Measurement report: Assessment of Asian emissions of ethane and propane with a chemistry transport model based on observations from the island of Hateruma

Adedayo R. Adedeji[1], Stephen J. Andrews[1], Matthew J. Rowlinson[1,2], Mathew J. Evans[1,2], Alastair C. Lewis[1,2], Shigeru Hashimoto[3], Hitoshi Mukai[3], Hiroshi Tanimoto[3], Yasunori Tohjima[3], and Takuya Saito[3]

[1]Wolfson Atmospheric Chemistry Laboratories, Department of Chemistry, University of York, York, UK
[2]National Centre for Atmospheric Science, Department of Chemistry, University of York, York, UK
[3]National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan

Correspondence: Adedayo R. Adedeji (adedayo.adedeji@york.ac.uk; rasaqdayo@gmail.com)

Abstract. The island of Hateruma is the southernmost inhabited island of Japan. Here we interpret observations of ethane (C_2H_6), and propane (C_3H_8), together with carbon monoxide (CO), nitrogen oxides (NO_x and NO_y) and ozone (O_3) made from the island in 2018 with the GEOS-Chem atmospheric chemistry transport model. We simulated the concentrations of these species within a nested grid centered over the site, with a model resolution of $0.5° \times 0.625°$. We use the Community Emissions Data System (CEDS) emissions dataset for anthropogenic emissions and add a geological source of C_2H_6 and C_3H_8. The model captured the seasonality of primary pollutants (CO, C_2H_6, C_3H_8) at the site - high concentrations in the winter months when oxidation rates are low and flow is from the north, and low concentrations in the summer months when oxidation rates are higher and flow is from the south. It also simulates many of the synoptic scale events with Pearson's correlation coefficients (r) of 0.74, 0.88 and 0.89 for CO, C_2H_6 and C_3H_8, respectively. Concentrations of CO are well simulated by the model (with a gradient of best fit between model and measurements of 0.91) but simulated concentrations of C_2H_6 and C_3H_8 are significantly lower than the observations (gradients of best fit between model and measurement of 0.57 and 0.41, respectively), most noticeably in the winter months. Simulated NO_x concentrations were underestimated but NO_y appear to be overestimated. The concentration of O_3 is moderately well simulated (gradient of best fit line of 0.76, with an r of 0.87) but there is a tendency to underestimate concentrations in the winter months. By switching off the model's biomass burning emissions we show that during winter biomass burning has limited influence on the concentration of compounds in the winter but can represent a more sizeable fraction in the summer. We also show that increasing the anthropogenic emissions of C_2H_6 and C_3H_8 within the domain by factors of 2.22 and 3.17, increases the model's ability to simulate these species in the winter months, consistent with previous studies.

1 Introduction

Trans-boundary pollution of the atmosphere is a major global issue (Itahashi et al., 2017; Shairsingh et al., 2019; Qu et al., 2021). This is especially true in East Asia where highly polluting regions can significantly impact the atmospheric composition

large distances downwind due to the transport of long-lived atmospheric pollutants. As well as the transport of primary pollutants, secondary production en route of pollutants such as ozone (O_3) and Particulate Matter (PM) can impact the concentration of these pollutants long distances from their sources (Griffith et al., 2020; Itahashi et al., 2017; Kim, 2019).

25   Ozone in the troposphere is produced from the oxidation of primary emitted compounds such as volatile organic compounds (VOCs), carbon monoxide (CO) and methane (CH_4) in the presence of oxides of nitrogen (NO_x). Thus increases in the emissions of these compounds can be expected to lead to increases in the concentration of O_3 downwind of the emission sites. The rapid industrialization in East Asia has led to increased emissions of these primary compounds and this in turn has, led to increases in the concentration of O_3 (Gaudel et al., 2018). Over the coming decades, there are predictions for this trend to

30   continue (Naja and Akimoto, 2004; Li et al., 2016). It is therefore important to measure the regional concentrations of atmospheric pollutants such as O_3 and its precursors (VOCs, NO_x, CO etc) in East Asian source regions and downwind over the coming years.

Finding suitable sites for making these long-term downwind measurements is however difficult. Sites need to be remote from local influences, yet sufficiently accessible for staff to visit for instrument maintenance and upgrades, and for data to

35   be transmitted back for processing etc. Hateruma island is a small island (12.7 km^2) located off the coasts of both Taiwan and Japan and so is subject to East Asian outflow (Figures 1(a) and 1(b)). It is the southernmost inhabited island (24.05° N, 123.80° E) in the Japanese archipelago, 500 km southwest of Japan's Okinawa Island, and is 250 km off the coast of Taiwan in the Pacific ocean (Yokouchi et al., 2011). As a part of its global monitoring effort, the Japanese National Institute for Environmental Studies' (NIES) Centre for Global Environmental Research (CGER) operates an atmospheric observatory on

40   the island to carry out atmospheric measurements. CGER has made measurements at the site of atmospheric constituents over a number of years. During the winter the site is characterised by northerly winds and elevated pollution influenced by emissions from East Asian countries, whereas in the summer the air comes from the south and is typically cleaner (Tohjima et al., 2002, 2010). Observations of atmspheric composition at Hateruma island have been used to analyse different problems. For instance, Yokouchi et al. (2006) and Shirai et al. (2010) used observations from Hateruma to assess national emissions

45   of HFCs and HCFCs from China, Korea, and Japan. Saito et al. (2010) continuously measured the atmospheric mixing ratios of perfluorocarbons (PFCs) at Hateruma Island and Cape Ochiishi since 2006, to infer their global and regional emissions. Tohjima et al. (2014) used observation of Carbon Monoxide (CO), Carbon Dioxide (CO_2) and Methane (CH_4) to assess changing emissions from China, and then went on to use a similar technique to assess the drop in CO_2 emissions from China due to COVID-19 regulation (Tohjima et al., 2020).

50   A wider range of observations are made at the site than have been previously published. These include measurements of CO, ethane (C_2H_6), propane (C_3H_8), nitrogen oxides (NO_x and NO_y) and O_3. Although these observations themselves are useful, the use of a chemical transport model (CTs) allows observations to be put into the context of our wider understanding of atmospheric emissions, deposition, transport and chemistry allowing.
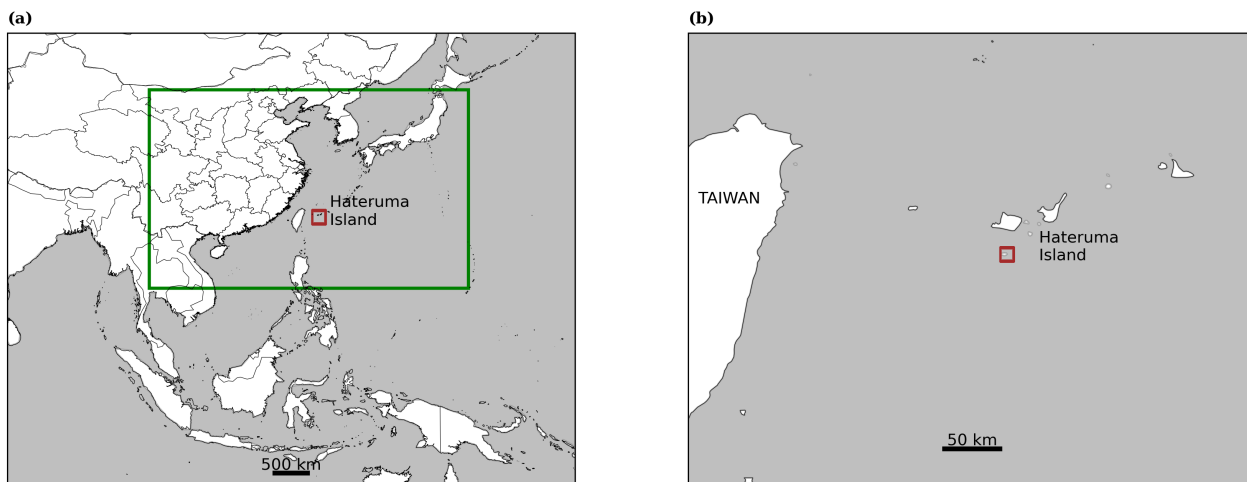
Figure 1. (a) The region of the high resolution modelling domain - spans from $14°$ N to $42°$ N and $100°$ E to $145°$ E, indicated by the green box;**(b)** zoomed in to show the location of the island of Hateruma - $24.05°$ N, $123.80°$ E, inside the red square (b).

Here we use a chemical transport model (GEOS-Chem) together with observations of the key atmospheric gases (CO, C_2H_6, C_3H_8, NO, NO_2, NO_y, and O_3) measured at the observatory in Hateruma in 2018 to evaluate the model's ability to simulate the long-range transport of key species to the island of Hateruma and our understanding of emissions in the region. We start with a description of the observations made at the site (Section 2) and provide a meteorological context to the observations using back trajectories. We then describe the GEOS-Chem model configuration in Section 3 and show an evaluation of the model performance in Section 4. Based on this evaluation we assess the sensitivity of the model to biomass burning emissions in Section 5, and in Section 6 we explore scaling anthropogenic Asian C_2H_6 and C_3H_8 emissions to better reflect the observations at the site. We draw conclusions in Section 7.

2   Observations and meteorological context

In this section, we describe the observations made at the site used for this study and the meteorological context of the air arriving at the site.

2.1   Observations

The observatory is located on the easterly corner of the island, with a disused airport and the populated area (with electricity generation from diesel engines and wind power) located to the west. We expect limited local anthropogenic emissions on the island because of its small size and population. The location of the population and power generation to the west of the observations also limits potential contamination at the site from most wind directions. A number of different observations are

made at the site but we are only concerned with a subset of the observations in this study. Table 1 gives the observations used in the study and the method used to make them. Observations are available hourly for most of 2018 with some missing periods.

Table 1. Physical and chemical variables measured hourly at Hateruma between 01 January to 31 December 2018 (see references for QA/QC information). Data capture indicates the number of hours when observations are available out of the 8760 for the year.

Variables	Method	Data Captured	Reference
Temperature	Pt100 (YOKOGAWA E734)	8751/8760	-
Wind speed	Wind speed and direction transmitter (YOKOGAWA WA7601)	8751/8760	Shirai et al. (2010)
Wind direction	Wind speed and direction transmitter (YOKOGAWA WA7601)	8751/8760	Shirai et al. (2010)
Relative humidity	Capacitance hygrometer (VISALA HUMICAP HMP155)	8751/8760	-
Carbon Monoxide (CO)	GC/RGD (Peak Performer 2, Peak Laboratories)	8604/8760	Tohjima et al. (2014)
Ethane (C_2H_6)	Preconcentration/GC/MS	5600/8760	Saito et al. (2010)
Propane (C_3H_8)	Preconcentration/GC/MS	5600/8760	Saito et al. (2010)
Nitrogen oxide (NO)	$NO/NO_2/NO_y$ analyzer (HSS-100, Sonoma Technology Inc.)	432/8760	-
Nitrogen dioxide (NO_2)	$NO/NO_2/NO_y$ analyzer (HSS-100, Sonoma Technology Inc.)	8088/8760	-
Nitrogen dioxide (NO_y)	$NO/NO_2/NO_y$ analyzer (HSS-100, Sonoma Technology Inc.)	8088/8760	-
Ozone (O_3)	UV absorption O_3 analyzer (Model 1100, Dylec Inc.)	8600/8760	-

Measurements of CO, ethane, and propane were made with outside air drawn from a main tower (sampling inlet: 36.5 m above ground and 46.5 m above sea level). CO was measured with a gas chromatograph/reduction gas detector (GC/RGD; Peak Performer 2, Peak Laboratories) (Tohjima et al., 2014). Ethane and propane were measured with an automated preconcentration/gas chromatography/mass spectrometry (GC/MS) instrument (7890B/5977B, Agilent Technologies) designed for hourly measurements of natural and anthropogenic halocarbons (Yokouchi et al., 2006; Saito et al., 2010). Ethane and propane measurements were calibrated using a gravimetrically prepared standard gas (Taiyo Nippon Sanso Co. Ltd.).

Air for $NO/NO_2/NO_y$ and O_3 measurements were drawn from a sub tower (sampling inlet: 14.8 m above ground and 24.8 m above sea level). Measurements of NO, NO_2, and NO_y were made by HSS-100 (Sonoma Technology Inc. (STI)) with a blue light converter for NO_2 (NO_x converter) and a molybdenum converter for NO_y (NO_y converter) (Galbally, 2020). The conversion efficiencies of the converters were daily monitored with NO_2 gas ($\approx$5 ppb) generated by a gas dilution system with gas phase titration (NO_x: 79$\pm$3%, NO_y: 98$\pm$1%). Detection limit of the NO measurement was about 2 ppt for 1 minute acquisition. Ozone concentration was observed by UV absorption method with an ozone monitor (Dylec 1100, Dylec Co., Ltd), which was calibrated with Standard Reference Photometer No. 35 (National Institute for Standard and Technology) located at NIES.

2.2 Meteorological context

Over the course of a year, the observatory is exposed to air masses from a wide range of locations. We calculate ten-day back trajectories for the site every hour from January to December in 2018 (52 weeks duration) using meteorological data

from NCEP Global Forecast System (National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2015) and the FLEXible PARTicle dispersion model (FLEXPART) (Stohl et al., 2003, 2005).

