

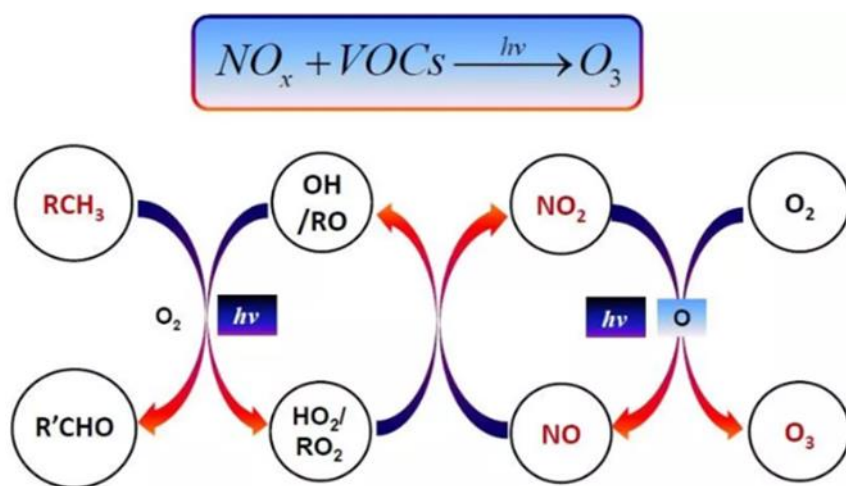
我国夏季对流层甲醛、臭氧与二氧化氮的分布与浓度关系分析

12032777 侯岳

一、研究背景

本项研究的研究对象为对流层三种重要的污染物臭氧、二氧化氮与甲醛。臭氧（O₃）是在一定的气象条件下，由大气中的挥发性有机化合物（VOC）、氮氧化物（NO_x）等通过一系列光化学反应生成的二次污染物，损害人体呼吸系统，影响植物生长；甲醛（HCHO）是对流层臭氧和二次有机气溶胶的重要前体物，一次来源包括室内装修材料的释放、农田排放、化石燃料的不完全燃烧等，二次来源为挥发性有机化合物在大气中经氧化作用生成，如甲烷、植物排放的异戊二烯等其丰度一定程度上代表 VOC 的浓度，甲醛会刺激人的眼睛、鼻子，引起呼吸系统疾病和癌症等一系列健康问题，被国际研究机构列为 I 类致癌物；二氧化氮（NO₂）的主要来源有汽车尾气、工业生产中锅炉的废气排放等，在大气中与 NO 互相转换，在臭氧的形成过程中具有重要作用，同时也是造成光化学污染和酸雨的重要因素。

三种污染物之间的反应机理可以由图一展示，臭氧参与大气中氮氧化物的互相转换，将 NO 氧化为 NO₂，但 NO₂ 会光解产生 O 原子和 O₂ 生成 O₃，转换过程无臭氧的净生成，但在 VOC 的氧化过程中生成的 HO₂、RO₂ 等自由基会抢先氧化 NO，导致消耗的臭氧减少，例如甲醛的生成和氧化过程都会产生自由基。其总体效果相当于 NO_x 和 VOC 在光照的条件下生成臭氧。甲醛和二氧化氮则为臭氧生成的前体物。



图一 O₃、VOC 与 NO_x 的反应

对于甲醛等污染物的观测，已有 GOME（搭载于 1996 年发射的 ERS2 卫星），SCIAMACHY（搭载于 2002 年发射的 ENVISAT 卫星），GOME-2A（搭载于 2007 年发射的 METOP1 卫星和 2013 年发射的 METOP-B 卫星），OMI（搭载于 2004 年发射的 AURA 卫星）等卫星探测器的发布产品，但并未达到很高的分辨率。

二、数据来源与研究目的

在这次研究中，我用到的数据是中国科学技术大学刘诚老师课题组利用 TROPOMI 卫星传感器反演的对流层 HCHO 与 NO₂ 的柱浓度产品，时间选取为 2019 年的 7 月，相对于以往的卫星探测器，TROPOMI 探测器具有更高的数据分辨率。，因此本项研究希望通过 TROPOMI 探测器的数据产品，将其可视化并观察中国夏季对流层甲醛与二氧化氮的分布特征，并评估卫星数据的质量。

