Supplement of Atmos. Chem. Phys., 22, 6539–6558, 2022
https://doi.org/10.5194/acp-22-6539-2022-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Age spectra and other transport diagnostics in the North American monsoon UTLS from SEAC^{4}RS in situ trace gas measurements

Eric A. Ray et al.

Correspondence to: Eric A. Ray (eric.ray@noaa.gov)

S1. Trace gas profiles and surface time series

Theta average profiles with 10 K resolution were calculated for every trace gas used in this study. A complicating factor in calculating averages for the shortest-lived species especially is the spread in mixing ratios of up to two orders of magnitude within a single theta bin. A small number of relatively high mixing ratios in a theta bin can skew the average high, such as seen for n-butane and i-butane in Figure S1. The initial theta average profiles for each trace gas in Figure S1 are shown by the dashed orange lines. These profiles are used in the first iteration of the calculation (Section 4.1). When the age spectra calculation is performed on the individual measurements (Section 4.2), those trace gases with elevated mixing ratios affected by pollution sources are identified, such as shown in Figures S8. The theta averages are then recalculated with scaling factors applied to the identified elevated mixing ratios resulting in new theta average profiles shown by the solid lines with symbols in Figure S1. The recalculated theta average profiles can have significantly lower mixing ratios for the short-lived trace gases but are nearly identical to the original theta average profiles for longer-lived trace gases such as chloroform ($CHCl_3$) that were less affected by pollution sources (Figure S1).

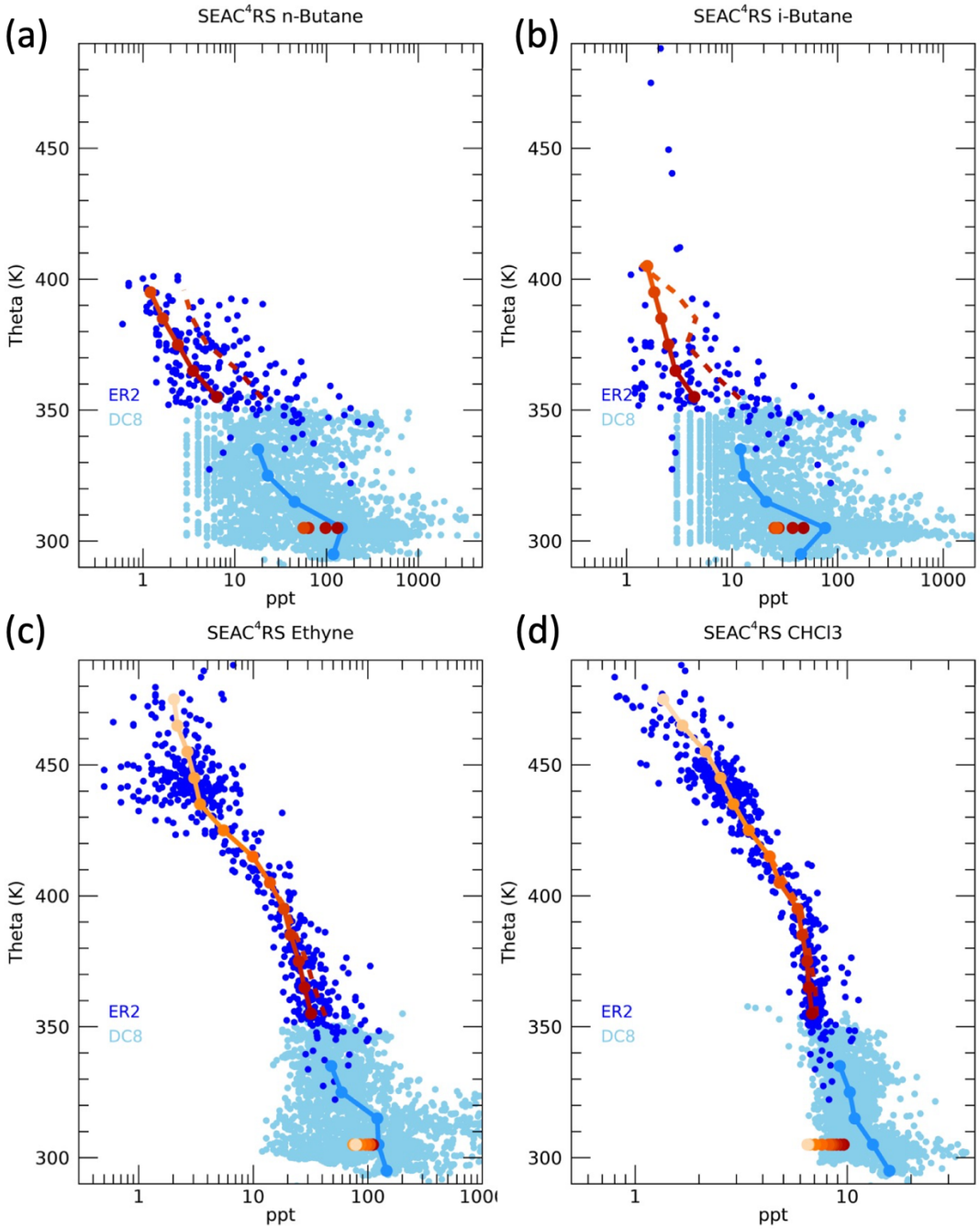

Figure S1. *Profiles of (a) n-butane, (b) i-butane, (c) ethyne and (d) $CHCl_3$ as measured by the WAS instrument on the ER-2 (dark blue) and on the DC-8 (light blue). The solid lines are mixing ratio averages for each 10 K theta layer and the orange symbols are BL mixing ratios based on the BL time series for each latitude and the derived age spectra and source latitudes for each layer. The orange shading represents the theta level where the darker shading is lower theta and the lighter shading higher theta. The dashed orange lines are the original theta average profiles before adjustments to account for elevated mixing ratios from pollution sources. The optimized BL mixing ratios χ_{io} for each theta level are shown by the orange shaded symbols at 305K.*

For some of the trace gases a scaling was applied to account for calibration offsets between WAS measurements and those from the surface networks. Calibration offsets are not unexpected between different measurement systems. The trace gases and scaling values applied to the WAS measurements (in parenthesis) are HCFC-141b (0.97), HFC-134a (0.95), HCFC-142b (0.991), halon 1211 (0.985), CCl_4 (1.02), CFC-11 (1.02), CFC-113 (0.985) and CFC-114

(1.028). These scaling values were obtained based primarily on the relative UT lifetimes shown in Table 1 and the $\mu_i^* - \tau$ relationships for the theta average profile in the 350-360K layer.

