An exploration of the discrepancy between observations and atmospheric ethane emission inventories: underestimations in key source emission categories

Jordan Y. Aljbour[1], Christopher Butenhoff[2], Andrew Rice[3], Alex Smith[4]

[1]Center for Climate and Aerosol Research, Portland State University, Portland, 97201, United States

Correspondence to: Jordan Y. Aljbour (jaljbour@pdx.edu)

Abstract: Ethane (**C_2H_6**) is one of the most abundant non-methane hydrocarbons within the atmosphere and is co-emitted with methane (**CH_4**) in key source categories. It has an ~2-month lifetime reacting with tropospheric hydroxyl radical (**OH**), and has a 100-year indirect global warming potential (**GWP**) of ~5.5. Ethane has anthropogenic, pyrogenic, biogenic, oceanic, and geologic emission sources with a budgetary composition relative to the order listed. Observations in atmospheric ethane concentrations show a negative trend of ethane mixing ratios from 1984 to 2010 over several regions globally, which has been associated with a net reduction in fugitive emissions from natural gas venting and flaring. Yet observations from 2010 and onward show a positive trend, particularly across the Northern Hemisphere, and additionally have displayed a decoupling in co-emission of ethane and methane. These recent trends in the atmospheric concentration, atmospheric distribution, emission sources, and decoupling from methane have yet to be adequately understood. With this work we've explored the discrepancy between current observational data and composite emission inventories for the global ethane budget. Our primary inquiry posits that the existing discrepancy can be attributed to underestimations in key source emission categories coupled with the exclusion of geologic and oceanic emission sources. Here we use emission scenarios using varied composite emission inventories within GEOS-Chem GCHP (**v13.0.2**) at a horizontal grid resolution of 4˚x5˚ and NOAA Global Monitoring Division (**GMD**) C_2H_6 observations. Emission scenarios include: scenario 1, Optimized Fossil Fuel (**OFF**), composed of the baseline Tzompa-Sosa 2010 C_2H_6 emission inventory with scaled anthropogenic fossil fuel emissions, the addition of geologic emissions, biogenic emissions (**MEGAN-MACC**), and oceanic emissions (**RETRO**); emission scenario 2, Optimized Biomass Burning (**OBB**), composed of the Tzompa-Sosa 2010 C_2H_6 emission inventory, GFED4 inventory with scaled biomass burning emissions, the addition of geologic emissions, biogenic emissions (**MEGAN-MACC**), and oceanic emissions (**RETRO**). By exploring more detailed emissions inventories we can better understand the recent trends in the global ethane budget, which may necessitate further re-evaluation of the contribution from key emission sources. A more thorough understanding remains a critical precursor in improving the efficacy of policies to reverse ongoing trends in atmospheric ethane and related NMVOCs.

1 – Introduction:

At present time, C_2H_6 atmospheric model simulations utilizing standard community emission inventories do not suffice in reproducing available observations—particularly in the Northern Hemisphere. Research has suggested that current observations can be reproduced in simulations utilizing a detailed atmospheric chemistry transport model, provided that emissions from biogenic, geologic, oceanic are accounted for; and anthropogenic fossil fuel emissions are scaled by a factor two to three times higher than is currently indicated standard inventories [*Dalsøren et al.*, 2018]. Furthering implications that due to the improved correspondence between observations and model simulations when emissions inventories are optimized—present emission inventories may require additional re-evaluation. Thereby we will begin by examining the standard methodology of aggregating C_2H_6 emission inventories from key emission sources, then discern the atmospheric mechanisms which contribute to the hemispheric C_2H_6 emission gradient, and conclude with dissecting the composition of the approximate global C_2H_6 budget.

C_2H_6 is primarily emitted during the production, processing, and transportation of natural gas sources—so much so that it has been passively considered as a tracer-indicator for fugitive emissions from the production of natural gas [*Schwietzke et al.*, 2014]. This indicative property also applies to geographic areas home to several CH_4 sources, as C_2H_6 is co-emitted from sources of fossil-fuel CH_4 emissions [*McKain et al.*, 2015]. This relationship in co-emission is the basis of inferring the relative C_2H_6 global budget based upon CH_4 emission estimates, as there exists an approximate $\frac{CH_4}{C_2H_6}$ ratio of $\frac{2000\ ppbv}{2\ ppbv}$ [*Tzompa-sosa et al.*, 2017]. This implicitly suggests that any existing discrepancy between C_2H_6 observations and model simulations extends to a potential discrepancy in the original CH_4 emission estimates [*Dalsøren et al.*, 2018]. Nevertheless, the method of yielding an approximate C_2H_6 emission budget from scaled CH_4 source emission estimates is the commonplace methodology for aggregating C_2H_6 emission inventories. So much so that the current baseline C_2H_6 emission inventory for fossil fuel and biofuel sources was composed by applying $\frac{CH_4}{C_2H_6}$ emission ratios to satellite-derived CH_4 emissions [*Turner et al.,* 2015; *Tzompa-Sosa et al.*, 2017].

C_2H_6, second only to CH_4, is one of the most abundant volatile organic compounds (**VOCs**) within the atmosphere. Existing observed mixing ratios in the lower troposphere range from ~0.2 ppbv above desolate geographic areas in the Southern Hemisphere to an observed maximum of ~1500 ppbv above natural gas and oil basins [*Helmig et al.*, 2014]. Given that the primary sink of C_2H_6 is tropospheric OH, C_2H_6 possesses significant seasonality and a seasonally dependent lifetime of ~2-months. This

coupled with the approximate $\frac{CH_4}{C_2H_6}$ ratio culminates in the potential instantaneous contribution of ~4-7% of the total OH loss with both species combined [*Tzompa-sosa et al.*, 2017]. This impact on OH extends to atmospheric distribution of several species, as the reaction and degradation of C_2H_6 leads to the production of tropospheric O3 and PAN. However, given that the majority of anthropogenic C_2H_6 sources reside within the Northern Hemisphere and that its ~2-month lifetime does not exceed the interhemispheric exchange rate [*Aydin et al.*, 2011]—the skewed hemispheric gradient intrinsically reduces the global ecological tropospheric impact.

