, used in the baseline simulation. The CEDS CMIP6 emission data[47] agree with previous emission data for ethane and propane used in atmospheric climate and air pollution studies (Figure 1b and Figure S1).

The ALT1 and ALT2 simulations, where the CEDS CMIP6 fugitive fossil fuel emissions are replaced with the new datasets[9,10,35,36], reproduce the inter-polar ethane ratio and the observed levels in Greenland (Figure 2b), Zeppelin Observatory at Svalbard (Figure 3a), and most other stations (Figure 3b, Figures S3-S9) much more closely. This is also the case for propane (Figure 3a, c, Figures S3-S9), for which fossil fuel emissions play an even larger role (Figure S1). A substantial improvement is found throughout the Arctic (Figure S5, Tables S2-S3). ALT1 performs better than ALT2. Both have positive mean biases, tending to overestimate episodes with high concentrations. We explore these and other episodes at Zeppelin (Arctic station with frequent sampling) in a systematic way (see Methods). Figure 4 shows that the episodes with the largest underestimation of ethane at Zeppelin in the baseline simulation occur for air masses originating from Eurasia. Fossil fuels are the dominant emission source in this region for most of the year, strongly suggesting that these are underestimated in the CMIP6 inventory. From Figure 4, it is also evident that ALT1 underestimates the fossil fuel emissions in northwestern Europe (i.e. mainly emissions from the North Sea) while it overestimates emissions from Russia. The ALT2 simulation mainly overestimates observed ethane levels at Zeppelin (Figure 3a). For this inventory, the fossil fuel emissions are likely overestimated both in the North Sea and Russia

147 (Figure 4). Overestimated emissions in ALT1 and ALT2 over Russia also seem likely from the
148 comparison with Tiksi station data (Figure S5, the only available station in Russia).

149 At mid-latitude stations in the U.S. and Canada, both ALT1 and ALT2 show good agreement
150 with measurements (Figure S6, Tables S2-S3). An exception is the highly oil and natural gas
151 influenced Southern Great Plains station (large underestimation) (Figure S6). Large emissions
152 from several nearby oil wells might not be fully resolved in the model averaging emissions over
153 the model grid scale. It is also a possibility that the ALT1 and ALT2 inventories underestimate
154 the emissions from nearby unconventional gas fields (e.g. Woodford, Barnett[49]) and oil wells.
155 The ALT1 and ALT2 simulations also improve agreement with measurements at non-Arctic
156 European stations compared to the baseline (Figure S7, Tables S2-S3), especially for ethane. At
157 most stations the ALT1 simulation is biased slightly low compared to the observations. ALT2
158 also performs better than the baseline, but overestimates, to varying degree, the measurements at
159 most European stations.

160 The lower latitude Cape Verde site also shows large improvements (ALT simulations vs.
161 baseline) during wintertime (Figure S8) when it is influenced by air passing over the Sahara[50]
162 from oil and gas fields in northern Africa and the Middle East. In the Southern Hemisphere, the
163 baseline simulation reproduces observed levels and seasonal patterns well and the alternative
164 simulations only result in minor differences (Figures 2-3 and Figure S9).

165 The ALT1 and ALT2 anthropogenic ethane emissions (excluding biomass burning) are slightly
166 smaller than the optimized anthropogenic emissions in other recent model studies[3,4] (Figure1b).
167 The optimized emissions in other studies are based on sensitivity simulations[3,4] finding that an
168 approximate doubling of anthropogenic emissions is needed to reproduce measurements at
169 Jungfraujoch[4] and a few other FTIR stations in the Northern Hemisphere[3]. In our study we
170 include natural geologic emissions and apply new detailed emission datasets for fugitive fossil
171 fuel emissions instead of performing an up-scaling of all anthropogenic emissions. Our model is
172 also run at higher spatial and temporal resolution and compared to a larger number of
173 measurement sites.

174 Overestimated atmospheric loss (i.e. too high hydroxyl levels) might lead to an underestimation
175 of observed ethane and propane levels. However, this cannot be a major cause for the
176 discrepancies in our baseline simulation. Scaling down tropospheric hydroxyl levels to the lower

range of model and observational based estimates (see Methods) improves the agreement slightly (Figure S10) but much of the underestimation in the mid-high Northern Hemisphere during wintertime persists. This seasonal pattern is expected as chemical loss at high northern latitudes is inefficient during wintertime when little sunlight and low water vapor concentrations result in low hydroxyl concentrations.

We have not included atmospheric oxidation of ethane and propane by halogens in the model simulations. The reasons and implications on uncertainty are discussed in the Supplementary. Inclusion of halogen chemistry would likely lead to slightly larger underestimation of ethane and propane in the baseline simulation, thereby supporting that emissions in standard community datasets are too low.

Our global total natural geologic emissions correspond with the best estimates from the study by ref. 40. Reported uncertainty ranges[39,40] are shown in Figure 1 and Figure S1. A recent study[51] suggests geologic methane emissions about 1/3 of that estimated by ref. 40. NMHC to methane emission ratios from geologic sources are uncertain and likely highly spatially and temporally variable. Therefore, findings suggesting lower methane flux in the far past do not necessarily imply lower preindustrial and present-day NMHC emissions. However, if the finding[51] was applied via downscaling to estimate ethane emissions, it then fits poorly with the findings of our preindustrial simulations (Figure 2a), unless we change the geographical distribution towards a larger fraction of emissions in the Northern Hemisphere. As noted (see Methods) the uncertainty in geographical distribution is large.