For CO_2 surface data we use CarbonTracker, version CT2019B, zonally averaged within the NAM longitudes for the NH extratropical source region and globally averaged for the tropical source region. For the NH extratropical region we used the 1°x1° resolution 'nam1x1' CarbonTracker files and for the tropical region we used the 3°x2° resolution 'glb3x2' files. CarbonTracker is a vertically resolved product and we used the lowest three levels to approximate the lower boundary layer sampling of the NOAA and AGAGE network measurements for the other trace gases in this study. The lowest three levels represent the atmosphere up to ~150 meters above the surface and are essentially always within the boundary layer in the NAM region. Examples of the CO_2 time series at different latitudes within the NAM region during 2013 are shown in Figure S2. We also performed the calculation with averages over the lowest two levels (up to ~50 meters) with very little change in the results.

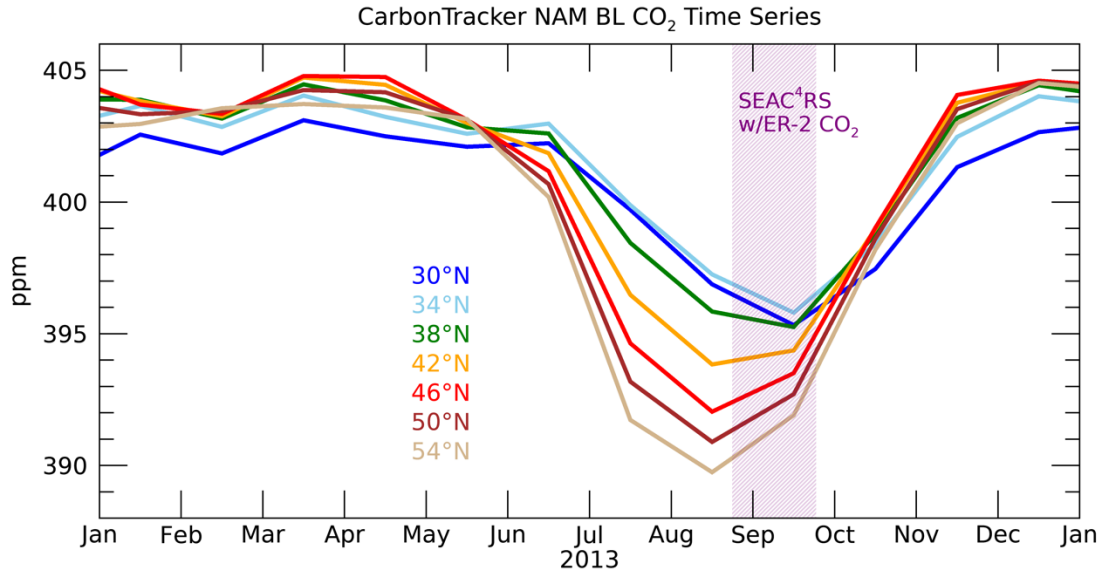

Figure S2. *Time series of CarbonTracker NAM CO_2 averaged over the longitudes 85°-112°W and the lowest two vertical levels. The dates of the SEAC⁴RS mission when ER-2 CO_2 measurements were available are indicated by the purple shading.*

S2. Partitioning of the age spectra between surface source regions

A fundamental aspect of the calculations performed here is the identification of surface source latitude regions to the NAM UTLS and the delineation between air that originated from the North American surface vs. from the tropical surface (Figures 1 and 3). As shown in Equations 4 and 5 we split the full age spectrum into parts that originated from each latitude region by a parameter *f(t)* that is a function of age such as shown in Figure S3. The basic assumption shown in Figure 3 is that the shortest time scales of transport from days to weeks are primarily due to convection from the local North American surface, while the longer time scales of transport come from the tropical surface. Thus, the parameter *f(t)* follows a Gaussian functional form such that for the youngest ages the North American surface source spectra compose greater than 90% of the total spectra and for ages near the value of t_f the tropical surface source spectra compose greater than 90% of the total spectra. The total spectra are evenly divided between the two source regions at an age of $t_f/2$.

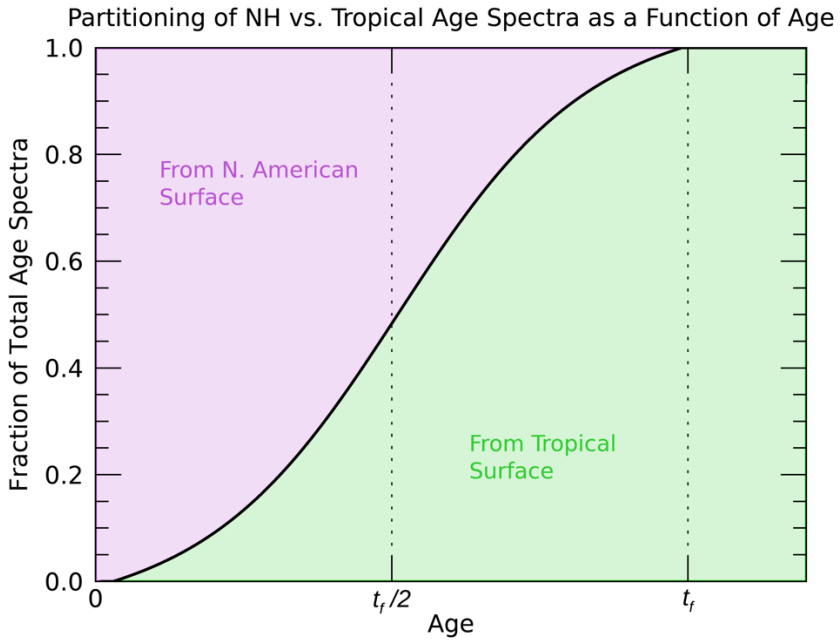

Figure S3. *Partitioning of the total age spectra between NA and tropical surface source regions as a function of age (solid black line is f(t) in Equations 4 and 5). The value of t_f represents the age above which the surface source is entirely from tropical latitudes.*