Figure 2 (right) shows the ratio of the time that these ten day trajectories spent over the different regions shown in Figure 2 (left). Over the year, most of the air is oceanic in origin from either the Northern or Southern Pacific (43% and 34%). However the air masses can also spend a significant fraction of the time over Inland China (3.6%), Russia (4.6%), Korea (1.7%), Japan, (3.9%), South East Asia (3.1%), the Philippines (1.2%) and the three East China city regions (3.7%).
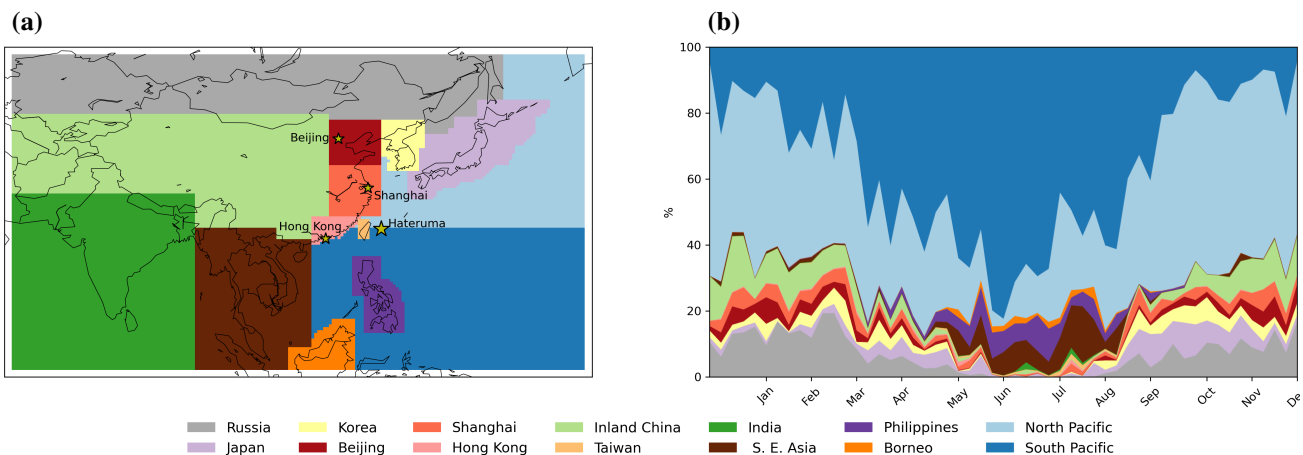
Figure 2. Trajectory analysis of air arriving at Hateruma; **(a)** location of the regions analysed; **(b)** weekly average percentage of time air masses spent over these regions

There is, however significant seasonal variation in air mass origin. The fraction that is oceanic (North and South Pacific) is lowest in the winter at around 60% with is more prevalent in the summer at around 80%. There is a significant shift from the North Pacific region in the winter to the South Pacific region in the summer. This pattern can be attributed to the annual meteorological cycle of this region characterised by the East Asian monsoon seasons (Shirai et al., 2010; Yokouchi et al., 2006, 2011, 2017). For the remaining 30% - 40% of the time, flow from China, Japan, Korea and Russia dominates in the winter, spring and autumn. During the summer months (April to August), the air originates from a southerly direction (Philippines, South East Asia, Borneo).

The site is therefore mainly subject to relatively clean oceanic air with occasional exposure to air from the Russia, Korea and China in the winter and Philippines, S.E Asia and Borneo in the Summer.

3   Chemical Transport Model configuration

We use a chemical transport model, GEOS-Chem version 12.7.1 (doi://10.5281/zenodo.3676008) to help analyse the measurements made at Hateruma. The model was first described by Bey et al. (2001) but hs had substantial improvements since then (www.geos-chem.org). We use a regional simulation with a spatial resolution of $0.5° \times 0.625°$ over the domain shown

in Figure 1(a). The model includes detailed HO_x–NO_x–VOC -ozone–halogen–aerosol tropospheric chemistry as described by Sherwen et al. (2016) and is driven by offline meteorology from the NASA Global Modelling and Assimilation Office (http://gmao.gsfc.nasa.gov, last access: 05 September 2021) forward processing product (GEOS-FP).

To generate restart files and boundary conditions, the model was run in a global configuration at $4°\times5°$ horizontal resolution for two years (1st January 2017 - 1st January 2019) with the first year considered as spin up. We then run the model in its regional configuration using the boundary conditions derived from the global simulation and output the result every hour. Figures A1 - A4 in the Appendix shows comparison between the model simulations run at the native meteorological resolution of $0.25°\times0.3125°$and the simulation run at $0.5°\times0.625°$. The differences between the two models are minimal, so we adopt the coarser $0.5°\times0.625°$resolution version as it is substantially faster to run with only a small degradation in performance.

We run the model in a slightly different emission configuration than its default. We use the the Community Emissions Data System (CEDS) emissions with applied monthly seasonality as described in Hoesly et al. (2018) and diurnal scale factor for all anthropogenic emissions, other than for those from aircraft where we use the Aviation Emissions Inventory Code (AEIC) emissions (Stettler et al., 2011). This contrasts with the model default configuration which uses Tzompa-Sosa et al. (2017) and Xiao et al. (2008) for anthropogenic ethane and propane emissions, but does makes our simulation consistent with recent CMIP model evaluations (Griffiths et al., 2021). The CEDS emissions were available at 0.1 degree resolution while Harmonized Emissions Component (HEMCO) module as described in Keller et al. (2014) and Lin et al. (2021) interpolates all emissions to the grid resolution ($0.5°\times0.625°$) of the model.

We also include geological emissions of ethane and propane that may represent a missing source (Etiope et al., 2019; Dalsøren et al., 2018). We assume a global total of 3.0 Tg year^{-1} of ethane and 1.7 Tg year^{-1} of propane (Dalsøren et al., 2018). We spatially distribute these emissions using a geological methane emission dataset (Etiope et al., 2018). We do not believe that these geological emissions are of specific importance for this region and only represent only 10-20% of the global emissions of both ethane and propane.

Other emissions include offline soil NO_x (Weng et al., 2020) and online lightning NO_x emissions (Murray et al., 2012). The PARANOX ship plume model by Holmes et al. (2014) which calculates the aging of emissions in ship exhaust plumes for NO_x,HNO_3 and O_3 was applied to the CEDS shipping emissions (Hoesly et al., 2018). Biomass burning emissions uses the Global Fire Emissions Database GFED 4.1 (Andreae and Merlet, 2001; Akagi et al., 2011; Giglio et al., 2013; Randerson et al., 2012; van der Werf et al., 2017). Biogenic emissions follow the estimation from the Model of Emissions of Gases and Aerosols from Nature (MEGAN) 2.1 (Guenther et al., 2012). The natural emissions of acetaldehyde follow the calculation from Millet et al. (2010). Natural sources of NH_3 are adapted from Global Emission InitiAtive (GEIA) as described by Bouwman et al. (1997), with the inclusion of the arctic seabird emissions (Riddick, 2012).

Table 2 gives the list of simulations performed and the sections in which these simulations are discussed.

Table 2. Summary of GEOS-Chem model simulations and reference sections.

Model simulation	Type	Section
1st	Base simulation	3 and 5
2nd	No-shipping emission simulation	3.7
3rd	North Asian biomass burning emission turned off	4
4th	South Asian biomass burning emission turned off	4
5th	Scaled anthropogenic emission simulation 1	5
6th	Scaled anthropogenic emission simulation 2	5

4 Model results and performance

140 In order to assess the veracity of the model simulation we compare the performance of the chemical transport model (GEOS-Chem), first against the meteorological observations made at the site and then against measurements of atmospheric chemical constituents. Since the model and observation data are available hourly and observation data capture is high in most cases, instances of missing data was dropped for both observation and model before comparison. Also, the observation data was adjusted from Japan local time (JST) to model default time (UTC). We compare hourly values from the observations and from
145 the model for the physical and chemical variables except for NO, NO_2 and NO_y where we compare local daylight averages due to high variability.

We use a number of standard metrics for describing the measurements and model (mean, median, standard deviation, 25th and 75th percentiles). When assessing the model performance we assess this in terms of the Root Mean Square Error (RMSE), mean bias and Pearson's correlation coefficient (r). We also calculate the slope of the best-fit lines using Orthogonal Distance
150 Regression (ODR).

In the case of wind direction, which is circular, not linear, the methodology for the estimation of RMSE and Bias follows Ferreira et al. (2008) and Carvalho et al. (2012).

4.1 Meteorological variables

Figure B1 in Appendix B shows the wind-rose representing observed and modelled wind speed and direction at the site. As the
155 populated part of the island lays to the west of the site this highlights that the likely local contamination is likely small. Figure 3 shows a comparison between the hourly observed and modelled time series for wind speed, wind direction, temperature and relative humidity at the observatory.

The wind direction (10 m above surface) is plotted in Figure 3 (a) as the incidence angle of the wind (North = $0°$ and $360°$). The observations are given with a resolution of $22.5°$ clockwise, which gives them a blocky appearance. Overall, the model
160 captured the observed wind direction well throughout the year with few discrepancies. The modelled wind direction has a low RMSE ($6.5°$) and a slightly positive bias ($6.5°$). This could be attributed to a small error in meteorological fields used in our simulations or to potentially a small error in the measurements. The seasonal variability of wind direction is well represented

in Figure 3 (top left), with the summertime characterised by winds coming from the south, and the winter having winds from a more northerly direction.

165   The modelled wind-speed have a correlation coefficient (r) of 0.90 and a RMSE of 2.17 m/s (Figure 3 (b)). However, the wind-speeds are lower in the model (mean of 6.43 +/- 3.12 (1σ) m/s), than in the observations (mean of 7.83 +/- 3.79 (1σ) m/s). This may reflect local reductions in wind speed over the island due to increased surface drag not being represented in the model due to its grid-resolution.

        Surface temperatures (2 m above surface), show a high degree of correlation (r = 0.92) and a RMSE of 1.75 °C, with the
170   mean observed surface temperature (24.53 +/- 4.27 (1σ) °C) close to that simulated (24.76 +/- 3.67 (1σ) °C). There is, however, significant hourly variation not captured by the model (Figure 3 (c)). This again could be due to resolution impacts at different scales since the island is smaller than the model grid box at Hateruma. Local heating and cooling at the observatory, and the island will not be represented in the model fields.

        The modelled surface relative humidity (RH at 2m above surface in Figure 3 (d)) is less well captured (r = 0.77; RMSE
175   = 11.26 %), with the modelled mean (73.09 +/- 9.89 (1σ) %) lower than the observed mean (82.00 +/- 10.32 (1σ) %). This difference could be attributed to the sub-grid issue mentioned earlier where the meteorology used in the model is on a fairly coarse scale while the observations are local.
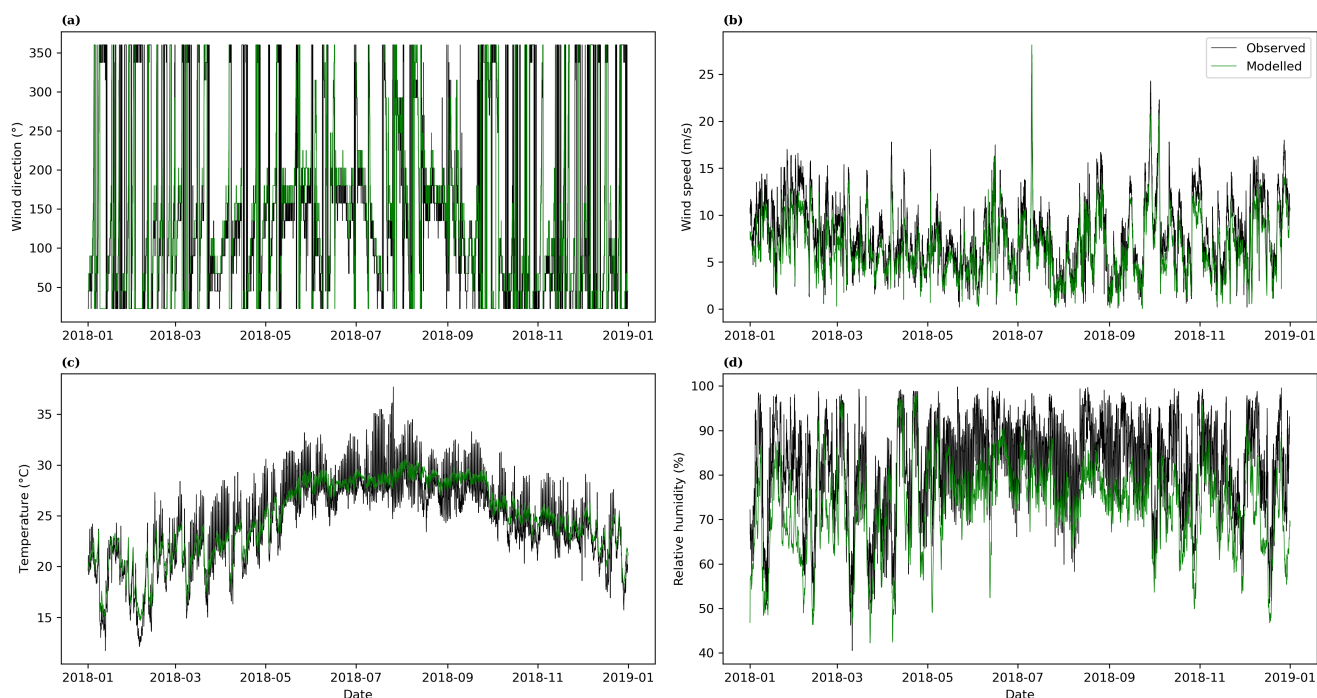
Figure 3. Observed and modelled **(a)** wind direction, **(b)** wind speed , **(c)** temperature and, **(d)** relative humidity at Hateruma

In general, the model performance compared to the meteorological observations is relatively good, reflecting the observational data assimilated into the GEOS-FP meteorological fields. We can now assess the model performance in simulating the concentration of the trace gas concentrations observed at the site.