与此同时，我还在中国环境监测总站的全国城市空气质量实时发布平台获取同时间段近地面臭氧浓度数据，一方面观察其分布特征，另一方面寻找对流层甲醛和二氧化氮与近地面臭氧浓度之间的关系，初步建立它们之间的线性关系，为通过甲醛和二氧化氮的卫星观测反演近地面臭氧浓度提供便利。

三、数据处理与可视化

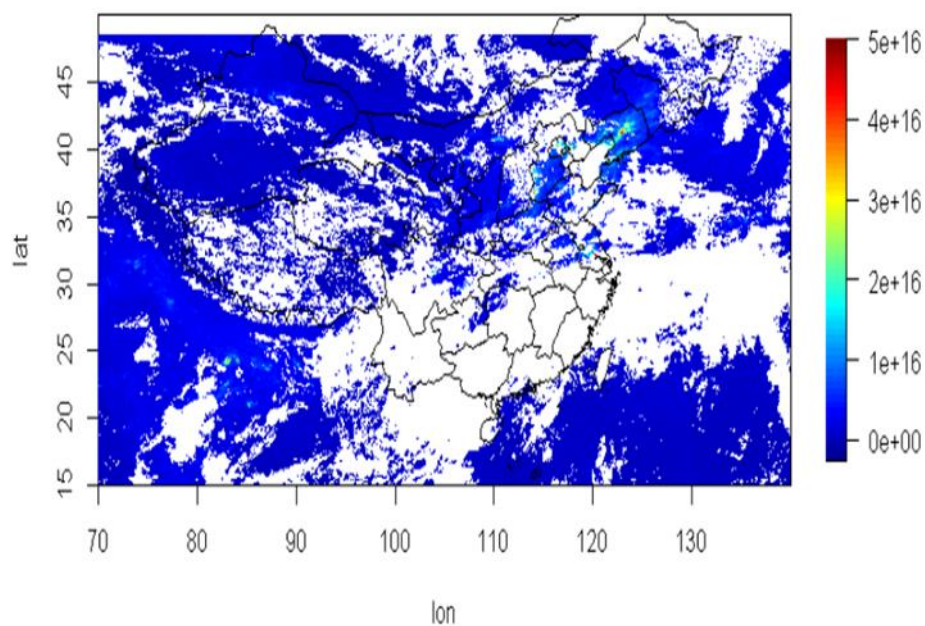
卫星数据文件为 nc 格式，数据量十分庞大。其中 NO₂ 每日一个数据产品，已经进行网格化处理，HCHO 则是每个扫描轨道一个产品，未经过网格化处理。产品包含每个像素点的对流层垂直柱浓度信息，并已经经过了去噪处理，因此部分像素点的数据缺失，研究的难点是数据的整理和可视化。