With about ~60% of global C_2H_6 emissions sourced from natural gas leakage events [*Xiao et al.*, 2008]—the remaining ~40% of global C_2H_6 emissions are commonly attributed to, in approximate equal partition, sources of biomass burning and biofuel consumption [*Rudolph et al.*, 1995]. In most standard community emission inventories biogenic, geologic, and oceanic emissions are considered negligible on a global scale [*Rudolph et al.*, 1995], but research has contested this determination by positing that key to the discrepancy between C_2H_6 observations and model simulations is the exclusion of geologic emissions [*Etiope & Ciccioli*, 2009]. Their results showcased that an estimated ~2-4 Tg $year^{-1}$ emission output of C_2H_6 can be attributed to geologic sources, such as volcanic seeps, onshore seeps, micro-seeps, marine seeps, and geothermal seeps [*Etiope & Ciccioli*, 2009]. This challenge to convention and the corresponding results serve as the basis of the deliberate inclusion of biogenic, geologic, and oceanic C_2H_6 emissions sources in this study. To consider the fact that the conventional global C_2H_6 budget, ~0.6FF%: ~0.2BF%: ~0.2BB%, has repeatedly yielded discrepancies between C_2H_6 observations and model simulations. The inclusion of additional source emission inventories appears warranted, which will both serve evaluate the efficacy of conventional and non-conventional global C_2H_6 budget compositions.

2 – Methods:

2.1 – Compilation of GEOS-Chem GCHP (*v13.0.2*):

GEOS-Chem is a global 3-D atmospheric chemistry model driven by meteorology data (**MERRA2**) sourced from the Goddard Earth Observing System (**GEOS**) of NASA's Global Modeling and Assimilation Office. This research utilizes GEOS-Chem GCHP (**v13.0.2**), which is the high-performance option within the most recent generation of GEOS-Chem atmospheric chemical composition modeling systems. It specializes in utilizing distributed-memory capability designed to enable efficient scaling across several cores and in turn composite finer resolution simulations. GEOS-Chem GCHP (**v13.0.2**), from now on GCHP, was compiled on the COEUS High-Performance Computing Cluster hosted through the Portland Institute for Computational Science and Portland State University. The

compilation process was long-delayed due to missing module dependencies required by GCHP and inconsistent ubiquity in choice of GNU-compiler (**GCCv9.2.0**) for the existing dependencies. This entailed delay, coupled with the unfortunately brief scope of the Center for Climate and Aerosol Research's REU summer research program, has rendered this research a work in progress. From this point on in this research manuscript keep this in mind, as well as that this work will continue and has been submitted to present at Advancing Earth and Space Science's Fall Meeting: 2021.

Implemented within GEOS-Chem is the Harvard-NASA Emission Component (**HEMCO**), which is a stand-alone software component for computing emissions within global atmospheric models. HEMCO determines emissions from a variety of different sources, regions, and species on a user-defined grid and can combine, overlay, and even update a set of data inventories with scale factors—as all specified by the user through the HEMCO configuration file (**HEMCO_Config.rc**). A user is able to incorporate new emission inventories with any given spatial and temporal resolution to HEMCO, and additionally can be accessed by the user without any necessary preprocessing of the data files or source code. This research utilizes HEMCO in the configuration of the varied emission scenarios compositions detailed in section 2.3, as all emission scaling and incorporation of additional source emission inventories were handled through HEMCO.

2.2 – Aggregation of Additional Source Emission Inventories:

All additional source emission inventories, such as: MEGAN-MACC, MEGANv2, POET, RETRO were accessed through Emissions of Atmospheric Compounds and Compilation of Ancillary Data (**ECCAD**). Given the user-functionality of HEMCO detailed in section 2.1, there was no necessary preprocessing of any of the additional source emission inventories accessed through ECCAD. Whereas significant preprocessing was necessary for the global geological CH_4 gridded emission inventory (**Etiope: GEOCH_4 (2018)**) produced by G. Etiope[6] and S. Schwietzke[6] [*Etiope et al.*, 2019].

Given that the global geologic data was explicitly for CH_4 emissions and was composed of several comma-separated value (**.csv**) datasets each corresponding to differing geologic emission sources—preprocessing in the form of scaling was necessary. Of the differing geologic emission sources each possess their own $\frac{CH_4}{C_2H_6}$ emission ratio to discern their respective C_2H_6 emission output [*Etiope & Ciccioli*, 2009]. After each sub-source: onshore seeps, micro-seeps, marine seeps, and geothermal seeps were scaled for C_2H_6 the new emission output values were reaggregated into a total geologic C_2H_6 emission inventory. Using the python modules pandas and xarray, the gridded total geologic C_2H_6 emission inventory was converted from .csv to the HEMCO compatible NetCDF (**.nc**) format. The

125 processed total geologic C_2H_6 emission inventory .nc file, along with the additional source emission inventory .nc files, were incorporated into GEOS-Chem via HEMCO.

2.3 – Baseline Emission Scenario, Optimized Fossil Fuel Emission Scenario (*OFF*), and Optimized Biomass Burning Emission Scenario (*OBB*):

The baseline emission scenario composition is conventional and includes the following emission
130 inventories: anthropogenic fossil fuel and biofuel source emissions reference the Tzompa-Sosa C_2H_6 (2010) emission inventory, and biomass burning source emissions reference the GFEDv4 (2018) emission inventory. The GCHP baseline emission scenario ran on a core count of ~54 at a resolution of ~4° x 5° (**C24**) and provided an output of six-global-monthly average species concentrations of C_2H_6 from 2018-01-01 to 2018-06-01. These baseline output files serve as the GCHP emission simulation data utilized in
135 the multi-variate timeseries analysis in section 2.4. As the baseline emission scenario, the outputs from this simulation will directly evaluate the conventional global C_2H_6 budget when compared with the NOAA C_2H_6 concentration observations. The primary indication of verifying the existence of any discrepancy is the relative correspondence between observations and model simulations.

The optimized fossil fuel emission scenario (**OFF**) is unconventional and includes the following
140 emission inventories: anthropogenic fossil fuel and biofuel source emissions reference the optimized Tzompa-Sosa C_2H_6 (2010) emission inventory, biomass burning source emissions reference the GFEDv4 (2018) emission inventory, biogenic source emissions reference the MEGAN-MACC (2010) emission inventory, geologic source emissions reference the Etiope $GEOCH_4$ (2018) scaled by the $\frac{CH_4}{C_2H_6}$ emission ratio for C_2H_6, and oceanic source emissions reference the RETRO (2000) emission inventory. Due to the
145 delay referenced to in section 2.1, the OFF simulation regrettably has yet to conclude and will be incorporated upon conclusion.

