

Supporting Information

**Identifying and Quantifying Source Contributions of
Air Quality Contaminants during Unconventional
Shale Gas Extraction**

Nur H. Orak^{1}, Matthew D. Reeder^{2,3}, Natalie J. Pekney³*

¹ *Corresponding author: Tel: +90-216-4140545, nho@alumni.cmu.edu
Marmara University, Department of Environmental Engineering

² Leidos Research Support Team, National Energy Technology Laboratory, Pittsburgh, PA

³U.S. Dept. of Energy National Energy Technology Laboratory, Pittsburgh, PA

Number of pages: 11

Number of Figures: 6

Number of Tables: 3

Structure map of the Marcellus formation

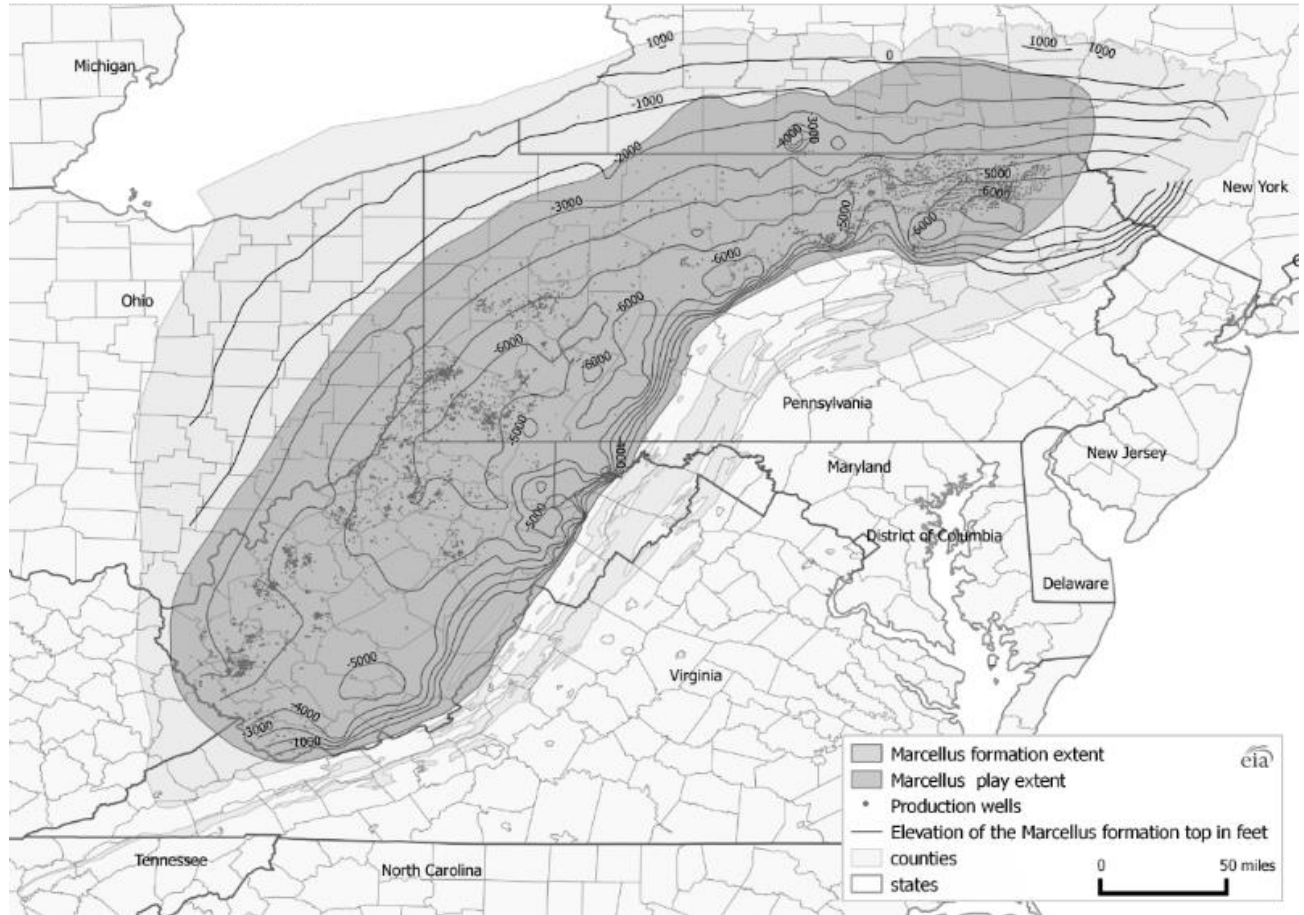


Figure S1. Structure map of the Marcellus formation (EIA 2015a).

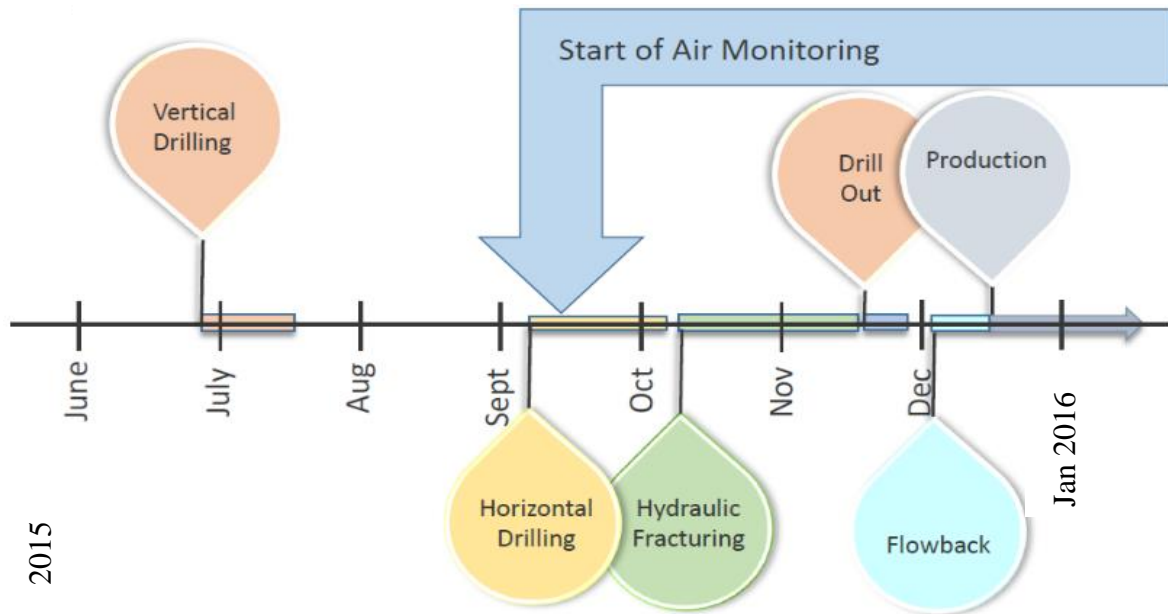


Figure S2. Unconventional natural gas production process activity diagram.

Table S1. List of Measured VOCs.

| | | |
|----------|----------|----------|
| Compound | Compound | Compound |
|----------|----------|----------|

