Assessing scales of variability for atmospheric composition field data relevant to future Decadal Survey satellite observations

Jim Crawford, Glenn Diskin, Glen Sachse, and Gao Chen — *NASA Langley Research Center*
Tom Ryerson, David Parrish, and John Holloway — *NOAA ESRL*
Andy Weinheimer — *NCAR*
Ron Cohen — *UC Berkeley*

Abstract. *Establishing appropriate specifications for satellite observations of atmospheric composition is a difficult and inexact task since neither models nor observations can provide both the resolution and spatial coverage required. Nevertheless, instrument specifications need to be supported by careful and detailed modeling analyses (e.g., OSSEs) and examination of available in situ observations. Despite shortcomings in temporal and spatial coverage, field observations are unique in capturing true atmospheric variability on scales down to and below those of satellite observations. Here we assess the spatial variability of field observations useful for establishing measurement requirements for future Decadal Survey observations (e.g., GEO-CAPE, GACM, and ASCENDS).*

Method. Spatial variability of in situ, airborne observations has been evaluated using a variogram approach outlined below.

The equations:

Classical Variogram Definition (Matheron, 1962)

$$2\gamma(h) \equiv \frac{1}{N(h)}\sum_{N(h)}(Z(s_i)-Z(s_j))^2$$

Where N is the number of data pairs separated by distance h; Z(s) is the variable of interest at a given location s; and locations s_i and s_j denote location pairs separated by distance h

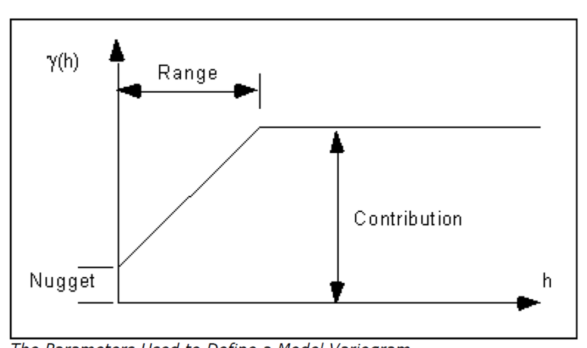
*Variogram Definition **used for this analysis** (also called a semimadogram)*

$$\gamma(h) \equiv \frac{1}{N(h)}\sum_{N(h)}|Z(s_i)-Z(s_j)|$$

Simply stated, it is the average difference for the variable of interest over a given distance. Future plans may include calculating other statistics (e.g., median and percentiles).

Basic variogram interpretation: *taken from*

http://www.ems-i.com/gmshelp/Interpolation/Interpolation_Schemes/Kriging/Variogram_Editor.htm



The Parameters Used to Define a Model Variogram.

- *The nugget represents a minimum variance. For this analysis, the nugget is likely dominated by the measurement uncertainty.*
- *The contribution (sometimes called the "sill") represents the average variance of points at such a distance away from the point in question that these is no correlation between the points.*
- *The range represents the distance at which there is no longer a correlation between the points.*
  For the airborne data analysis presented here, the distance (h) is considered to represent satellite resolution and the variogram (γ(h)= average difference) to be an indication of expected sub-grid variability for a given resolution.

Data filtering and assumptions:

Data assessed for all pairs below a given altitude (e.g., 1 km, 2 km)

Data pairs with distances of up to 100 km included

Data pairs must span less than 15 minutes (roughly equivalent to 100 km at a flight speed of 100 m/s) which minimizes differences that may be attributed to chemistry (especially for NO2) and transport.

Assessed variables measured at 1 hz (roughly 100 m resolution for NOAA P-3 and 150 m for NASA DC-8)

Data pairs restricted to daylight conditions as defined by solar zenith angles of 70 degrees or less

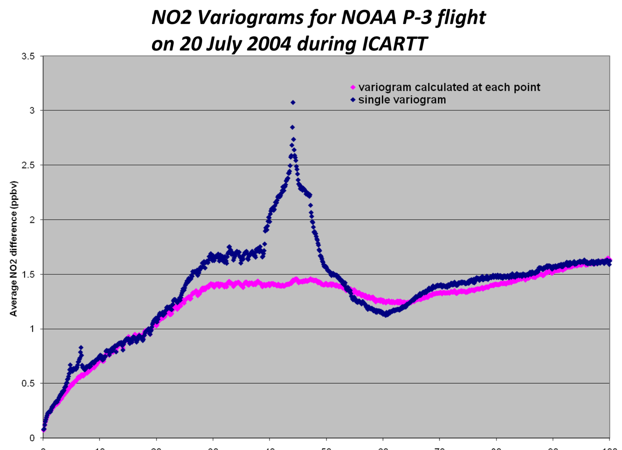
Data are assumed to be isotropic (i.e., vector direction between data pairs is not important)

