Methanem Comparison

Jerry Wu

2025-01-09

```
# Load in trailer Methane
Trailer <- readRDS('TrailerProcessed-20240601.rds')

trailer_comps_wind <- Trailer %>% select(time_utc, ch4, co, co2_ppm, nox, benzene, wdr_deg, wsp_ms) %>%
  mutate(day = as.Date(format(as.POSIXct(time_utc), '%Y-%m-%d'))) %>%
  rename('co2' = 'co2_ppm')
```

```
# Load in VNF data
vnf <- readRDS('pb-vnf_20230501-20240601.rds')

vnf <- vnf %>%
  mutate(across(where(is.numeric),  ~ na_if(.x, 999999))) %>% # replace 999999 as missing
  filter(!(is.na(temp_bb) | is.na(methane_eq))) # keep those not missing temperature

vnf <- vnf %>%
  filter(temp_bb >= 1600)

loving_lonlat <- c(-104.1089, 32.2961)
distance_km_lov <- function(long, lati){
  start <- c(long, lati)
  distGeo(start, loving_lonlat) / 1000
}

vnf <- vnf %>%
  mutate(distToLovi = mapply(distance_km_lov, lon, lat))
```

```
# Preprocessing
vnf_200km <- vnf %>%
  filter(distToLovi <= 200)

radius <- c(5, 10, 20, 50, 100)

trailer_compounds <- c('ch4', 'co', 'co2', 'nox', 'benzene')

# Compute daily average
trailer_daily <- trailer_comps_wind %>%
  select(-time_utc) %>%
  group_by(day) %>%
  summarise(across(!wdr_deg, ~mean(.x, na.rm=T)),
            wdr_deg = as.numeric(mean(circular(wdr_deg, units = "degrees"), na.rm=T))) %>%
  mutate(wdr_deg = if_else(wdr_deg < 0, wdr_deg+360, wdr_deg))
```

```
## Warning: There were 2 warnings in `summarise()`.
```

```
## The first warning was:
## i In argument: `wdr_deg = as.numeric(mean(circular(wdr_deg, units = "degrees"),
##   na.rm = T))`.
## i In group 47: `day = 2023-05-31`.
## Caused by warning in `mean.circular()`:
## ! No observations (at least after removing missing values)
## i Run `dplyr::last_dplyr_warnings()` to see the 1 remaining warning.
```

```r
# Compute average measurement from 6pm to 6am
trailer_night_avg <- trailer_comps_wind %>%
  filter(hour(ymd_hms(time_utc)) <= 6 ) %>%
  select(-time_utc) %>%
  group_by(day) %>%
  summarise(across(everything(), ~mean(.x, na.rm=T))) %>%
  mutate(wdr_deg = if_else(wdr_deg < 0, wdr_deg+360, wdr_deg))

# Compute flare angle
angles <- tibble(st_sfc(st_point(loving_lonlat), crs = 4326),
                 vnf_200km[,c('lon', 'lat')] %>%
                   st_as_sf(coords = c('lon', 'lat')) %>%
                   st_set_crs(4326)) %>%
                pivot_longer(cols = everything()) %>%
                pull(value) %>% # extract coordinates only
                st_geod_azimuth() %>%
                set_units('degrees') %>% # convert to degrees
                drop_units()
angles <- angles[c(T, F)] # keep only odd index, valid pairs
angles <- if_else(angles < 0, angles + 360, angles)
vnf_200km$angle <- angles

corr_result <- tibble(radius = numeric(),
                      trailer_compound = character(),
                      flare_count = numeric(),
                      daily_corr = numeric(),
                      nightly_corr = numeric())

vnf_trailer_full <- tibble(date = vnf_200km %>%
    filter(distToLovi <= max(radius)) %>% pull(date) %>% unique()) %>%
    left_join(trailer_daily,
              join_by(date == day)) %>%
    left_join(trailer_night_avg,
              join_by(date == day), suffix = c('.day', '.night'))

for (r in radius) {
  # Filter for flares within radius r
  temp <- vnf_200km %>%
    filter(distToLovi <= r)