4.2 Carbon monoxide

Figure 4 (a) (left) shows the time series of CO measured and modelled at the site, with Figure 4 (a) (right) showing the correlation between the hourly modelled and measured values. Concentrations are highest in the winter months reflecting transport from North Asia (China, Korea, Japan) and the North Pacific (Figure 2), and lower oxidation by OH in the winter months (Schlesinger and Bernhardt, 2020). They are lower in the summer reflecting transport from southern regions (further away and typically have lower emissions), and increased oxidation due to enhanced OH concentrations in the summer (Schlesinger and Bernhardt, 2020). Koike et al. (2006) measured CO in northern Japan and reported similar seasonality in CO concentrations over the year.

Over the year, the model simulates CO concentration well with a mean value of 146 +/- 77 (1σ) ppbv compared to mean observations of 136 +/- 62 (1σ) ppbv. The median CO concentrations calculated and measured are 127 ppbv and 122 ppbv, respectively. The 25th and 75th percentile modelled is 89 ppbv and 180 ppbv compared to 89 ppbv and 169 ppbv observed. There is a good degree of correlation (r = 0.74) between the model and measurements with a RMSE of 53 ppbv, and the line of best fit having a gradient of 0.91.

4.3 Ethane

Similar to CO, the highest observed ethane (C_2H_6) concentrations occur in the winter with substantially lower concentration occurring in the summer (Figure 4 (b) - left). The model underestimates the mean concentrations (972 +/- 541 (1σ) pptv) compared to the observations (1188 +/- 842 (1σ) pptv)) with a RMSE of 481 pptv. Median concentrations are similarly calculated to be 942 pptv, with a 25th to 75th percentile spread of 552-1386 ppbv, while the observations are 1134 pptv with a spread of 368-1825 pptv. This underestimate in modelled ethane is most noticeable in the winter months. However, in the summer the model can overestimate. This leads to two populations in the model - measurement scatter plot (Figure 4 (b) - right). With a strongly correlated but overestimated population at observed concentrations below 500 pptv (in the summer), and another population, with underestimated concentration above this (in winter). The correlation coefficient between the model and measurement is relatively high at 0.88 but the line of best fit between model and measurement is dominated by the wintertime underestimate to give a slope of 0.57.

4.4 Propane

Figure 4 (c) (left) shows the comparison between the measured and modelled time series for propane (C_3H_8). A similar pattern to C_2H_6 is seen with high winter time concentrations and much lower values in the summer. Unlike C_2H_6 though, a summer time overestimate is not evident. The model captures much of the variability in the C_3H_8 concentration with a correlation
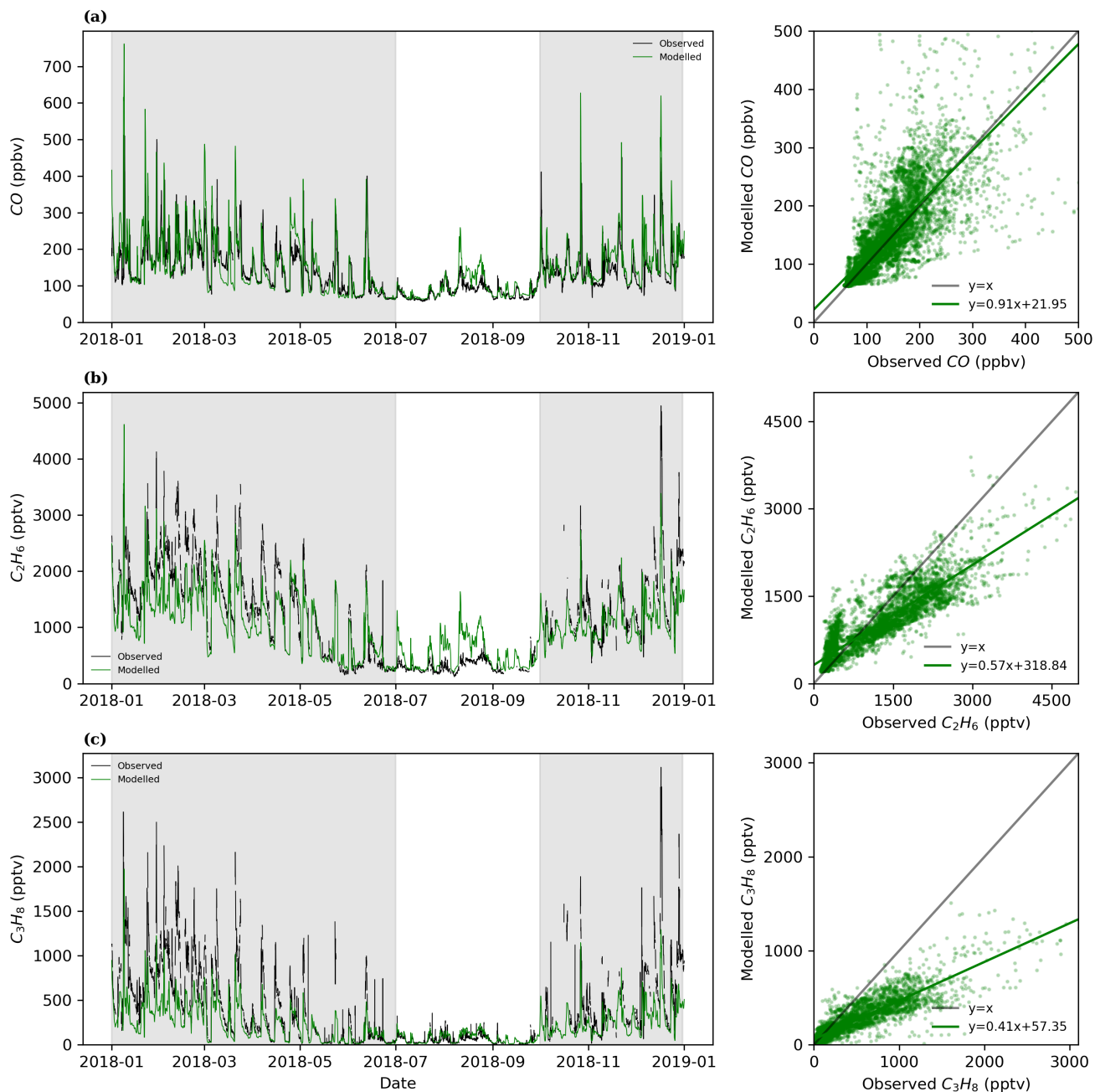
Figure 4. Modelled and measured hourly **(a)** CO, **(b)** C_2H_6, and **(c)** C_3H_8 concentrations calculated for Hateruma site

coefficient of 0.89 (Figure 4 (c)- right). The mean concentration of C_3H_8 is 391 +/- 432 (1σ) pptv for the measurements and
210   210 +/- 198 (1σ) pptv for the model. The median C_3H_8 concentration calculated and measured is 156 pptv (25th and 75th

percentiles of 69 pptv and 335) and 221 pptv (25th and 75th percentiles of 62 pptv and 611 pptv) respectively. The RMSE is 320 pptv and the gradient of the line of best fit between calculated results and measurements is 0.41. Thus, similar to C_2H_6, the model substantially underestimates C_3H_8 concentrations.

4.5 NO_x

215   We evaluate here the model's performance in simulating NO and NO_2 (collectively known as NO_x). The short lifetime of NO and NO_2 (on the order of minutes to hours during the day) makes evaluation against model data difficult due to large and rapid variations. During the day NO concentrations are high due to its photolytic NO_2 source, whereas, at night, away from recent emissions, its concentrations are effectively zero. For this evaluation, we compare mean day values (06:00 local to 18:00 local) between the model and observation. Figure 5 (a) shows the comparison between the measured and modelled daytime mean for
220   the NO concentration. The mean observed value (7 +/- 28 (1σ) pptv) is substantially lower than the modelled value (21 +/- 41 (1σ) pptv). The model also shows very little skill in the day to day variability with an r of 0.03 and a gradient of the best fit line of 0.05.

Figure 5 (b) (left) compares the time series of daily average NO_2 between the measured and modelled, with Figure 5 (b) (right) showing the relationship as a scatter plot and best-fit line. The model simulates NO_2 better than NO. The mean NO_2
225   concentration measured is 260 +/- 185 (1σ) pptv while the modelled NO_2 mean concentration is 182 +/- 323 (1σ) pptv. The median NO_2 concentration observed is 198 pptv with a 25th and 75th percentile range 127 and 335 pptv, whereas that modelled is and 121 pptv, with a 25th to 75th percentile range of 97 pptv and 152 pptv. The correlation coefficient between observation and measurement is though low (0.36) with RMSE of 322 pptv, and a gradient of the best-fit line of 0.65.

Generally, the model performance for the NO_x species is poorer than for other species. This likely reflects a number of
230   problems. It is difficult to make observations at these low concentrations (Reed et al., 2016) and for many days the NO observations are below the detection limit. It also likely reflects the short lifetime of NO_x making it susceptible to local chemical and emissions processes which the model can not resolve. Summer NO and NO_2 observed at Hateruma is higher than model estimates which is similar to the observations by Han et al. (2019) over Mongolia where thermal decomposed PAN during the warmer season contributes to higher NO_2 levels in remote regions. However, in general the modelled NO_x at
235   Hateruma (dominated by NO_2) underestimates the observed values.

4.6 NO_y

Figure 5 (c) (left) shows the daily average measured and modelled time-series for gas phase reactive nitrogen species (NO_y). We define here NO_y for the model as NO + NO_2 + 2×N_2O_5 + HNO_2 + HNO_3 + PAN (which will tend to marginally underestimate the true modelled NO_y as it misses halogen nitrates and some other minor consisituents such as NO3 and some
240   organic nitrates). We assumed that none of the particulate nitrate is measured. Figure 5 (c) (right) shows the scatter plot and line of best-fit between measurements and calculations. The mean observed NO_y concentration is 683 +/- 698 (1σ) pptv while the mean modelled NO_y concentration is 1173 +/- 1340 (1σ) pptv. The correlation coefficient between observation and model calculations is 0.79 with RMSE of 1035 pptv. The gradient of the best-fit line is 1.55. Thus the model appears to over estimate
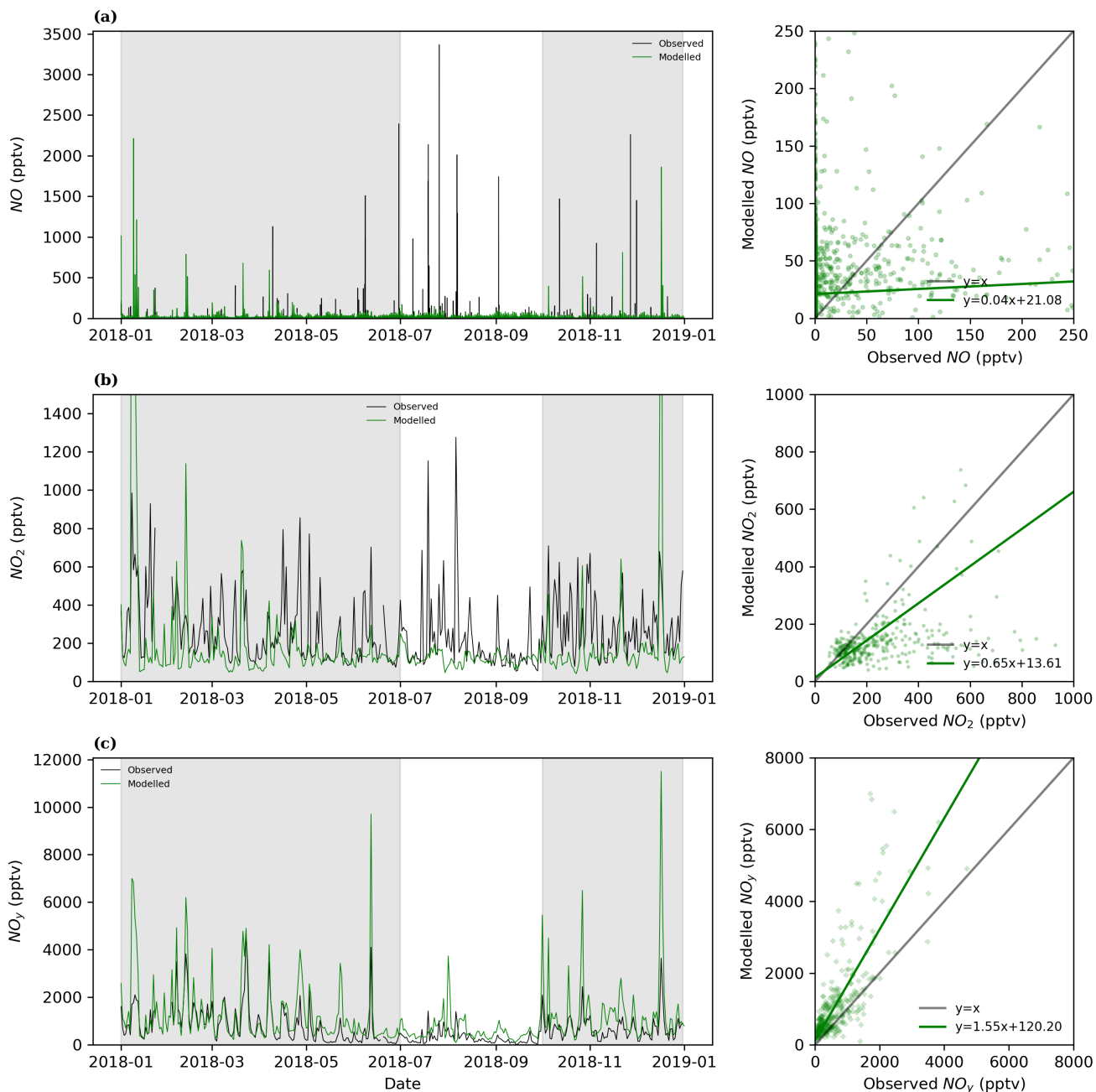
Figure 5. Modelled and measured mean daytime **(a)** NO, **(b)** NO_2, and **(c)** NO_y concentrations calculated for Hateruma