Data are assumed to represent a well-mixed boundary layer (i.e., vertical separation between data pairs is not a discriminator)

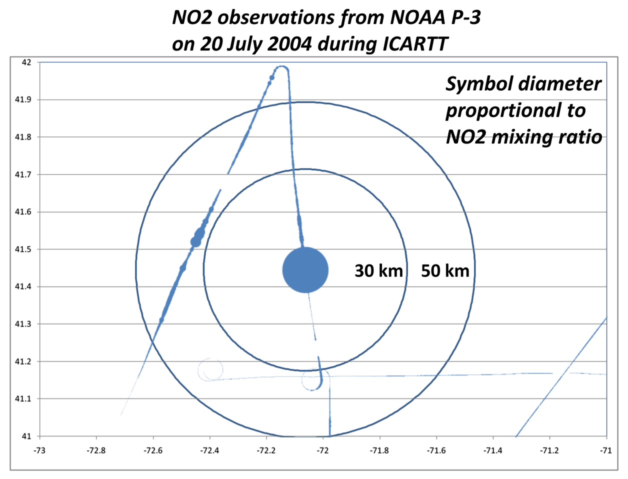
A final caveat concerning the influence of flight patterns on variogram behavior:

In general, variograms for individual flights are well behaved, but there are exceptions. Since research flights are often targeted and do not represent systematic geographic coverage, situations can arise when certain distances are overemphasized and plume gradients are overrepresented. In the case shown below, there is a large concentration of data pairs in proximity to a large NO2 plume at distances of 30-50 km, leading to an overemphasized NO2 difference for these distances. This problem is mitigated by computing the variogram at each point rather than a single variogram for the entire flight.

NO2 Variograms for NOAA P-3 flight on 20 July 2004 during ICARTT

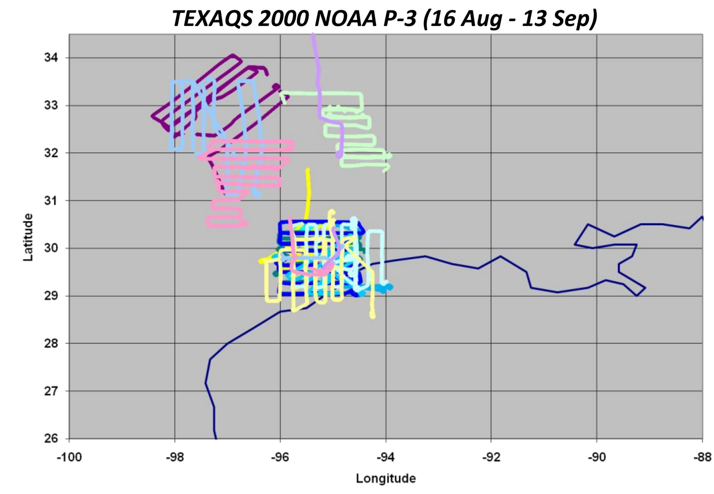


NO2 observations from NOAA P-3 on 20 July 2004 during ICARTT

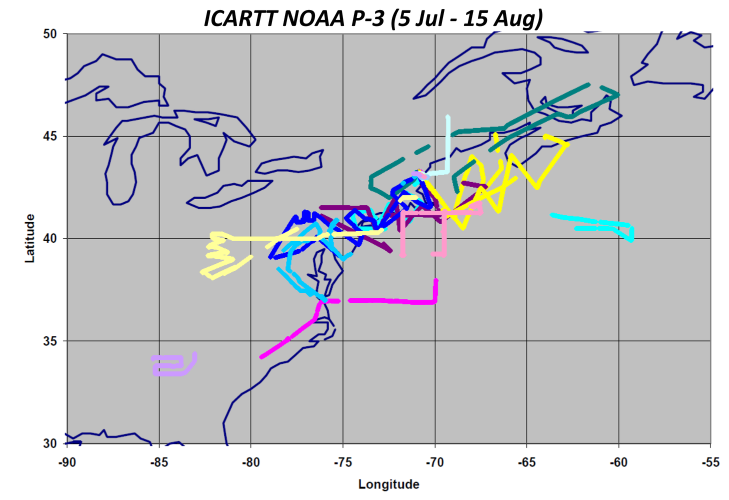


Distribution of Flight Data Collected in the Boundary Layer (below 2 km): Variogram analyses were conducted on observations collected during four campaigns