The optimized biomass burning emission scenario (**OBB**) is unconventional and includes the following emission inventories: anthropogenic fossil fuel and biofuel source emissions reference the Tzompa-Sosa C_2H_6 (2010) emission inventory, biomass burning source emissions reference the optimized
150 GFEDv4 (2018) emission inventory, biogenic source emissions reference the MEGAN-MACC (2010) emission inventory, geologic source emissions reference the Etiope $GEOCH_4$ (2018) scaled by the $\frac{CH_4}{C_2H_6}$ emission ratio for C_2H_6, and oceanic source emissions reference the RETRO (2000) emission inventory. Due to the delay referenced to in section 2.1, the OBB simulation regrettably has yet to conclude and will be incorporated upon conclusion. The remainder of this manuscript will focus nearly entirely on the
155 comparative analysis of NOAA's C_2H_6 concentration observations and GCHP's baseline C_2H_6 simulated

concentration outputs such to discern their correspondence. **See Table – 2 on page 14 of 18 for reference**.

2.4 – Analysis of Emission Scenarios vs. Observations at Key Observation Sites:

The analysis of both NOAA's C_2H_6 concentration observations and GCHP's baseline C_2H_6 simulated concentration outputs took place within the python environment. The following modules were utilized in the analysis: os, numpy, netCDF4, pandas, xarray, and matplotlib. Each of the site dependent NOAA C_2H_6 observation files were read into python and were subsequently plotted in a set of subplots corresponding to each site from 20180101 to 20180601. **See output as Figure – 1 on page 16 of 18.**

A similar order of operations occurred for the GCHP six-global-monthly average species concentrations of C_2H_6 from 2018-01-01 to 2018-06-01, but additional input information was required for the set of subplots to correspond to the observation sites. This was done by specifying that each site dependent GCHP *pandas.dataframe()* corresponds the NOAA observation site coordinates to the nearest grid cell within the GCHP outputs. **See output as Figure – 2 on page 17 of 18.**

In order to plot both sets of site dependent NOAA and GCHP data in a method akin to their respective individual subplots, the format of the time variables for both sets of data needed to match. Given the sets of NOAA data were formatted in '*yyyymmdd*', by using pandas.to_datetime() the formatting was changed to '*yyyy-mm-dd*' to match with the sets of GCHP data. Once conducted, the same order of operations compiled properly—plotting both NOAA's C_2H_6 concentration observations and GCHP's baseline C_2H_6 simulated concentration outputs on the same set of subplots from 2018-01-01 to 2018-06-01. **See output as Figure – 3 on page 18 of 18.**

3 – Results & Discussion:

3.1 – Trends in Observational Data (NOAA):

The three trends in the observational data with most significance are the quantitative visualization of the hemispheric gradient, strong seasonality, and relative zonal concentration gradient of atmospheric C_2H_6. The hemispheric gradient is best visualized by the scale of each subplot's y-axis, as the scale directly corresponds to the magnitude of observed atmospheric C_2H_6. Of the observation sites seen in Figure – 1, the following are located within the Northern Hemisphere: ALT, AMY, BRW, HFM, KUM, LEF, MHD, MLO, NWR, SUM, and THD; whereas the following sites are located within the Southern Hemisphere: CGO, PSA, SMO, and SPO. Collectively the Northern Hemispheric sites utilize an average

scale range of ~750pptv to ~2100pptv, while the Southern Hemispheric sites utilize an average range of ~110pptv to ~230pptv. The hemispheric differential in scale magnitude showcases the hemispheric gradient of atmospheric C_2H_6, as the higher-concentrated Northern Hemisphere emits more C_2H_6 compared to the lower-concentrated Southern Hemisphere.

The trend in seasonality of atmospheric C_2H_6 is showcased by contextualizing the relative atmospheric C_2H_6 concentration in the timeseries with the hemisphere the site is located. Of the Northern Hemispheric sites, the start of the timeseries coincides with the first day of Winter. All showcase relatively-linear decreases in atmospheric C_2H_6 concentration as Winter transitions to Spring (**2018-03-01**) and continues to decrease until gradually increasing as Spring transitions to Summer (**2018-06-01**). Of the Southern Hemispheric sites, the start of the timeseries coincides with the first day of Summer, nearly all showcase relatively-linear increases in atmospheric C_2H_6 concentration as Summer transitions to Autumn (**2018-03-01**) and continues to increase until gradually slowing said increase as Autumn transitions to Winter (**2018-06-01**). Thereby the annual seasonality of atmospheric C_2H_6 is hemispherically-relative—resembling the appearance of an imperfect concave-up parabola in the Northern Hemisphere and an imperfect concave-down parabola in the Southern Hemisphere.

The trend in the average relative zonal concentration gradient of atmospheric C_2H_6 is discernable by comparing the scale magnitude of the observations with the coordinates of the observation site. By partitioning the latitudinal coordinates of each observation site into zonal ranges of parallels; the average scale magnitude of atmospheric C_2H_6 concentration at each site within each zone corresponds to the relative zonal concentration gradient. For the Northern Hemisphere the first zonal range of 0˚N—20˚N is home to KUM and MLO, making the average relative zonal concentration ~450pptv to ~1250pptv. The second zonal range of 20˚N—50˚N is home to AMY, HFM, LEF, NWR, and THD, making the average relative zonal concentration ~1000pptv to ~2300pptv. The third zonal range of 50˚N—90˚N is home to ALT, BRW, MHD, and SUM, making the average relative zonal concentration ~750pptv to ~2250pptv. For the Southern Hemisphere the first zonal range of 0˚S—20˚S is home to SMO, making the average relative zonal concentration ~200pptv to ~350pptv. The second zonal range of 20˚S—50˚S is home to CGO, making the average relative zonal concentration ~100pptv to ~250pptv. The third zonal range of 50˚S—90˚S is home to PSA and SPO, making the average relative zonal concentration ~125pptv to ~275pptv. **See Table – 1 on page 14 of 18, Table – 3 on page 15 of 18, and Figure – 1 on page 16 of 18 for reference.**

3.2 – Trends in Emission Scenarios (GCHP):

The three trends in the simulation data with most significance are the verification of the quantitative visualization of the hemispheric gradient, strong seasonality, and relative zonal concentration gradient of atmospheric C_2H_6 showcased in the observational data. As demonstrated in section 3.1, the hemispheric gradient is best visualized by the scale of each subplot's y-axis, as the scale directly corresponds to the magnitude of simulated atmospheric C_2H_6. Of the observation sites seen in Figure – 2, the following are located within the Northern Hemisphere: ALT, AMY, BRW, HFM, KUM, LEF, MHD, MLO, NWR, SUM, and THD; whereas the following sites are located within the Southern Hemisphere: CGO, PSA, SMO, and SPO. Collectively the Northern Hemispheric sites utilize an average scale range of ~1255pptv to ~1928pptv, while the Southern Hemispheric sites utilize an average range of ~93pptv to ~145pptv. The simulated hemispheric differential in scale magnitude verifies the hemispheric gradient of atmospheric C_2H_6 posited in section 3.1, as the higher-concentrated Northern Hemisphere emits more C_2H_6 compared to the lower-concentrated Southern Hemisphere.