We tested versions of *f(t)* with t_f from 30 to 200 days to find the partitioning with the lowest difference quantity *D*, as described in the optimization section below, at each theta level from 350-470 K. Below 380 K the shortest t_f time scales result in the lowest (best) values of *D*, while above 420 K the longest t_f time scales give the best results. Based on this, we were able to optimize the total *D* over all theta levels by choosing t_f values that range from 50 days below 380 K to 150 days above 420 K with a transition layer from 380-400 K where t_f increased from 60 to 100 days. What this means physically is that there is a greater proportion of older air (~months) from the NH surface in the LS compared to the UT. This older air could have been convectively injected over previous months or transported by large scale Rossby wave mixing into the tropical upper troposphere during the spring where it could then have been isentropically mixed into the LS. This relatively old air from the NH surface would of course have to be transported back into the NH UT but would be diluted enough that it would be difficult to detect in the UT with this method.

S3. Surface source latitude distributions

The peaks of L_{TR} and L_{NA} are separated by 4° latitude and have half widths of 10° (Figure S4). In the tropics the peak latitudes (y_{pTR}) range from -4°S to 24°N and in NA the peaks (y_{pNA}) range from 26°-54°N. Note that there is overlap in L_{TR} and L_{NA} distributions from 20-30°N so that surface air in this range could be considered to have come from either or both regions depending on the optimal solution. To facilitate the calculation, we interpolated the $\chi_i(y, t - t')$ time series of each trace gas to 1° latitude resolution.

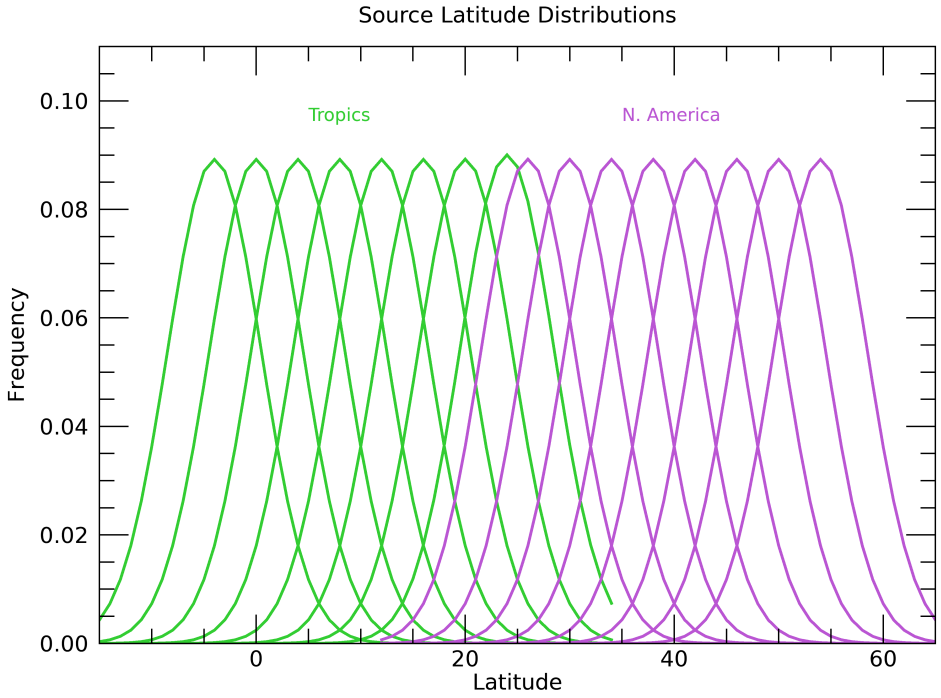

Figure S4. *Distributions of the surface source latitudinal distributions* L_{TR} and L_{NA}.

S4. Optimization method for average theta profiles

The optimum age spectra and surface source regions at each theta level were chosen based on the minimum absolute differences between both the measured and calculated BL fractions (μ^*, μ) and CO_2 mixing ratios (χ_{CO2}, χ_{CO2o}). For the BL fraction differences we used the absolute value of the symmetric signed percentage bias (SSPB) (Morley et al., 2018) quantity ($aSSPB_\mu$), and for CO_2 mixing ratios we used the absolute value of the mixing ratio differences ($a\chi_{CO2diff}$). The minimization of a combined difference quantity, $D = aSSPB_\mu + 10(a\chi_{CO2diff})$, was used to choose the optimum set of K, y_{TR} and y_{NA} indices. Since the units are not the same for the two difference quantities that compose D, we chose to scale the CO_2 differences in order to better match the $aSSPB_\mu$ values and to enhance the sensitivity of the CO_2 surface source dependency of D. Figure S3 shows the range of $aSSPB_\mu$, $10(a\chi_{CO2diff})$ and D for the 380 K level. For each value of K, which represents a single total age spectrum, the range in differences is due to the different source regions y_{TR} and y_{NA}. The $aSSPB_\mu$ values in the light blue shading have a much narrower range for each K value compared to the $10(a\chi_{CO2diff})$ values in the yellow shading. This highlights the strength of using the combined difference quantity D in the optimization since the BL fractions from the set of 20 trace gases used here are sensitive to the age spectra choice and relatively insensitive to the BL source regions, while CO_2 is mostly the opposite. For CO_2, very small minimum differences can be found for a wide range of K due to fact that CO_2 in the BL has such a large spread in latitude that a good solution can be found for most age spectra by adjusting the surface source region.