NO_y concentrations by roughly 50% despite not including some of the NO_y species and the potential for some particulate NO_y being sampled by the observations. The model performance is better than for NO_x. However, whereas the NO_x appears to be underestimated in the model, the NO_y is overestimated. HNO_3 and PAN are dominant in the model estimated NO_y levels at Hateruma. Thus, NO_y overestimated and NO_x underestimate could be related to NO_2 formation and loss pathways (Han et al., 2019).

A simulation without shipping emissions within the high resolution domain, reduces NO_y concentrations by 25%. Thus, even switching off these emissions entirely, does not compensate fully for the model overestimate and makes the model underestimation of NO_x worse. It is therefore unclear why the model overestimates the NO_y concentrations. This could indicate an excessive emissions in the region, too long a modelled NO_y lifetime, or, some difficulty in the model chemistry. Further work will be necessary to understand NO_x and NO_y in the region.

4.7 Ozone

Figure 6 (left) shows the measured and modelled timeseries for O_3. In general, the model performs well (RMSE of 8 ppbv; mean measured O_3 of 38 +/- 16 (1σ) ppbv; mean model O_3 of 36 +/- 14 (1σ) ppbv) and simulates much of the variability with a correlation coefficient (r) of 0.87. However, the line of best fit between observed and model results is low, with a gradient of 0.76 (Figure 6 - right). The median O_3 concentration calculated is 35 ppbv (25th and 75th percentiles of 22 ppbv and 46 ppbv) and whereas that measurement is 40 ppbv (25th and 7th percentiles of 24 ppbv and 49 ppbv).

During the summer, there are periods of low O_3, reaching 6.5 ppbv. The model fails to capture these low concentrations simulating a minimum of 10 ppbv. Conversely, in the winter months, the model reproduces much of the variability but has a tendency to underestimate the observed concentrations. Overall, this leads to overestimates at the low concentrations and underestimates at the high concentrations, and thus a reduced gradient in the scatter plot (Figure 6).
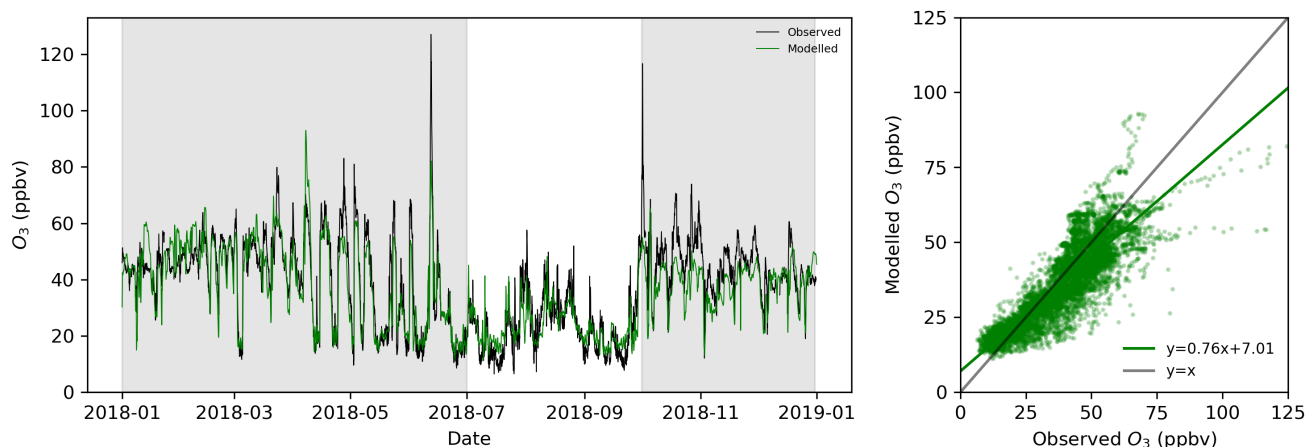
Figure 6. Modelled and measured hourly O_3 concentrations calculated for Hateruma site

4.8 Summary

265 In general the model has some skill at picking out the variations in meteorological history of the air masses arriving at the site, reflected in the generally high correlation coefficients between the model and the measurements. This mainly reflects the quality of the NASA generated meteorological field used to drive the model and their use of Data Assimilation to provide globally resolved meteorological parameterization. The success in simulating the absolute concentrations is more varied. Table 3 summarizes the mean model and measured concentrations for spring (February-April), summer (May-July), autumn (August-

270 October), winter (November-January).

Table 3. Mean and standard deviations in concentrations of measured and modelled species over different seasons between 1 January 2018 - 31 December 2018.

	Spring		Summer		Autumn		Winter	
	Measured	Modelled	Measured	Modelled	Measured	Modelled	Measured	Modelled
CO (ppbv)	177 +/- 56	178 +/- 73	99 +/- 49	100 +/- 54	110 +/- 45	120 +/- 51	160 +/- 61	187 +/- 85
C_2H_6 (pptv)	1854 +/- 611	1382 +/- 441	458 +/- 415	554 +/- 366	579 +/- 476	698 +/- 359	1638 +/- 691	1249 +/- 468
C_3H_8 (pptv)	588 +/- 416	314 +/- 195	105 +/- 187	79 +/- 86	154 +/- 242	118 +/- 117	646 +/- 473	332 +/- 212
NO (pptv)	3 +/- 9	20 +/- 28	16 +/- 49	19 +/- 6	6 +/- 18	17 +/- 16	6 +/- 17	30 +/- 74
NO_2 (pptv)	276 +/- 160	169 +/- 166	222 +/- 178	132 +/- 42	229 +/- 190	123 +/- 74	316 +/- 193	296 +/- 602
NO_y (pptv)	1183 +/- 882	1497 +/- 1296	383 +/- 518	839 +/- 1078	440 +/- 432	890 +/- 1080	751 +/- 572	1467 +/- 1678
O_3 (ppbv)	47 +/- 13	46 +/- 14	28 +/- 18	27 +/- 12	33 +/- 16	29 +/- 11	43 +/- 7	40 +/- 9

There is a clear seasonality in the observations at Hateruma and this is similar to other oceanic sites in East Asia. Such seasonality is reported at Cape Ochi-ishi (lat 43°10′N, long 145°30′E) and Oki island (lat 36°17′N, long of 133°11′E) where Asian outflows amplifies the winter levels of the measured hydrocarbons (Tohjima et al., 2002; Sharma et al., 2000). While NO_x has a very short atmospheric lifetime and its formation and loss is related to NO_y species such as HNO_3 and PAN, it is

275 reported to have different seasonality in polluted and remote regions (Han et al., 2019).

In Table 3, carbon monoxide concentrations are relatively well simulated in all seasons with a small bias high in the Autumn and Winter. Ethane is underestimated in the spring and winter, but overestimates in the summer and autumn. Propane concentrations are underestimated in all seasons. NO_y observations are in overestimate by the model, whereas NO and NO_2 concentrations are underestimated.

280 Although the largest model failures are for NO_x and NO_y species, there is uncertainty about the accuracy of these measurements at such low concentrations. Thus, we focus on the hydrocarbons in which we have more confidence in. The model underestimates C_2H_6 and C_3H_8 concentrations especially in the spring and winter. We therefore conduct a number of model experiments to investigate these observations. We first run simulations to assess the role of biomass burning in controlling the concentrations of these species at the site, we then assess how much the Asian anthropogenic source of these compounds would

285 have to increase by, to give agreement between the model and the measurements in the winter months.

5   Biomass burning sources

In order to understand the impact of biomass burning on the composition of the air arriving at the site we conduct additional simulation switching off the biomass burning emissions from North Asia and South Asia separately as shown in Figure B2. Two global simulations were run switching off the northern or southern biomass burning to generate new boundary conditions and then these were used for the two regional simulations which again switched off the biomass burning in either the North or South of Asia.

Figures 7 (a)-(f) show the time series of the percentage change in modelled concentrations of C_2H_6, C_3H_8, CO, O_3, NO_x and NO_y at the site when biomass burning emissions in either the northern or southern domain are switched off. Given the short lifetime of NO_y and NO_x, the site is less sensitive to biomass burning NO_x than CO, C_2H_6 and C_3H_8, and Figures 7 (d) and (f) show the correlation between O_3 and NO_x - when less ozone is formed due to switching off emissions, more NO_x is accumulated.

During the winter months the contributions from the North and South Asia are relatively small, however in the summer, both sources can contribute significantly (20-35 %) to the modelled C_2H_6, C_3H_8 and CO. The contributions are smaller for O_3 maximizing at 10 % during events in the summer time. There is sharp a shape spike around the 10th of August where CO, C_2H_6 and C_3H_8 levels dip by around 40-50 % due to switching off North Asian biomass burning emissions. After this sharp peak, the next decline from 16th August was more sustained and due to South Asian biomass burning emission being turned off. These sharp dips correspond to the passing of typhoons in the area which can rapidly draw air from different direction for short period of time. Figure B3 shows the back-trajectory analysis for this period when the wind direction rapidly changed. There is also a tendency for increased Asian fire activity in the summer which explains why the strong and rapidly changing wind would bring in more emissions.
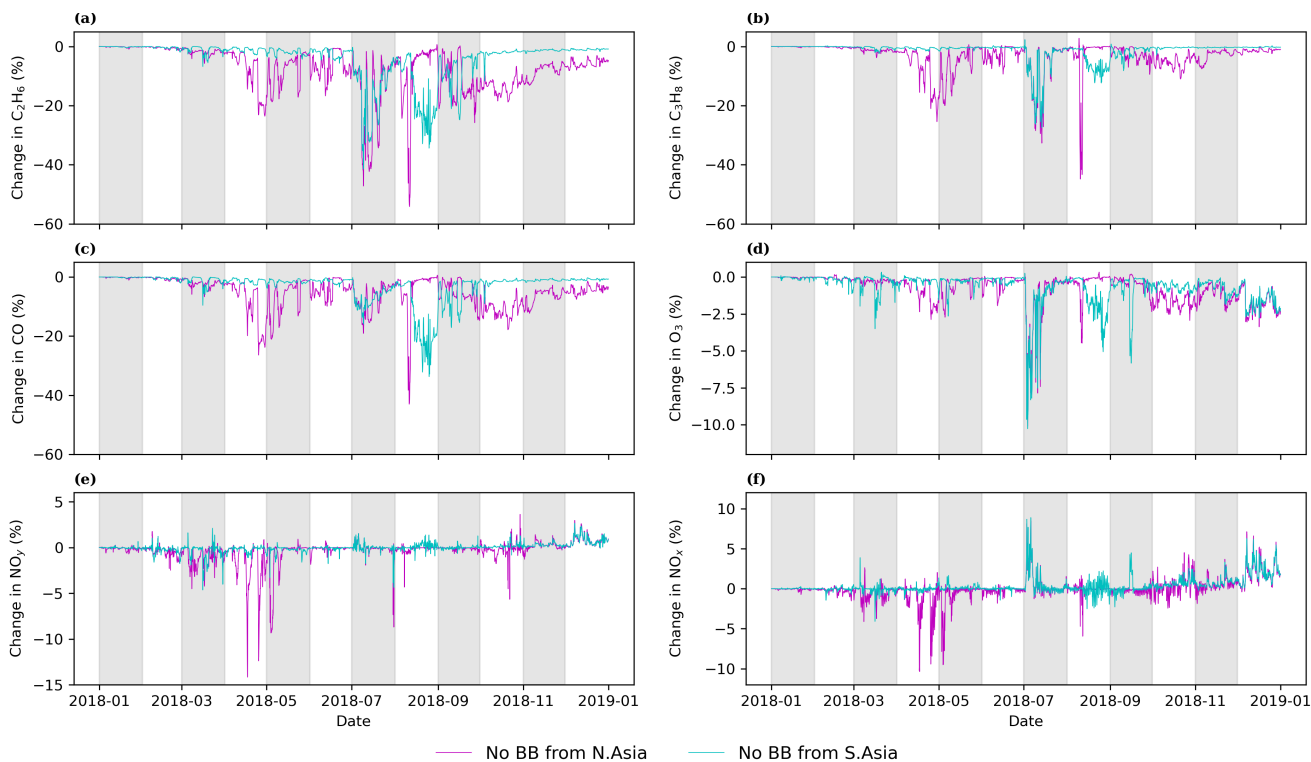
Figure 7. Percentage change in **(a)** C_2H_6, **(b)** C_3H_8, **(c)** CO, **(d)** O_3, **(e)** NO_y and **(f)** NO_x concentrations between standard and simulations without north Asian or south Asian biomass burning

The periods of model overestimate in summer ethane, notably in August (Figure 4 (b)), correspond to periods with a high fraction of modelled ethane biomass burning ($\sim$ 30%). However, the model overestimate (see Figure 4 (b)) is much large than this ($\sim$ 100 %). Thus, even if the biomass burning is switched off all together, the model overestimate during the summer would still remain. This suggests that the model overestimate of C_2H_6 in the summer is not primarily related to the modelled representation of biomass burning but likely results in uncertainties in the anthropogenic emissions of ethane south of Hateruma.