The relative uncertainties in global total emissions in the new fugitive fossil fuel datasets (applied in the ALT1, ALT2 inventories) are about half of those used in CEDS CMIP6 (baseline inventory) (see Supplementary). Another improvement is substantially reduced uncertainties in geographical emission distributions (see Methods). Based on the comparisons in the previous paragraphs the correct global total anthropogenic ethane emissions seem to be close to the levels in the ALT1 and ALT2 simulations (rectangle, triangle Figure S11). These levels are greater than the upper cap of the baseline uncertainty bar in Figure S11, i.e. the baseline inventory likely underestimates emissions. The total emissions in the baseline inventory are close to the mean and median of those in eight other standard community emissions datasets suggesting that applying

these will also result in too low modeled ethane concentrations. For the ALT1 and ALT2 inventories, only total emissions near the lower end of the uncertainty range (Figure S11) reproduce the observed levels. Various model uncertainties widen the possible emission range but major ones associated to the OH sink have relatively small impacts on modeled ethane concentrations. Model uncertainties do therefore not change our conclusions regarding under- and overestimation in the different emission inventories. Using the alternative emission datasets (ALT1 or ALT2) instead of the standard community emission inventories greatly improves the comparison with observations. Due to sparse observation coverage in some world regions, and the uncertainty ranges of the emissions and atmospheric chemical loss of ethane and propane, we do not provide an overall performance ranking between the emission datasets ALT1 and ALT2.

Impacts on other atmospheric constituents

The higher ethane and propane in the ALT1 and ALT2 simulations compared to the baseline simulation impact the greenhouse gas methane and major surface pollutants. The impacts on tropospheric methane are moderate, leading to 0.5 % (ALT1) and 0.7 % (ALT2) higher methane lifetimes due to lower tropospheric hydroxyl. Modeled baseline ozone mixing ratios are compared to surface measurements in Figure S12 for the period June-August when ozone photochemistry is most active in the Northern Hemisphere. The model reproduces the gradients between regions with high photochemical production and cleaner background areas. At many stations the model is at, or within, a few ppb (or percent) of the measurements. In regions with high levels of surface ozone in the baseline, particularly the Middle East and eastern Asia, ozone is 5-13 % (3-11 ppb) higher in spring/summer (Figure S13) in ALT1 and ALT2. In these regions, the ozone production is more sensitive to the amount of NMHCs since high NOx is present. Surface ozone differences in other regions are generally small (0-5 % or 0-3 ppb). If fossil emissions of other related NMHCs are underestimated as well (e.g. butane, pentane etc.), the impacts on ozone and other air pollutants will be larger. Impacts on the air pollutants NO_2, PAN and CO are discussed in the Supplementary.

Methane constitutes the largest share of hydrocarbons emitted from fossil sources, and a recent study[52] suggests underestimation of fossil (geologic + fossil fuel) methane emissions in previous estimates. Compared to previous inventories the much higher fossil fuel ethane and propane emissions in the new datasets[9,10,35,36] used in ALT1 and ALT2 in this study are mainly due to higher NMHC to methane emission ratios. The improved agreement with ice-core ethane measurements for the simulation with geologic emissions supports the idea that there is a considerable geologic methane emission source[52,53]. As for ethane and propane, geologic emissions of methane have been neglected in many model studies. In accordance with ref. 52 we suggest a need for more studies evaluating the reported level of fossil (natural geologic + anthropogenic fossil fuel) methane emissions in current emission inventories. Understanding the contribution from different natural and anthropogenic emission sources is a critical precursor to design efficient measures to reverse ongoing atmospheric ethane, propane, and methane increasess[4,54,55].

References

1 Rudolph, J. The tropospheric distribution and budget of ethane. *Journal of Geophysical Research: Atmospheres* **100**, 11369-11381 (1995).

2 Rosado-Reyes, C. M. & Francisco, J. S. Atmospheric oxidation pathways of propane and its by-products: Acetone, acetaldehyde, and propionaldehyde. *Journal of Geophysical Research: Atmospheres* **112**, D14310 (2007).

3 Franco, B. *et al.* Evaluating ethane and methane emissions associated with the development of oil and natural gas extraction in North America. *Environmental Research Letters* **11**, 044010 (2016).

4 Helmig, D. *et al.* Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production. *Nature Geosci* **9**, 490-495 (2016).

5 Stein, O. & Rudolph, J. Modeling and interpretation of stable carbon isotope ratios of ethane in global chemical transport models. *Journal of Geophysical Research: Atmospheres* **112**, D14308 (2007).

6 Thompson, A., Rudolph, J., Rohrer, F. & Stein, O. Concentration and stable carbon isotopic composition of ethane and benzene using a global three-dimensional isotope inclusive chemical tracer model. *Journal of Geophysical Research: Atmospheres* **108**, 4373 (2003).

7 Emmons, L. K. *et al.* The POLARCAT Model Intercomparison Project (POLMIP): overview and evaluation with observations. *Atmos. Chem. Phys.* **15**, 6721-6744 (2015).

8 Li, M. *et al.* Mapping Asian anthropogenic emissions of non-methane volatile organic compounds to multiple chemical mechanisms. *Atmos. Chem. Phys.* **14**, 5617-5638 (2014).

9 Schwietzke, S., Griffin, W. M., Matthews, H. S. & Bruhwiler, L. M. P. Global Bottom-Up Fossil Fuel Fugitive Methane and Ethane Emissions Inventory for Atmospheric Modeling. *ACS Sustainable Chemistry & Engineering* **2**, 1992-2001 (2014).

10 Höglund-Isaksson, L. Bottom-up simulations of methane and ethane emissions from global oil and gas systems 1980 to 2012. *Environmental Research Letters* **12**, 024007 (2017).