  # For those flares, get average methane_eq then join with trailer data
  # NOTE: since we have a n-to-1 mapping between flares and trailer,
  # it is difficult to get a single wind difference value for each day.
  # Instead, I will check if there exists a flare in a similar direction as wind
  flare_is_from_wd <- temp %>%
    left_join(trailer_daily,
```

```
            join_by(date == day)) %>%
    left_join(trailer_night_avg,
              join_by(date == day), suffix = c('.day', '.night')) %>%
    group_by(date) %>%
    summarise(flare_wd_day = sum(abs(angle - wdr_deg.day) <= 30),
              flare_wd_night = sum(abs(angle - wdr_deg.night) <= 30),
              flare_count = length(unique(vnf_id))) %>%
    rename(setNames(c('flare_wd_day', 'flare_wd_night', 'flare_count'), paste0(c('flare_wd_day_', 'flare

  temp <- temp %>%
    select(date, methane_eq, angle) %>%
    group_by(date) %>%
    summarise(avg_methane_eq = mean(methane_eq)) %>%
    rename(setNames('avg_methane_eq', paste0('avg_methane_eq_', r)))

  merged <- temp %>%
    left_join(trailer_daily,
              join_by(date == day)) %>%
    left_join(trailer_night_avg,
              join_by(date == day), suffix = c('.day', '.night')) %>%
    left_join(flare_is_from_wd, join_by(date)) %>%
    rename(setNames(paste0('avg_methane_eq_', r), 'ch4_vnf'))

  vnf_trailer_full <- vnf_trailer_full %>%
    left_join(temp, join_by(date)) %>%
    left_join(flare_is_from_wd, join_by(date))

  corr <- tibble(radius = r,
                 trailer_compound = trailer_compounds,
                 flare_count = merged %>%
                   pull(paste0('flare_count_', r)) %>%
                   sum(),
                 daily_corr = sapply(paste0(trailer_compounds, '.day'),
                              function(x) cor(merged$ch4_vnf, merged[[x]],
                                              use = 'complete')),
                 nightly_corr = sapply(paste0(trailer_compounds, '.night'),
                              function(x) cor(merged$ch4_vnf, merged[[x]],
                                              use = 'complete')))

  corr_result <- rbind(corr_result, corr)
}

knitr::kable(corr_result %>% arrange(trailer_compound, radius), digits = 3)
```

radius	trailer_compound	flare_count	daily_corr	nightly_corr
5	benzene	12	0.321	0.663
10	benzene	95	0.124	0.108
20	benzene	370	-0.056	-0.025
50	benzene	2746	-0.025	0.021
100	benzene	8842	0.039	0.060
5	ch4	12	0.586	0.670
10	ch4	95	0.158	0.212
20	ch4	370	-0.022	0.025

radius	trailer_compound	flare_count	daily_corr	nightly_corr
50	ch4	2746	0.002	0.000
100	ch4	8842	0.070	0.041
5	co	12	0.400	0.427
10	co	95	0.065	0.079
20	co	370	-0.027	-0.077
50	co	2746	-0.055	-0.062
100	co	8842	-0.032	-0.058
5	co2	12	0.373	0.455
10	co2	95	0.106	0.144
20	co2	370	0.020	0.032
50	co2	2746	0.045	0.048
100	co2	8842	-0.004	0.006
5	nox	12	0.564	0.561
10	nox	95	0.078	0.039
20	nox	370	-0.107	-0.098
50	nox	2746	-0.009	0.003
100	nox	8842	0.041	0.057

```
ch4_nox_5km <- lm(avg_methane_eq_5 ~ ch4.night + co.night + co2.night + nox.night + benzene.night + wsp
summary(ch4_nox_5km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_5 ~ ch4.night + co.night + co2.night +
##     nox.night + benzene.night + wsp_ms.night + flare_wd_night_5 +
##     flare_count_5, data = vnf_trailer_full)
##
## Residuals:
##         8        50        55        81       181       183       221       259
##  0.019435 -0.005276  0.001591 -0.006968 -0.007752 -0.001109 -0.003215  0.010966
##       272       302       317       362
##  0.001335 -0.013823 -0.004850  0.009665
##
## Coefficients: (1 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       5.543e-01  1.546e+00   0.359    0.738
## ch4.night        -4.207e-06  1.928e-05  -0.218    0.838