首先直接对单个文件的数据进行读取，并用 image.plot 函数直接作图，代码截图如图二所示：

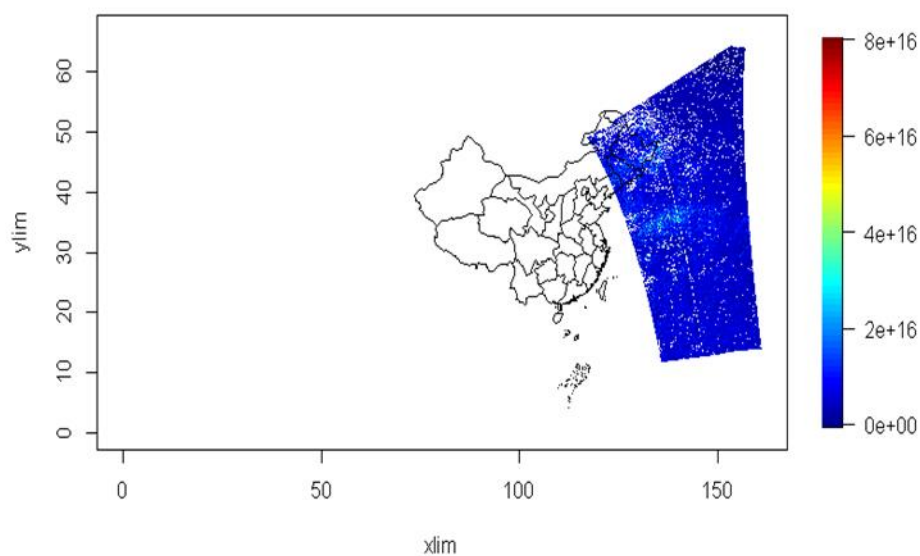
```
setwd("G:/TROPOMI_NO2")
ex.nc <- open.nc("TROPOMI_NO2_VCD_2019-06-01.nc")
print.nc(ex.nc)
lat <- var.get.nc(ex.nc,"lat")
lon <- var.get.nc(ex.nc,"lon")
vcd <- var.get.nc(ex.nc,"tropo_vcd")
close.nc(ex.nc)
image.plot(lon,lat,vcd)
map('china',add=T,lwd=0.75,col="black")
setwd("G:/TROPOMI_HCHO")
ex.nc2 <- open.nc("TRHCHO_20190602T024059_20190602T042229_08466_01_010000_20190602T060441.he5")
print.nc(ex.nc2)
lati <- var.get.nc(ex.nc2,"latitude")
long <- var.get.nc(ex.nc2,"longitude")
Tvcd <- var.get.nc(ex.nc2,"TropoVCD")
lati[which(lati<0)] <- 0
long[which(long<0)] <- 0
Tvcd[which(Tvcd<0)] <- 0
image.plot(long,lati,Tvcd)
map('china',add=T,lwd=0.75,col="black")
```

图二 部分代码

得到单日（或单个轨道）的 NO₂ 与甲醛的浓度分布图如图三和图四所示：



图三 NO2 单日分布图



图四 HCHO 单次轨道扫描分布图

通过图三和图四可见单日或单次扫描都不能很好的显示污染物在我国分布情况，前者数据缺失太严重，后者则是无法覆盖到整个中国范围，且在覆盖范围内也有数据缺失，因此我选择对数据进行月平均处理。对于 NO2，代码如下：

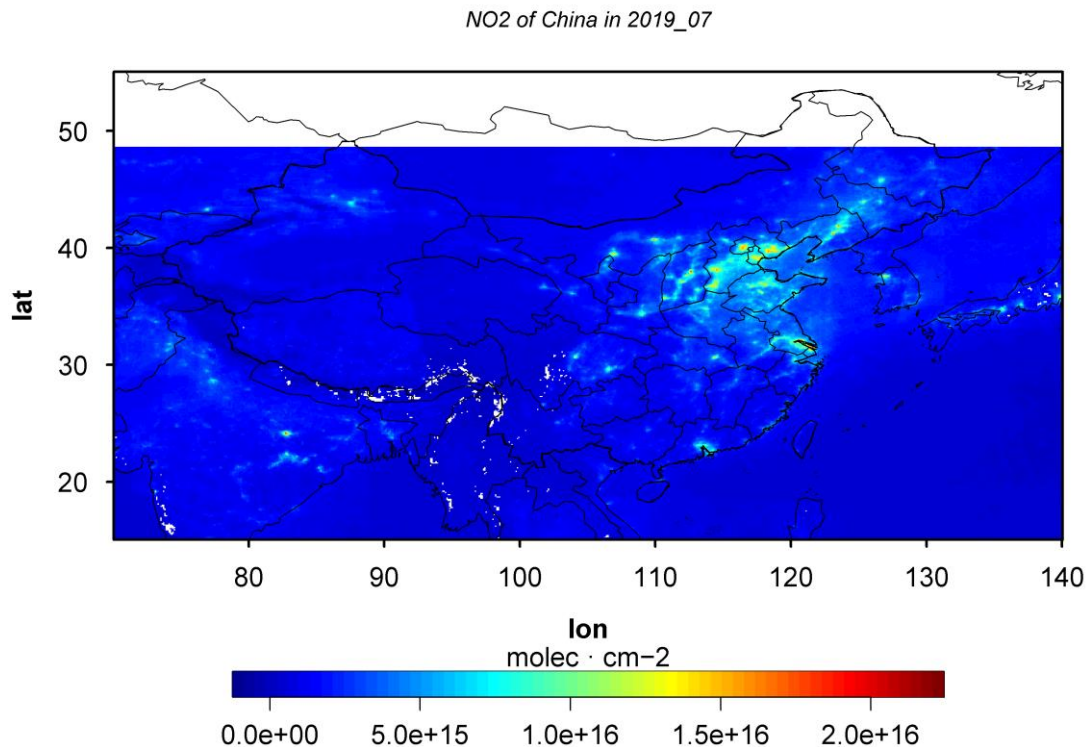
```

#读取文件
temp_07 <- list.files(pattern="TROPOMI_NO2_VCD_2019-07.*")
#设置经纬度网格
latt <- c(15:55)
longg <- c(70:140)
latt2 <- latt/10
longg2 <- longg/10
#读取每个文件的NO2值，并且设置经纬度分辨率为0.1，并求平均值
no2_july <- array(0,dim=c(31,700,400))
for (k in 1:31) {
  ex.nc <- open.nc(temp_07[k])
  lat <- var.get.nc(ex.nc,"lat")
  lon <- var.get.nc(ex.nc,"lon")
  vcd <- var.get.nc(ex.nc,"tropo_vcd")
  close.nc(ex.nc)
  for (i in 1:700) {
    for (j in 1:400) {
      no2_july[k,i,j] <- mean(vcd[which(lon>=70+(i-1)/10 & lon<70+i/10),
                             which(lat>=15+(j-1)/10 & lat<15+j/10)],na.rm=T)
    }
  }
}
no2_mean <-matrix(0,700,400)
for (i in 1:700){
  for(j in 1:400){
    no2_mean[i,j]=mean(no2_july[,i,j],na.rm=T)
  }
}
}

