Application of AERMOD model with incomplete data in Quito

Leonardo Soto
CICAM · Escuela Politécnica Nacional
21.5.2020

The aim of this document is to present a way to use different data sources to simulate the air pollutants dispersion with AERMOD. Terrain variables are obtained from satellite mapping, government agencies and literature. For meteorological data, REMMAQ and NCEP FNL Operational Model Global Tropospheric Analyses data are used as meteorological inputs. Finally, this report presents some insights to troubleshot other incomplete data difficulties.

Introduction

Chronic exposure to air pollutants is a global health problem. The WHO estimated that the atmospheric pollution led to 4,2 millions of premature deaths every year and 91% of populations lived in place where the WHO regulations are not meet (WHO, 2016). Nowadays. air pollution is one of the main environmental problems that all urban zones are tackling.

In Ecuador, cities as Quito, Cuenca and Guayaquil are the only ones which includes an air quality management on their municipality agenda. Quito has the Metropolitan Atmospheric Monitoring Network (REMMAQ) with 9 air stations and 11 meteorological stations all over its urban zone. Some of the measured parameters are $PM_{2.5}$ NO_x, CO, SO_2 y O_3, wind speed, wind direction, temperature, and radiation. A REMMAQ report states that in 2019, the air was good 12% of days, acceptable 78% of days and warning 10% of the days (QuitoInforma, 2020). In Cuenca, the Air Quality Improvement Corporation (CUENCAIRE) has monitoring equipment and has established a baseline for the air quality. Guayaquil counts with an emissions inventory. The rest of the cities does not have a system for air quality evaluation. This is high concern especially in cities with oil industry activity, cement industries, wildfires and volcanic activity.

To improve the available capacity to respond against air pollution it is necessary to implement tools for air quality evaluation like monitoring, modeling and inventorying. Since monitoring is not feasible in many cities, some others like inventorying and modeling could be a practicable alternative. However, modeling is relatively less implemented compared with inventorying. Modeling requires a high grade of specialization and usually it is labor intensive. In Europe there are some models as Polyphemus.., in USA there are AERMOD, CALPUFF and ICSS. In South America there are efforts to establish a regional model like MODELAR (Modelo Regulatório de Qualidade do Ar) in Brazil.

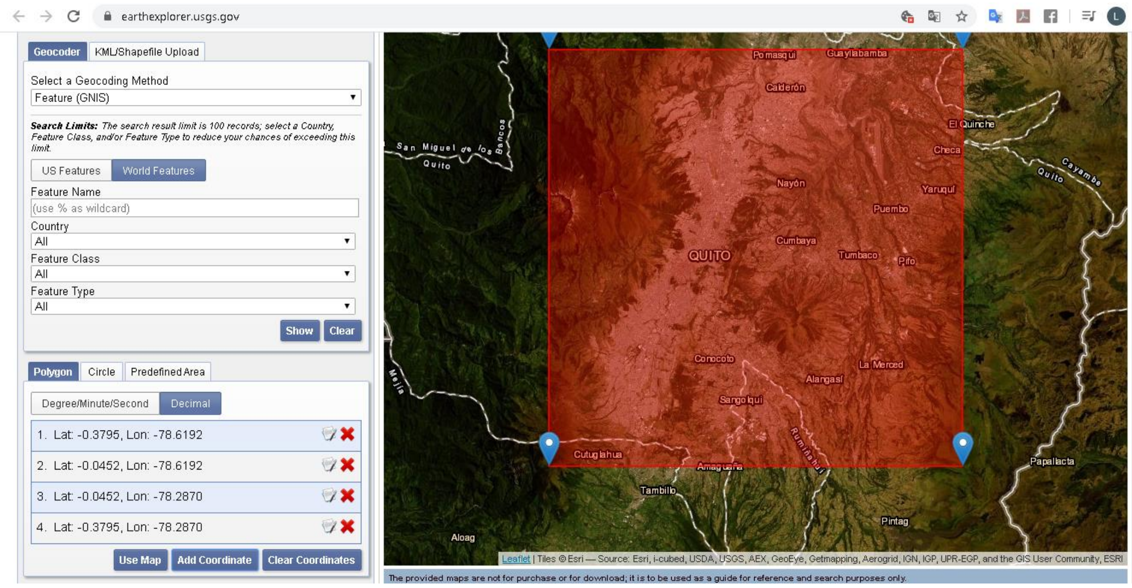
One of the most extended models is AERMOD and it has been used for regulatory purposes in some countries besides USA (SEA, 2012). It offers a validated way of estimate concentrations of pollutants even at 50 km away from the source. Some research applications include evaluation of dispersion of pollutants from industrial clusters, chimneys, urban calefaction, roads, and parking lots. AERMOD is a gaussian model that incorporate the effects of planetary boundary layer turbulence, complex terrain, and downwash. The source code and software is access-free and can be downloaded from https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models. However, one of the main difficulties of using AERMOD outside of US is the lack of adequate data required by the model. The principal inconvenient is supplying meteorological data which is usually provided by radiosondes and meteorological stations and towers.