11 Huang, G. *et al.* Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970–2012. *Atmos. Chem. Phys.* **17**, 7683-7701 (2017).

12 Xiao, Y. *et al.* Global budget of ethane and regional constraints on U.S. sources. *Journal of Geophysical Research: Atmospheres* **113**, D21306 (2008).

13 Zavala-Araiza, D. *et al.* Reconciling divergent estimates of oil and gas methane emissions. *Proceedings of the National Academy of Sciences* **112**, 15597-15602 (2015).

14 Karion, A. *et al.* Aircraft-Based Estimate of Total Methane Emissions from the Barnett Shale Region. *Environ. Sci. Technol.* **49**, 8124-8131 (2015).

15 Kort, E. A. *et al.* Four corners: The largest US methane anomaly viewed from space. *Geophysical Research Letters* **41**, 6898-6903 (2014).

16 Kort, E. A. *et al.* Fugitive emissions from the Bakken shale illustrate role of shale production in global ethane shift. *Geophysical Research Letters* **43**, 4617-4623 (2016).

17 Peischl, J. *et al.* Quantifying atmospheric methane emissions from the Haynesville, Fayetteville, and northeastern Marcellus shale gas production regions. *Journal of Geophysical Research: Atmospheres* **120**, 2119-2139 (2015).

18 Pétron, G. *et al.* A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin. *Journal of Geophysical Research: Atmospheres* **119**, 6836-6852 (2014).

19 Swarthout, R. F. *et al.* Impact of Marcellus Shale Natural Gas Development in Southwest Pennsylvania on Volatile Organic Compound Emissions and Regional Air Quality. *Environ. Sci. Technol.* **49**, 3175-3184 (2015).

20 Vinciguerra, T. *et al.* Regional air quality impacts of hydraulic fracturing and shale natural gas activity: Evidence from ambient VOC observations. *Atmos. Environ.* **110**, 144-150 (2015).
21 Schneising, O. *et al.* Remote sensing of fugitive methane emissions from oil and gas production in North American tight geologic formations. *Earth's Future* **2**, 548-558 (2014).
22 Moore, C. W., Zielinska, B., Pétron, G. & Jackson, R. B. Air Impacts of Increased Natural Gas Acquisition, Processing, and Use: A Critical Review. *Environ. Sci. Technol.* **48**, 8349-8359 (2014).
23 Roest, G. & Schade, G. Quantifying alkane emissions in the Eagle Ford Shale using boundary layer enhancement. *Atmos. Chem. Phys.* **17**, 11163-11176 (2017).
24 Pozzer, A. *et al.* Simulating organic species with the global atmospheric chemistry general circulation model ECHAM5/MESSy1: a comparison of model results with observations. *Atmos. Chem. Phys.* **7**, 2527-2550 (2007).
25 González Abad, G. *et al.* Ethane, ethyne and carbon monoxide concentrations in the upper troposphere and lower stratosphere from ACE and GEOS-Chem: a comparison study. *Atmos. Chem. Phys.* **11**, 9927-9941 (2011).
26 Angelbratt, J. *et al.* Carbon monoxide (CO) and ethane (C_2H_6) trends from ground-based solar FTIR measurements at six European stations, comparison and sensitivity analysis with the EMEP model. *Atmos. Chem. Phys.* **11**, 9253-9269 (2011).
27 Lin, M., Holloway, T., Carmichael, G. R. & Fiore, A. M. Quantifying pollution inflow and outflow over East Asia in spring with regional and global models. *Atmos. Chem. Phys.* **10**, 4221-4239 (2010).
28 Helmig, D. *et al.* Climatology and atmospheric chemistry of the non-methane hydrocarbons ethane and propane over the North Atlantic. *Elementa* **3** (2015).
29 Pozzer, A. *et al.* Observed and simulated global distribution and budget of atmospheric C_2-C_5 alkanes. *Atmos. Chem. Phys.* **10**, 4403-4422 (2010).
30 Helmig, D. *et al.* Reconstruction of Northern Hemisphere 1950-2010 atmospheric non-methane hydrocarbons. *Atmos. Chem. Phys.* **14**, 1463-1483 (2014).
31 Simpson, I. J. *et al.* Long-term decline of global atmospheric ethane concentrations and implications for methane. *Nature* **488**, 490-494 (2012).
32 Hausmann, P., Sussmann, R. & Smale, D. Contribution of oil and natural gas production to renewed increase in atmospheric methane (2007–2014): top–down estimate from ethane and methane column observations. *Atmos. Chem. Phys.* **16**, 3227-3244 (2016).
33 Aydin, M. *et al.* Recent decreases in fossil-fuel emissions of ethane and methane derived from firn air. *Nature* **476**, 198-201 (2011).
34 Zeng, G. *et al.* Trends and variations in CO, C_2H_6, and HCN in the Southern Hemisphere point to the declining anthropogenic emissions of CO and C_2H_6. *Atmos. Chem. Phys.* **12**, 7543-7555 (2012).
35 Schwietzke, S., Griffin, W. M., Matthews, H. S. & Bruhwiler, L. M. P. Natural Gas Fugitive Emissions Rates Constrained by Global Atmospheric Methane and Ethane. *Environ. Sci. Technol.* **48**, 7714-7722 (2014).
36 Sherwood, O. A., Schwietzke, S., Arling, V. A. & Etiope, G. Global Inventory of Gas Geochemistry Data from Fossil Fuel, Microbial and Biomass Burning Sources, Version 2017. *Earth Syst. Sci. Data* **2017**, 639-656 (2017).
37 Stohl, A. *et al.* Black carbon in the Arctic: the underestimated role of gas flaring and residential combustion emissions. *Atmos. Chem. Phys.* **13**, 8833-8855 (2013).
38 McLinden, C. A. *et al.* Space-based detection of missing sulfur dioxide sources of global air pollution. *Nature Geosci* **9**, 496-500 (2016).