```

图五 部分代码

首先创立 700*400 的网格，对应经度 70° - 140° ，纬度 15° - 55° ，以 0.1° 为分辨率，对于每个像素格子，筛选出经纬度符合再进行月份平均，并做图，得出 2019 年 7 月的 NO₂ 分布图：



图六 2019 年 7 月 NO₂ 月平均分布图

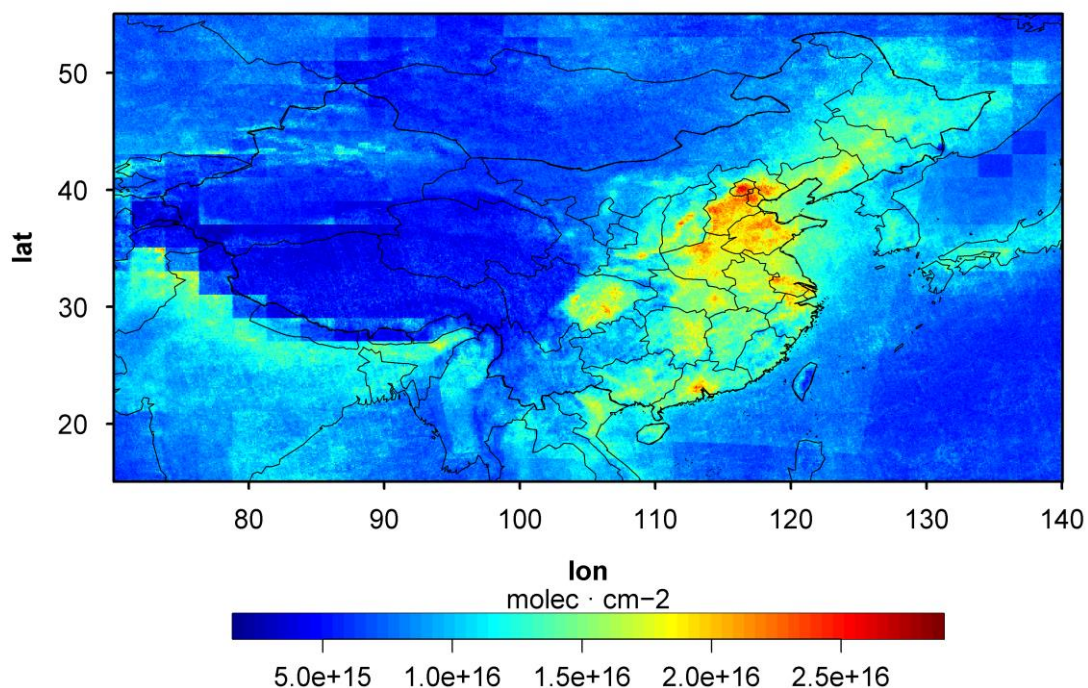
可以看到我国 NO₂ 大部分地区浓度较低，但是华北地区、长三角、珠三角、东北地区的 NO₂ 浓度较高，尤其是京津冀地区，是 NO₂ 的重灾区。