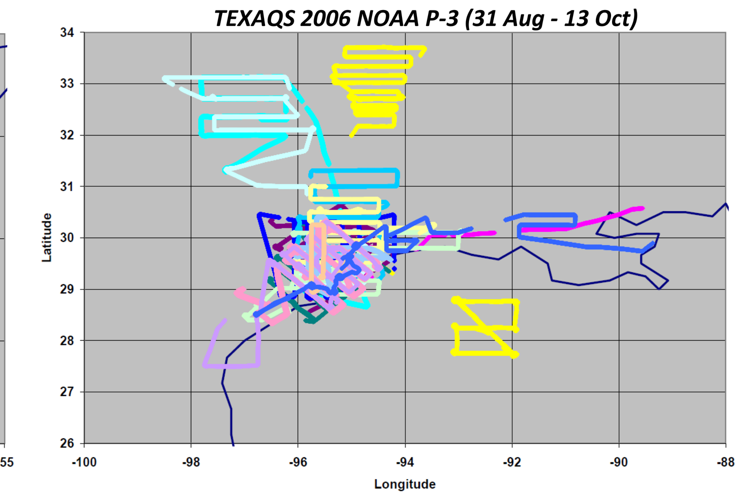
TEXAQS 2000 NOAA P-3 (16 Aug - 13 Sep)



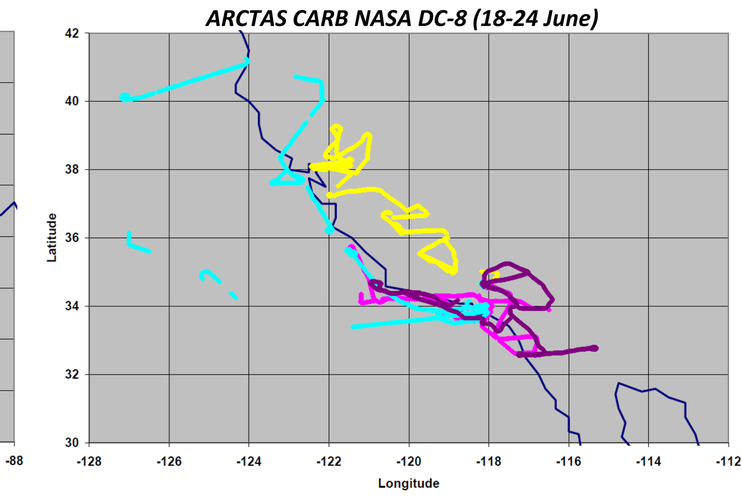
ICARTT NOAA P-3 (5 Jul - 15 Aug)



TEXAQS 2006 NOAA P-3 (31 Aug - 13 Oct)

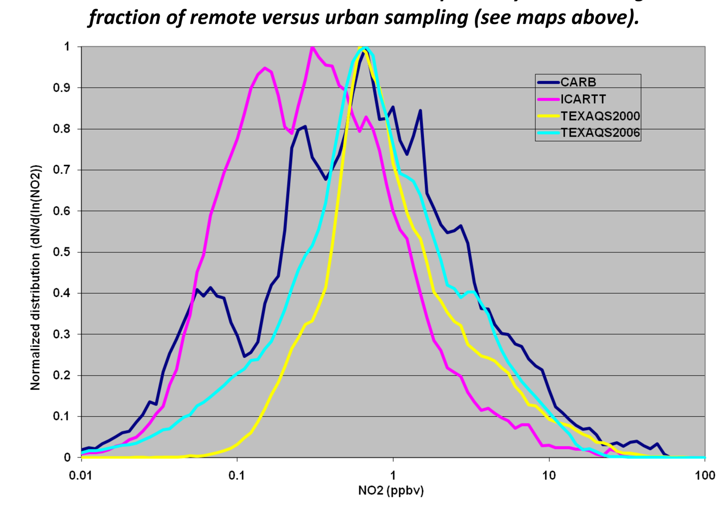


ARCTAS CARB NASA DC-8 (18-24 June)