As demonstrated in section 3.1, the trend in seasonality of atmospheric C_2H_6 is showcased by contextualizing the relative atmospheric C_2H_6 concentration in the timeseries within the hemisphere the site is located. Of the Northern Hemispheric sites, the start of the timeseries coincides with the first day of Winter, all showcase relatively-linear decreases in atmospheric C_2H_6 concentration as Winter transitions to Spring (**2018-03-01**) and continues to decrease until gradually increasing as Spring transitions to Summer (**2018-06-01**). Of the Southern Hemispheric sites, the start of the timeseries coincides with the first day of Summer, nearly all showcase relatively-linear increases in atmospheric C_2H_6 concentration as Summer transitions to Autumn (**2018-03-01**) and continues to increase until gradually slowing said increase as Autumn transitions to Winter (**2018-06-01**). Thereby the annual seasonality of atmospheric C_2H_6 is hemispherically-relative—resembling the appearance of an imperfect concave-up parabola in the Northern Hemisphere and an imperfect concave-down parabola in the Southern Hemisphere. The results utilizing the GCHP six-global-monthly average species concentrations of C_2H_6 verify the posited findings in section 3.1 with marginal discrepancies.

As demonstrated in section 3.1, the trend in the average relative zonal concentration gradient of atmospheric C_2H_6 is discernable by comparing the scale magnitude of the observations with the coordinates of the observation site. By partitioning the latitudinal coordinates of each observation site into zonal ranges of parallels; the average scale magnitude of atmospheric C_2H_6 concentration at each site within each zone corresponds to the relative zonal concentration gradient. For the Northern Hemisphere the first zonal range of 0˚N—20˚N is home to KUM and MLO, making the average relative zonal concentration ~600pptv to ~1100pptv. The second zonal range of 20˚N—50˚N is home to AMY, HFM, LEF, NWR, and THD, making the average relative zonal concentration ~1680pptv to ~2360pptv. The

third zonal range of 50˚N—90˚N is home to ALT, BRW, MHD, and SUM, making the average relative zonal concentration ~1050pptv to ~1800pptv. For the Southern Hemisphere the first zonal range of 0˚S—20˚S is home to SMO, making the average relative zonal concentration ~120pptv to ~220pptv. The second zonal range of 20˚S—50˚S is home to CGO, making the average relative zonal concentration ~90pptv to ~120pptv. The third zonal range of 50˚S—90˚S is home to PSA and SPO, making the average relative zonal concentration ~80pptv to ~120pptv. The results utilizing the GCHP six-global-monthly average species concentrations of C_2H_6 verify the posited findings in section 3.1 with moderate discrepancies. **See Table – 1 on page 14 of 18, Table – 3 on page 15 of 18, and Figure – 2 on page 17 of 18 for reference.**

3.3 – Comparison of Observational Data with Baseline Emission Scenario:

Discrepancies exist between the NOAA observational and GCHP simulated datasets, as the baseline emission scenario consistently lacks correspondence with the observational concentrations of atmospheric C_2H_6 at each of the key sites. The following metric will be utilized in determining the correspondence between the sets of data: significant under/over-estimation corresponding to an average ± ~500pptv to ~1000pptv (Northern Hemisphere) or ~50pptv to ~100pptv (Southern Hemisphere) discrepancy, moderate under/over-estimation corresponding to an average ± ~250pptv to ~500pptv (Northern Hemisphere) or ~25pptv to ~50pptv (Southern Hemisphere) discrepancy, and marginal under/over-estimation corresponding to an average ± ~125pptv to ~250pptv (Northern Hemisphere) or ~12.5pptv to ~25pptv (Southern Hemisphere) discrepancy.

The correspondence for sites within the Northern Hemisphere generally displays a significant underestimation from the baseline emission scenario, with the exception of AMY and NWR which saw significant overestimations. In the case of AMY, the overestimation may be due to continental emissions from Asia driving the emission inventory beyond the scope of observations. While in the case of NWR, the overestimation may be due to a combination of continental emissions from North America driving the emission inventory beyond the scope of observations and the influence of upper-tropospheric winds on transporting atmospheric C_2H_6 when taking observations. While the correspondence for sites within the Southern Hemisphere generally displays a moderate underestimation from the baseline emission scenario. Moreover, the correspondence for each zonal range in the Northern Hemisphere are inconsistent, as the first zonal range of 0˚N—20˚N, home to KUM and MLO saw a moderate underestimation by the baseline emission scenario. While the second zonal range of 20˚N—50˚N, home to AMY, HFM, LEF, NWR, and THD primarily saw moderate underestimations by the baseline emission scenario. With the third zonal range of 50˚N—90˚N, home to ALT, BRW, MHD, and SUM saw significant underestimations by the

baseline emission scenario. As for the correspondence for each zonal range in the Southern Hemisphere are too inconsistent, as the first zonal range of 0˚S—20˚S, home to SMO saw a significant underestimation by the baseline emission scenario. While the second zonal range of 20˚S—50˚S, home to CGO saw a significant underestimation by the baseline emission scenario. With the third zonal range of 50˚S—90˚S, home to PSA and SPO saw significant underestimations by the baseline emission scenario. **See Table – 1 on page 14 of 18, Table – 3 on page 15 of 18, and Figure – 3 on page 18 of 18 for reference.**

4 – Conclusions:

By using GEOS-Chem GCHP (**v13.0.2**) to conduct the baseline emission scenario for C_2H_6 detailed in section 2.3, we contrast the yielded global monthly-average species concentration of C_2H_6 with NOAA's C_2H_6 concentration observations from 2018-01-01 to 2018-06-01. Resulting in the determination of existing discrepancies between the NOAA C_2H_6 observations and produced baseline C_2H_6 emission scenario, as the baseline emission scenario consistently lacks correspondence with the observational concentrations of atmospheric C_2H_6 at each of the key sites. Ranging in scale of inconsistency, the baseline C_2H_6 emission scenario on average significantly underestimates *(± ~500pptv to ~1000pptv (**Northern Hemisphere**) or ~50pptv to ~100pptv (**Southern Hemisphere**)*) the C_2H_6 concentrations at nearly all of the key sites—positing that additional research is necessary in working to replicate current C_2H_6 concentration observations.