## co.night          5.829e-05  3.121e-04   0.187    0.861
## co2.night        -1.188e-03  3.668e-03  -0.324    0.762
## nox.night         1.463e-04  1.673e-04   0.874    0.431
## benzene.night     2.204e-02  3.805e-02   0.579    0.594
## wsp_ms.night     -2.328e-03  9.297e-03  -0.250    0.815
## flare_wd_night_5 -5.467e-04  2.231e-02  -0.024    0.982
## flare_count_5            NA         NA      NA       NA
##
## Residual standard error: 0.01548 on 4 degrees of freedom
##   (358 observations deleted due to missingness)
## Multiple R-squared:  0.5813, Adjusted R-squared:  -0.1513
## F-statistic: 0.7934 on 7 and 4 DF,  p-value: 0.6307
```

```
ch4_nox_10km <- lm(avg_methane_eq_10 ~ ch4.night + co.night + co2.night + nox.night + benzene.night + ws
summary(ch4_nox_10km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_10 ~ ch4.night + co.night + co2.night +
##     nox.night + benzene.night + wsp_ms.night + flare_wd_night_10 +
##     flare_count_10, data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.036735 -0.021280 -0.009354  0.010693  0.170610
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -1.121e-01  6.560e-01  -0.171    0.865
## ch4.night          2.128e-05  1.381e-05   1.541    0.128
## co.night          -3.306e-05  1.225e-04  -0.270    0.788
## co2.night          3.623e-04  1.565e-03   0.231    0.818
## nox.night         -1.737e-04  2.346e-04  -0.740    0.462
## benzene.night     -2.012e-02  1.898e-02  -1.060    0.293
## wsp_ms.night      -5.468e-03  4.187e-03  -1.306    0.196
## flare_wd_night_10  6.318e-04  1.013e-02   0.062    0.950
## flare_count_10     5.083e-03  6.879e-03   0.739    0.463
##
## Residual standard error: 0.03782 on 66 degrees of freedom
##   (295 observations deleted due to missingness)
## Multiple R-squared:  0.0881, Adjusted R-squared:  -0.02243
## F-statistic: 0.7971 on 8 and 66 DF,  p-value: 0.6072
```

```
ch4_nox_20km <- lm(avg_methane_eq_20 ~ ch4.night + co.night + co2.night + nox.night + benzene.night + w
summary(ch4_nox_20km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_20 ~ ch4.night + co.night + co2.night +
##     nox.night + benzene.night + wdr_deg.night + wsp_ms.night +
##     flare_wd_night_20 + flare_count_20, data = vnf_trailer_full)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.05508 -0.02882 -0.01084  0.01154  0.30197
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -5.235e-01  5.368e-01  -0.975   0.3307
## ch4.night          1.489e-05  1.010e-05   1.474   0.1422
## co.night          -2.064e-04  1.151e-04  -1.793   0.0745 .
## co2.night          1.480e-03  1.273e-03   1.163   0.2462
## nox.night         -2.926e-04  2.158e-04  -1.356   0.1766
## benzene.night     -9.661e-03  1.568e-02  -0.616   0.5385
## wdr_deg.night     -1.572e-04  7.772e-05  -2.022   0.0445 *
## wsp_ms.night       5.069e-04  2.963e-03   0.171   0.8644
## flare_wd_night_20 -8.752e-03  6.606e-03  -1.325   0.1867
## flare_count_20     4.709e-03  3.266e-03   1.442   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.05089 on 194 degrees of freedom
##   (166 observations deleted due to missingness)
## Multiple R-squared:  0.06007,    Adjusted R-squared:  0.01647
## F-statistic: 1.378 on 9 and 194 DF,  p-value: 0.2005
```

```
ch4_nox_50km <- lm(avg_methane_eq_50 ~ ch4.night + co.night + co2.night + nox.night + benzene.night + w
summary(ch4_nox_50km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_50 ~ ch4.night + co.night + co2.night +
##     nox.night + benzene.night + wdr_deg.night + wsp_ms.night +
##     flare_wd_night_50 + flare_count_50, data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.042946 -0.016265 -0.004685  0.010797  0.121250
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -1.844e-01  1.576e-01  -1.170   0.2430