39 Nicewonger, M. R., Verhulst, K. R., Aydin, M. & Saltzman, E. S. Preindustrial atmospheric ethane levels inferred from polar ice cores: A constraint on the geologic sources of atmospheric ethane and methane. *Geophysical Research Letters* **43**, 214-221 (2016).

40 Etiope, G. & Ciccioli, P. Earth's Degassing: A Missing Ethane and Propane Source. *Science* **323**, 478-478 (2009).

41 Aikin, A. C., Herman, J. R., Maier, E. J. & McQuillan, C. J. Atmospheric chemistry of ethane and ethylene. *Journal of Geophysical Research: Oceans* **87**, 3105-3118 (1982).

42 Patra, P. K. *et al.* Observational evidence for interhemispheric hydroxyl-radical parity. *Nature* **513**, 219-223 (2014).

43 Strode, S. A. *et al.* Implications of carbon monoxide bias for methane lifetime and atmospheric composition in chemistry climate models. *Atmos. Chem. Phys.* **15**, 11789-11805 (2015).

44 Rigby, M. *et al.* Role of atmospheric oxidation in recent methane growth. *Proceedings of the National Academy of Sciences* **114**, 5373-5377 (2017).

45 Naik, V. *et al.* Preindustrial to present-day changes in tropospheric hydroxyl radical and methane lifetime from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP). *Atmos. Chem. Phys.* **13**, 5277-5298 (2013).

46 Søvde, O. A. *et al.* The chemical transport model Oslo CTM3. *Geosci. Model Dev.* **5**, 1441-1469 (2012).

47 Hoesly, R. M. *et al.* Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emission Data System (CEDS). *Geosci. Model Dev. Discuss.* **2017**, 1-41 (2017).

48 Stohl, A., Forster, C., Frank, A., Seibert, P. & Wotawa, G. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* **5**, 2461-2474 (2005).

49 Smith, M. L. *et al.* Airborne Ethane Observations in the Barnett Shale: Quantification of Ethane Flux and Attribution of Methane Emissions. *Environ. Sci. Technol.* **49**, 8158-8166 (2015).

50 Carpenter, L. J. *et al.* Seasonal characteristics of tropical marine boundary layer air measured at the Cape Verde Atmospheric Observatory. *Journal of Atmospheric Chemistry* **67**, 87-140 (2010).

51 Petrenko, V. V. *et al.* Minimal geological methane emissions during the Younger Dryas–Preboreal abrupt warming event. *Nature* **548**, 443-446 (2017).

52 Schwietzke, S. *et al.* Upward revision of global fossil fuel methane emissions based on isotope database. *Nature* **538**, 88-91 (2016).

53 Etiope, G., Lassey, K. R., Klusman, R. W. & Boschi, E. Reappraisal of the fossil methane budget and related emission from geologic sources. *Geophysical Research Letters* **35**, L09307 (2008).

54 Rigby, M. *et al.* Renewed growth of atmospheric methane. *Geophysical Research Letters* **35**, L22805 (2008).

55 Dlugokencky, E. J. *et al.* Observational constraints on recent increases in the atmospheric CH4 burden. *Geophysical Research Letters* **36**, L18803 (2009).

56 Tzompa-Sosa, Z. A. *et al.* Revisiting global fossil fuel and biofuel emissions of ethane. *Journal of Geophysical Research: Atmospheres* **122**, 2493-2512 (2017).

Corresponding author

Stig B. Dalsøren. E-mail: stig.dalsoeren@hi.no

Acknowledgements

This research is funded by the Research Council of Norway through the MOCA (Methane Emissions from the Arctic Ocean to the Atmosphere: Present and Future Climate Effects) project (Grant no. 225814). Furthermore, the conclusions of the paper is largely supported and strengthened by the use of globally distributed observational data and we acknowledge all data providers and the great efforts of EMEP, ACTRIS, NOAA ESRL/INSTAAR, and The World Data Centre for Greenhouse Gases (WDCGG) under the WMO-GAW program to make long-term measurements public and available. The Horizon 2020 research and innovation program ACTRIS-2 Integrating Activities (grant agreement No. 654109) is acknowledged for the work with quality assurance and quality control of NMHC data in Europe. The VOC observations within the NOAA-INSTAAR GGGRN are supported in part by the U.S. National Oceanic and Atmospheric Administration's Climate Program Office's AC4 Program, and quality control was supported in part by the U.S. National Science Foundation grant 1108391.

The GLOGOS dataset was kindly provided by CGG Geoconsulting. CGG also provided us with a derived product from the Global Offshore Seepage Database (GOSD) indicating where offshore seepage occurs.