然后对 HCHO 数据也进行月平均，但是由于原始 HCHO 数据未进行网格化，因此在月平均处理上和 NO₂ 有所不同，在处理 NO₂ 数据时，对于特定的网格点，去寻找坐落在此网格点的数据；而在处理 HCHO 时，则是对所有数据点进行扫描，将其划分到所属的网格点上，相比于之前的方法极大提高了运算效率（我曾按照处理 NO₂ 的思路来处理 HCHO 但是程序运行了两天都没有结果，最终改变思路后 10 分钟就可以运算出来）。关键代码如图七所示：

```
#读取文件
tempHCHO_07 <- list.files(pattern="TRHCHO_201907.*")
#读取每个文件的HCHO值，并且设置经纬度分辨率为0.1，并求平均值
hcho_july <- array(0,dim=c(700,400))
hcho_n <- matrix(0,700,400)
hcho_mean <- matrix(0,700,400)
for (k in 1:125) {
  ex.nc <- open.nc(tempHCHO_07[k])
  lati <- var.get.nc(ex.nc,"latitude")
  long <- var.get.nc(ex.nc,"longitude")
  Tvcd <- var.get.nc(ex.nc,"TropoVCD")
  close.nc(ex.nc)
  for (i in 1:nrow(Tvcd)){
    for (j in 1:ncol(Tvcd)){
      if(lati[i,j]>=15 & lati[i,j]<55 & long[i,j] >=70 &
        long[i,j]<140 & !is.na(Tvcd[i,j])){
        c=ceiling(lati[i,j]*10)-150
        d=ceiling(long[i,j]*10)-700
        hcho_july[d,c]<- hcho_july[d,c]+Tvcd[i,j]
        hcho_n[d,c]<- hcho_n[d,c]+1
      }
    }
  }
}
for (i in 1:700){
  for (j in 1:400){
    hcho_mean[i,j]<-hcho_july[i,j]/hcho_n[i,j]
  }
}
}__
```

图七 部分代码

然后得到 2019 年 7 月甲醛的分布图如图八所示：

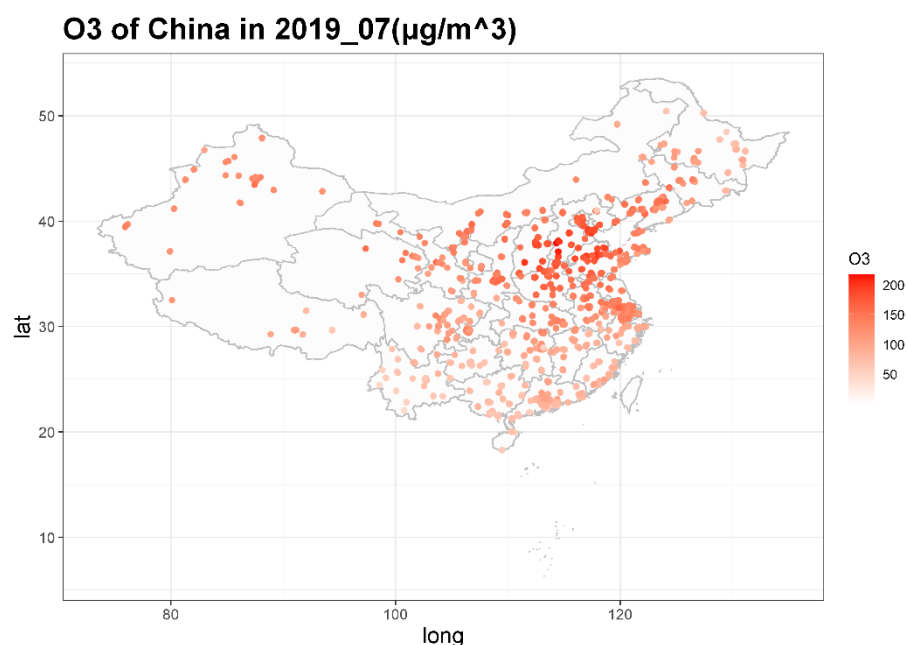


图八 2019 年 7 月 HCHO 分布图

从图中可以看到甲醛的空间分布特征与二氧化氮类似，均呈现东部高西部低的趋势，但与二氧化氮不同的是我国东部地区夏季甲醛浓度整体偏高，不像 NO₂ 那样只分布城市附近。