This lack of correspondence with the baseline C_2H_6 emission scenario coupled with the implicit relationship between $\frac{CH_4}{C_2H_6}$—any existing discrepancies between C_2H_6 observations and model scenarios extends to potential discrepancies in the original CH_4 emission estimates utilized. Thereby, we suggest that revaluation of the reported level of fossil CH_4 emissions in current emission inventories may be necessary. For a more adequate understanding of the emission contribution from different natural and anthropogenic emission sources is a critical precursor in the design of improving the efficacy of policies set to reverse ongoing trends in the budget of atmospheric C_2H_6 and CH_4.

Thereby we conclude by reemphasizing the present importance of revaluating global C_2H_6 emission inventories, as this work, and many before it, showcases it is warranted. Whether that be via revisiting conventional global C_2H_6 budget compositions with mind for optimization or to implement nonconventional global C_2H_6 budget compositions—the inconsistent correspondence between C_2H_6 observations and baseline emission scenarios remains a looming uncertainty in our understanding of the atmospheric C_2H_6 burden.

Conflicts of Interests:

The authors declare that they have no conflicts of interest.

Acknowledgements:

This research was sponsored by the National Science Foundation: Division of Atmospheric and Geospace Sciences (*Grant Award Number: 1950702*).

Supplementary Materials & Data Availability:

Supplementary datasets of observed global C_2H_6 concentrations measured by NOAA and INSTAAR are available to the scientific community and may be obtained by contacting the National Oceanic and Atmospheric Administration's Global Monitoring Laboratory and the University of Colorado-Boulder's Institute of Arctic and Alpine Research. The additional source emission inventories, such as: MEGAN-MACC, MEGANv2, POET, RETRO can be accessed through Emissions of Atmospheric Compounds and Compilation of Ancillary Data (**ECCAD**). Whereas the Global Geological CH_4 Gridded Emission Inventory produced by G. Etiope[6] and S. Schwietzke[6] can be accessed upon request through the National Oceanic and Atmospheric Administration's Global Monitoring Laboratory. All of which may be directly obtained by contacting the corresponding author.

References:

1. H. Angot et al. Temporary pause in the growth of atmospheric ethane and propane in 2015-2018. *Atmospheric Chemistry and Physics* ***Discuss/Preprint***, (2021).
2. M. Aydin et al. Recent decreases in fossil-fuel emissions of ethane and methane derived from firn air. *Nature* ***476***, 198—201 (2011).
3. L. Bindle et al. Grid-Stretching Capability for the GEOS-Chem 13.0.0 Atmospheric Chemistry Model. *Geoscientific Model Development* ***Discuss/Preprint***, (2021).
4. S. Dalsøren et al. Discrepancy between simulated and observed ethane and propane levels explained by underestimated fossil emissions. *Nature Geoscience* ***11***, 178—184 (2018).
5. S. Eastham et al. GEOS-Chem High Performance (GCHP v11-02c): a next-generation implementation of the GEOS-Chem chemical transport model for massively parallel applications. *Geoscientific Model Development* ***11***, 2941—2953, (2018).
6. G. Etiope et al. Gridded maps of geological methane emissions and their isotopic signature. Earth System Science Data 11, 1—22, (2019).

7. G. Etiope et al. Reappraisal of the fossil methane budget and related emission from geologic sources. *Geophysical Research Letters* ***35***, (2008).

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

9. G. Etiope and S. Schwietzke. Global geological methane emissions: An update of top-down and bottom-up estimates. Elementa Science of the Anthropocene 7, 47—56, (2019).

10. B. Franco 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).

11. B. Franco et al. Retrieval of ethane from ground-based FTIR solar spectra using improved spectroscopy: Recent burden increase above Jungfraujoch. *Journal of Quantitative Spectroscopy & Radiative Transfer* ***160***, 36—49, (2015).

12. D. Helmig et al. Reversal of global atmospheric ethane and propane trends largely due to US oil and natural gas production. *Nature Geoscience* ***9***, 490—499 (2016).

13. M. S. Long et al. Development of a grid-independent GEOS-Chem chemical transport model (v9-02) as an atmospheric chemistry module for Earth system models. *Geoscientific Model Development* ***8***, 595—602, (2015).

14. McKain et al. Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proceedings of the National Academy of Sciences of the United States of America* ***112***, 1941—1946, (2015).

15. A. Milkov and G. Etiope. Revised genetic diagrams for natural gases based on a global dataset of >20,000 samples. *Organic Geochemistry* ***125***, 109—120, (2018).

16. M. Nicewonger et al. 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).

17. A. Pozzer et al. Observed and simulated global distribution and budget of atmospheric C2-C5 alkanes. *Atmospheric Chemistry and Physics* ***10***, 4403—4422 (2010).

18. J. Rudolph. The tropospheric distribution and budget of ethane. *Journal of Geophysical Research* ***100***, (1995).

19. S. Schwietzke et al. Natural gas fugitive emissions rates constrained by global atmospheric methane and ethane. *Environmental Science and Technology* ***48***, 7714—7722, (2014).

20. I. Simpson et al. Long-term decline of global atmospheric ethane concentrations and implications for methane. *Nature* ***488***, 490—494 (2012).

21. K. Sindelarova et al. Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics* ***14***, 9317—9341, (2014).

22. Y. Sun et al. Reduction in C2H6 from 2015-2020 over Hefei, Eastern China points to air quality improvement in China. *Atmospheric Chemistry and Physics **21***, 11759—11779, (2021).

23. P. Tans. A note on isotopic ratios and the global atmospheric methane budget. *Global Biogeochemical Cycles **11***, 77—81 (1997).

24. A. J. Turner et al. Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmospheric Chemistry and Physics **15***, 7049—7069, (2015).