We use 130 K values with a range of 1-1000 m^2/s evenly spaced in log10(K). For the lowest theta layer in the calculation (350-360 K), the K range is unrestricted in the optimization. Above this layer, the range of K values is restricted to those less than the optimal K value found in the layer below (K_{opt}) minus a small offset. This restriction ensures that the mean and modal ages get older with higher theta as expected. The offset applied between layers varies from 15-50 days in mean age and 0.5-10 days in modal age.

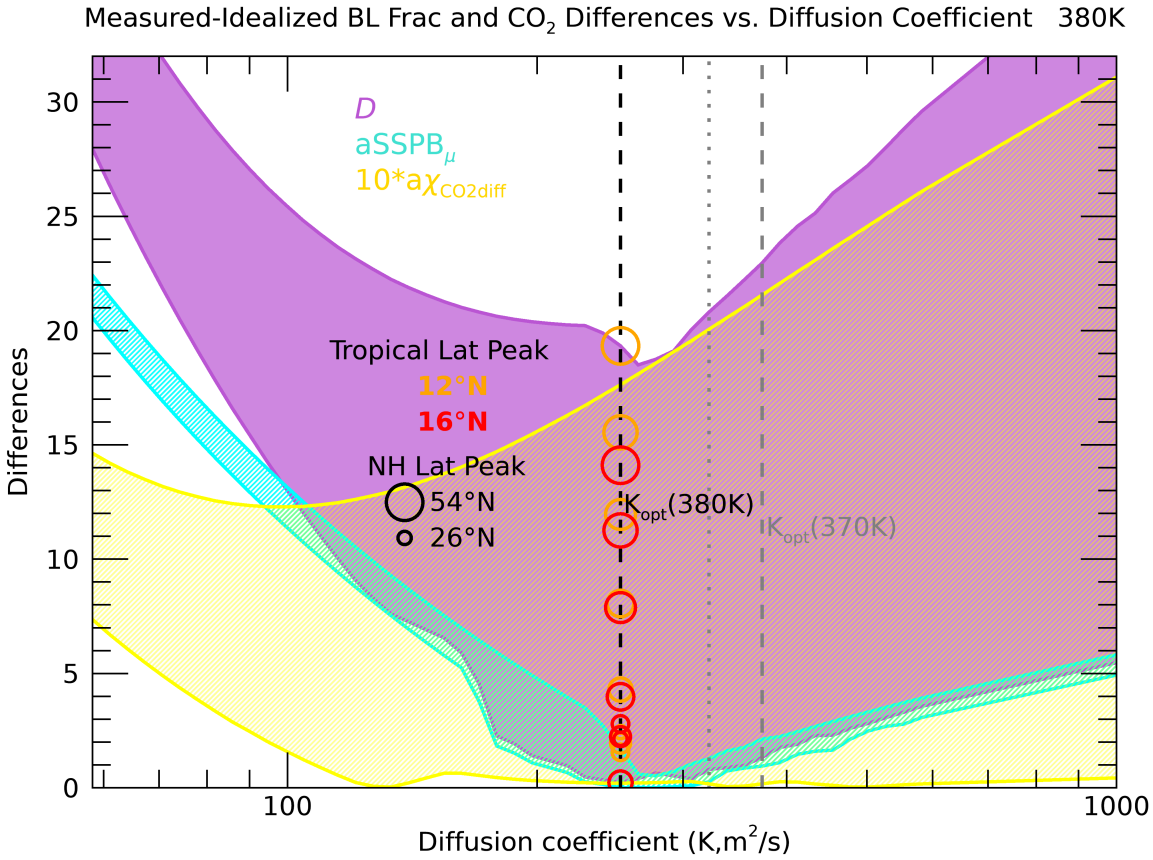

Figure S5. *Differences as a function of the diffusion coefficient for the 380-390 K layer averages used in the optimization method (see text for details). The optimal value of K for this theta layer is indicated by the dashed black line and values of D for different surface source latitudes at this K value are shown by the open circles (colored by the tropical latitude and sized by the NH extratropical latitude as shown in the legend.*

In the example shown in Figure S5 for the 380-390K layer, the minimum value of D corresponds with K_{opt}=300 m^2/s. In the layer below, 370-380K, the value of K_{opt}=480 m^2/s (dashed grey line) and the maximum K in the optimization for the 380-390K layer is 420 m^2/s (dotted grey line). For this layer, the restriction of K does not affect the optimization since the minimum value of D is at a significantly lower value of K than the cutoff value.

An additional constraint on the optimizations was to limit the tropical source peak latitude in each layer to be either equal to or 4° south of the tropical peak from the layer below. For the initial lowest theta layer, no restriction was made on the tropical peak location. This constraint is based on the idealized transport shown in the Figure 3 schematic, which is consistent with the results of previous modeling studies on the sources of air to the UTLS (Birner and Boenisch, 2011; Orbe et al., 2015). These studies show that higher theta levels in the UTLS are generally influenced by more equatorial latitudes compared to lower theta levels. In Figure S5, the D values at K_{opt} are shown as open circles for y_{TR} = 12°N and 16°N since for the layer below the optimum value of y_{TR} = 16°N. The NA source latitude peaks are not constrained for any of the theta layers.