## ch4.night         -8.374e-07  3.490e-06  -0.240   0.8105
## co.night          -4.932e-05  2.772e-05  -1.779   0.0762 .
## co2.night          6.295e-04  3.755e-04   1.676   0.0946 .
## nox.night         -7.454e-06  8.673e-05  -0.086   0.9316
## benzene.night      1.828e-03  5.798e-03   0.315   0.7527
## wdr_deg.night     -3.474e-05  3.181e-05  -1.092   0.2756
## wsp_ms.night       8.400e-04  1.071e-03   0.784   0.4335
## flare_wd_night_50 -3.015e-05  8.257e-04  -0.037   0.9709
## flare_count_50     5.611e-04  3.626e-04   1.548   0.1227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02325 on 323 degrees of freedom
##   (37 observations deleted due to missingness)
## Multiple R-squared:  0.02962,    Adjusted R-squared:  0.002581
## F-statistic: 1.095 on 9 and 323 DF,  p-value: 0.3656
```

```
ch4_nox_100km <- lm(avg_methane_eq_100 ~ ch4.night + co.night + co2.night + nox.night + benzene.night +
summary(ch4_nox_100km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_100 ~ ch4.night + co.night + co2.night +
##     nox.night + benzene.night + wsp_ms.night + flare_wd_night_100 +
##     flare_count_100, data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.039011 -0.013483 -0.003297  0.007836  0.155490
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.512e-02  1.450e-01   0.587   0.5575
## ch4.night          -2.826e-07  3.244e-06  -0.087   0.9306
```

```
## co.night           -4.034e-05  2.521e-05  -1.600   0.1104
## co2.night          -3.261e-05  3.447e-04  -0.095   0.9247
## nox.night           6.806e-05  8.204e-05   0.830   0.4073
## benzene.night       7.135e-03  5.247e-03   1.360   0.1747
## wsp_ms.night        9.815e-04  9.892e-04   0.992   0.3218
## flare_wd_night_100  2.758e-04  1.899e-04   1.452   0.1473
## flare_count_100     2.123e-04  1.056e-04   2.010   0.0451 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02217 on 350 degrees of freedom
##   (11 observations deleted due to missingness)
## Multiple R-squared:  0.04554,    Adjusted R-squared:  0.02372
## F-statistic: 2.087 on 8 and 350 DF,  p-value: 0.03635
```

```
# Regress concentration of CH4 against wind and flare
ch4_count_20km <- lm(ch4.night ~ wdr_deg.night + wsp_ms.night + flare_wd_night_20 + flare_count_20, data
summary(ch4_count_20km)
```

```
##
## Call:
## lm(formula = ch4.night ~ wdr_deg.night + wsp_ms.night + flare_wd_night_20 +
##     flare_count_20, data = vnf_trailer_full)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -775.79 -242.93  -96.54  123.05 3157.71
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3373.7595   137.1194  24.605   <2e-16 ***
## wdr_deg.night       -1.3716     0.6067  -2.261   0.0248 *
## wsp_ms.night      -174.4919    17.8219  -9.791   <2e-16 ***
## flare_wd_night_20  114.4799    52.7115   2.172   0.0310 *
## flare_count_20      31.5666    26.4048   1.195   0.2333
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 425.3 on 205 degrees of freedom
##   (160 observations deleted due to missingness)
## Multiple R-squared:  0.3816, Adjusted R-squared:  0.3696
## F-statistic: 31.63 on 4 and 205 DF,  p-value: < 2.2e-16
```

```
# Regress concentration of Benzene against wind and flare
benzene_count_20km <- lm(benzene.night ~ wdr_deg.night + wsp_ms.night + flare_wd_night_20 + flare_count
summary(benzene_count_20km)
```

```
##
## Call:
## lm(formula = benzene.night ~ wdr_deg.night + wsp_ms.night + flare_wd_night_20 +
##     flare_count_20, data = vnf_trailer_full)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.63930 -0.19088 -0.05765  0.15160  0.95451
##
```

```
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1332508  0.0898176  12.617  < 2e-16 ***
## wdr_deg.night      -0.0012480  0.0003968  -3.145  0.00191 **
## wsp_ms.night       -0.1322361  0.0116736 -11.328  < 2e-16 ***
## flare_wd_night_20   0.0849643  0.0342975   2.477  0.01406 *
## flare_count_20      0.0378353  0.0172192   2.197  0.02913 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2763 on 203 degrees of freedom
##   (162 observations deleted due to missingness)
## Multiple R-squared:  0.4618, Adjusted R-squared:  0.4512
## F-statistic: 43.55 on 4 and 203 DF,  p-value: < 2.2e-16
```