The main idea of this work is to show a way to apply the AERMOD model when data is incomplete as it happens in most part of the world. It focusses on data processing.

Methodology

Study area

For this example, the study area is Quito, Ecuador. The area covers a square region 37 km x 37 km, from 765000 E 9958000 N 17 S to 802000 E 9995000 N 17 S. A smaller area is recommended.



Terrain

A digital elevation model database provided by USGS for South America is accessible through the link http://dds.cr.usgs.gov/srtm/version2_1/SRTM3/South_America/ . The file is type 3-arc second from the Shuttle Radar Topography Mission (SRTM). The resolution of this product is about 90 m.

Surface characteristics like albedo, Bowen ration and surface roughness length can be estimated according to available land use maps (e.g. http://geoportal.agricultura.gob.ec/geonetwork), satellite images (e.g. Google Earth) and literature. Those characteristics are related to the meteorological station location (1 km radius for surface roughness length, and 10km x 10km for albedo and Bowen ratio). This information could be specified by angular sector and meteorological season.

Surface meteorology

To comply with the meteorology input data, it is possible to use observations and results from models like NCEP GFS or WRF. In case of Quito, REMMAQ is a network that provides data that can be used in AERMET pre-processor. This includes wind speed and direction, temperature, humidity, pressure, and radiation. This data is accessible in http://190.11.24.212/reportes/ReporteDiariosData.aspx.

Data was processed in an Excel spreadsheet arranging column according to the SCRAM format http://www.webmet.com/MetGuide/SCRAMSurface.html. Other format options are CD144, SAMSON, HUSWO, ISHD y TD-3280.

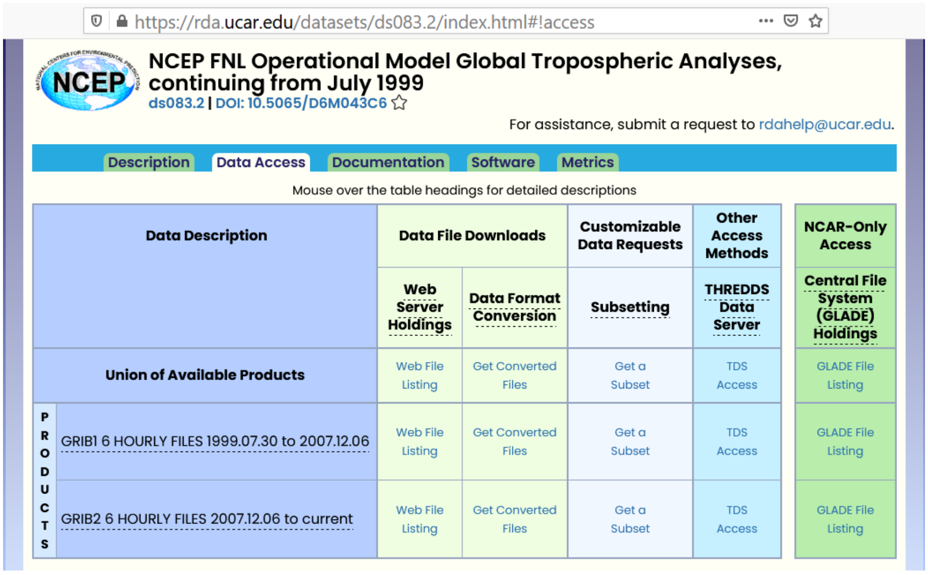
As there is not cloudiness data in REMMAQ, an alternative is to search for historical observations. Considering that finding observations could be difficult, an estimate for cloudiness could be get from radiation data (e.g. maximum radiation correspond to minimum cloudiness). In other case, radiation could be used directly in AERMET by configuring the ONSITE pathway and READ keyword. However, for this example cloudiness will be used.