As mentioned earlier, for the theta layer in Figure S5 the variability in D is primarily driven by the NA source regions and this is shown by the spread in both the red and orange circles representing both tropical source latitudes. This spread is almost entirely due to CO_2 differences from the large summer extratropical latitudinal gradients. The larger size circles represent higher NA source latitudes and since these have the largest D values, they are not compatible with the measurements. The lowest value of D corresponds with a tropical peak of 16°N, as for the layer below, and a NA peak source latitude of 38°N.

The large range of differences in Figure S5 makes it somewhat difficult to clearly see the minimum difference, so in Figure S6 the y-axis is expanded and all of the layers are shown with the corresponding minimum values and source latitudes. In the lowest two layers the peak latitude combination of 20°N and 48°N have the clear lowest differences driven by the low CO_2 differences shown in the light purple squares. The CO_2 differences shown here are 10 times the actual mixing ratio differences and so are less than 0.1 ppmv in nearly every layer.

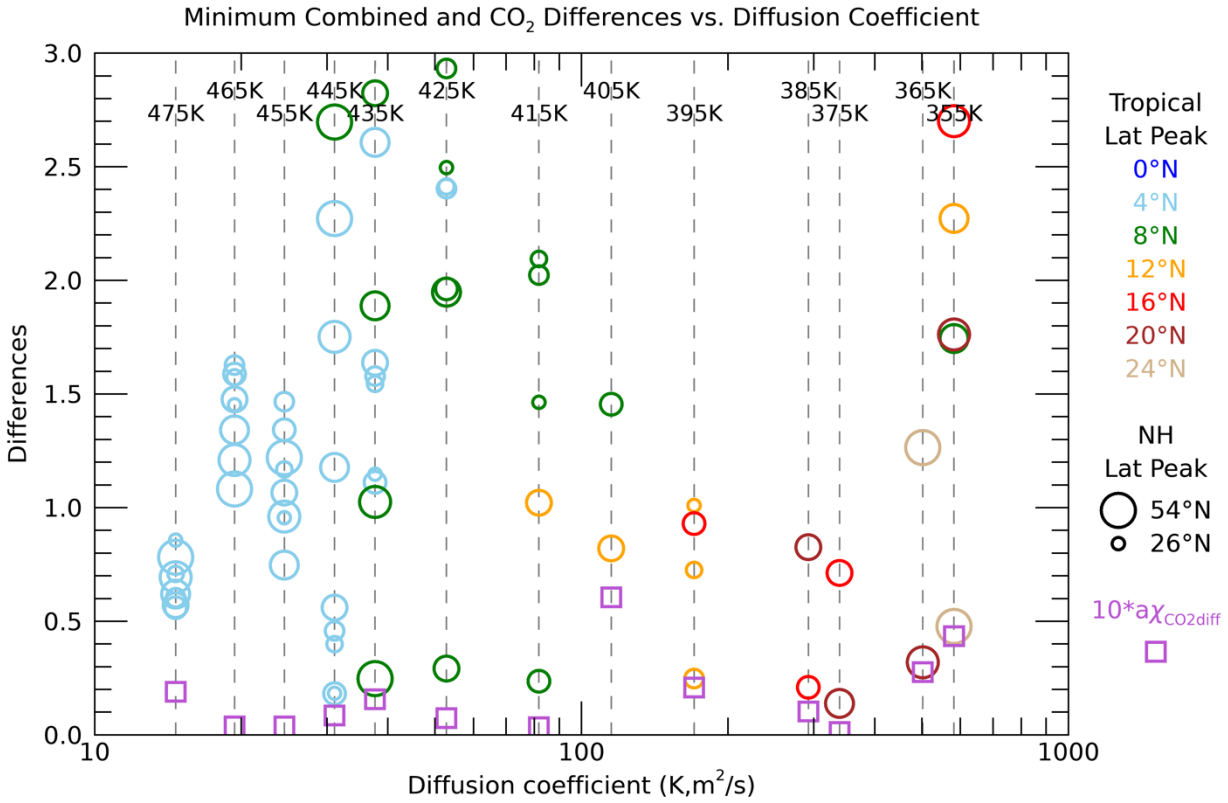

Figure S6. *Combined differences (D) and 10 times the CO_2 mixing ratio differences* ($10*a\chi_{CO2diff}$) *as a function of the diffusion coefficient K for all of the theta layers. Each dashed line indicates the optimum value of K for the theta layer labeled at the top and the associated values of D are indicated by the circles colored by the tropical peak latitude and sized by the NA peak latitude. The CO_2 differences are indicated by the purple squares.*