总体上看，TROPOMI 的数据成像质量较高，可以达到很高的空间分辨率，因此可以用来研究小尺度范围内空气质量的空间分布。但是单日数据误差较大，例如 NO₂ 出现了大面积的数据缺失情况，因此可能不适合用来研究某一地点的污染物浓度时间序列。

而对于地面观测站点的臭氧浓度数据，原数据为 csv 格式，我提取出了一千多个站点每个站点的每日最大八小时滑动平均浓度作为当日的 O3 浓度，将数据重新整理，取得月平均，并利用 ggplot 函数进行画图，得到地面站点的 O3 浓度散点图（由于是单站数据，因此没有画填充图）。如图九所示：



图九 2019 年 7 月臭氧分布散点图

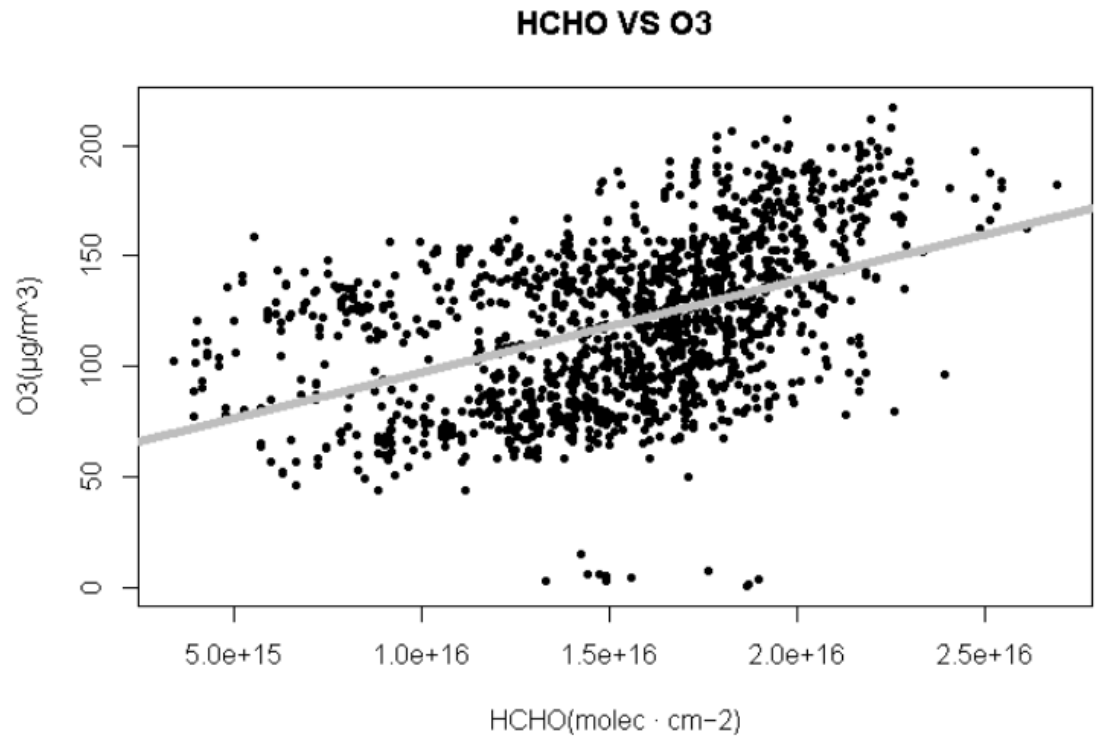
从图中可以看到华北地区臭氧浓度要远高于其他地区。总体上看，和甲醛及二氧化氮的分布情况类似。

