README file for the dataset associated with the research article titled: Seasonality and source apportionment of non-methane volatile organic compounds at Boulder Reservoir between 2017 and 2019

Research Article Authors: Ilana B. Pollack[1], Detlev Helmig[2], Katelyn O'Dell[1], and Emily V. Fischer[1]

Author Affiliations: [1]Colorado State University, [2]Boulder Air LLC, Boulder

Research Article Abstract:

Multiyear observations of 13 non-methane volatile organic compounds (NMVOCs) were collected at Boulder Reservoir in the Colorado Northern Front Range Metropolitan Area (NFRMA). We separate abundances of NMVOCs into source contributions using two approaches that have been applied to prior NFRMA datasets. Positive matrix factorization (PMF) analysis identifies five NMVOC factors in winter, spring, and fall that correspond to long-lived and short-lived NMVOCs from regional oil and gas production, traffic, local shorter-lived alkenes, and regional anthropogenic background. In summer, there is an additional biogenic NMVOC factor dominated by isoprene. Over all seasons, the PMF model indicates that 57-84% (range includes uncertainties) of C_2-C_5 alkanes are attributed to oil and natural gas activities. Ethyne is largely from traffic with contributions ranging from 45±6% in winter to 75±21% in summer. Ethene and propene are associated with a potentially separate source of shorter-lived alkenes that we cannot identify. The largest contributing sectors to the observed hazardous air pollutants (HAPs) differ substantially by species and season. Benzene is attributed to oil and natural gas production, traffic and other industrial activities. Toluene is predominantly attributed to regional anthropogenic activities in all seasons. Of the HAPs quantified in this dataset, hexane stands out as largely attributed to oil and natural gas production. The largest differences in the attribution of sources using the two different approaches are for benzene and toluene. Consistent with prior analyses, this work shows that the NFRMA is more strongly influenced by oil and natural gas sources than many other U.S. urban regions.

Keywords: speciated volatile organic compounds, source apportionment, multivariate regression, positive matrix factorization, oil and natural gas production, Colorado Northern Front Range Metropolitan Area

Data contacts: Ilana B. Pollack (primary), ipollack@rams.colostate.edu, 970-491-8605
Emily V. Fischer, evf@rams.colostate.edu, 970-491-8587

Data Description:

Dataset containing 13 non-methane volatile organic compounds (NMVOCs) and meteorological parameters collected at Boulder Reservoir (40.0704 N, 105.2198 W, 1604 m asl), located in Boulder, Colorado between 4 April 2017 and 31 December 2019. Data are reported at the local mid time that the NMVOC sample was collected which is reported on Mountain Standard Time (MST). A flag for smoke impacted samples is also included.

Format of Data Files:

Data are contained in a single, comma-delimited CSV file with 17 columns. The date string is in the format MM/DD/YYYY HH:MM in military time. The date and time represent the local mid time of each sample on Mountain Standard Time (MST). Mixing ratios for the 13 NMVOCs are reported as 3-decimal floating-point number for each species; units are in parts per billion by volume (ppb). Also included are columns for a smoke impacted sample flag. A smoke flag of 0 indicates that the sample was smoke free; a value of 1 indicates that the sample was smoke impacted. Two additional columns report the wind direction in degrees (north = 0^{o} and wind speed (in units of m s^{-1}) are averaged over the start and stop time of the VOC sample.

Detection limits:

NMVOC detection limits were estimated from the smallest, reliably identifiable peak areas and the compound carbon response factors. Determined values were usually in the 10-40 ppt range. When there was no quantifiable peak in the chromatograms, results are replaced with a value equal to half the estimated detection limit. Individual NMVOC detection limits are listed here:

Species	***MDL***	***MDL/2***
ethane	0.035	0.018

propane	0.022	0.011
i-butane	0.017	0.009
n-butane	0.016	0.008
i-pentane	0.014	0.007
n-pentane	0.014	0.007
hexane	0.014	0.007
ethyne	0.045	0.023
ethene	0.035	0.018
propene	0.023	0.012
benzene	0.012	0.006
toluene	0.014	0.007
isoprene	0.014	0.007

Please note that for odd ppb detection limits (such as 0.035 and 0.017), the value for half the detection limit is rounded to the nearest ppb (e.g., half of 0.017 ppb is replaced with 0.009 ppb).

Data Filtering:

The monitoring site can be influenced by occasional local pollution events, such as from cars driving to the site or idling at the turnaround of the access road. We have also on a few occasions encountered emissions from practice drills at the nearby Boulder County Regional Fire Training Center. The data were not filtered for any such samples.

Those interested in using these data are encouraged to contact Dr. Ilana B. Pollack (ipollack@rams.colostate.edu) and Dr. Emily V. Fischer (evf@rams.colostate.edu) for more information.

Funding for this work was provided by Boulder County.