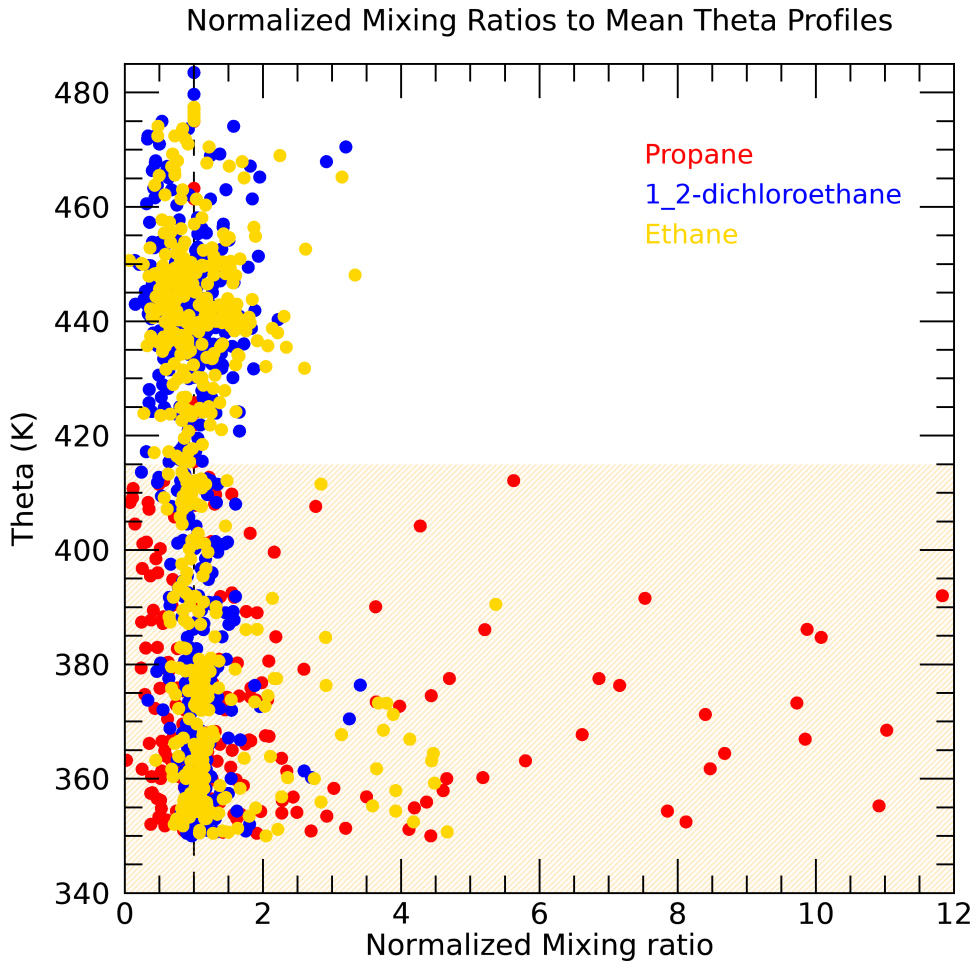

Figure S7. *Profiles of propane, 1,2-dichloroethane and ethane mixing ratios normalized to the mean theta profile for each trace gas (χ_{inorm}). The orange shading represents the theta range where these values are used to scale the BL mixing ratios for trace gases in the calculation of age spectra from individual measurement locations.*

S5. Scaling Factors

To address the outlier measurements with relatively large mixing ratios we normalize the individual mixing ratios with the mean profiles of each trace gas by $\chi_{inorm}(\boldsymbol{x}) = \chi_i(\boldsymbol{x})/\chi_i(z)$ where $\boldsymbol{x}$ refers to an individual measurement location (Figure S7). All of the χ_{inorm} profiles have a similar pattern of variability with the largest spread of values below 420 K and a much smaller spread above that level. Our assumption is that the spread below 420 K is primarily driven by NA source region variability since transport alone cannot cause such high values, and above this level the spread is primarily driven by transport variability since the NA has a relatively small influence at those locations. Thus, we define a scaling factor for $\theta(\boldsymbol{x})$ < 420K,

$$S_{inorm}(\boldsymbol{x}) = \begin{cases} 1 & , \ \chi_{inorm}(\boldsymbol{x}) < 1.2 \\ \chi_{inorm}(\boldsymbol{x}) & , \ \chi_{inorm}(\boldsymbol{x}) \geq 1.2 \end{cases} . \quad \text{(S1)}$$

That is, for a trace gas with a mixing ratio 20% or more larger than the theta layer average value, we scale χ_{io} by S_{inorm} so that ideally all $\mu^*(K, z, y_{TR}, y_{NH}) < 1$. We find initial estimates of $\mu_i^*(\boldsymbol{x})$ using the optimized χ_{io} BL values from the theta average layer that encompasses the individual measurement location.

After the initial scaling we address remaining $\mu^*(\boldsymbol{x})$ > 1 outliers with an additional scaling factor

$$S_{i\mu}(\boldsymbol{x}) = \begin{cases} 1 & , \ \mu_i^*(\boldsymbol{x}) \leq 1 \\ \mu_i^*(\boldsymbol{x})/\mu_i^*(z) & , \ \mu_i^*(\boldsymbol{x}) > 1 \end{cases} \quad \text{(S2)}$$

where $\mu^*(z)$ is the BL fraction for the theta average layer (Figure 6a) that encompasses the measurement location $\boldsymbol{x}$. A combined scaling factor defined as $S_i = S_{inorm} * S_{i\mu}$ is then applied to all χ_{io} and an initial optimization is performed as was done for the theta average profiles in the previous section. For each location the trace gas lifetimes are initialized with the theta layer average values $\tau_i(z)$ and adjusted following the optimization to the $\tau_i(\boldsymbol{x})$ values.

After this initial optimization we compare the $\tau_i(\boldsymbol{x})$ to $\tau_i(z)$ values and make adjustments to χ_{io} for any trace gas with $\tau_i(\boldsymbol{x})$ that is a factor of 4 smaller or larger than $\tau_i(z)$. That is,

$$S_{i\tau}(\boldsymbol{x}) = \begin{cases} 1 & , \quad \tau_i(z)/4 \leq \tau_i(\boldsymbol{x}) \leq 4 * \tau_i(z) \\ \mu_i^*(\boldsymbol{x})/\mu_i^*(z) & , \ \tau_i(\boldsymbol{x}) < \tau_i(z)/4 \ or \ \tau_i(\boldsymbol{x}) > 4 * \tau_i(z) \end{cases}. \quad \text{(S3)}$$

The reason we make this adjustment is that even with the scaling described above there are still outliers in $\mu_i^*(\boldsymbol{x})$ that are apparent in the $\mu(\tau)$ relationships. We do not want to overly constrain the possible transport variability revealed by the majority of individual measurements within a theta layer so the initial scaling steps are only intended to identify clear outliers and broadly bring the trace gases into $\mu(\tau)$ alignment. Outliers that remain following the initial optimization will generally not affect the best fit and value of *K*, but will appear as outliers in the profiles of $\tau_i(\boldsymbol{x})$. The overall scaling factor is then defined as $S_i = S_{inorm} * S_{i\mu} * S_{i\tau}$. After the scaling factor $S_{i\tau}$ is applied to χ_{io} the optimization is performed again to calculate the final transport parameters for each individual location.