| | | |
|--------------------|---------------------|------------------------|
| Acetylene | Trans-2-pentene | Methylcyclohexane |
| n-Butane | Propane | Methylcyclopentane |
| 1-Butene | Propylene | 2-Methylheptane |
| Cis-2-butene | Benzene | 3-Methylheptane |
| Trans-2-butene | Cyclohexane | 2-Methylhexane |
| Cyclopentane | n-Decane | 3-Methylhexane |
| 2,2-Dimethylbutane | m-Diethylbenzene | n-Nonane |
| 2,3-Dimethylbutane | p-Diethylbenzene | n-Propylbenzene |
| Ethane | 2,3-Dimethylpentane | Styrene |
| Ethylene | 2,4-Dimethylpentane | Toluene |
| 1-Hexene | n-Dodecane | 1,2,3-Trimethylbenzene |
| Isobutane | Ethyl Benzene | 1,2,4-Trimethylbenzene |
| Isopentane | o-Ethyltoluene | 1,3,5-Trimethylbenzene |
| Isoprene | m-Ethyltoluene | 2,2,4-Trimethylpentane |
| n-Pentane | p-Ethyltoluene | 2,3,4-Trimethylpentane |
| 1-Pentene | n-Heptane | n-Undecane |
| 2-Methylpentane | n-Hexane | o-Xylene |
| 3-Methylpentane | Isopropylbenzene | m-Xylene |
| Cis-2-pentene | n-Octane | p-Xylene |