Upper meteorology

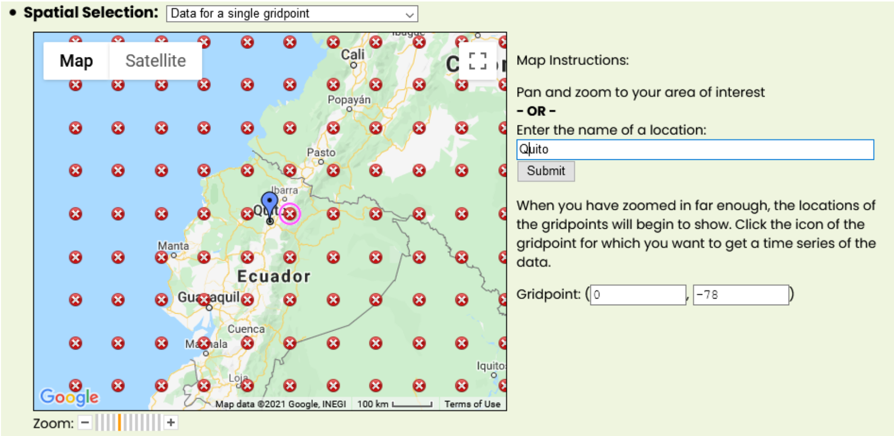
Data can be obtaining from radiosondes or meteorological towers. However, those instruments are not commonly used in many countries, and other options should be considered. There are also some commercial providers, but those files usually are expensive and could not be as representative as observations and specific simulations. Simulation models are an attractive alternative to comply this requirement. The data obtained from the Weather Research and Forecast (WRF) can be processed and used directly in the AERMOD through the Mesoscale Model Interface (MMIF). This alternative will be explained in detail in a next report. The data to be used herein is the NCEP FNL Operational Model Global Tropospheric Analyses (can be used initialize WRF), continuing from July 1999. T DOI: 10.5065/D6M043C6. The grid is 1° x 1° and data is generated every 6 hours.

The procedure to retrieve the meteorological data from NCEP FNL is described below:

1. Register as a user at https://rda.ucar.edu/
2. Log in and get a subset of GRIB1 or GRIB2 group according to the dates of simulation from https://rda.ucar.edu/datasets/ds083.2/index.html#!access . For this example, GRIB1 will be used as the year to study is 2006[1].
3. Select all available parameters and verticals levels for a single gridpoint and netCDF4 format. Some of the variables and units are shown in Table No.