25. Z. Tzompa-sosa et al. Addressing the Global Ethane Budget within GEOS-Chem. *Department of Atmospheric Science: Colorado State University & School of Engineering and Applied Sciences: Harvard University*, (2017).

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

27. G. R. van der Werf et al. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires during 1997-2009. *Atmospheric Chemistry and Physics **10***, 11707—11735, (2010).

28. G. R. van der Werf et al. Global fire emissions estimates during 1997-2016. *Earth System Science Data **9***, 697—720, (2017).

29. Y. Xiao et al. Constraints on Asian and European sources of methane from CH4-C2H6-CO correlations in Asian outflow. *Journal of Geophysical Research **109***, (2004).

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

31. J. Zhuang et al. Enabling High-Performance Cloud Computing for Earth Science Modeling on Over a Thousand Cores: Application to the GEOS-Chem Atmospheric Chemistry Model. *Journal of Advances in Modeling Earth Systems **12***, (2020).

NOAA Observation Sites					
Site	**Nearest Municipality**	**Country**	**Lat**	**Lon**	**Elevation**
ALT	Alert	Canada	82.5 N	62.3 W	210 m asl
AMY	Anmyeon-do	South Korea	36.5389 N	126.3295 E	47.0 m asl
BRW	Barrow, AK	USA	71.3 N	156.6 W	8 m asl
CGO	Cape Grim	Australia	40.682 S	144.688 E	164 m asl
HFM	Harvard Forest	USA	42.5 N	72.2 W	340 m asl
KUM	Cape Kumukahi, HI	USA	19.516 N	154.811 W	3 m asl
LEF	NEAR Worcester, WI	USA	45.6 N	90.27 W	868 m asl
MHD	Mace Head	Ireland	53.3 N	9.9 W	42 m asl
MLO	Mauna Loa, HI	USA	19.5362 N	155.5763 W	3397 m asl
NWR	Niwot Ridge, CO	USA	40.1 N	105.5 W	3475 m asl
PSA	Palmer Station	Antarctica	64.6 S	64.0 W	10 m asl
SMO	NEAR Vaifanua	American Samoa	14.247 S	170.564 W	77 m asl
SPO	Not Specified	South Pole	89.98 S	24.8 W	2837 m asl
SUM	Summit	Greenland	72.6 N	38.4 W	3200 m asl
THD	Trinidad Head, CA	USA	41.0 N	124.1 W	120 m asl

Table – 1. Legend for all of the NOAA observation sites.

Emission Scenario Compositions		
Baseline Emission Scenario:		
Source Emission Categories	**Emission Inventory**	**Additional Notes**
Anthropogenic Fossil Fuel	Tzompa-Sosa: C_2H_6 (**2010**)	Baseline – N/A
Biofuel	Tzompa-Sosa: C_2H_6 (**2010**)	Baseline – N/A
Biomass Burning	GFEDv4 (**2018**)	Baseline – N/A
*Optimized Fossil Fuel Emission Scenario (**OFF**):*		
Source Emission Categories	**Emission Inventory**	**Additional Notes**
Anthropogenic Fossil Fuel	Tzompa-Sosa: C_2H_6 (**2010**)	Incremental scaling: starting at x1.4
Biofuel	Tzompa-Sosa: C_2H_6 (**2010**)	Not Optimized – N/A
Biomass Burning	GFEDv4 (**2018**)	Not Optimized – N/A
Biogenic	MEGAN-MACC (**2010**)	Additional emission inventory

Geologic	Etiope: GEOCH$_4$ (**2018**)	Additional emission inventory
Oceanic	RETRO (**2000**)	Additional emission inventory
*Optimized Biomass Burning Emission Scenario (**OBB**):*		
Source Emission Categories	**Emission Inventory**	**Additional Notes**
Anthropogenic Fossil Fuel	Tzompa-Sosa: C2H6 (**2010**)	Not Optimized – N/A
Biofuel	Tzompa-Sosa: C_2H_6 (**2010**)	Not Optimized – N/A
Biomass Burning	GFEDv4 (**2018**)	Incremental scaling: starting at x1.2
Biogenic	MEGAN-MACC (**2010**)	Additional emission inventory
Geologic	Etiope: GEOCH$_4$ (**2018**)	Additional emission inventory
Oceanic	RETRO (**2000**)	Additional emission inventory

Table – 2. Detailed composition of each emission scenario with additional information regarding scaling.

C_2H_6 Hemispheric & Zonal Gradient Statistics		
*NOAA: C_2H_6 Observational Data (**2018**):*		
Northern Hemispheric Gradient	**Southern Hemispheric Gradient**	
~750pptv to ~2100pptv	~110pptv to ~230pptv	
*GCHP: C_2H_6 Simulated Data (**2018**):*		
Northern Hemispheric Gradient	**Southern Hemispheric Gradient**	
~1255pptv to ~1928pptv	~93pptv to ~145pptv	
*NOAA: C_2H_6 Observational Data (**2018**):*		
Zonal Range	**Weighed Observation Sites**	**Average Zonal Concentration Gradient**
90˚N—50˚N	ALT, BRW, MHD, SUM	~750pptv to ~2250pptv
20˚N—50˚N	AMY, HFM, LEF, NWR, THD	~1000pptv to ~2300pptv
0˚N—20˚N	KUM, MLO	~450pptv to ~1250pptv
0˚S—20˚S	SMO	~200pptv to ~350pptv
20˚S—50˚S	CGO	~100pptv to ~250pptv
50˚S—90˚S	PSA, SPO	~125pptv to ~275pptv
*GCHP: C_2H_6 Simulated Data (**2018**):*		
Zonal Range	**Weighed Observation Sites**	**Average Zonal Concentration Gradient**
90˚N—50˚N	ALT, BRW, MHD, SUM	~1050pptv to ~1800pptv

20˚N—50˚N	AMY, HFM, LEF, NWR, THD	~1680pptv to ~2360pptv
0˚N—20˚N	KUM, MLO	~600pptv to ~1100pptv
0˚S—20˚S	SMO	~120pptv to ~220pptv
20˚S—50˚S	CGO	~90pptv to ~120pptv
50˚S—90˚S	PSA, SPO	~80pptv to ~120pptv

Table – 3. Tabulated hemispheric and zonal gradient statistics table, composed of the yielded results discussed in section 3.1 and 3.2.