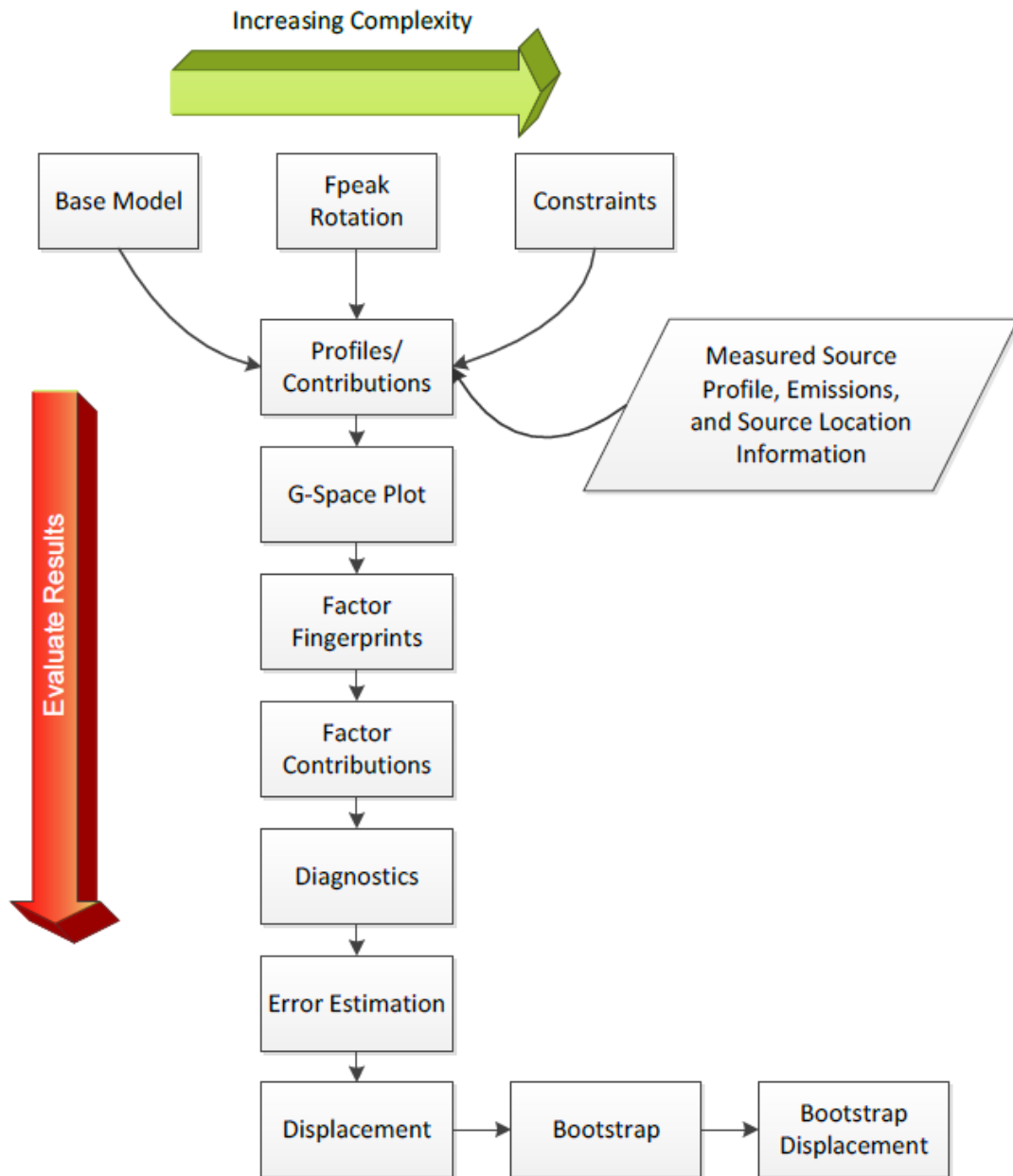


Figure S3. PMF Evaluation Process, adapted from Norris and Duvall 2014.

Table S2. Evaluation of PMF solutions for Drilling through production phases.

| Diagnostic | 2 factors | 3 factors | 4 factors |
|---|-----------|-----------|-----------|
| Q_{expected} | 45,594 | 42,319 | 39,044 |
| Q | 561,925 | 354,642 | 213,876 |
| Q_{robust} | 353,713 | 248,477 | 161,946 |
| $Q_{\text{robust}}/Q_{\text{expected}}$ | 12.3 | 5.9 | 5.5 |
| DISP %dQ | -0.001 | 0 | -0.015 |
| DISP swaps | 0 | 0 | 0 |
| Factors with BS mapping | 99% | 95% | 74% |
| BS-DISP % cases with swaps | 0 | 0 | 3 |

The base run was automatically selected by the program based on the lowest Q_{robust} . Since finding a rotationally unique solution is rare, it is acceptable to observe an increasing Q_{value} due to the F_{peak} rotation with a less than 5% change in Q (dQ).

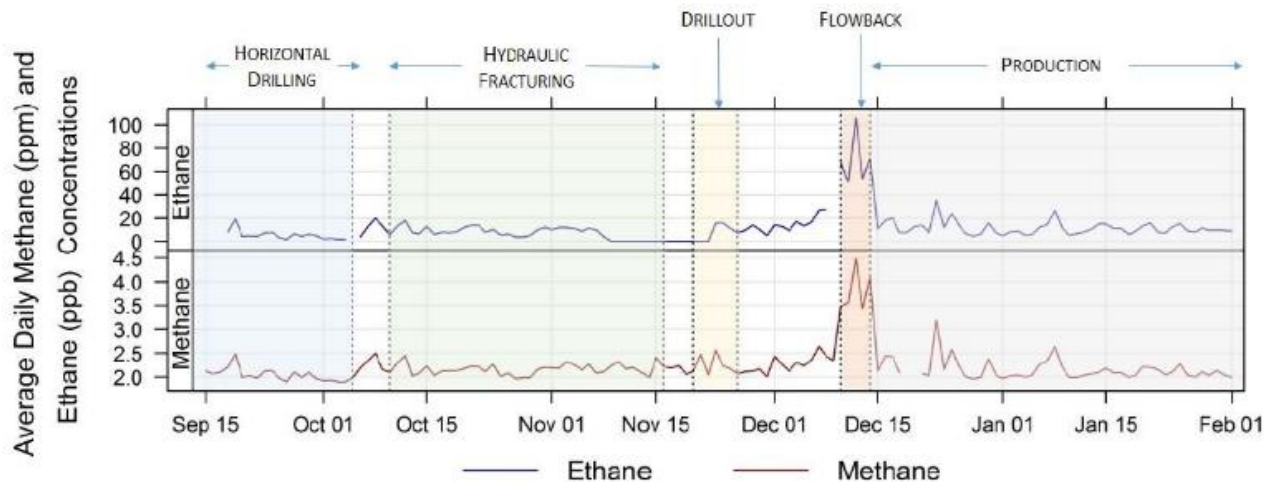


Figure S4. Ethane and methane 24-hour average concentrations through different phases of well pad activity.