During the wintertime the model underestimates the observed ethane and propane. This is a period when the contribution from biomass burning is relatively low (<10%). Thus the potential for the underestimate in wintertime ethane and propane to be related to errors in the biomass burning appears to be small. Instead we now evaluate how much the Asian anthropogenic source of ethane and propane would have to be increased by, to fit the observed concentrations.

6   Anthropogenic emissions

The period between October to June (Figures 4 - 6) is characterised by elevated levels of pollution at the site due to flow from Russia, China, Korea and Japan (Figure 2). It is also a period when the model substantially underestimates the concentrations

of ethane and propane. In this section we explore what scaling factor would have to be applied to the Asian anthropogenic emissions of ethane and propane to better fit those observations.

Figures 8 (a) and (d) show modelled ethane (top) and propane (bottom) plotted against observations in our base model. We separate the plot into two periods here as northern-flow (blue July-September) and southern-flow (red October-June). It is obvious from the comparison that during the northern-flow period, the model-measured gradient is lower than expected (0.62 for ethane and 0.41 propane) whereas in the southern-flow period they are substantially overestimate for ethane (3.97) and to a smaller extend so for propane (1.51).

Given the site's largest exposure to relatively recent north Asian emissions in the northern-flow period, we focus on this. We multiply the Asian anthropogenic emission of ethane and propane within the high-resolution domain by a first rough estimate of a correction factor "A" (1.75 for ethane and 2.58 for propane which are roughly the reciprocal the model - measurement gradients show in Figure 8 i.e. 1./0.62 and 1./0.41) aimed at making the gradient of the best fit lines in the northern-flow period to become 1. This would represent a correction factor if all of the ethane and propane observed at the site were from within the simulation domain in Asia, and the oxidation lifetimes were correct.
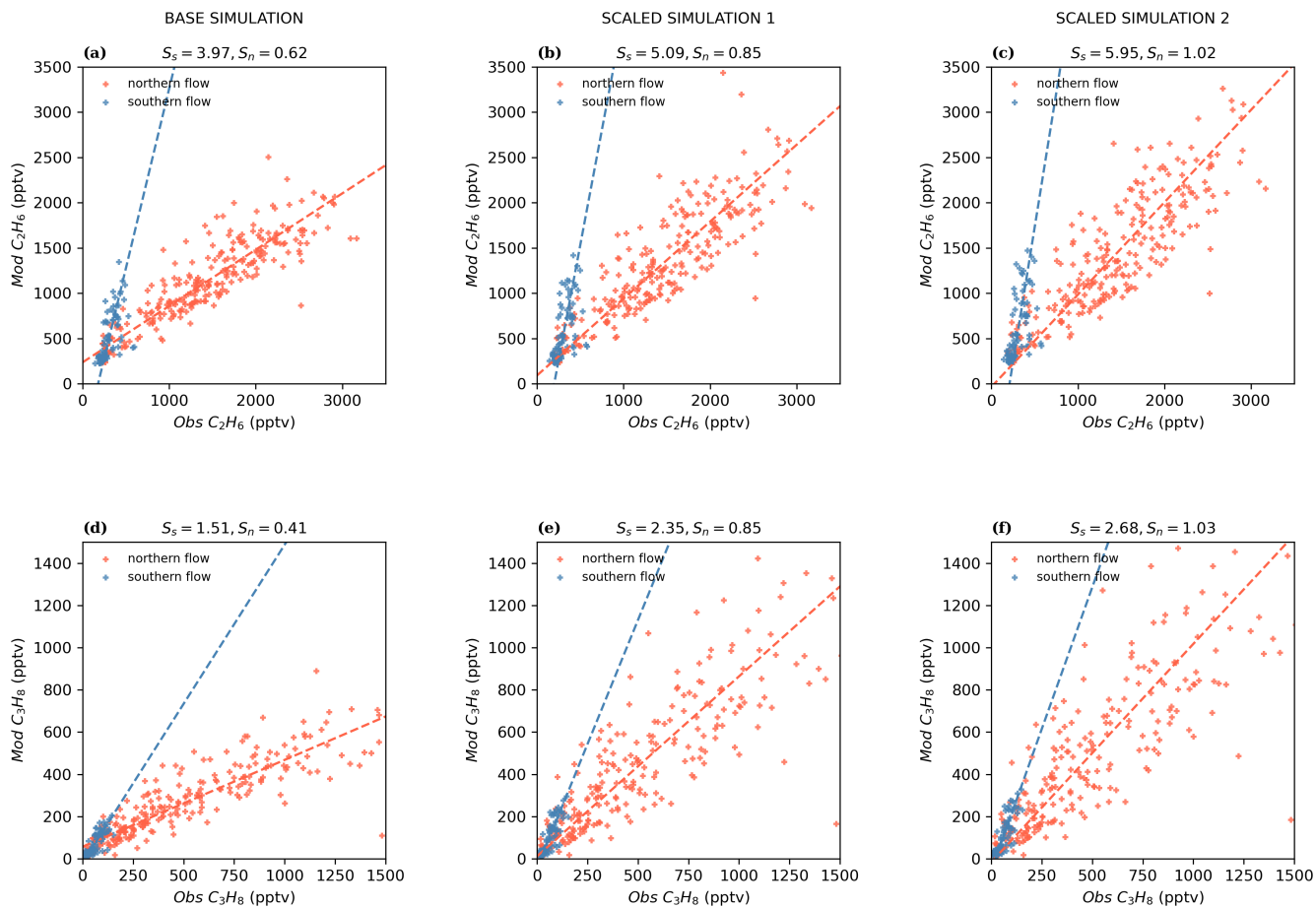
Figure 8. C_2H_6 and C_3H_8 modelled plotted against measurements in base simulation ((**a**) and (**d**)), scaled simulations with initial correction factor ((**b**) and (**e**)) and optimized correction factor ((**c**) and (**f**)). Data is split into a southern-flow period (July-September) in blue and northern-flow period (October-June) in red. S_s and S_n give the slope of the line of best fit for the timeseries in southern and northern flow periods respectively.

The annual simulation was re-run with the anthropogenic emissions of ethane and propane increased within the "Asian" high resolution domain (Figure 8 (b) and (e)). Assuming a linear response, these two simulations allows the concentration of ethane and propane in the base simulation to be represented as $[X]=[X]_{Asian}+[X]_{Rest\,of\,the\,world}$, and the concentra-

335 tions in the simulation with increase emissions to be represented as $[X]=1.75x[X]_{Asian}+[X]_{Rest\,of\,the\,world}$ for ethane and $[X]=2.58x[X]_{Asian}+[X]_{Rest\,of\,the\,world}$ for propane. The base concentrations at each time-point can then be decomposed to give an "Asian" component (although this is only the component within the higher resolution domain) and a Rest of the world component. A scaling can then be applied which optimizes the fit between the measurement and the model by finding an optimized multiplier for the Asian emission - $A_{optimized}$.

340   The value of $A_{optimized}$ is found to be 2.22 and 3.17, for ethane and propane, respectively if the best fit line is optimized (Figure 8). This is consistent with the value found by optimizing the ratio of mean concentrations in the period when the wind flow is northerly (calculated as 2.24 and 3.14, for ethane and propane, respectively). Re-running the model with these increases on the ethane and propane emissions within the domain gives an ethane and propane measurement versus model scatter plots in Figures 8 (c),(f), a time series shown in Figure 9 and a correlation between the whole dataset in 10. The correlation coefficient

345   (r) slightly reduced after the scaling (0.88 to 0.84, and 0.89 to 0.86 for C_2H_6 and C_3H_8, respectively).
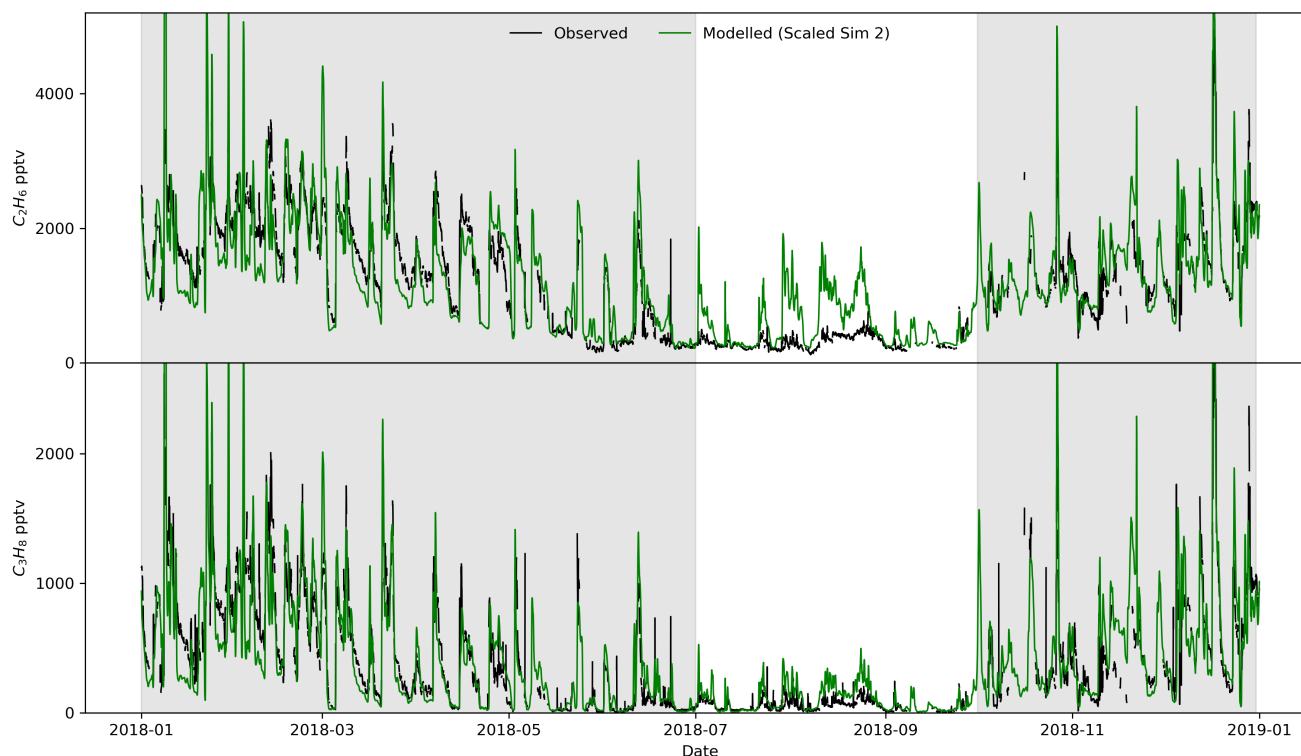
Figure 9. Timeseries of hourly C_2H_6 (top) and C_3H_8 (bottom) concentrations observed and with the scaled (2nd simulation) anthropogenic emissions.