Author contributions

S.B.D., G.M., and Ø.H. designed the study with input from A.S., C.L.M., and I.P. S.B.D performed the simulations with the OsloCTM3 model, analyzed the model results, and performed the comparisons with measurement data. Ø.H. and G.M. provided assistance with the analysis and comparison studies. I.P performed the simulations with the Flexpart model and I.P. and A.S. analyzed the output. S.SC. and L.H.I. provided the new emission datasets for fugitive fossil fuel emissions. S.B.D. developed gridded inventories for geologic emissions. C.L.M, D.H., S.R., S.SA., N.S., K.A.R., L.J.C., A.C.L., S.P., and M.W. provided the observational data for ethane and propane. S.B.D. led the writing of the manuscript in close collaboration with G.M. and Ø.H. All authors contributed to the writing and review of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Figure captions

Figure 1: Global total sectoral ethane emissions in this study and other studies. a) Global total ethane emissions in preindustrial simulations in this study (black symbols) compared to emissions in other studies (blue and yellow symbols and bars). *See Methods section for details. $^{\delta}$Box model study[39]. †Bottom up inventory[40]. **b)** Global total ethane emissions in the year 2011 baseline and alternative (ALT1, ALT2) simulations in this study (black symbols) compared to emissions in other studies. The inventories and estimates cover the year 2000 and onwards. The closest year to our simulation year 2011 is chosen for inventories not covering that year. Bottom up inventories (yellow symbols and bars)$^{\#}$: Fossil fuel, biofuel, agriculture, waste: CEDS CMIP6[47] (used in baseline in this study), HTAPv2, Edgar 3.2 FT, RETRO, POET, CMIP5 (average of MACCITY, ACCMIP, RCP2.6, RCP4.5, RCP6, RCP8.5) (as reported and referenced in ECCAD database: http://eccad.aeris-data.fr/), ARCTAS (as reported by ref. 7), EDGAR4.3.2 (as reported by ref. 11), and new inventory in ref. 56. Biomass burning: GFEDv4 (as used in baseline in this study), GICC, ACCMIP, POET, GFASv1.2, MACCITY, RETRO, RCP2.6, RCP4.5, RCP8.5 (as reported and referenced in ECCAD database: http://eccad.aeris-data.fr/), and FINN (as reported by ref. 3,4). Oceans: RETRO (used in baseline in this study). Vegetation: MEGAN-MACC (used in baseline this study), and MEGANv2 (as reported and referenced in ECCAD database: http://eccad.aeris-data.fr/). Natural geologic: As reported by ref. 40 (median estimate used in baseline in this study). Top-down estimates from box models (dotted blue bars)$^{\Phi}$: Fossil fuel, biofuel, agriculture, waste: ref. 31,33,39. Biomass burning: ref. 39. Geologic: ref. 39. Optimized emissions in 3D model studies (brown symbols)*: Ref. 3,4.

Figure 2: Observed and modeled annual mean ethane mixing ratios and inter-polar ratios. a) Observed[39] and modeled (this study) preindustrial inter-polar ratio and mixing ratios at Summit, Greenland and at the West Antarctic Ice Sheet (WAIS) site. Observation error bars are the reported ± 2 standard errors in ref. 39. **b)** Observed and modeled inter-polar ratio and mixing ratios in 2011. For the Antarctic, the closest station (South Pole) with data to the WAIS measurement site (no data for 2011) was used. See Table 1 for more information about the simulations.

Figure 3: Comparison of year 2011 modeled and observed ethane and propane at surface sites. a) Comparison of modeled and observed year 2011 ethane (upper row) and propane (lower row) at the Zeppelin station. A selection of comparisons for other sites is shown in the Supplementary. **b)** Comparison of modeled (background colors) and observed surface (color-filled circles) of ethane (ppt) for the year 2011. Model data for the lowest model layer were used. Stations with less than 6 samples within the 3 months averaging period were excluded from the comparison. Mountain stations typically sampling free tropospheric air and situated in areas where the model resolution does not resolve the terrain were also excluded. Details on the applied observation datasets are provided in the Methods section. Maps for the ALT2 simulation are shown in the Supplementary. **c)** Same for propane.

Figure 4: **Footprints at Zeppelin.** Yearly mean Flexpart footprints (see Methods for details on approach) for ethane at Zeppelin. Upper left: Baseline simulation. Episodes with Largest Underestimation (LU). Upper right panel: ALT1 simulation. Episodes with LU. Lower left: ALT1 simulation. Episodes with Largest Overestimation (LO). Lower right: ALT2 simulation. Episodes with LO. The unit ns/kg is proportional to the residence time in a given volume of air.

Methods

Models:

OsloCTM3:

We use the OsloCTM3 model[46] to simulate the preindustrial (year 1750) and current (year 2011) distributions of atmospheric ethane and propane. The model is run with approximately 1.1°x1.1° (T159) horizontal resolution. To spin up the model 15 months simulations were made with coarse resolution (2.2°x2.2°) followed by 4 months with 1.1°x1.1° (T159) resolution. After the spin up a set of simulations (Table 1) were made. A coupled tropospheric and stratospheric 60 layer (surface-0.1 hPa) version is used with 100 chemical active species affecting atmospheric oxidation capacity. OsloCTM3 was described and evaluated by ref. 46 and was used in several studies related to atmospheric oxidation capacity[57].

The OsloCTM3 simulations are driven with 3-hourly year 2011 meteorological forecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS) model (see ref. 46 for details). These data are 36-hours forecasts produced with 12 hours of spin-up starting from an ERA-Interim analysis at noon on the previous day.

Flexpart:

To investigate the origin of air masses observed at the Zeppelin station, we use version 9.2 of the LPDM FLEXPART (FLEXible PARTicle dispersion model)[48]. The model is driven with 3-hourly operational meteorological analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) with 91 vertical levels and a horizontal resolution of 1°×1°. Computational particles released from the location of the Zeppelin station are tracked 20 days back in time in FLEXPART's "retroplume" mode. The model output consists of an emission sensitivity, the surface footprint of which is used here to identify source regions related to OsloCTM3 under- and overestimation of observed ethane concentrations.