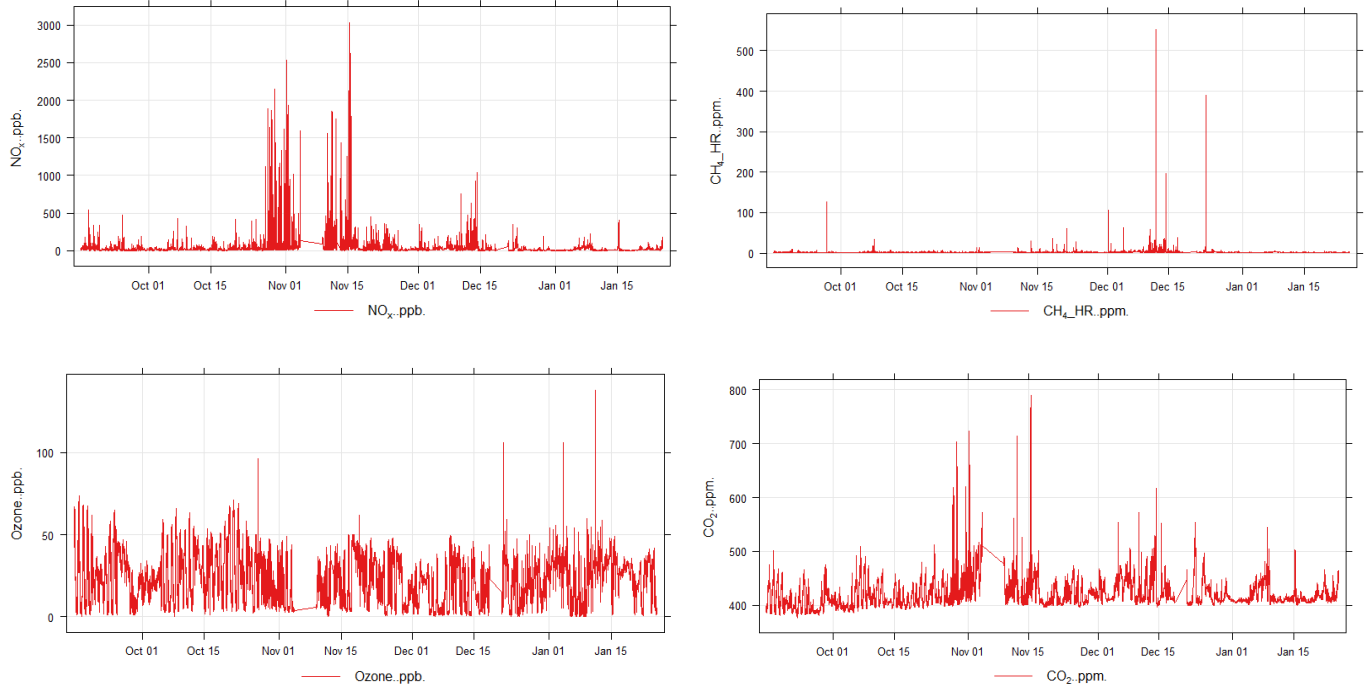


Figure S5. NO_x , CH_4 , O_3 , and CO_2 concentration time series at well pad development site.

