Methanem Comparison

Jerry Wu

2025-01-09

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

trailer_ch4_CO_NOx_wind <- Trailer %>% select(time_utc, ch4, co, nox, wdr_deg, wsp_ms) %>%
  mutate(day = as.Date(format(as.POSIXct(time_utc), '%Y-%m-%d')))
```

```r
# 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))
```

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

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

trailer_compounds <- c('ch4', 'co', 'nox')
```

```r
# Compute daily average
trailer_daily <- trailer_ch4_CO_NOx_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.
# Compute average measurement from 6pm to 6am
trailer_night_avg <- trailer_ch4_CO_NOx_wind %>%
  filter(hour(ymd_hms(time_utc)) <= 6 |
           hour(ymd_hms(time_utc)) >= 18) %>%
  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(),
                      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,
```

```r
                        join_by(date == day), suffix = c('.day', '.night')) %>%
    group_by(date) %>%
    summarise(flare_wd_day = any(abs(angle - wdr_deg.day) <= 15),
              flare_wd_night = any(abs(angle - wdr_deg.night) <= 15),
              flare_count = n()) %>%
    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,
                 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	daily_corr	nightly_corr
5	ch4	0.586	0.643
10	ch4	0.158	0.168
20	ch4	-0.022	0.022
50	ch4	0.002	0.004
100	ch4	0.070	0.047
5	co	0.400	0.344
10	co	0.065	0.078
20	co	-0.027	-0.052
50	co	-0.055	-0.082
100	co	-0.032	-0.082
5	nox	0.564	0.561
10	nox	0.078	0.042
20	nox	-0.107	-0.103

radius	trailer_compound	daily_corr	nightly_corr
50	nox	-0.009	-0.013
100	nox	0.041	0.044

```r
ch4_nox_5km <- lm(avg_methane_eq_5 ~ ch4.night + co.night + nox.night + wdr_deg.night + wsp_ms.night + 
summary(ch4_nox_5km)
```

```
## 
## Call:
## lm(formula = avg_methane_eq_5 ~ ch4.night + co.night + nox.night +
##     wdr_deg.night + wsp_ms.night + flare_wd_night_5 + flare_count_5,
##     data = vnf_trailer_full)
## 
## Residuals:
##         8         50         55         81        181        183        221
##  0.0110944  0.0044419  0.0102751 -0.0108310 -0.0045607 -0.0003115 -0.0033735
##       259        272        302        317        362
##  0.0076135 -0.0064150 -0.0068703 -0.0056236  0.0045607
## 
## Coefficients: (1 not defined because of singularities)
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -4.173e-02  5.132e-02  -0.813    0.453
## ch4.night          -1.098e-05  1.720e-05  -0.639    0.551
## co.night            3.345e-04  1.999e-04   1.674    0.155
## nox.night          -1.886e-04  3.662e-04  -0.515    0.629
## wdr_deg.night      -4.254e-05  1.317e-04  -0.323    0.760
## wsp_ms.night        2.090e-02  1.167e-02   1.792    0.133
## flare_wd_night_5TRUE 3.189e-03  1.150e-02   0.277    0.793
## flare_count_5             NA         NA      NA       NA
## 
## Residual standard error: 0.01092 on 5 degrees of freedom
##   (358 observations deleted due to missingness)
## Multiple R-squared:  0.7393, Adjusted R-squared:  0.4265
## F-statistic: 2.363 on 6 and 5 DF,  p-value: 0.1818
```

```r
ch4_nox_10km <- lm(avg_methane_eq_10 ~ ch4.night + co.night + nox.night + wdr_deg.night + wsp_ms.night 
summary(ch4_nox_10km)
```

```
## 
## Call:
## lm(formula = avg_methane_eq_10 ~ ch4.night + co.night + nox.night +
##     wdr_deg.night + wsp_ms.night + flare_wd_night_10 + flare_count_10,
##     data = vnf_trailer_full)
## 
## Residuals:
##      Min        1Q    Median        3Q       Max
## -0.039310 -0.020108 -0.007328  0.009765  0.172277
## 
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         6.430e-02  4.712e-02   1.365    0.177
## ch4.night           1.436e-05  1.587e-05   0.905    0.369
## co.night           -1.102e-04  1.741e-04  -0.633    0.529
## nox.night          -6.485e-05  4.298e-04  -0.151    0.881
```

```
## wdr_deg.night          8.998e-07  1.181e-04   0.008    0.994
## wsp_ms.night          -5.738e-03  3.749e-03  -1.531    0.131
## flare_wd_night_10TRUE -7.637e-04  1.499e-02  -0.051    0.960
## flare_count_10         4.941e-03  6.909e-03   0.715    0.477
##
## Residual standard error: 0.03788 on 67 degrees of freedom
##   (295 observations deleted due to missingness)
## Multiple R-squared:  0.07086,    Adjusted R-squared:  -0.02621
## F-statistic:  0.73 on 7 and 67 DF,  p-value: 0.6472
```