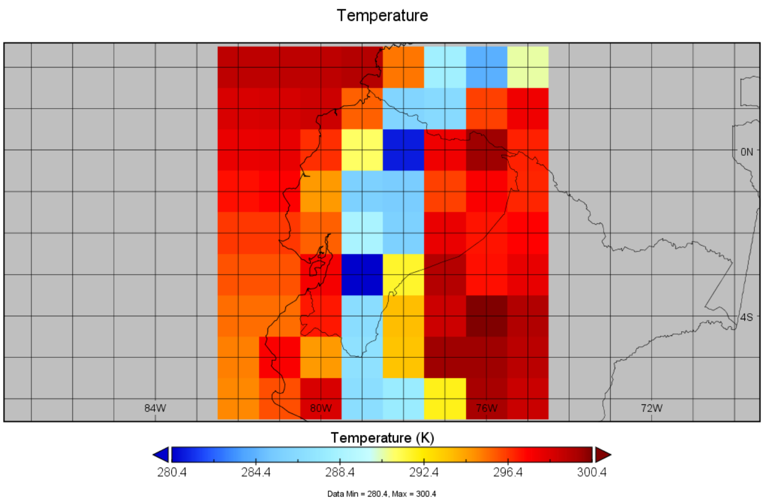


[1] The year 2006 was chosen to compare results with a previous study (Ruiz, 2013).



4. Download and unzip the files.

The obtained data can be visualized with many programs. One of these programs is Panoply.



Surface temperature graph in Ecuador on 2006-01-01 00h00 using Panoply

Table No. GRIB Variables and units

Parameter Code	Short Name	Description	Units
1	PRES	Pressure	Pa
2	PRMSL	Pressure reduced to MSL	Pa
7	HGT	Geopotential height	gpm
10	TOZNE	Total ozone	Dobson
11	TMP	Temperature	K
13	POT	Potential temperature	K
15	TMAX	Maximum temperature	K
16	TMIN	Minimum temperature	K
27	GPA	Geopotential height anomaly	gpm
33	UGRD	u-component of wind	$m\ s^{-1}$
34	VGRD	v-component of wind	$m\ s^{-1}$
39	VVEL	Vertical velocity (pressure)	$Pa\ s^{-1}$
41	ABSV	Absolute vorticity	s^{-1}
51	SPFH	Specific humidity	$kg\ kg^{-1}$
52	RH	Relative humidity	%