Table S3. F_{peak} Summary.

| Drilling through Production Phases | | | | | |
|---|--------------------|-------------------|----------------------|----------------|----------|
| Strength | dQ (Robust) | Q (Robust) | % dQ (Robust) | Q (Aux) | Q |
| -1.0 | 10.0 | 23047 | 0 | 7.7 | 329063 |
| -0.5 | 5.0 | 230042 | 0 | 3.7 | 329064 |
| 0.5 | 4.0 | 230041 | 0 | 4.5 | 329066 |
| 1.0 | 6.0 | 230043 | 0 | 7.4 | 329068 |

For both time period results, Q values did not vary significantly with F_{peak} values of -1.0, -0.5, 0.5, and 1.0); therefore, we can consider all four model results for evaluation. Factor profiles and contributions were examined to determine the impact of the rotation by comparing to the base run results. As a result, for baseline conditions some optimization is gained using an F_{peak} of 1.0. There is a small deviation in species for the three factors. Furthermore, F_{peak} -rotated factor fingerprints were compared with the base model outputs. The optimized distribution of pollutants in the three factors provides more interpretable source profiles with respect to marker species. Also, there is small improvement with the source profiles for the well pad drilling through production phases.

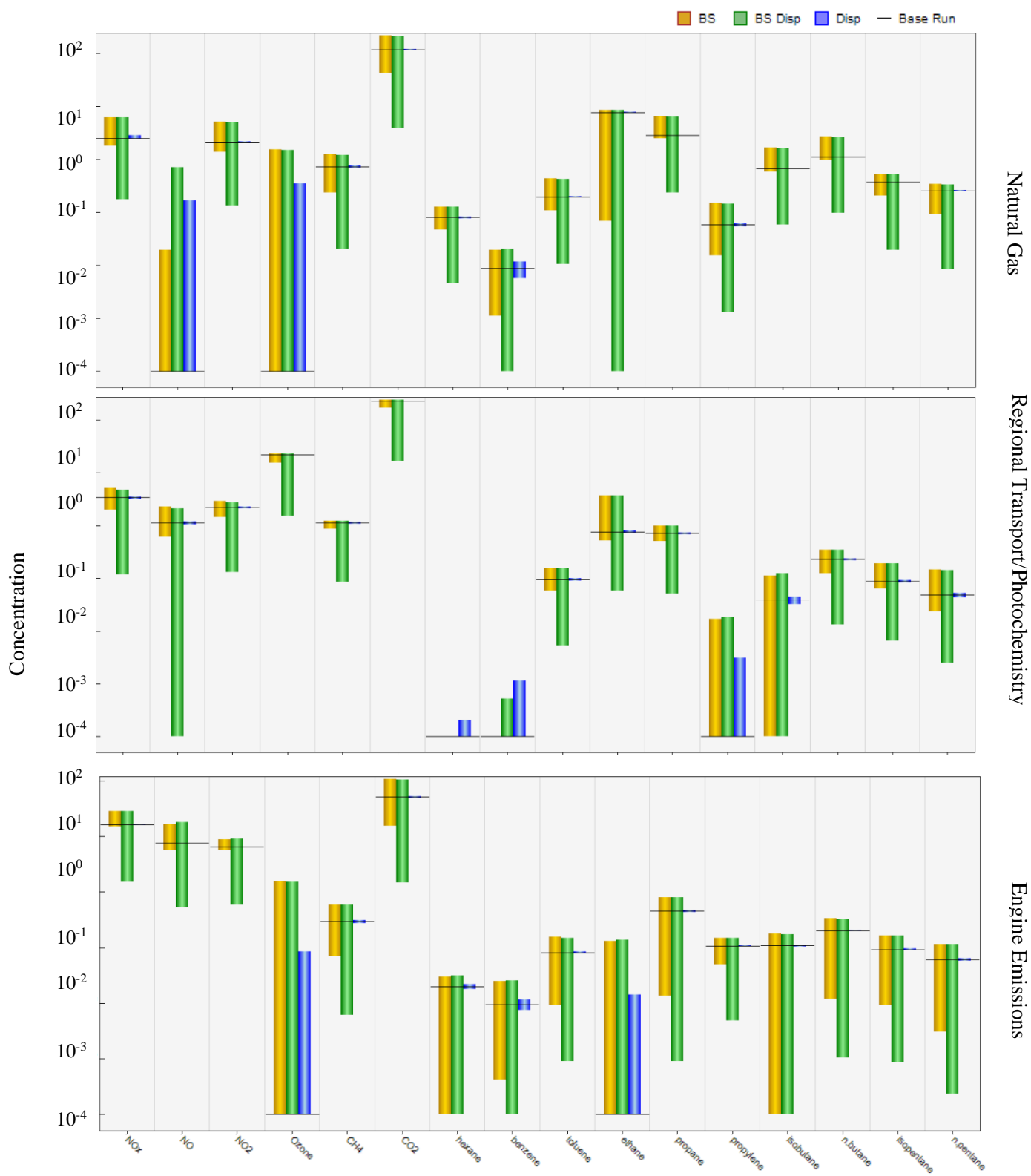


Figure S6. Error estimation summary plot of range of concentration by pollutants in each factor, active phase.

Methods/rationale for uncertainty calculations

All parameters have instrument error (2 x detection limit) as the base uncertainty.

Uncertainty is added based on the number of measurements included in the average (as standard error). For the TEOM and VOCs data, where only two measurements are used and the standard error becomes “range/2”, an additional factor is calculated to account for the time-weighted averages.

In special cases, where there is missing data before or after a given measurement (as a result of instrument malfunction, power failure, etc), an additional uncertainty is added based on the number of minutes of available data relative to the total number of minutes possible.

TEOM

Instrument error (2*detection limit) = 0.2

+

Standard error (for n=2)

$$= \text{range}/2$$