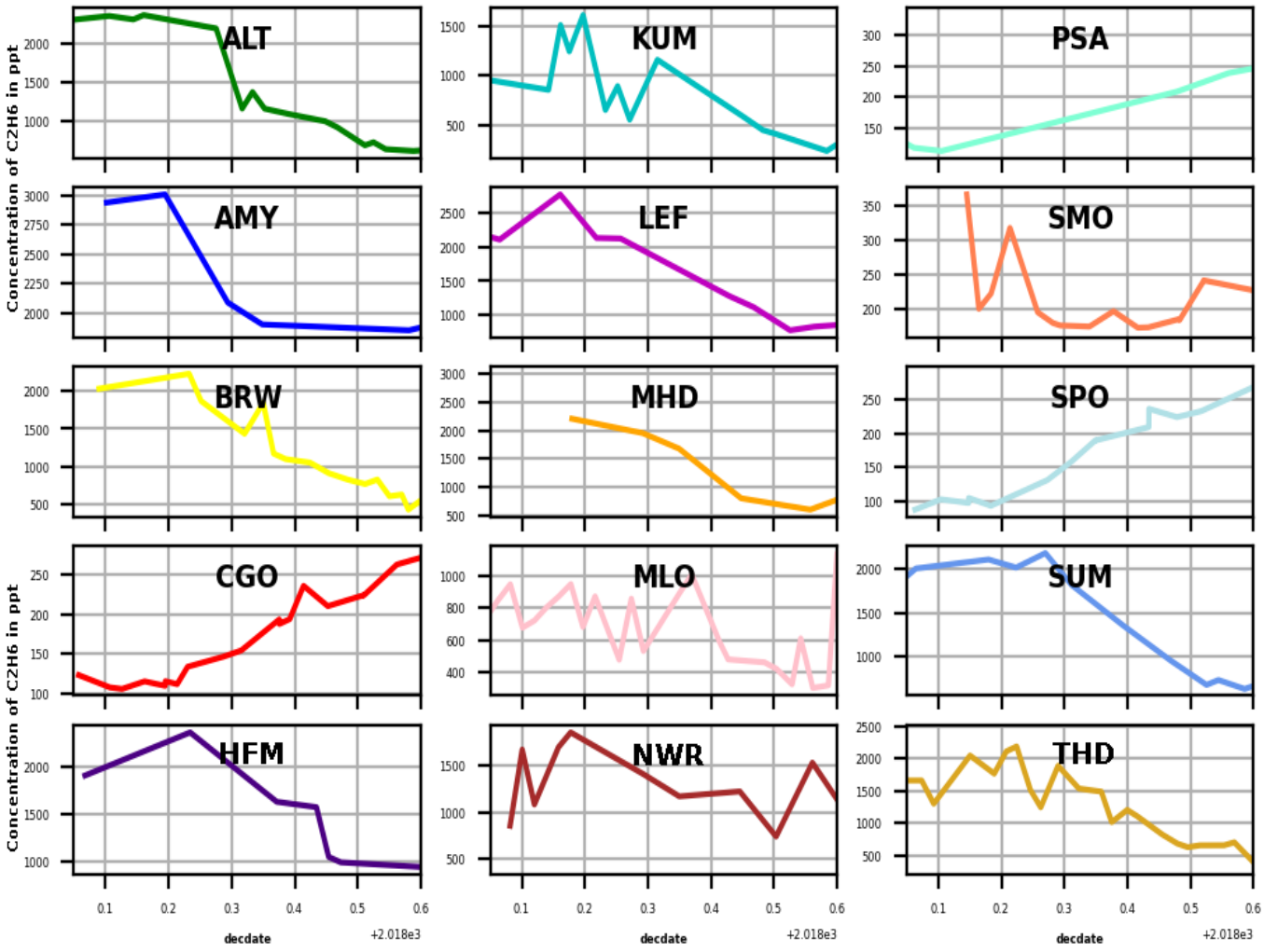

Figure – 1. Series of subplots of NOAA's C_2H_6 concentration observations (**C_2H_6_C**) at each observation sites from 2018-01-01 to 2018-06-01. The y-axis denotes relative concentration of C_2H_6 in parts per trillion (**pptv**) and the x-axis denotes the allotted timeframe.