The format for upper meteorological input in AERMET to be used here is FSL (Forecast Systems Laboratory). The arranging of variables is described in https://ruc.noaa.gov/raobs/fsl_format-new.html. There are some ways to extract the netCDF4 data and compile it in FSL format. One way is using a spreadsheet, but this could be very labor intensive and time consuming. Another way is using some programming language (e.g. R or Phyton). Here, the following R script for processing a folder of nc files is proposed:

```
[R]
rm(list=ls())
library(fields)
library(ncdf4)
library(dplyr)
library(gdata)
library(readr)
setwd("D:/R/")
lista <- list.files(pattern = "\\.nc$")
N <- length(lista)
for(i in 1:N) {
 f <- nc_open(lista[i], verbose = FALSE, write = FALSE)

  #READING
  ymd = ncvar_get(f,"initial_time0")
  lev = ncvar_get(f,"lv_ISBL4")
  hgt = ncvar_get(f,"HGT_3_ISBL")
  tmp = ncvar_get(f,"TMP_3_ISBL")
  r_h = ncvar_get(f,"R_H_3_ISBL")

  #MISSING VALUES
  r_h <- unknownToNA(x=r_h, unknown=0)
  u_w = ncvar_get(f,"U_GRD_3_ISBL")
  v_w = ncvar_get(f,"V_GRD_3_ISBL")


  #EXTRACTION
  año <- substr(ymd,9,10)
  mes <- substr(ymd,1,2)
  dia <- substr(ymd,4,5)
  hora <- substr(ymd,13,14)

  lev.vec <- as.vector(lev)[-(1:13)]
  hgt.vec <- as.vector(hgt)[-(1:13)]
  tmp.vec <- as.vector(tmp)[-(1:13)]
  r_h.vec <- as.vector(r_h)[-(1:8)]
  u_w.vec <- as.vector(u_w)[-(1:13)]
  v_w.vec <- as.vector(v_w)[-(1:13)]

  niveles<-length(lev.vec)

  tab <- data.frame(cbind(lev.vec,hgt.vec,tmp.vec,r_h.vec,u_w.vec,v_w.vec))
  names(tab) <- c("lev","hgt","tmp","r_h","u_w","v_w")

  #CALCULATIONS
  tmp.vec <- (tmp.vec-273.15)
  w_s.vec <- sqrt(u_w.vec^2+v_w.vec^2)
  t=atan(v_w.vec/u_w.vec)*180/3.1416
  tab1 <- data.frame(cbind(u_w.vec,v_w.vec,t))
  names(tab1) <- c("u_w","v_w","t")
  tab1<-data.frame(mutate(tab1,w_d = case_when ( u_w.vec<=0 & v_w.vec>0 ~ 270-t,
                                        u_w.vec<0 & v_w.vec<=0 ~ 270-t,
                                        u_w.vec>=0 & v_w.vec<=0 ~ 90-t,
                                        TRUE ~ 90-t )))

  w_d.vec <- tab1$w_d
  tdw.vec <- tmp.vec-(100-r_h.vec)/5

  #FORMAT

  lev.vec <- lev.vec*10
  hgt.vec <- round(hgt.vec)
  tmp.vec <- round(tmp.vec*10)
  tdw.vec <- round(tdw.vec*10)
```

```
  w_d.vec <- round(w_d.vec)
  w_s.vec <- round(w_s.vec*10)

  #COMPILATION

  tab2 <- data.frame(cbind(lev.vec, hgt.vec,tmp.vec,tdw.vec,w_d.vec,w_s.vec))
  tab2 <- tab2[order(-lev.vec),]
  names(tab2) <- c(paste(año,mes,dia,sep=""), paste(hora,"   ",niveles,sep=""),"","","","")
  colnames(tab2) <- formatC(colnames(tab2), width = 7, flag = " ")
  csvfile <-paste(lista[i],".txt",collapse="")
  write.fwf(tab2, csvfile, width = rep(7,ncol(tab2)), sep="")
  rm(f)   rm(lev)   rm(hgt)   rm(tmp)   rm(r_h)   rm(u_w)   rm(v_w)   rm(lev.vec)   rm(hgt.vec)
  rm(tmp.vec)   rm(r_h.vec)   rm(u_w.vec)   rm(v_w.vec)   rm(tab)   rm(w_s.vec)   rm(t)
  rm(tab1)   rm(w_d.vec)   rm(tdw.vec)   rm(tab2)

}

txts <- list.files(pattern = "\\.txt$")

combined=NULL

for (file in txts){

  temp_dataset <-read.table(file, header=FALSE, sep="\t")
  combined<-rbind(combined, temp_dataset)
  rm(temp_dataset)
}

#EXPORT

write.fwf(combined, file = "output.txt", sep="")

rm(combined)
```

output.txt - Notepad

File Edit Format View Help

```
   5000   5866    -33   -149    255     62
06123106    13
  10000    144    256    178    291     14
   9750    359    242    164    283     13
   9500    580    230    152    270     12
   9250    806    215    137    280     11
   9000   1038    199    121    300     14
   8500   1521    169     91    326     22
   8000   2031    137     59    318     24
   7500   2571    111     33    297     31
   7000   3145     86     10    282     44
   6500   3756     61     11    275     54
   6000   4410     32    -28    275     53
   5500   5112      1    -81    268     53
   5000   5870    -39   -153    252     62
06123112    13
  10000    141    264    180    324      9
   9750    356    251    167    329      6
   9500    577    239    155    315      3
   9250    803    222    138    304      4
   9000   1036    206    122    309      6
   8500   1519    180     96    356     14
   8000   2030    146     62    356     15
   7500   2572    120     36    321     22
   7000   3148     96     26    291     25
   6500   3760     59     13    296     30
   6000   4412     24    -14    300     28
   5500   5112    -11    -59    300     43
   5000   5866    -53   -143    288     47
06123118    13
  10000     96    340    236    265     24
   9750    312    326    222    263     24
   9500    534    314    210    258     25
   9250    762    299    195    252     22
   9000    997    285    181    245     19
   8500   1488    251    147    246     17
   8000   2009    214    110    276     19
   7500   2562    172     68    277     32
   7000   3146    121     37    282     40
   6500   3763     72     34    284     34
   6000   4418     32     -2    287     37
   5500   5121      0    -60    283     31
   5000   5878    -44   -114    259     41
```

Ln 20441, Col 43 | 100% | Windows (CRLF) | UTF-8

Figure No. FSL format data obtained by processing nc files with R

Additional changes for the plain text could be done with any text editor software (e.g. Notepad+). Once the input files are compiled, AERMET can be executed.