For the northern-flow period (October to June), when the air is predominantly off Northern Asia, the modelled ethane now has a mean of 1443 +/- 737 (1σ) pptv, compared to an observed mean of 1449 +/- 725 (1σ) pptv. The propane simulation has a mean of 493 +/- 418 (1σ) pptv compared to an observed mean of 494 +/- 407 (1σ) pptv. During the southern-flow period (July - September), the model now significantly overestimates the ethane (mean simulated= 587 +/- 352 (1σ) pptv, mean observed=

350   308 +/- 99 (1σ) pptv) and propane (mean simulated = 105 +/- 88 (1σ) pptv, mean observed= 60 +/- 37 (1σ) pptv).
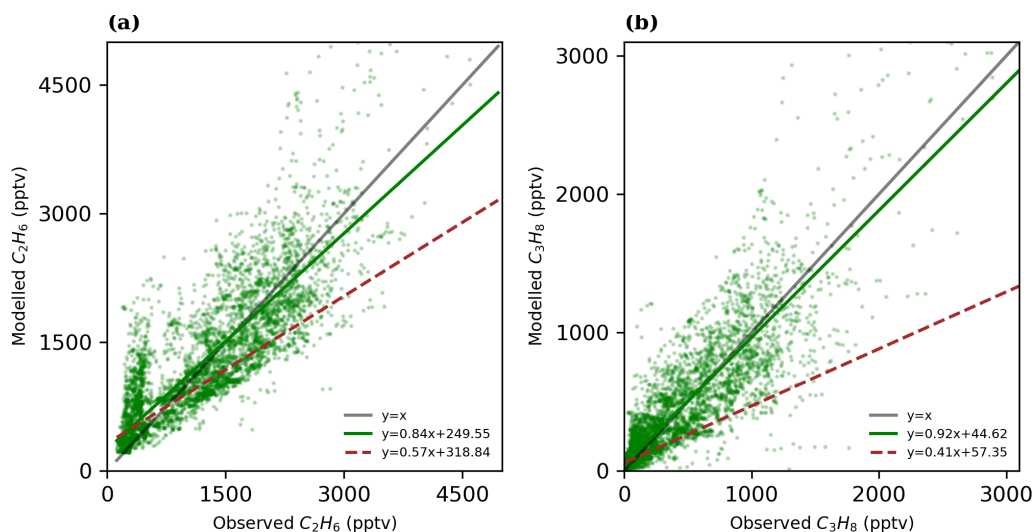
Figure 10. Scatter plot showing hourly **(a)** C_2H_6 and **(b)** C_3H_8 concentrations between observation concentrations and model calculations with scaling - 2nd simulation. Yellow dashed line shows best fit line before scaling.

The improvement in model performance during the northern-flow and the degradation in performance during the southern-flow towards a non-uniformity in the correction factor. This could indicate a difference in the seasonality of the emissions than those used here Hoesly et al. (2018) with a significant reduction in the summer time emissions compared to the winter. Or a different reason for the increase in north Asia than south Asia which is primarily sampled in the summer.

355   Nonetheless, the observations here suggest a large increases in anthropogenic emissions of ethane and propane from north Asia in our high resolution simulation domain. Large increases in global ethane and propane from those found in emissions inventories have been suggested before. Dalsøren et al. (2018) found it necessary to increase CEDS anthropogenic ethane emissions by a factor of roughly 2, and CEDS propane emission by a factor of nearly 3 to get agreement between their model and observations. These are very similar to the ratios found here. Tzompa-Sosa et al. (2017) found it necessary to substantially

360   increase anthropogenic ethane source to fit observations. Mo et al. (2020) also measured vertical profiles of VOCs (including C_2H_6 and C_3H_8) and reported that emission fluxes were up to 3 times larger than in the Multi-resolution Emission Inventory for China (MEIC) inventory estimates. It seems likely that current estimates of north Asian emissions of ethane and propane are currently underestimated.

    Figures 11 (a) and (b) show the impact of the large increase in Asian ethane and propane anthropogenic emissions on the

365   concentrations of average hydroxyl radical and ozone in the region in July, respectively. This is the period of highest ozone concentration over China. The effect on the OH concentration in July was small and patchy with up to a 2% reduction in some places. On the other hand, O_3 concentration increased slightly maximizing over Beijing where that amounts of a 2 % (around 1 ppbv) increase.
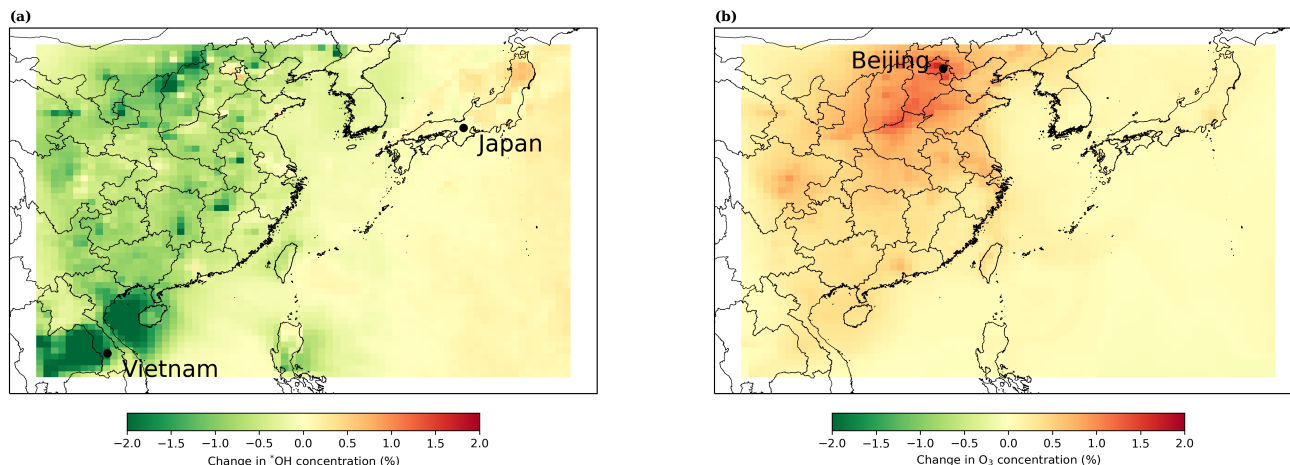
Figure 11. Percentage changes in average July concentrations of **(a)** OH and **(b)** O_3 after increasing Asian ethane and propane anthropogenic emissions.

7 Conclusions

370 Measurement were made of a number of compounds in the air over the island of Hateruma (CO, C_2H_6, C_3H_8, NO, NO_2, NO_y and O_3). We show that the site mainly subject to clean Pacific air throughout the year. If the site is subjected to polluted air masses, these are more likely to have originated from north Asia (Russia, China, Japan, Korea etc) in the winter months and from south Asia (Philippines, Borneo and other regions in Southeast Asia including Vietnam, Indonesia, Peninsula Malaysia, Thailand) in the summer months. This gives the site a significant seasonal cycle in the concentrations of pollutants considered

375 in this study except NO_x.

We have compared these observations to the output of the GEOS-Chem model, run in a regional configuration. We find that CO and O_3 are well simulated, the model overestimates NO_y observations, but tends to underestimates NO, NO_2, ethane and propane.

The underestimates in ethane and propane during the wintry-months are unlikely to be reconciled by increases in biomass

380 burning emissions, but could be reconciled by substantial (factors of 2 - 3) increases in the Asian anthropogenic source of these compounds (consistent with previous studies). These large increase in emissions have negligible influence on the hydroxyl radical concentrations and very little impact on the ozone in the region cities such as Beijing where there is up to 1 ppbv change in ozone concentration in July.

We do not believe the overestimates in the NO_y can be attributed to problems with the simulation of shipping NO_x, as

385 even switching off the shipping NO_x emissions doesn't remove the over estimate. Also, biomass burning emissions can only contribute up to 10 % of NO_x levels estimated at the site, therefore the overestimates likely lies in the emissions of NO_x from industrial regions close, issues with the chemistry or in the observations.

The site is located at an unusual location, which is subject to air masses from a large number of locations and so can be used to understand emissions of compounds from a number of location. Future plans to enhance the measurement capability at the site with a GC-MS system should allow an evaluation of a large number of different organic compounds, further enhancing the capacity to understand sources of pollution in the region.

Code and data availability. Hateruma data will be available from Global Environmental Database (https://db.cger.nies.go.jp/ged/en/index.html). GEOS-Chem version 12.7.1 was used in this project DOI: 10.5281/zenodo.3676008 and input data files for GEOS-Chem can be downloaded from http://geoschemdata.wustl.edu

Appendix A: Comparison of simulation resolutions
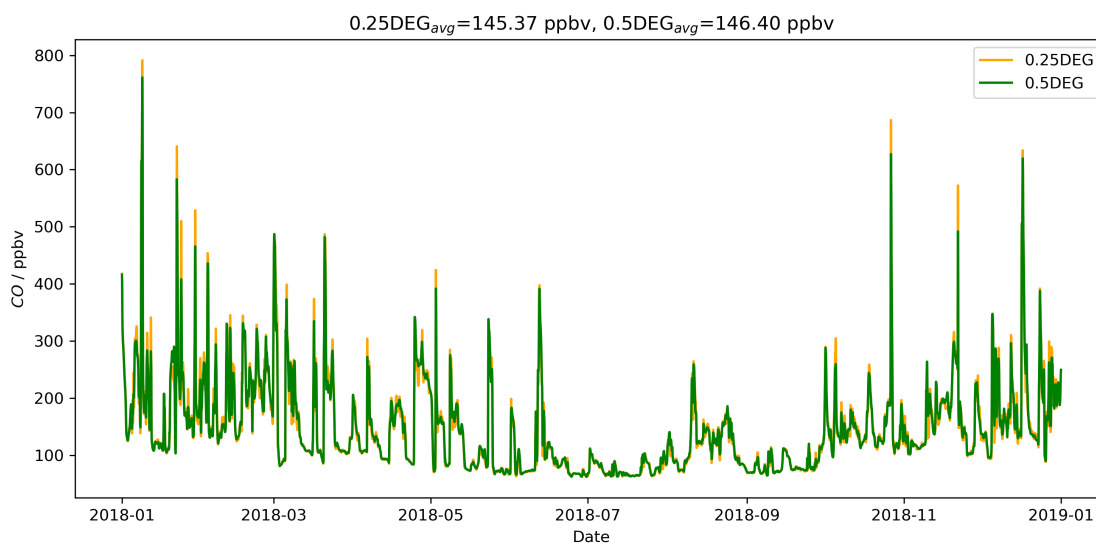
Figure A1. Modelled timeseries comparing CO simulation at $0.25° \times 0.3125°$ to $0.5° \times 0.625°$ resolution at the site
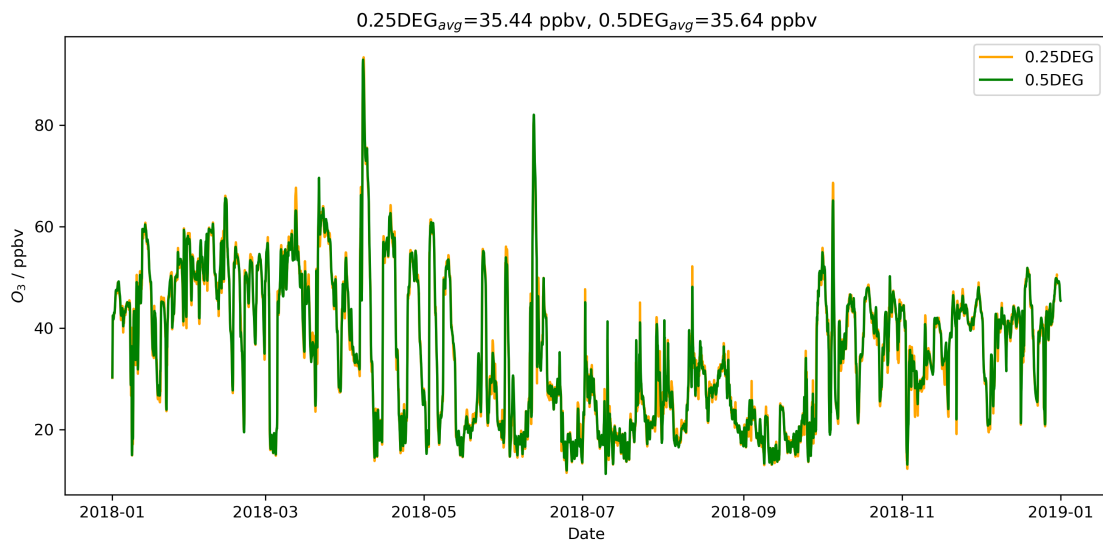
Figure A2. Modelled timeseries comparing O_3 simulation at $0.25° \times 0.3125°$ to $0.5° \times 0.625°$ resolution at the site
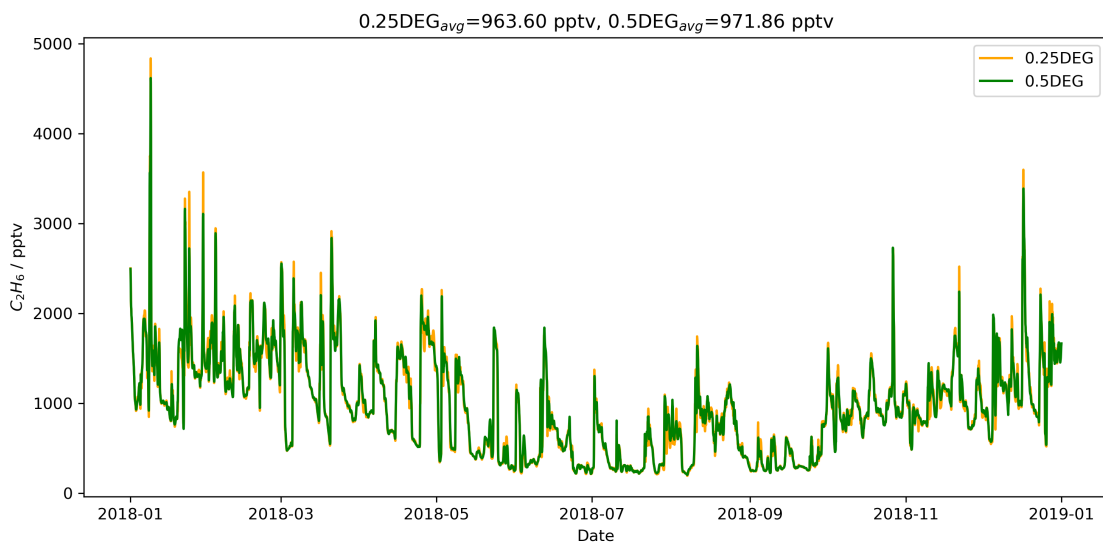
Figure A3. Modelled timeseries comparing C_2H_6 simulation at $0.25° \times 0.3125°$ to $0.5° \times 0.625°$ resolution at the site
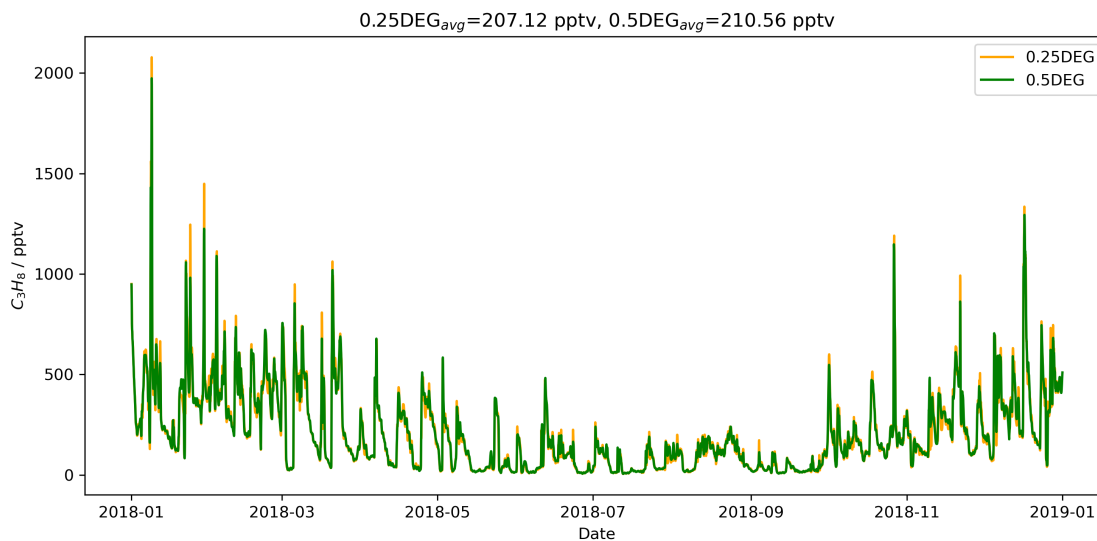
Figure A4. Modelled timeseries comparing C_3H_8 simulation at $0.25° \times 0.3125°$ to $0.5° \times 0.625°$ resolution at the site