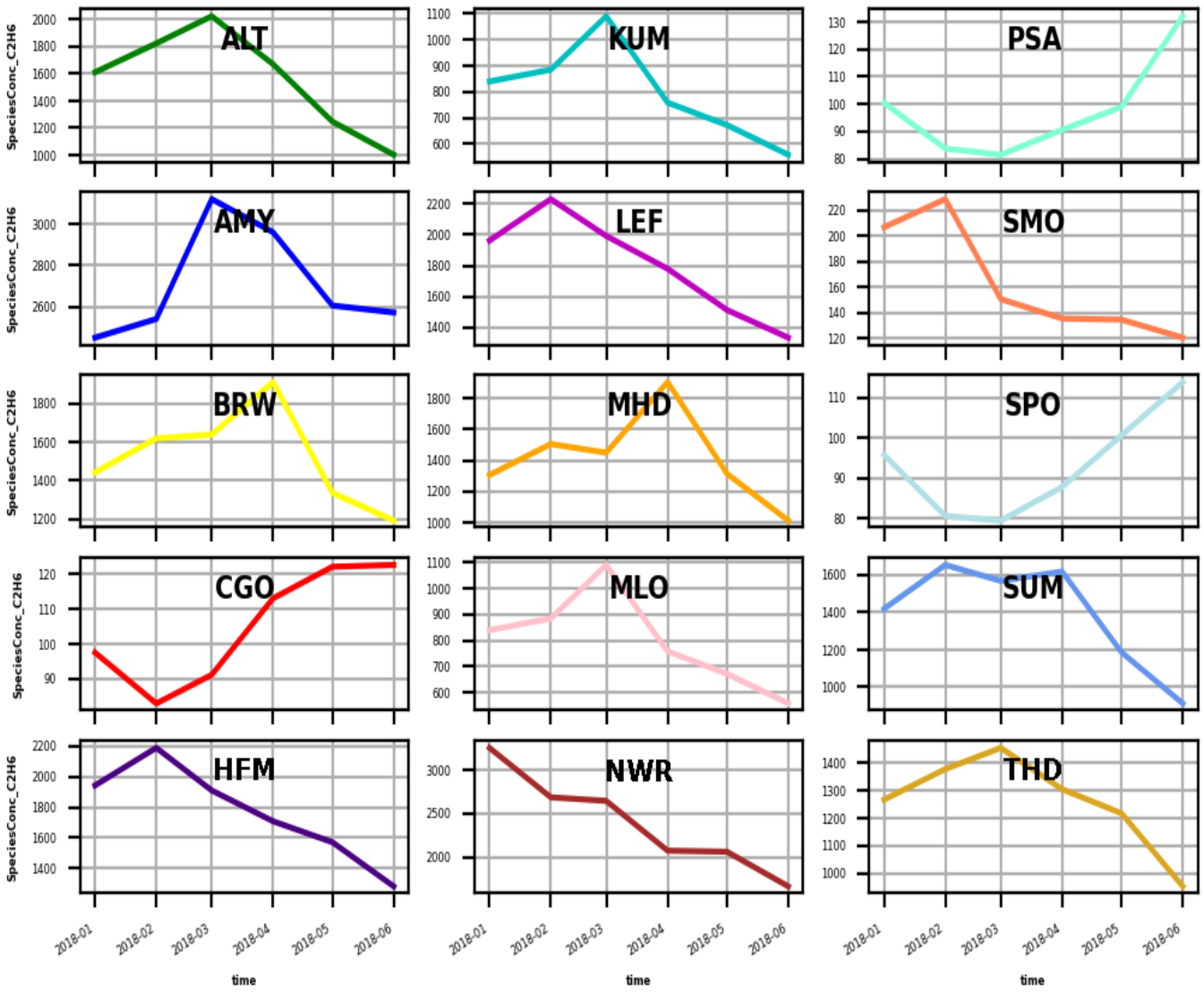

Figure – 2. Series of subplots of GCHP's baseline emission scenario C_2H_6 simulated concentration outputs (**SpeciesConc_C_2H_6**) at each observation sites from 2018-01-01 to 2018-06-01. The y-axis denotes relative concentration of C_2H_6 in parts per trillion (**pptv**) and the x-axis denotes the allotted timeframe.

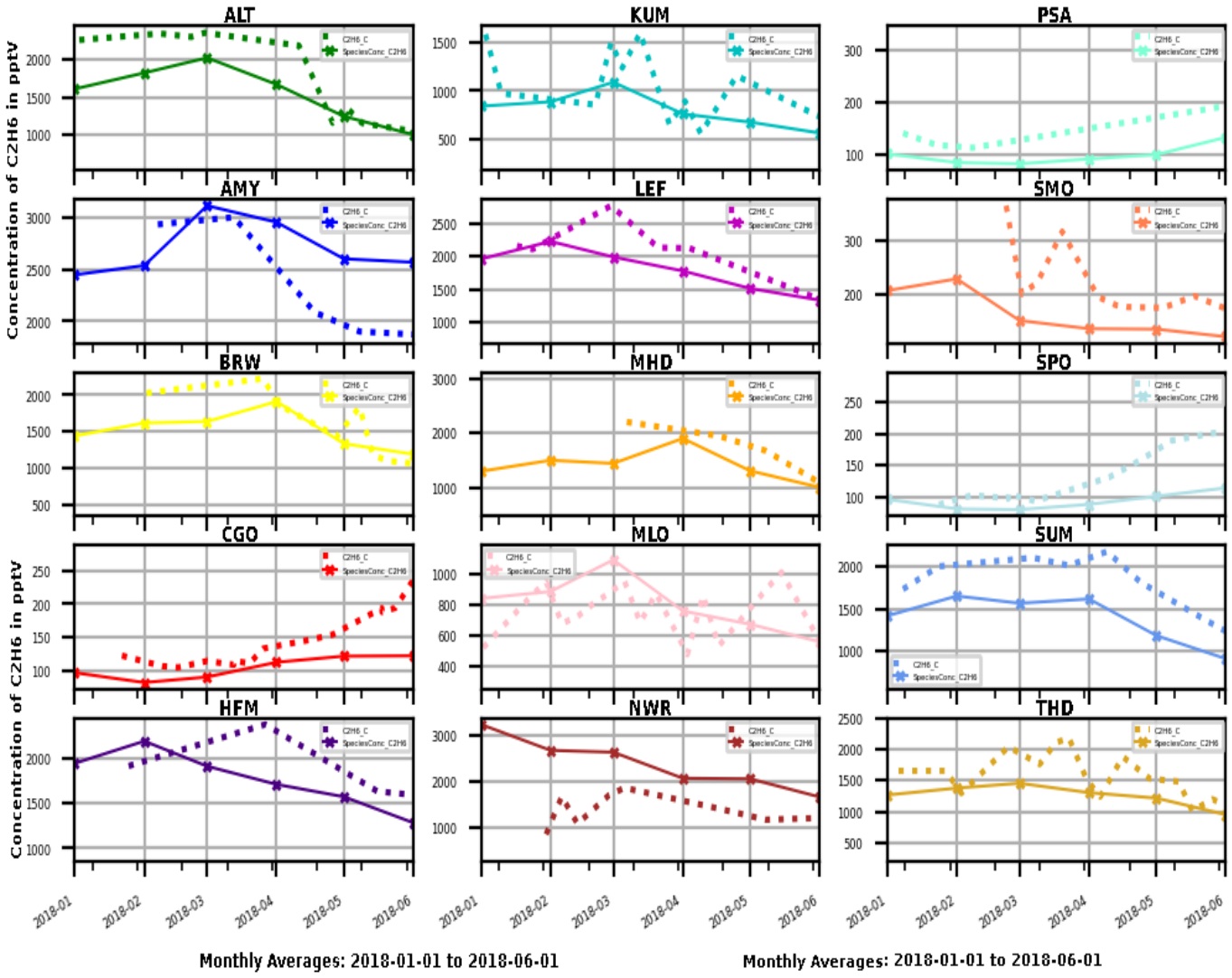

Figure – 3. Series of subplots of NOAA's C_2H_6 concentration observations (**C_2H_6_C**) vs GCHP's baseline emission scenario C_2H_6 simulated concentration outputs (**SpeciesConc_C_2H_6**) at each observation sites from 2018-01-01 to 2018-06-01. On each subplot, the NOAA C_2H_6 concentration observations (**C_2H_6_C**) correspond to the dotted line-type with no data-markers, while the GCHP baseline C_2H_6 simulated concentrations (**SpeciesConc_C_2H_6**) correspond to the solid line-type with x-style data-markers. The y-axis denotes relative concentration of C_2H_6 in parts per trillion (**pptv**) and the x-axis denotes the allotted timeframe.