Emissions and model simulations

Baseline emission inventories all constituents

For anthropogenic SOx, NH_3, CO, NOx, and NMHCs emissions, we use the Community Emissions Data System (CEDS) Project emission inventory[47] for the years 1750 and 2011. CEDS is the state of the art dataset currently used in the Coupled Model Inter-comparison Project Phase 6 (CMIP6). For biomass burning, we use GFEDv4[58] year 2011 emissions and the historical biomass burning dataset for CMIP6 for 1750[59]. Sulfur emissions from other sources are described in ref. 60. Non-methane volatile organic compounds (NMVOCs) emissions from vegetation and oceans are neglected (set to zero) in some studies but not in this study. Biogenic emissions of CO and NMVOCs from vegetation are set to year 2010 (last year covered by dataset) emissions from MEGAN-MACC[61] both for the preindustrial and year 2011 simulations. For NO_x from soil and CO and NMHCs from the oceans we use the year 2000 emissions in the RETRO inventory[62]. Lightning NO_x emissions are described in ref. 46. For natural NH_3 sources we use emissions from GEIA[63] for 1990. Methane emissions are described in ref. 57.

Geologic emissions of ethane and propane

For ethane and propane we include geologic emissions suggested by ref. 40 in the baseline emission inventories. For both 1750 and 2011 the global total emissions are set to the medians (3 Tg/yr for ethane, 1.7 Tg/yr for propane) of the ranges (2-4 Tg/yr ethane, 1-2.4 Tg/yr propane) estimated by ref. 40. Their study splits the emissions into six main geologic sources: Mud volcanoes, gas seeps, microseepage, submarine seeps, geothermal manifestations and volcanoes. The geographical distribution of geologic emissions has not been gridded to files suitable for atmospheric modeling studies. Here we use several datasets to develop gridded inventories. The emissions from gas seeps (macroseepage) and mud volcanoes are distributed in accordance with the GLOGOS dataset which lists more than 2000 terrestrial (onshore) seeps from 87 countries. In addition to site locations, GLOGOS provides measured or estimated (visually) fluxes of methane for a few sites and methane, ethane and propane concentrations in the gas for some more sites. However, the majority of sites lack such information and we therefore distribute the emissions evenly on the sites to obtain the global total macroseepage emissions estimates from ref. 40. For submarine seeps we used a derived product (see acknowledgement) from the Global Offshore Seepage Database (GOSD) indicating where offshore seepage occurs. We scale the density map from this dataset to obtain the global total emissions from marine seepage in ref. 40. We assume zero emissions from marine seepage in grid-boxes fully covered by sea-ice. For

emissions from microseepage we use the CGG Robertson Tellus Sedimentary Basins of the World Map to distribute the emissions from ref. 40. Microseepage, which is diffuse exhalation from soil in petroleum basins, is the largest geologic source but also the most uncertain one[40,64] in terms of magnitude and geographical distribution. We spread the emissions evenly over the world's petroleum basins, which represents the potential area for such diffusion and we likely to some degree overestimate the geographical extent. We assume that permafrost or thick ice- and snow-layers hinder diffusion. The northward and southward extent of emissions is therefore limited to 66° N and 60° S to account for this in a simplified way. The emissions from volcanoes and geothermal sources are gridded using the geographical distribution for SO_2 emissions[60] for such activity. Geologic seepage (macroseepage, some microseepage, and marine seepage) occur at many of the places where current oil, gas, and coal extraction take place.

Baseline and alternative year 1750 preindustrial ethane emissions and simulations

The preindustrial baseline simulation (Table 1) includes geologic emissions of the magnitude suggested by Etiope et al.[40] and within the range found by Nicewonger et al.[39] (Figure 1a). In Nicewonger et al. an observed inter-polar ethane asymmetry requires a certain combination of emissions from biomass burning and natural geologic sources (Figure 1a). Preindustrial biomass burning emissions are particularly uncertain, and the magnitude in the inventory (CEDS CMIP6 year 1750) in our baseline simulation is high compared to the range obtained by ref. 39. We therefore perform an additional simulation using a biomass burning inventory (CMIP5 1850[65]) with lower emissions (Table 1, Figure 1a) and different geographical distribution. Transport to high southern latitudes has inter-annual variability. To check the sensitivity on our results we do a simulation with meteorological input data for a different year (Table 1).

Baseline year 2011 ethane and propane emissions and simulation

Figure 1b shows global total sectoral ethane emissions used in the year 2011 baseline simulation in this study compared to emissions used in other inventories. Based on the preindustrial simulations we include geologic emissions in our baseline simulation for current conditions (year 2011). Except for the box model optimized inventory for 2000-2010 in Nicewonger et al.,

569 geologic emissions were not included in any other model studies. Our applied emissions for
570 biomass burning (GFEDv4 year 2011) are in the mid-range compared to other inventories. The
571 other anthropogenic emissions (sum of fossil fuel, biofuel, agriculture, and waste in Figure 1b) in
572 the CEDS CMIP6 data[47] applied in this study are quite close to the median of the estimates in
573 other inventories. Due to likely trends in anthropogenic emissions over the last decades[3,4,31,33],
574 different basis years partly explain the large range of the emission estimates in various
575 inventories presented in Figure 1b. The other reason is the large uncertainties in existing
576 inventories due to incomplete approaches and datasets (discussed in main text and section
577 below). Summing up all sectors in Figure 1b our total emissions are in the upper range of other
578 studies. This is mainly due to the inclusion of geologic emissions.