Appendix B: Other figures
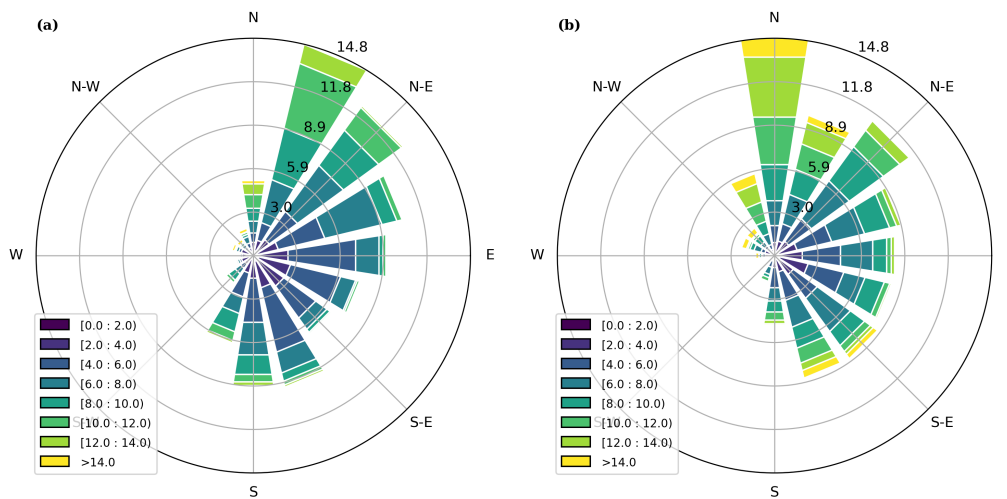
Figure B1. Windrose for **(a)** observed and **(b)** modelled wind speed (ms^{-1}) and direction at Hateruma

Global$_{mean}$=0.16x10^{-12} kg s^{-1}; N.Asia$_{mean}$=0.07x10^{-12} kg s^{-1}; S.Asia$_{mean}$=0.15x10^{-12} kg s^{-1}
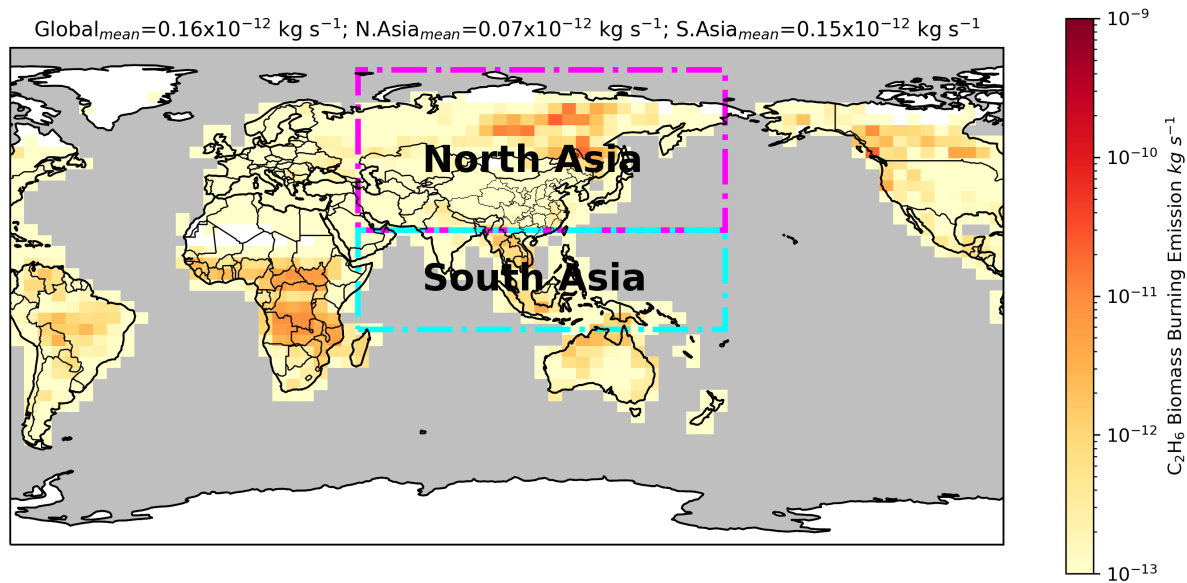
Figure B2. Annual mean biomass burning emission flux for C_2H_6 in 2018 (GFED). Regions indicated are the North and South Asia region which had biomass burning emissions switched off in the simulations. The Asia box extends from $46°$E to $180°$E; N.Asia box covers $82°$N to $24°$N while S.Asia spans $-12°$N to $24°$N
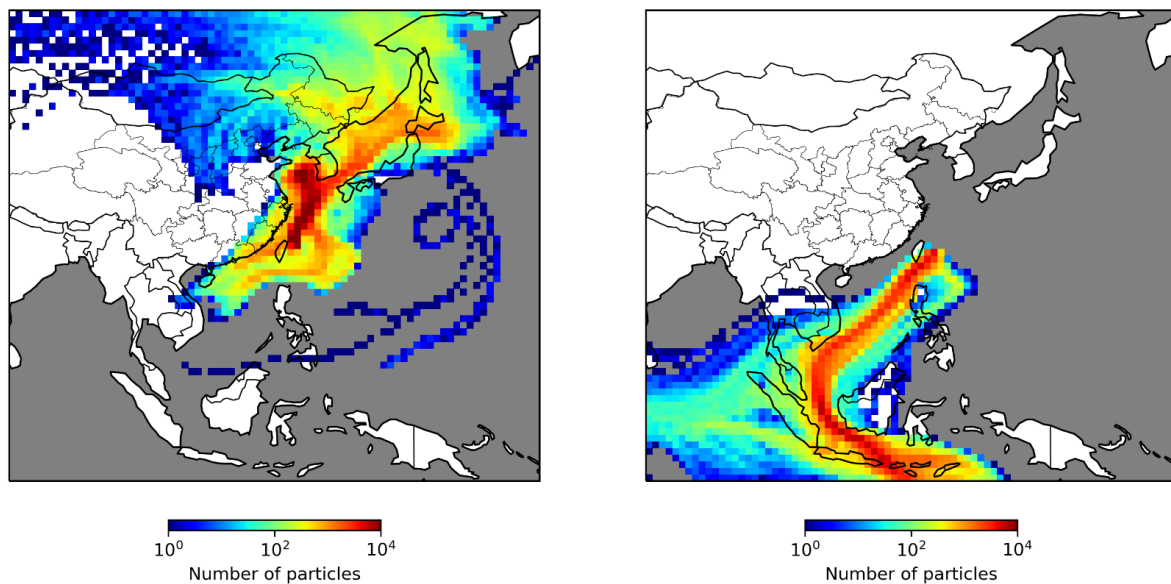
Figure B3. Back-trajectory analysis showing air masses arriving at hateruma between **(a)** August 10 -11 and (b) August 16 - 17

Author contributions. AA ran the GEOS-Chem simulations and made the analysis and visualization of the model outputs. MR contributed the FLEXPART model outputs and visualization. ME, SA, AL, and TS developed the project. ME assisted in the analysis and interpretation of all model outputs. TS provided the ethane and propane observations. YT contributed the CO data while SH, HM and HT provided the

400   $NO/NO_2/NO_y$ and O_3 observations used in model validation. Paper was written by AA and ME with contributions from all co-authors

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. This project was undertaken on the Viking Cluster, which is a high performance compute facility provided by the University of York. We are grateful for computational support from the University of York High Performance Computing service, Viking and the Research Computing team.

405   We would like to thank National Environment Research Council (NERC/UKRI) for providing support for this research through grant (Grant Reference: NE/S012273/1). We also thank National Centre for Atmospheric Science (NCAS) for their support.

This work was supported by JSPS under the National Centre for Atmospheric Science Joint Research Program implemented in association with UKRI (JRP-LEAD with UKRI). Grant number : [JSPSJRP20181708].

We thank staff members of the Global Environmental Forum Foundation (GEFF) for their great help in running the instruments at Hateruma

410   station.

References

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, Atmospheric Chemistry and Physics, 11, 4039–4072, https://doi.org/10.5194/acp-11-4039-2011, 2011.

Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochemical Cycles, 15, 955–966, https://doi.org/10.1029/2000GB001382, 2001.

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., Li, Q., Liu, H. Y., Mickley, L. J., and Schultz, M. G.: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106, 23 073–23 095, https://doi.org/10.1029/2001JD000807, 2001.

Bouwman, A. F., Lee, D. S., Asman, W. A. H., Dentener, F. J., Van Der Hoek, K. W., and Olivier, J. G. J.: A global high-resolution emission inventory for ammonia, Global Biogeochemical Cycles, 11, 561–587, https://doi.org/https://doi.org/10.1029/97GB02266, 1997.

Carvalho, D., Rocha, A., Gómez-Gesteira, M., and Santos, C.: A sensitivity study of the WRF model in wind simulation for an area of high wind energy, Environmental Modelling Software, 33, 23–34, https://doi.org/10.1016/j.envsoft.2012.01.019, 2012.

Dalsøren, S. B., Myhre, G., Hodnebrog, Ø., Myhre, C. L., Stohl, A., Pisso, I., Schwietzke, S., Höglund-Isaksson, L., Helmig, D., Reimann, S., Sauvage, S., Schmidbauer, N., Read, K. A., Carpenter, L. J., Lewis, A. C., Punjabi, S., and Wallasch, M.: Discrepancy between simulated and observed ethane and propane levels explained by underestimated fossil emissions, 11, 178–184, https://doi.org/10.1038/s41561-018-0073-0, 2018.

Etiope, G., Ciotoli, G., Schwietzke, S., and Schoell, M.: Global geological CH_4 emission grid files, https://doi.org/10.25925/4J3F-HE27, 2018.

Etiope, G., Ciotoli, G., Schwietzke, S., and Schoell, M.: Gridded maps of geological methane emissions and their isotopic signature, Earth System Science Data, 11, 1–22, https://doi.org/10.5194/essd-11-1-2019, 2019.

Ferreira, A. P., Castanheira, J. M., Rocha, A., and Ferreira, J.: Estudo de sensibilidade das previsões de superfície em Portugal pelo WRF face à variação das parametrizações físicas, 2008.