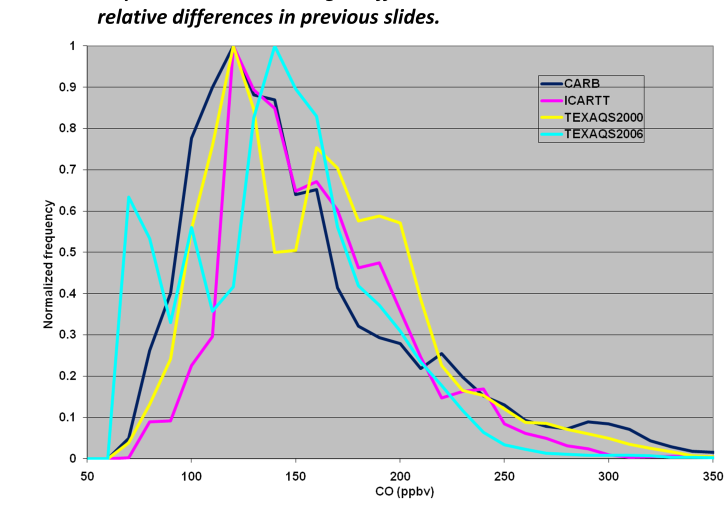


Normalized Frequency Distributions for NO2, CO, and Ozone

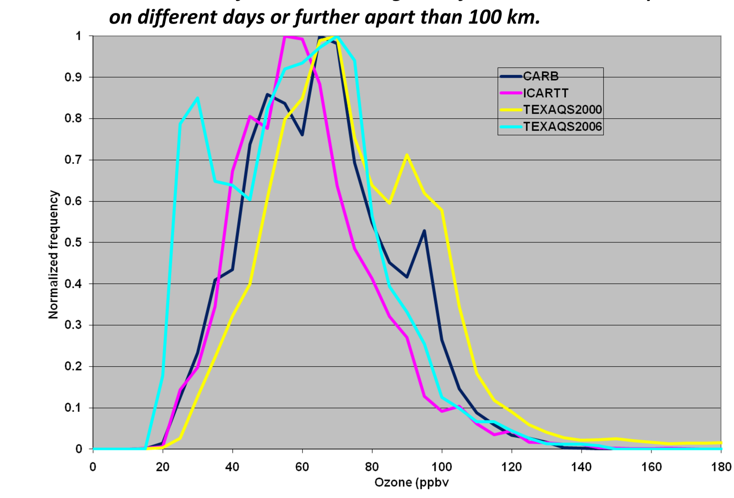
NO2: CARB and TEXAQS2000 data have the largest number of extreme values (10 ppbv or more) which helps corroborate the larger relative NO2 differences observed on the previous slide. The ICARTT distribution is skewed low primarily due to a larger fraction of remote versus urban sampling (see maps above).



CO: CARB data again shows the broadest distribution which helps corroborate the larger differences in both absolute and relative differences in previous slides.

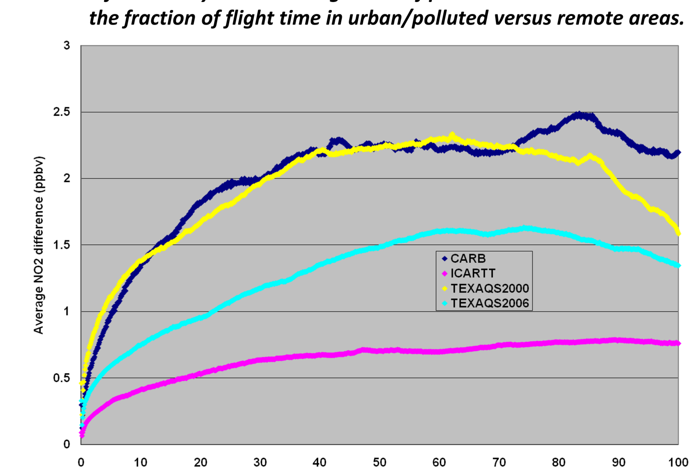


Ozone: TEXAQS2000 and CARB distributions have larger fraction of high ozone values. Some multimodal behavior is evident, but this will not influence the variograms if the modes are sampled on different days or further apart than 100 km.