579

580 *Alternative year 2011 ethane and propane emissions and simulations*

581 We also perform simulations with alternative ethane and propane emissions from the energy
582 sector (Table 1). In these we replace emissions from the energy sector in the CEDS CMIP6
583 inventory with two new datasets[9,10,35,36] for upstream (fuel production, gathering, and processing)
584 and downstream (transmission and distribution) emissions from oil, gas and coal systems. These
585 studies take much better into account that the emission factors from venting and flaring of
586 associated gas released during extraction vary considerably across different oil, coal and gas
587 fields in the world. The studies used novel approaches to quantify and attribute methane and
588 NMHC emissions, combining field measurements and country specific information from
589 published sources with observed flaring of associated gas from satellite images, to arrive at
590 country-specific emissions from flows of associated gas. Two simulations are performed with
591 these datasets. In what we refer to as the ALT1 simulation emissions from oil and gas from ref.
592 10 are combined with coal emissions from ref. 9 (updated with data from ref. 36) to obtain a
593 complete inventory for the energy sector. In the simulation entitled ALT2 we use oil, gas, and
594 coal emissions from ref. 9,35 updated with data from ref. 36. The dataset does not include
595 propane emissions and we use the global mean propane to ethane emission ratio from the ref. 10
596 datasets to obtain propane emissions for the ALT2 simulation. In ALT2 we use the low estimate
597 for natural gas from ref. 9 since it was shown to be the most likely[35]. Table 2 provides an
598 overview of the fugitive fossil fuel emissions in the ALT1 and ALT2 inventories. Due to

599 substantial geographical overlap it is likely that some emissions from geologic seepage are
600 included[35] in fugitive fossil fuel inventories. To avoid double-counting, we subtract ethane and
601 propane emissions from the oil, gas, and coal grids. In the absence of a well-established gridded
602 emission inventory of geologic seeps, we subtract geologic seepage ethane and propane
603 emissions from oil, gas, and coal grids in equal parts, i.e., one third each. By reducing with
604 amounts corresponding to all emissions from geologic seepage, the resulting inventories (ALT1
605 and ALT2) could be regarded as lower estimates of emissions from current oil, gas, and coal
606 activity relative to ref. 10 and ref. 9.

607

608 *Uncertainties in baseline and alternative anthropogenic emission inventories*

609 The uncertainty calculations for the baseline inventory and alternative (ALT1, ALT2) inventories
610 are discussed in the Supplementary in relation to Figure S11.

611

612 *Sensitivity simulation on atmospheric sink*

613 Oxidation by hydroxyl (OH) in the troposphere is the main sink for ethane and propane[1,2,41]. The
614 uncertainty for the reaction rates are rather small, 15-20 % at 298 K[2,66] . For hydroxyl
615 concentrations, the uncertainty is larger. Our global averaged tropospheric OH (1.35×10^6
616 molecules/cm^3) in our 2011 baseline simulation is on the high side compared to other model
617 studies[45]. The same holds for the hemispheric OH ratio (1.55) compared to model and
618 observational based estimates[42,44]. The modeled global average methane lifetime, which is highly
619 dependent on the modeled OH concentration, is also low compared to observational based
620 estimates[67,68]. We therefore did a sensitivity study scaling down the global mean OH
621 concentration to 1×10^6 molecules/cm^3. The scaling was done separately for the hemispheres to
622 also get a hemispheric ratio of 1. This is in the lower range of model and observation based
623 estimates both for the OH concentration and hemispheric ratio and was done to see to what
624 degree lower OH concentrations in the Northern Hemisphere would improve the comparison
625 with observations. The OH concentration was scaled down only in the chemical reactions with
626 ethane and propane and not in reactions with other atmospheric constituents.

Measurement data

More and more high quality measurements of NMHC have become available through coordination by the World Meteorological Organization (WMO) Global Atmospheric Watch (GAW) program. Participating networks include ACTRIS (Aerosol, Clouds, and Trace gases Research Infrastructure, the European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases (http://www.actris.eu)), EMEP (The European Monitoring and Evaluation Program) and NOAA ESRL/INSTAAR (National Oceanic & Atmospheric Administration - Earth System Research Laboratory/Institute for Arctic and Alpine Research). To reveal strengths and discrepancies in model performance and evaluate emission inventories we compare the model results to surface ethane and propane observations for the year 2011. We use data from surface sites reported to the World Data Center for Greenhouse Gases (WDCGG) (http://ds.data.jma.go.jp/gmd/wdcgg/) following the NOAA/INSTAAR and WMO-GAW measurement protocols, and from EMEP complying with ACTRIS recommendations. EMEP and ACTRIS data were downloaded from EBAS (http://ebas.nilu.no), and are also accessible through the ACTRIS data portal (http://actris.nilu.no). These data follow the ACTRIS quality assurance procedures and protocols. Generally, the ACTRIS data have higher time resolution (up to 2 hour). Intercomparison exercises[69-71] have shown consistency of NOAA/INSTAAR and EMEP/ACTRIS measurement sites within the data quality objective of +/-10% of the WMO Global Atmospheric Watch program[72,73].

A subset of data from 96 sites from a total of 132 sites (66 for both ethane and propane) are shown in the comparisons for 2011. The observations at the given location, altitude and time were compared to output for the closest model grid box, level, and time. Figure 3 shows all sites and Figure S4 shows the locations of the subset of sites selected for detailed comparison to model results. Criteria for selection were data quality assurance, access to continuous time series with few gaps, coverage of different regions, and site characteristics (e.g. elevation, topography, and influence of pollution episodes) likely captured by the resolution of a global model.

A comparison between modelled and observed ozone is performed in the Supplementary and the applied ozone measurement data are presented there.

Table 1: Overview of the simulations performed with OsloCTM3.