四、数据对比与分析

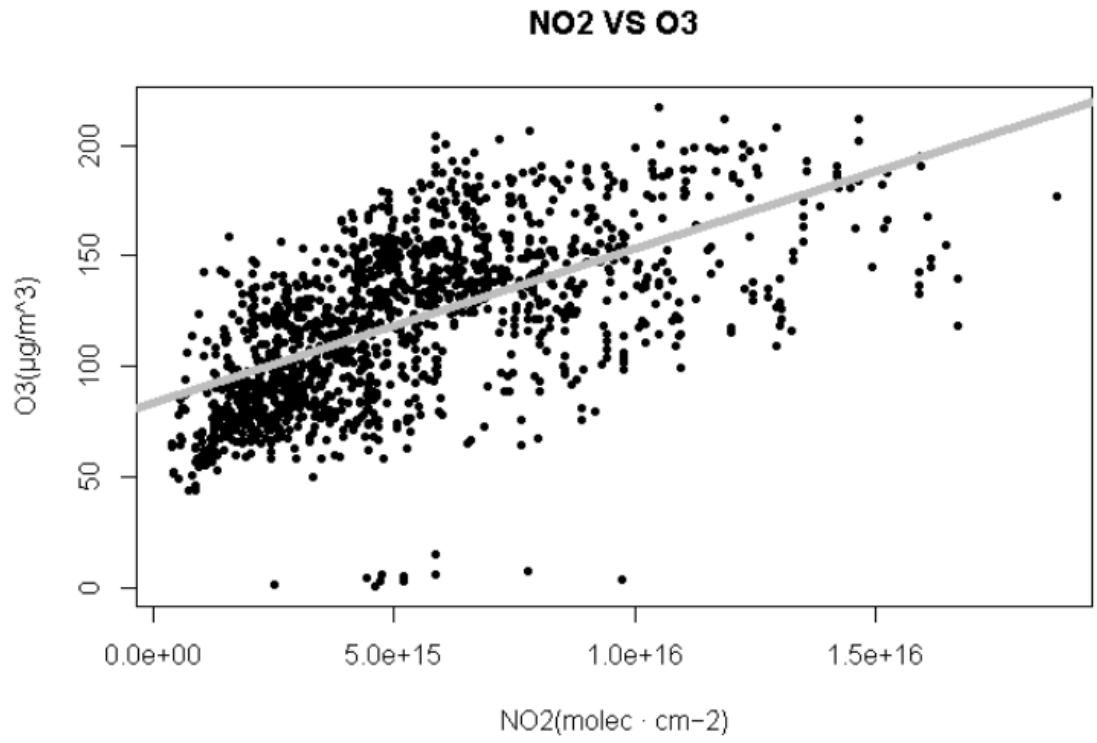
在得到了甲醛、二氧化氮的卫星观测数据和臭氧的地面观测数据后，利用地面观测站点所在的单个网格的柱浓度平均值作为该站点卫星观测柱浓度，以地面臭氧浓度为因变量，对流层甲醛、二氧化氮柱浓度为自变量，进行相关性检验，并进行简单的线性拟合，寻找臭氧对甲醛和二氧化氮的依赖性。

首先是对数据的合并，读取各个站点的经纬度，并寻找到该站点在 HCHO 与 NO2 数据矩阵中所在的网格坐标并读取 NO2 与 HCHO 数据，最终将得到的 HCHO 与 NO2 浓度数组和 O3 浓度合并为一个 data frame。