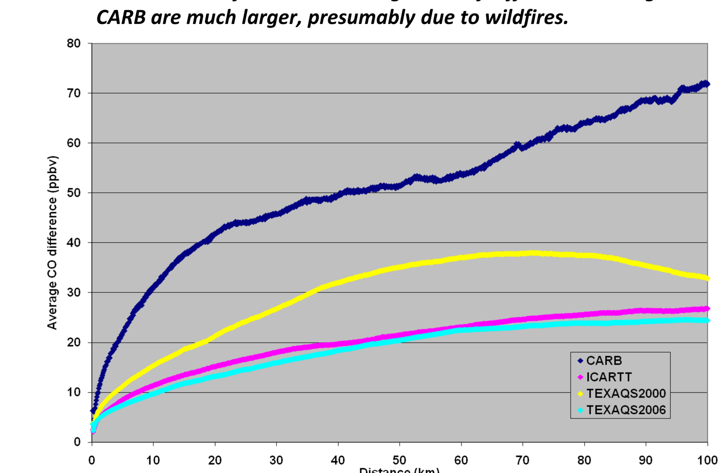


Variograms for NO2, CO, and Ozone

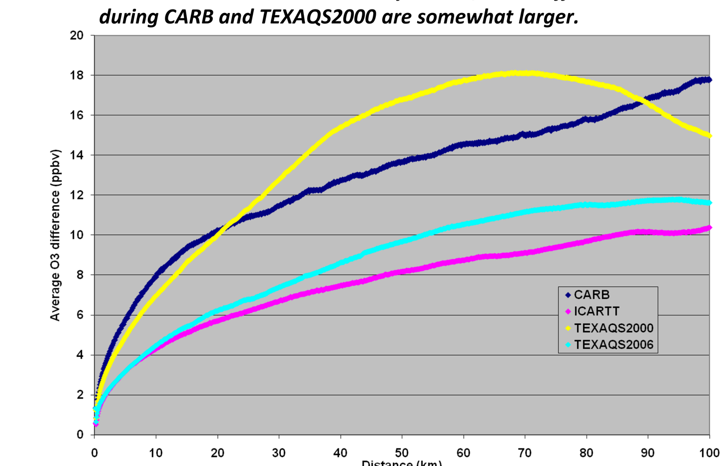
NO2: Basic behavior is similar for all four campaigns, although magnitudes differ. Interpreting magnitude is difficult since it is influenced by both the magnitude of pollution encountered and the fraction of flight time in urban/polluted versus remote areas.



CO: These variograms are less likely to level off compared to the NO2 variograms. While the phenomenon could be related to lifetime, it indicates that gradients for CO persist over greater distances than for NO2. The magnitude of differences during CARB are much larger, presumably due to wildfires.

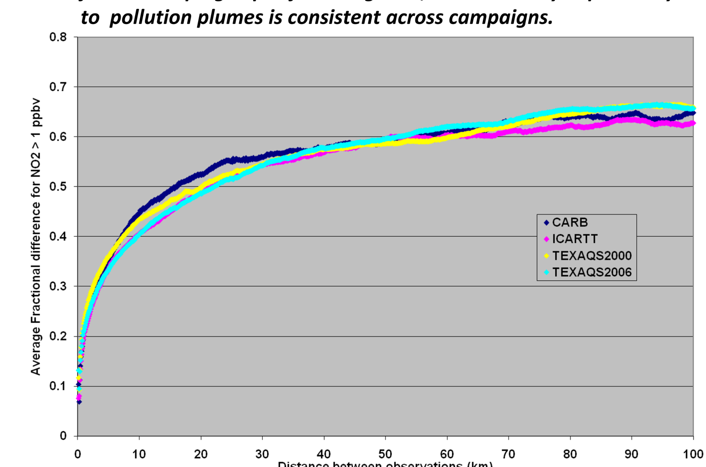


Ozone: These variograms behave similarly to the CO variograms. ICARTT and TEXAQS2006 are very similar, while differences during CARB and TEXAQS2000 are somewhat larger.

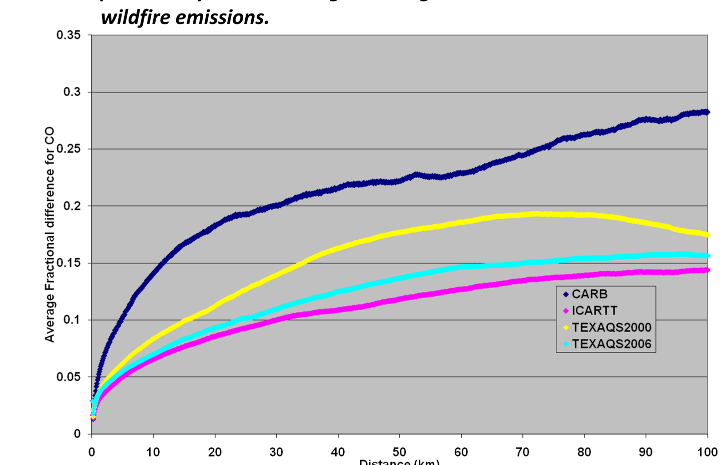


Normalized Variograms for NO2, CO, and Ozone: ***Rather than the absolute difference, these variograms are based on the absolute difference divided by the larger of the two values***