Profiles of S_{iNH} for propane and ethane are shown in Figure S8. The scaling of propane can be as large as 70 but typically ranges from 2-20, whereas the scaling of ethane ranges from 2-7 with only one value over 10. We can compare the scaling factors among different trace gases to approximate emission factors, which are often used to identify pollution sources of air measured in the troposphere (e.g. Swarthout et al., 2013). Figure S9 shows the scaling factors for the 10 shortest lived trace gases relative to propane scaling factors. Most of the scaling factors are less than the value for propane at an individual measurement location with the exception of i-butane and n-butane, which often have larger scaling factors than propane. Previous studies have identified emission ratios of various volatile organic compounds such as propane and ethane from different sources. The emission ratio of ethane to propane is ~1.5-2 and n-butane to propane is ~0.5-1 from natural gas sites and urban areas [e.g. Swarthout et al., 2013]. While there are a few ethane scale factors larger than propane, the average scaling factor ratio of ethane to propane is ~0.5. And similarly for n-butane, while there are some scale factors less than propane, the average scale factor ratio of n-butane to propane is ~1.2. Although there is potential to identify specific pollution sources in the UTLS with this technique, it does not appear that we can do so with this data set.

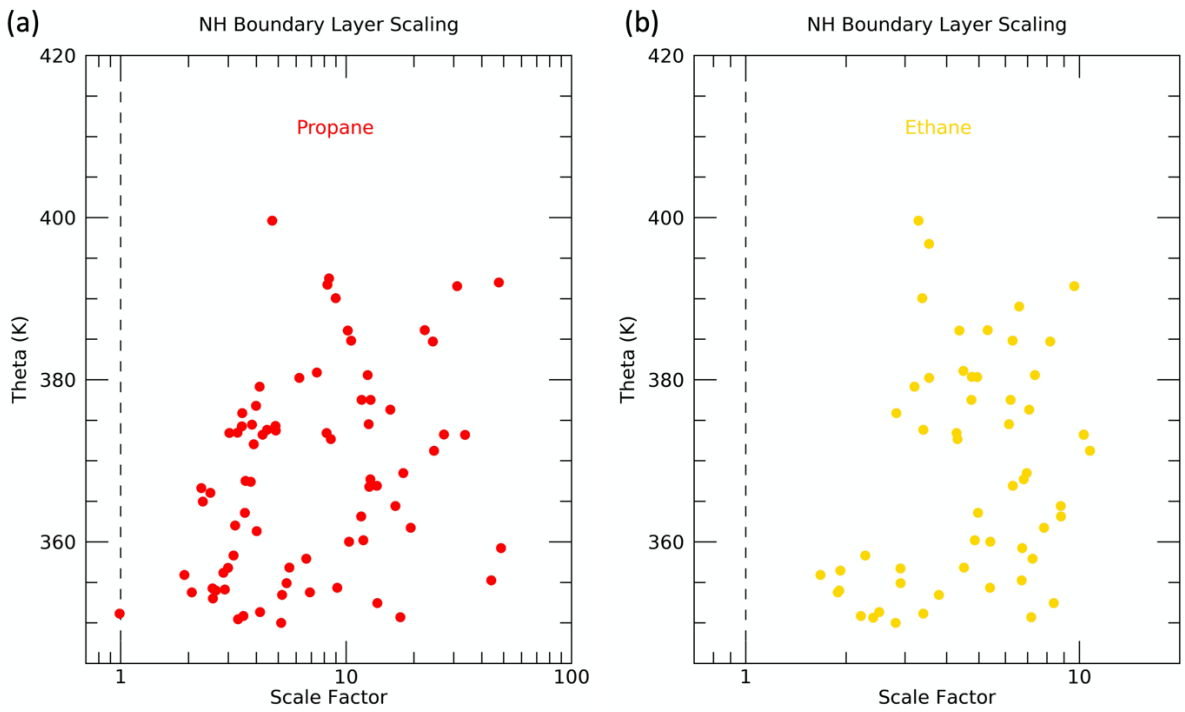

Figure S8. *Profiles of the NA BL scaling factors S_{iNA} for (a) propane and (b) ethane.*