```r
ch4_nox_20km <- lm(avg_methane_eq_20 ~ ch4.night + co.night + nox.night + wdr_deg.night + wsp_ms.night
summary(ch4_nox_20km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_20 ~ ch4.night + co.night + nox.night +
##     wdr_deg.night + wsp_ms.night + flare_wd_night_20 + flare_count_20,
##     data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.057344 -0.029069 -0.011730  0.009176  0.298495
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            7.574e-02  4.036e-02   1.877   0.0620 .
## ch4.night              2.338e-05  1.432e-05   1.632   0.1042
## co.night              -2.042e-04  1.598e-04  -1.278   0.2027
## nox.night             -6.034e-04  3.792e-04  -1.591   0.1131
## wdr_deg.night         -1.581e-04  9.281e-05  -1.703   0.0901 .
## wsp_ms.night           2.595e-03  2.743e-03   0.946   0.3453
## flare_wd_night_20TRUE -6.354e-03  1.066e-02  -0.596   0.5518
## flare_count_20         3.543e-03  3.205e-03   1.105   0.2704
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05111 on 199 degrees of freedom
##   (163 observations deleted due to missingness)
## Multiple R-squared:  0.04321,    Adjusted R-squared:  0.009551
## F-statistic: 1.284 on 7 and 199 DF,  p-value: 0.26
```

```r
ch4_nox_50km <- lm(avg_methane_eq_50 ~ ch4.night + co.night + nox.night + wdr_deg.night + wsp_ms.night
summary(ch4_nox_50km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_50 ~ ch4.night + co.night + nox.night +
##     wdr_deg.night + wsp_ms.night + flare_wd_night_50 + flare_count_50,
##     data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.04353 -0.01552 -0.00293  0.01001  0.12096
##
## Coefficients:
```

```
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           7.577e-02  1.380e-02   5.490 8.06e-08 ***
## ch4.night             6.191e-06  4.525e-06   1.368   0.1722
## co.night             -6.451e-05  4.182e-05  -1.543   0.1238
## nox.night            -5.431e-05  1.462e-04  -0.371   0.7105
## wdr_deg.night        -7.033e-05  3.741e-05  -1.880   0.0610 .
## wsp_ms.night          1.467e-03  1.022e-03   1.436   0.1519
## flare_wd_night_50TRUE -6.057e-03  3.408e-03  -1.777   0.0765 .
## flare_count_50        6.718e-04  3.134e-04   2.143   0.0328 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02303 on 329 degrees of freedom
##   (33 observations deleted due to missingness)
## Multiple R-squared:  0.03206,    Adjusted R-squared:  0.01147
## F-statistic: 1.557 on 7 and 329 DF,  p-value: 0.1474
```

```r
ch4_nox_100km <- lm(avg_methane_eq_100 ~ ch4.night + co.night + nox.night + wdr_deg.night + wsp_ms.night
summary(ch4_nox_100km)
```

```
##
## Call:
## lm(formula = avg_methane_eq_100 ~ ch4.night + co.night + nox.night +
##     wdr_deg.night + wsp_ms.night + flare_wd_night_100 + flare_count_100,
##     data = vnf_trailer_full)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.040271 -0.013533 -0.003920  0.007869  0.160492
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            5.352e-02  1.395e-02   3.838 0.000147 ***
## ch4.night              7.771e-06  4.526e-06   1.717 0.086865 .
## co.night              -6.802e-05  4.129e-05  -1.647 0.100368
## nox.night              6.003e-05  1.480e-04   0.406 0.685302
## wdr_deg.night         -5.386e-07  4.006e-05  -0.013 0.989280
## wsp_ms.night           2.001e-03  9.805e-04   2.040 0.042050 *
## flare_wd_night_100TRUE 4.466e-03  3.665e-03   1.218 0.223896
## flare_count_100        2.606e-04  1.125e-04   2.317 0.021062 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02351 on 356 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.05216,    Adjusted R-squared:  0.03352
## F-statistic: 2.799 on 7 and 356 DF,  p-value: 0.007571
```

```r
# 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
## -485.39 -184.50  -72.31  131.76 1825.16
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          2784.4756   106.1110  26.241  < 2e-16 ***
## wdr_deg.night          -0.2343     0.5425  -0.432   0.6662
## wsp_ms.night         -103.8084    14.9286  -6.954 4.62e-11 ***
## flare_wd_night_20TRUE  123.1431    64.4082   1.912   0.0573 .
## flare_count_20          15.5806    19.3256   0.806   0.4210
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 314.1 on 206 degrees of freedom
##   (159 observations deleted due to missingness)
## Multiple R-squared:  0.2578, Adjusted R-squared:  0.2434
## F-statistic: 17.89 on 4 and 206 DF,  p-value: 1.273e-12
```