Results

A simulation for SO_2 on 2006 shows high concentrations of SO_2 (max 350 ug/m3) around the principal combustion sources. Dispersion plumes extended over the geographical valleys and reach over 30 km distances (10 ug/m3) as shown in Figure No.

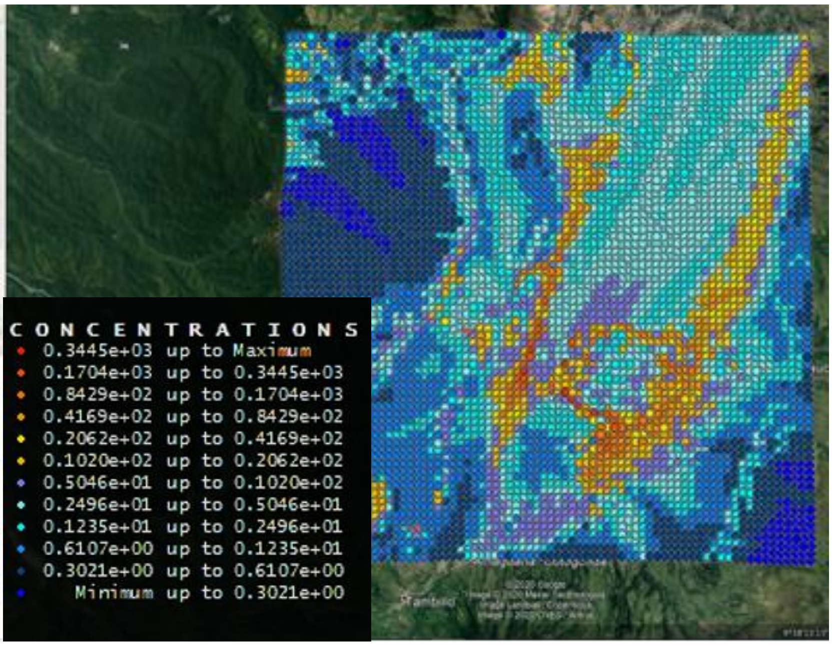


Figure No. SO_2 24h maximum concentrations simulated with AERMOD. Year: 2006. Google Earth.

No distinguishable correlation can be stated for the daily average measured data of El Camal station and simulated data for the corresponding receptor. This could be explained by the presence of other sources of SO_2 as other industries and diesel combustion vehicles (e.g. buses in Quito uses high Sulphur combustible). Simulated values were less than measured values.