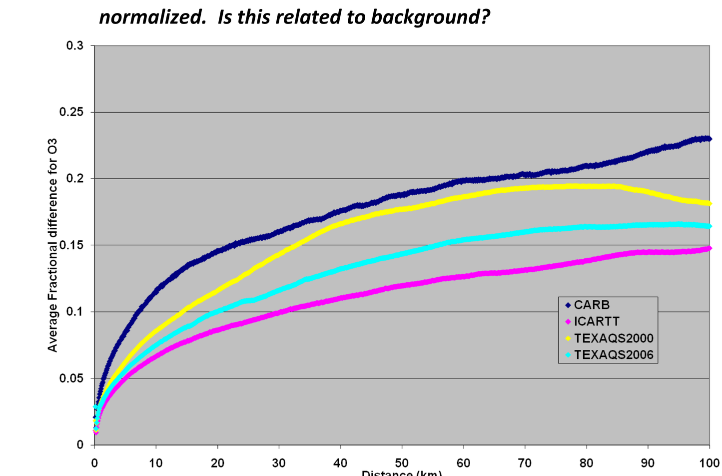
NO2: Here, variograms have been calculated for the fractional difference in NO2 for values in excess of 1 ppbv. The similarity in these curves suggests that despite the differences in magnitude for the campaign-specific variograms, the variability in proximity to pollution plumes is consistent across campaigns.



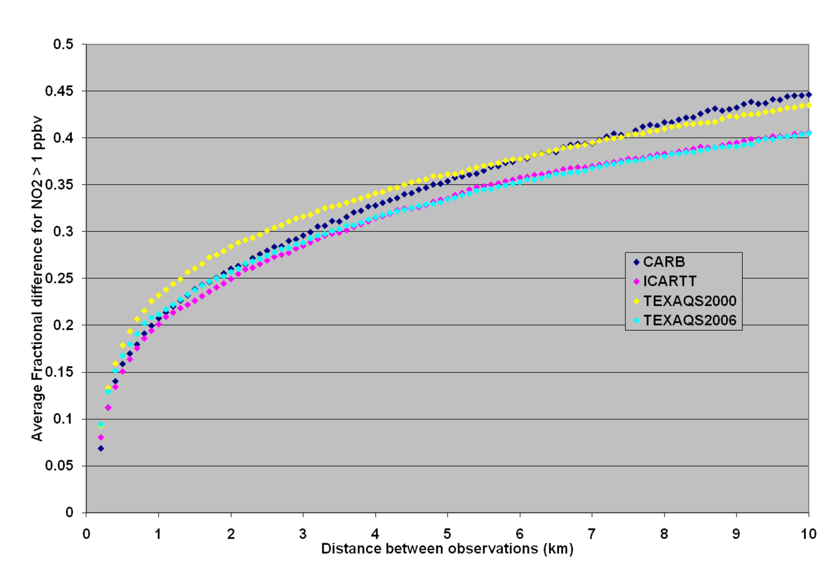
CO: The higher relative variability in the CARB data is presumably due to the higher background associated with wildfire emissions.



Ozone: CARB continues to stand out even when data is normalized. Is this related to background?



NO2 variability at GEO-CAPE scales: ***Normalized NO2 variograms suggest that the largest gradients exist at scales of 3 km and that expected gradients double from 3 to 10 km***



Interesting Behavior for ratioed variograms: ***When taking the ratio of O3 and CO variograms, no significant gradient exists for distances greater than ~10km. Values range from 0.25 to 0.5 and bracket the often cited O3:CO ratio of 0.3 expected from aged anthropogenic emissions. It is interesting that TEXAQS ratios are the same even though actual variograms have different magnitudes. While more analysis is needed, one might presume that CARB values are low due to fire influence on CO which is associated with less NOx and subsequent O3 production. Similarly, Houston may be high due to petrochemical emissions leading to an uncharacteristically high O3:CO ratio. It is also interesting to note the inverse behavior in the CARB ratio over short 0-3 km distances where plume gradients dominate.***