首先,对全国范围内的一千多个站点,分别画了 O3 关于 NO2、O3 关于 HCHO 的散点图以观察它们之间的相关性, 如图十、图十一所示。



图十 臭氧-甲醛散点图



图十一 臭氧-二氧化氮散点图

从两个散点图中可以看到，臭氧和甲醛、臭氧和二氧化氮均呈较为明显的正相关关系，将二者分别进行简单的线性拟合，臭氧与甲醛拟合结果如下：

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.576e+01  3.387e+00   16.46  <2e-16 ***
total_data$HCHO 4.149e-15  2.115e-16   19.62  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 32.65 on 1459 degrees of freedom
Multiple R-squared:  0.2088,    Adjusted R-squared:  0.2082
F-statistic: 385 on 1 and 1459 DF,  p-value: < 2.2e-16
```

拟合结果表明臭氧与甲醛相关性较强，但二者显然不是简单的线性关系，因此 Rsquare 较小，O₃/HCHO 的斜率为 4.149e-15。

二氧化氮与臭氧拟合结果如下：

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.360e+01  1.473e+00   56.76  <2e-16 ***
total_data$NO2 6.983e-15  2.406e-16   29.02  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 29.23 on 1457 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared:  0.3663,    Adjusted R-squared:  0.3658
F-statistic: 842.1 on 1 and 1457 DF,  p-value: < 2.2e-16
```

相对于臭氧与甲醛的拟合，臭氧与二氧化氮的拟合 Rsquare 更大，拟合结果更好一些，O₃/NO₂ 的斜率为 6.983e-15。

根据前面的分布图我们可以知道，不同地区臭氧、二氧化氮和甲醛的浓度有显著差异，为了更精确的探究不同地区三者浓度之间的关系，我分别选取了华北地区和华南地区的观测站点数据，并以臭氧浓度为因变量，以甲醛和二氧化氮浓度为自变量，分别进行了 Two-way ANOVA 检验和多元线性拟合，可以探究在这两个典型区域臭氧对于两种前体物的敏感性。

对于华北地区：

```
Call:
lm(formula = O3 ~ HCHO + NO2, data = huabei)

Residuals:
    Min       1Q   Median       3Q      Max
-141.551   -9.358    1.716   12.608   38.574

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.702e+01  9.466e+00  10.249  < 2e-16 ***
HCHO         2.834e-15  6.073e-16   4.666 5.14e-06 ***
NO2          1.843e-15  5.550e-16   3.320 0.00104 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.36 on 237 degrees of freedom
Multiple R-squared:  0.2845,    Adjusted R-squared:  0.2784
F-statistic: 47.11 on 2 and 237 DF,  p-value: < 2.2e-16
```

可以看到O3相对于甲醛和二氧化氮的斜率分别为2.834e-15和1.843e-15，其中HCHO的p值相对于NO2更小，说明华北地区臭氧浓度对甲醛的敏感性更高，但和NO2也有较强的正相关关系。

华南地区的结果：

```
Call:
lm(formula = O3 ~ HCHO + NO2, data = huanan)

Residuals:
    Min       1Q   Median       3Q      Max
-19.988  -6.630  -1.274   7.398  29.159

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.919e+01  3.335e+00  14.750  < 2e-16 ***
HCHO         1.988e-15  2.590e-16   7.673 4.79e-13 ***
NO2          6.802e-16  4.184e-16   1.626   0.105
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10 on 229 degrees of freedom
Multiple R-squared:  0.3321,    Adjusted R-squared:  0.3263
F-statistic: 56.94 on 2 and 229 DF,  p-value: < 2.2e-16
```

可以看到华南地区NO2的p值较大，为0.105，且斜率为6.802e-16，明显小于HCHO，说明华南地区相对于华北地区，臭氧对NO2的敏感性大大降低。

五、总结

此次大作业简单处理了一下TROPOMI的卫星观测数据，对甲醛和二氧化氮的对流层柱浓度进行了可视化处理，发现我国甲醛、臭氧和二氧化氮的分布呈现东部高、西部低的趋势，TROPOMI卫星的数据成像质量较好，可以用于小尺度范围内的污染物浓度空间分布分析。

经过对甲醛、二氧化氮和臭氧的浓度关系探究，发现我国臭氧浓度和甲醛、二氧化氮浓度的空间分布呈现较为明显的正相关关系。分别对华北地区和华南地区进行了多元线性拟合，发现华北和华南地区，臭氧对甲醛的依赖性更强，且华南地区臭氧对NO2的依赖性很弱。

此项分析还有很大的局限性，包括卫星观测数据本身具有极大的误差，进行月平均处理并不能完全消除这些误差。并且臭氧浓度和甲醛、二氧化氮浓度之间也并非简单的线性关系，甲醛和二氧化氮二者本身也存在一定的相关性等。

代码和二次处理的数据已经上传至我的github：
https://github.com/Houyue99/ESE5023_Assignments