Galbally, I.: Nitrogen oxides (NO, NO2, NOy) measurements at Cape Grim A technical manual, 2020.

Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., Clerbaux, C., Coheur, P.-F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados-Muñoz, M. J., Hannigan, J. W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J. L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M. G., Schwab, J., Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D. W., Thouret, V., Thompson, A. M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H. M., Vigouroux, C., Xu, X., Zeng, G., and Ziemke, J.: Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation, Elementa: Science of the Anthropocene, 6, https://doi.org/10.1525/elementa.291, 39, 2018.

Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317–328, https://doi.org/10.1002/jgrg.20042, 2013.

Griffith, S. M., Huang, W.-S., Lin, C.-C., Chen, Y.-C., Chang, K.-E., Lin, T.-H., Wang, S.-H., and Lin, N.-H.: Long-range air pollution transport in East Asia during the first week of the COVID-19 lockdown in China, Science of The Total Environment, 741, 140 214, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.140214, 2020.

450 Griffiths, P. T., Murray, L. T., Zeng, G., Shin, Y. M., Abraham, N. L., Archibald, A. T., Deushi, M., Emmons, L. K., Galbally, I. E., Hassler, B., Horowitz, L. W., Keeble, J., Liu, J., Moeini, O., Naik, V., O'Connor, F. M., Oshima, N., Tarasick, D., Tilmes, S., Turnock, S. T., Wild, O., Young, P. J., and Zanis, P.: Tropospheric ozone in CMIP6 simulations, Atmospheric Chemistry and Physics, 21, 4187–4218, https://doi.org/10.5194/acp-21-4187-2021, 2021.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases
455 and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, Geoscientific Model Development, 5, 1471–1492, https://doi.org/10.5194/gmd-5-1471-2012, 2012.

Han, K., Lee, S., Yoon, Y., Lee, B., and Song, C.: A model investigation into the atmospheric NOy chemistry in remote continental Asia, Atmospheric Environment, 214, 116 817, https://doi.org/https://doi.org/10.1016/j.atmosenv.2019.116817, 2019.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M.,
460 Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750-2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geoscientific Model Development, 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Holmes, C. D., Prather, M. J., and Vinken, G. C. M.: The climate impact of ship NO_x emissions: an improved estimate accounting for plume chemistry, Atmospheric Chemistry and Physics, 14, 6801–6812, https://doi.org/10.5194/acp-14-6801-2014, 2014.

465 Itahashi, S., Uno, I., Osada, K., Kamiguchi, Y., Yamamoto, S., Tamura, K., Wang, Z., Kurosaki, Y., and Kanaya, Y.: Nitrate transboundary heavy pollution over East Asia in winter, Atmospheric Chemistry and Physics, 17, 3823–3843, https://doi.org/10.5194/acp-17-3823-2017, 2017.

Keller, C. A., Long, M. S., Yantosca, R. M., Da Silva, A. M., Pawson, S., and Jacob, D. J.: HEMCO v1.0: a versatile, ESMF-compliant component for calculating emissions in atmospheric models, Geoscientific Model Development, 7, 1409–1417, https://doi.org/10.5194/gmd-
470 7-1409-2014, 2014.

Kim, M. J.: The effects of transboundary air pollution from China on ambient air quality in South Korea, Heliyon, 5, e02 953, https://doi.org/https://doi.org/10.1016/j.heliyon.2019.e02953, 2019.

Koike, M., Jones, N. B., Palmer, P. I., Matsui, H., Zhao, Y., Kondo, Y., Matsumi, Y., and Tanimoto, H.: Seasonal variation of carbon monoxide in northern Japan: Fourier transform IR measurements and source-labeled model calculations, Journal of Geophysical Research:
475 Atmospheres, 111, https://doi.org/10.1029/2005JD006643, 2006.

Li, J., Yang, W., Wang, Z., Chen, H., Hu, B., Li, J., Sun, Y., Fu, P., and Zhang, Y.: Modeling study of surface ozone source-receptor relationships in East Asia, Atmospheric Research, 167, 77–88, https://doi.org/https://doi.org/10.1016/j.atmosres.2015.07.010, 2016.

Lin, H., Jacob, D. J., Lundgren, E. W., Sulprizio, M. P., Keller, C. A., Fritz, T. M., Eastham, S. D., Emmons, L. K., Campbell, P. C., Baker, B., Saylor, R. D., and Montuoro, R.: Harmonized Emissions Component (HEMCO) 3.0 as a versatile emissions component for
480 atmospheric models: application in the GEOS-Chem, NASA GEOS, WRF-GC, CESM2, NOAA GEFS-Aerosol, and NOAA UFS models, Geoscientific Model Development, 14, 5487–5506, https://doi.org/10.5194/gmd-14-5487-2021, 2021.

Millet, D. B., Guenther, A., Siegel, D. A., Nelson, N. B., Singh, H. B., de Gouw, J. A., Warneke, C., Williams, J., Eerdekens, G., Sinha, V., Karl, T., Flocke, F., Apel, E., Riemer, D. D., Palmer, P. I., and Barkley, M.: Global atmospheric budget of acetaldehyde: 3-D model analysis

and constraints from in-situ and satellite observations, Atmospheric Chemistry and Physics, 10, 3405–3425, https://doi.org/10.5194/acp-10-3405-2010, 2010.

Mo, Z., Huang, S., Yuan, B., Pei, C., Song, Q., Qi, J., Wang, M., Wang, B., Wang, C., Li, M., Zhang, Q., and Shao, M.: Deriving emission fluxes of volatile organic compounds from tower observation in the Pearl River Delta, China, Science of The Total Environment, 741, 139 763, https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.139763, 2020.

Murray, L. T., Jacob, D. J., Logan, J. A., Hudman, R. C., and Koshak, W. J.: Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD017934, 2012.

Naja, M. and Akimoto, H.: Contribution of regional pollution and long-range transport to the Asia-Pacific region: Analysis of long-term ozonesonde data over Japan, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/https://doi.org/10.1029/2004JD004687, 2004.

National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce: NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, https://doi.org/10.5065/D65D8PWK, 2015.

Qu, Z., Wu, D., Henze, D. K., Li, Y., Sonenberg, M., and Mao, F.: Transboundary transport of ozone pollution to a US border region: A case study of Yuma, Environmental Pollution, 273, 116 421, https://doi.org/https://doi.org/10.1016/j.envpol.2020.116421, 2021.

Randerson, J. T., Chen, Y., van der Werf, G. R., Rogers, B. M., and Morton, D. C.: Global burned area and biomass burning emissions from small fires, Journal of Geophysical Research: Biogeosciences, 117, https://doi.org/10.1029/2012JG002128, 2012.

Reed, C., Evans, M. J., Di Carlo, P., Lee, J. D., and Carpenter, L. J.: Interferences in photolytic NO_2 measurements: explanation for an apparent missing oxidant?, Atmospheric Chemistry and Physics, 16, 4707–4724, https://doi.org/10.5194/acp-16-4707-2016, 2016.

Riddick, S.N.;Dragosits, U. T. F. S. M.: Global ammonia emissions from seabirds, https://doi.org/10.5285/c9e802b3-43c8-4b36-a3a3-8861d9da8ea9, 2012.

Saito, T., Yokouchi, Y., Stohl, A., Taguchi, S., and Mukai, H.: Large Emissions of Perfluorocarbons in East Asia Deduced from Continuous Atmospheric Measurements, Environmental Science & Technology, 44, 4089–4095, https://doi.org/10.1021/es1001488, pMID: 20441144, 2010.

Schlesinger, W. H. and Bernhardt, E. S.: Chapter 11 - The Global Carbon and Oxygen Cycles, in: Biogeochemistry (Fourth Edition), edited by Schlesinger, W. H. and Bernhardt, E. S., pp. 453–481, Academic Press, fourth edition edn., https://doi.org/10.1016/B978-0-12-814608-8.00011-6, 2020.

Shairsingh, K. K., Jeong, C.-H., and Evans, G. J.: Transboundary and traffic influences on air pollution across two Caribbean islands, Science of The Total Environment, 653, 1105–1110, https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.11.034, 2019.

Sharma, U. K., Kajii, Y., and Akimoto, H.: Measurement of NMHCs at Oki Island, Japan,: An evidence of long range transport, Geophysical Research Letters, 27, 2505–2508, https://doi.org/10.1029/2000GL011500, 2000.

Sherwen, T., Schmidt, J. A., Evans, M. J., Carpenter, L. J., Großmann, K., Eastham, S. D., Jacob, D. J., Dix, B., Koenig, T. K., Sinreich, R., Ortega, I., Volkamer, R., Saiz-Lopez, A., Prados-Roman, C., Mahajan, A. S., and Ordóñez, C.: Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem, Atmospheric Chemistry and Physics, 16, 12 239–12 271, https://doi.org/10.5194/acp-16-12239-2016, 2016.

Shirai, T., Yokouchi, Y., Sugata, S., and Maksyutov, S.: HCFC-22 flux estimates over East Asia by inverse modeling from hourly observations at Hateruma monitoring station, Journal of Geophysical Research: Atmospheres, 115, 2010.

Stettler, M., Eastham, S., and Barrett, S.: Air quality and public health impacts of UK airports. Part I: Emissions, Atmospheric Environment, 45, 5415–5424, https://doi.org/10.1016/j.atmosenv.2011.07.012, 2011.

Stohl, A., Forster, C., Eckhardt, S., Spichtinger, N., Huntrieser, H., Heland, J., Schlager, H., Wilhelm, S., Arnold, F., and Cooper, O.: A backward modeling study of intercontinental pollution transport using aircraft measurements, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002862, 2003.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmospheric Chemistry and Physics, 5, 2461–2474, https://doi.org/10.5194/acp-5-2461-2005, 2005.

Tohjima, Y., Machida, T., Utiyama, M., Katsumoto, M., Fujinuma, Y., and Maksyutov, S.: Analysis and presentation of in situ atmospheric methane measurements from Cape Ochi-ishi and Hateruma Island, Journal of Geophysical Research: Atmospheres, 107, ACH 8–1–ACH 8–11, https://doi.org/https://doi.org/10.1029/2001JD001003, 2002.

Tohjima, Y., Mukai, H., Hashimoto, S., and Patra, P. K.: Increasing synoptic scale variability in atmospheric CO_2 at Hateruma Island associated with increasing East-Asian emissions, Atmospheric Chemistry and Physics, 10, 453–462, https://doi.org/10.5194/acp-10-453-2010, 2010.

Tohjima, Y., Kubo, M., Minejima, C., Mukai, H., Tanimoto, H., Ganshin, A., Maksyutov, S., Katsumata, K., Machida, T., and Kita, K.: Temporal changes in the emissions of CH_4 and CO from China estimated from CH_4 / CO_2 and CO / CO_2 correlations observed at Hateruma Island, Atmospheric Chemistry and Physics, 14, 1663–1677, https://doi.org/10.5194/acp-14-1663-2014, 2014.

Tohjima, Y., Patra, P. K., Niwa, Y., Mukai, H., Sasakawa, M., and Machida, T.: Detection of fossil-fuel CO2 plummet in China due to COVID-19 by observation at Hateruma, 10, https://doi.org/10.1038/s41598-020-75763-6, 2020.

Tzompa-Sosa, Z. A., Mahieu, E., Franco, B., Keller, C. A., Turner, A. J., Helmig, D., Fried, A., Richter, D., Weibring, P., Walega, J., Yacovitch, T. I., Herndon, S. C., Blake, D. R., Hase, F., Hannigan, J. W., Conway, S., Strong, K., Schneider, M., and Fischer, E. V.: Revisiting global fossil fuel and biofuel emissions of ethane, Journal of Geophysical Research: Atmospheres, 122, 2493–2512, https://doi.org/10.1002/2016JD025767, 2017.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth System Science Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Weng, H., Lin, J., Martin, R., Millet, D. B., Jaeglé, L., Ridley, D., Keller, C., Li, C., Du, M., and Meng, J.: Global high-resolution emissions of soil NO x, sea salt aerosols, and biogenic volatile organic compounds, Scientific Data, 7, 1–15, https://doi.org/10.1038/s41597-020-0488-5, 2020.

Xiao, Y., Logan, J. A., Jacob, D. J., Hudman, R. C., Yantosca, R., and Blake, D. R.: Global budget of ethane and regional constraints on U.S. sources, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009415, 2008.

Yokouchi, Y., Taguchi, S., Saito, T., Tohjima, Y., Tanimoto, H., and Mukai, H.: High frequency measurements of HFCs at a remote site in east Asia and their implications for Chinese emissions, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL026403, 2006.

Yokouchi, Y., Saito, T., Ooki, A., and Mukai, H.: Diurnal and seasonal variations of iodocarbons (CH_2ClI, CH_2I_2, CH_3I, and C_2H_5I) in the marine atmosphere, Journal of Geophysical Research: Atmospheres, 116, 2011.

Yokouchi, Y., Saito, T., Zeng, J., Mukai, H., and Montzka, S.: Seasonal variation of bromocarbons at Hateruma Island, Japan: implications for global sources, Journal of Atmospheric Chemistry, 74, 171–185, 2017.