...multiplied by a factor to account for the averages being weighted based on the number of minutes at each concentration (quadratic equation):

$$= (\text{range}/2) * (ax^2 + bx + c)$$

...where x is the number of minutes for the measurement within the target hour (for example: for an original data point at “02:41”, x = 41) and a,b,c are constants that were calculated based on the following criteria:

-Maximum added uncertainty for measurements at x=30 (“02:30” “03:30” etc)

-No added uncertainty for measurements at x=0/60 (“02:00” “03:00” etc)

+

Factor related to the number of minutes used in the average:

60 minute sample = no added uncertainty

Largest increase in uncertainty for samples where the adjacent sample is missing (before or after):

$$= ((60/\text{minutes used}) - 1) * (\text{hourly averaged concentration}^{0.5})$$

VOCs

Instrument error (2*detection limit) = 0.2

+

Standard error (for n=2)

$$= \text{range}/2$$

...multiplied by a factor to account for the averages being weighted based on the number of minutes at each concentration (quadratic equation):

$$= (\text{range}/2) * (ax^2 + bx + c)$$

...where x is the number of minutes for the measurement within the target hour (for example: for an original data point at "02:41", x = 41) and a,b,c are constants that were calculated based on the following criteria:

-Maximum added uncertainty for measurements at x=30 ("02:30" "03:30" etc)

-No added uncertainty for measurements at x=0/60 ("02:00" "03:00" etc)

+

Factor related to the number of minutes used in the average:

40 minute sample (sample collection starts and is injected within the same hour) = no added uncertainty

35 minute sample (start and injection in consecutive hours) – slight increase in uncertainty

Largest increase in uncertainty for samples where the adjacent sample is missing (before or after):

$$((40/\text{minutes used})-1)*(\text{hourly averaged concentration}^{0.5})$$

NO/NO2/Ozone/SO2

Instrument error (2*detection limit): NO/NO2=0.1 Ozone= 0.8

+

Standard error

$$=\text{standard deviation}/(\text{number of measurements}^{0.5})$$

+

Number of measurements in average, relative to the maximum available (42)

$$=((42/\text{number of measurements})-1)*(\text{hour averaged concentration}^{0.5})$$