Year	Simulation	Simulation name	Inventories anthropogenic and natural emissions*	Geologic emissions	Oxidation chemistry	Meteorological year
1750	Preindustrial	1750 Baseline	Baseline	Yes	Interactive	2011
	Preindustrial sensitivity	1750 NOGEO	Baseline	No	Interactive	2011
		1750 MET	Baseline	Yes	Interactive	2014
		1750 BBURN	Baseline, but CMIP5 biomass burning	Yes	Interactive	2011
2011	Baseline	2011 Baseline	Baseline	Yes	Interactive	2011
	Baseline sensitivity	OH	Baseline	Yes	Interactive, OH field scaled down in reactions with ethane and propane**	2011
	Alternative	ALT1	ALT1	Yes	Interactive	2011
		ALT2	ALT2	Yes	Interactive	2011

*The baseline, ALT1 and ALT2 emission inventories are described in the above section.

****** The scaling of OH is described in the above section.

Table 2: Overview of the new fugitive fossil fuel emission inventories used in the alternative year 2011 simulations.

Inventory	Natural gas	Oil	Coal
ALT1 ethane	ref. 10	ref. 10	ref. 9 updated with data from ref. 36
ALT1 propane	Same as above	Same as above	Same as above
ALT2 ethane	ref. 9,35 updated with data from ref. 36	ref. 9 updated with data from ref. 36	ref. 9 updated with data from ref. 36
ALT2 propane	Same as above using propane to ethane ratio from ref. 10	Same as above using propane to ethane ratio from ref. 10	Same as above using propane to ethane ratio from ref. 10

Code availability. We have opted not to make the computer codes associated with this paper available, because replication of our results does not require access to the computer codes.

Data availability. The ethane and propane surface measurement data used in this study are freely available. We use data reported to the World Data Center for Greenhouse Gases (WDCGG) (http://ds.data.jma.go.jp/gmd/wdcgg/). EMEP and ACTRIS data are available from EBAS (http://ebas.nilu.no), and are also accessible through the ACTRIS data portal (http://actris.nilu.no). The sites used for detailed comparison with model results are listed in Table S1. The sites can easily be found by name or map search in the databases. The new emission datasets for fugitive fossil fuel emissions are available upon request to Lena Höglund-Isaksson (hoglund@iiasa.ac.at) and Stefan Schwietzke (stefan.schwietzke@noaa.gov). The gridding used to develop geologic ethane and propane emissions suitable for atmospheric modeling studies was based on commercial datasets owned by CGG geoconsulting. CGG should be contacted and consulted for the task of gridding geologic emissions.

References

57 Dalsøren, S. B. *et al.* Atmospheric methane evolution the last 40 years. *Atmos. Chem. Phys.* **16**, 3099-3126 (2016).

58 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**, 11707-11735 (2010).

59 van Marle, M. J. E. *et al.* Historic global biomass burning emissions based on merging satellite observations with proxies and fire models (1750-2015). *Geosci. Model Dev. Discuss.* **2017**, 1-56 (2017).

60 Berglen, T., Berntsen, T., Isaksen, I. & Sundet, J. A global model of the coupled sulfur/oxidant chemistry in the troposphere: The sulfur cycle. *J. Geophys. Res.: Atmos.* **109**, 27 pp (2004).

61 Sindelarova, K. *et al.* Global dataset of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmos. Chem. Phys. Discuss.* **14**, 10725-10788 (2014).

62 Schultz, M. *et al.* Emission data sets and methodologies for estimating emissions.REanalysis of the TROpospheric chemical composition over the past 40 years. A long-term global modeling study of tropospheric chemistry funded under the 5th EU framework programme. EU-Contract No. EVK2-CT-2002-00170 (2008).

63 Bouwman, A. F. *et al.* A global high-resolution emission inventory for ammonia. *Global Biogeochem. Cycles* **11**, 561-587 (1997).

64 Etiope, G. & Klusman, R. W. Microseepage in drylands: Flux and implications in the global atmospheric source/sink budget of methane. *Global and Planetary Change* **72**, 265-274 (2010).

65 Lamarque, J. F. *et al.* Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmos. Chem. Phys.* **10**, 7017-7039 (2010).

66 Atkinson, R. Kinetics of the gas-phase reactions of OH radicals with alkanes and cycloalkanes. *Atmos. Chem. Phys.* **3**, 2233-2307 (2003).

67 Prinn, R. G. *et al.* Evidence for variability of atmospheric hydroxyl radicals over the past quarter century. *Geophysical Research Letters* **32**, L07809 (2005).
68 Prather, M. J., Holmes, C. D. & Hsu, J. Reactive greenhouse gas scenarios: Systematic exploration of uncertainties and the role of atmospheric chemistry. *Geophysical Research Letters* **39**, L09803 (2012).
69 Rappenglück, B. *et al.* The first VOC intercomparison exercise within the Global Atmosphere Watch (GAW). *Atmos. Environ.* **40**, 7508-7527 (2006).
70 Hoerger, C. C. *et al.* ACTRIS non-methane hydrocarbon intercomparison experiment in Europe to support WMO GAW and EMEP observation networks. *Atmos. Meas. Tech.* **8**, 2715-2736 (2015).
71 Plass-Dülmer, C., Schmidbauer, N., Slemr, J., Slemr, F. & D'Souza, H. European hydrocarbon intercomparison experiment AMOHA part 4: Canister sampling of ambient air. *Journal of Geophysical Research: Atmospheres* **111**, D04306 (2006).
72 Schultz, M. G., H. Akimoto, J. Bottenheim, B. Buchmann, I.E. Galbally, S. Gilgle, et al. The Global Atmosphere Watch reactive gases measurement network. *Elem. Sci. Anth.* **3** (2015).
73 Helmig, D. *et al.* Volatile Organic Compounds in the Global Atmosphere. *Eos, Transactions American Geophysical Union* **90**, 513-514 (2009).