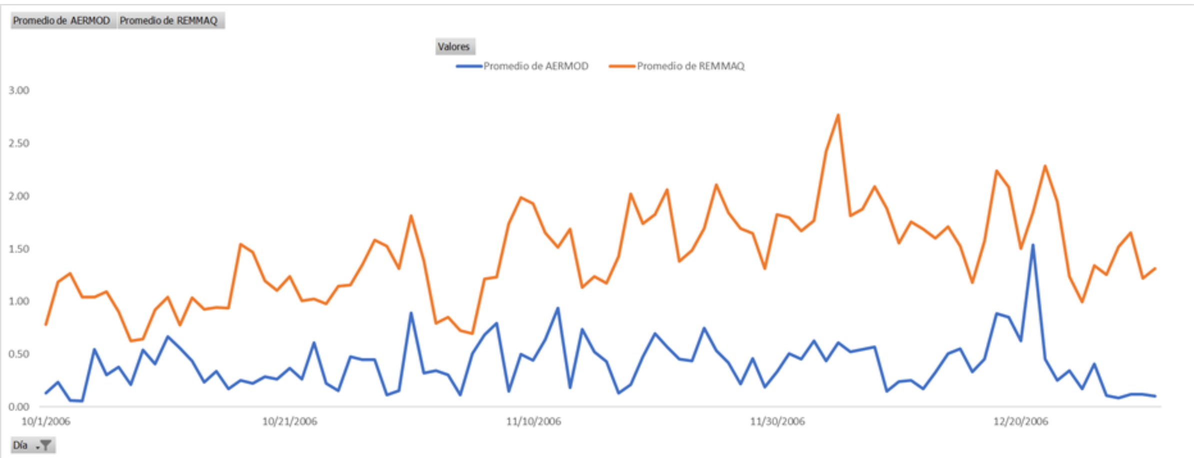


Figure No. . *SO_2 Daily average: simulated vs REMMAQ data. El Camal station, fourth quarter of 2006.*

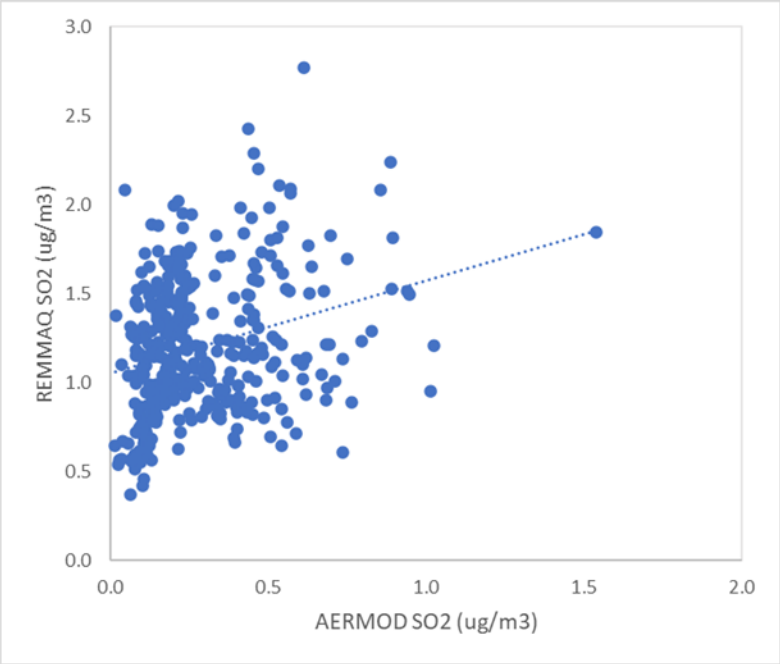


Figure No. . *SO2 Daily average: simulated vs REMMAQ data. El Camal station, 2006*

Conclusions

No available input data required by the AERMOD model could be estimated by other means. Elevation and characteristic of terrain can be obtained from satellite mappings, public institutions, and literature. Meteorological data can be supplied by the NCEP FNL Operational Model Global Tropospheric Analyses as an alternative for locations without instrumental capacity and infrastructure to provide measured data.

More revisions are required to ensure that results obtained by FNLs could be comparable to those obtained by measured data.

References

SEA (2012) Guía para el uso de modelos de calidad del aire en el SEIA. Available on https://www.sea.gob.cl/sites/default/files/migration_files/guias/Guia_uso_modelo_calidad_del_aire_seia.pdf Viewed on May 18, 2021.

Quito informa (2020) El 90% de los días, durante 2019, la Calidad del Aire en Quito se mantuvo bajo Norma. Available on http://www.quitoinforma.gob.ec/2020/01/11/el-90-de-los-dias-durante-2019-la-calidad-del-aire-en-quito-se-mantuvo-bajo-norma/ Viewed on May 18, 2021.

WHO (2016). Calidad del aire y salud. Available on https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health. Viewed on May 19, 2021.

Ruiz (2013). Modelación de la dispersión de las emisiones de SO2 y NOx de las termoeléctricas en la ciudad de Quito en el software CALPUFF. Available on http://bibdigital.epn.edu.ec/handle/15000/8069 Viewed on May 19,2021