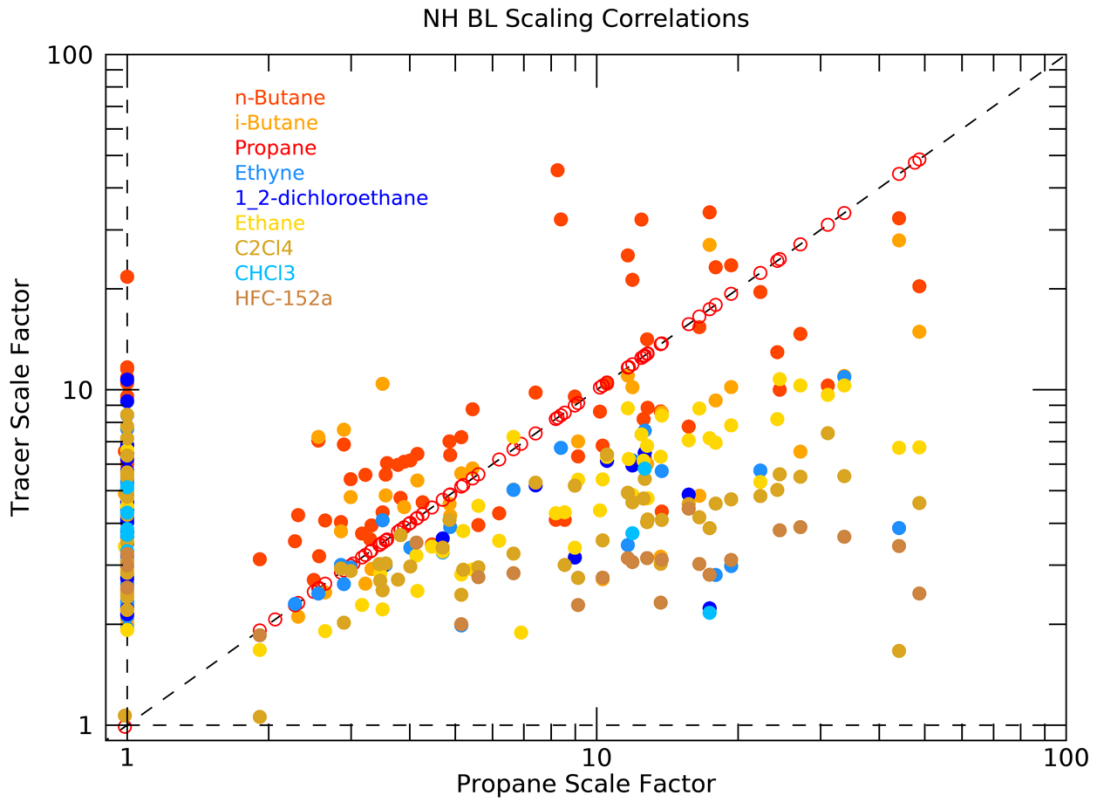

Figure S9. *Comparisons of the NA BL scaling factors (S_{iNA}) for propane with those of nine other trace gases. The one-to-one line is indicated by the dashed line and the open symbols are the individual propane scaling factors.*

As a check on the size of the S_{iNA} values, we first find the NA BL mixing ratios for each measurement location by dividing the optimized χ_{io} values by the NH source fraction as was done for the scaling factors. Then we multiply these values by the S_{iNH} values to estimate the 'polluted' NH BL mixing ratios. Figure S10 shows profiles of propane and ethane from both the ER-2 and DC-8 as well as the adjusted BL mixing ratios. For propane, the adjusted mixing ratios range from 200-10,000 ppt, which is encompassed by the range of mixing ratios measured in the lower troposphere by the DC-8. Similarly for ethane, the adjusted NH BL mixing ratios lie within the range of those measured by the DC-8. The DC-8 measurements are only from the time of the mission so this does not necessarily represent the full spread of mixing ratios over the previous months that may have acted as a source to the UTLS. But it is at least an indication that the scaling factors used in the calculation are not much larger than any measurements in the lower troposphere.

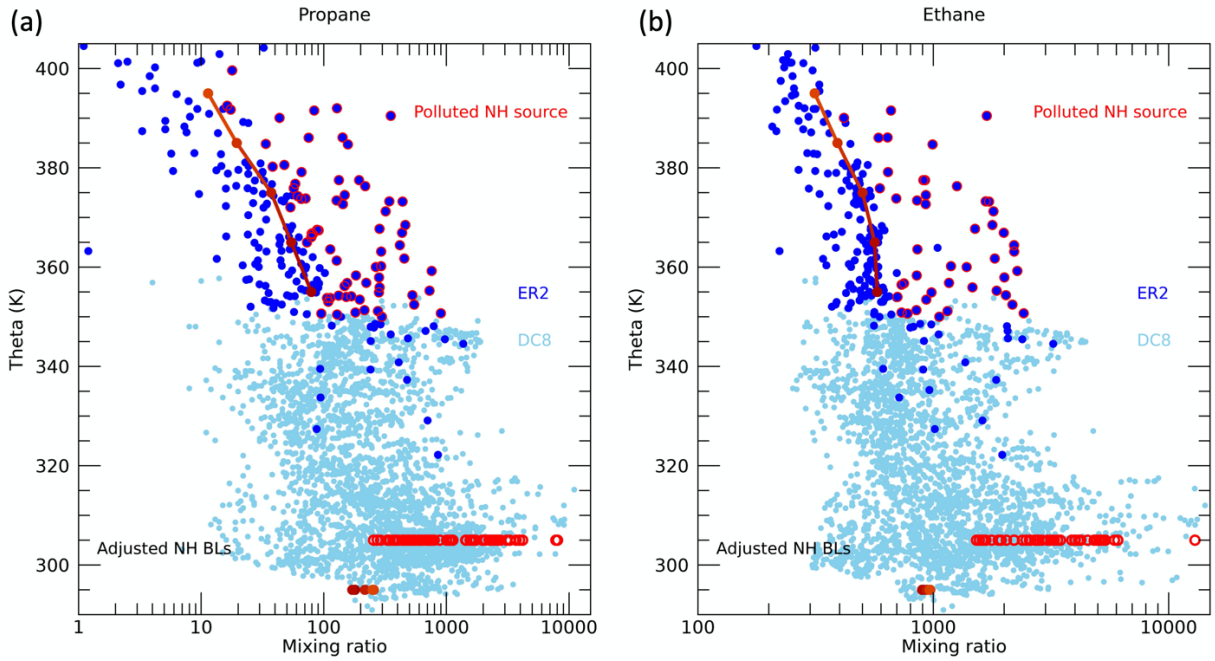

Figure S10. *Profiles of (a) propane and (b) ethane mixing ratios as measured during SEAC⁴RS by WAS on the ER-2 (blue circles) and DC-8 (light blue circles). Measurement locations with values of $S_{iNA} > 0$ are indicated by the red circles around the blue circles and the associated scaled NH boundary layer mixing ratios are indicated by the open red circles at 305 K.*

References

Morley, S. K., Brito, T. V., and Welling, D. T., Measures of model performance based on the log accuracy ratio, Space Weather, 16, 69-88, doi:10.1002/2017SW001669, 2018.

Swarthout, Robert F., Rachel S. Russo, Yong Zhou, Andrew H. Hart and Barkley C. Sive (2013), Volatile organic compound distributions during the NACHTT campaign at the Boulder Atmospheric Observatory: Influence of urban and natural gas sources, *J. Geophys. Res.*, *118*, 10,614-10,637, doi:10.1002/jgrd.50722.