Producing new CAMS aerosol input for climate files

Ana Šljivić, Ján Mašek

February, 2024

1 Extracting data from CAMS database

On the Copernicus webpage:
https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4-monthly?tab=overview, monthly averaged fields from the CAMS global reanalysis (EAC4) can be extracted as grib or netcdf file format. Data are available for the period 2003 - 2022.
Vertically integrated mass [kg/m^2] for 11 aerosol species has been downloaded with the purpose to include them to climatological files. Spatial resolution is 0.75° x 0.75°, giving 241 latitudes and 480 longitudes for the whole globe.
We extracted the following variables:

Vertically integrated mass of dust aerosol (0.03 - 0.55 μm)

Vertically integrated mass of dust aerosol (0.55 - 9 μm)

Vertically integrated mass of dust aerosol (9 - 20 μm)

Vertically integrated mass of hydrophilic black carbon aerosol

Vertically integrated mass of hydrophilic organic matter aerosol

Vertically integrated mass of hydrophobic black carbon aerosol

Vertically integrated mass of hydrophobic organic matter aerosol

Vertically integrated mass of sea salt aerosol (0.03 - 0.5 μm)

Vertically integrated mass of sea salt aerosol (0.5 - 5 μm)

Vertically integrated mass of sea salt aerosol (5 - 20 μm)

Vertically integrated mass of sulphate aerosol

Vertically integrated mass of sulphur dioxide

After choosing the variables, there are options to choose year (from 2003 - 2022) and months. We chose all years and all months to be in one grib file.
After filling the required form, one should press "Submit form". To get the data, user has to be logged in.

2 Averaging

Detailed description of the grib content can be seen by using the command `grib_dump`. Some of the descriptors are `shrotName` and `indicatorOfParameter` which are listed in Table 1 (Section 3).
In the downloaded grib file are included 20 years of monthly values for all 11 aerosol types. First we have to separate it to grib files for each month and for each aerosol type to make averaging easier.

Separating was done by using a command:

```
grib_filter rule.filter name_of_the_grib.grib
```

where file `rule.filter` contains:

```
write "[file]_[month]_[shortName].grib";
```

Result is 132 grib files, one for each aerosol type and for each month. Each grib file contains 20 fields, one per year. The script for averaging over years is using module `pygrib` for opening and reading grib data. After averaging, they are saved in txt file named `<aerosol><N>`, where `aerosol` is aerosol type and `N` is number of the month.
Bash script is going through all aerosol types and all months, calling averaging python script `camscms.py`.
Averaging script has three arguments: name of the grib file, aerosol type and month. Output of the script is txt file with 241 lines and 480 columns containing averaged values for specific month and aerosol type over 20 years.

Bash script is:

```
for aerosol in $(echo "aermssdus aermssdum aermssdul aermsssss aermssssm aermsssssl aermssbchphil
aermssbchphob aermssomhphil aermssomhphob aermsssu"); do
  for month in 1 2 3 4 5 6 7 8 9 10 11 12; do
    python camscms.py mars_2003-2022.grib_${month}_${aerosol}.grib ${aerosol} ${month}
  done
done
```

while `camscms.py` script for averaging is:

```
import pygrib
import numpy as np
import sys
name = str(sys.argv[1])
aerosol = str(sys.argv[2])
month = str(sys.argv[3])
output=aerosol+month
grbs = pygrib.open(name)
grball=np.empty((241,480,20))
i=0
for grb in grbs:
  grball[:,:,i]=grb.values
  i=i+1
ave=np.average(grball, axis=2)
np.savetxt(output, ave, delimiter=' ', newline='\n')
```

To get values of latitudes (241) and longitudes (480), it should be added:

```
grb=grbs[2]
lat, lon = grb.latlons()
lat1d=lat[:,1]
lon1d=lon[1,:]
np.savetxt("lat", lat1d, delimiter=' ', newline='\n')
np.savetxt("lon", lon1d, delimiter=' ', newline='\n')
```

3 Converting ASCII to binary files

Next step is to convert 132 ASCII files (11 aerosols x 12 months) to 12 binary files which will be read in the `eincli12.F90` subroutine (see Section 4).
Script for converting is on kazi1: `/home/mma157/wrk/cams4e923/convert_cams4e923.pl`.
Variable `$DATA` is the location of the input ASCII files. Script is merging all 11 aerosol types for each month in one file. The order should be the same as is hardcoded in the `eincli12.F90` (Table 1). Result is 12 `aero_cams.mm` binary files, where `mm` is month.

Table 1: List of aerosol types

shortName	indicatorOfParameter	EINCLI12 name	description
aermssssss	19	SURFAEROCMS.SS1	sea salt, small
aermssssm	20	SURFAEROCMS.SS2	sea salt, medium
aermssssl	21	SURFAEROCMS.SS3	sea salt, large
aermssdus	43	SURFAEROCMS.DD1	desert dust, small
aermssdum	44	SURFAEROCMS.DD2	desert dust, medium
aermssdul	45	SURFAEROCMS.DD3	desert dust, large
aermssomhphil	62	SURFAEROCMS.OM1	organic matter, hydrophilic
aermssomhphob	61	SURFAEROCMS.OM2	organic matter, hydrophobic
aermssbchphil	78	SURFAEROCMS.BC1	black carbon, hydrophilic
aermssbchphob	77	SURFAEROCMS.BC2	black carbon, hydrophobic
aermsssu	87	SURFAEROCMS.SU	sulphates

The input data have latitude ordering from 90 to -90 degrees (N-S) and longitude from 0 to 360 degrees (E-W). Data will be processed during the e923 procedure, thus we have to adjust the data according to what the interpolation routine einter2.F90 expects as input. Convention is the following: Globe is divided to rectangles, with boundaries spaced regularly from equator and Greenwich meridian. Longitude ordering is from 0 to 360 deg (E-W), latitude from -90 to 90 degrees (S-N). Data values are not given at rectangle vertices, but at their centers. In our case with grid interval 0.75 deg, the longitudes are 0.375, 1.125, . . . , 359.625 deg, latitudes -89.625, ..., -0.375, 0.375, 88.875, 89.625.
The input file is ordered from N to S, but latitude reversal is done during reading of the file in eincli12.F90. However, interpolation from CAMS grid to e923 grid has to be done during preparation of the binary files. It is done in the script and output binary files use 4 bytes per value in big-endian order, having 480 longitudes x 240 latitudes x 11 aerosol types. They are thus portable between different machines.

4 Tools - script on belenos

Procedures described in Sections 2 and 3 (averaging and converting ASCII to binary files) can be easily performed by running one script on belenos: `/home/gmap/mrpm/sljivica/camscms/tools/grib2bin_ave.sh`.
Input is one grib file (obtained as described in Section 1) and output are 12 binary files named `aero_cams.mm`. In the script, user should adapt the name of the input grib file and the location where output files should be created. Except of the input grib file and the script, in the running directory should also be located scripts: `camscms.py, run.sh, convert_cams4e923.pl`.

5 Code changes

Starting point was Piotr Sekula's pack on belenos: `/home/gmap/mrpe/sekulap/pack/cy46t1_bf.07_ver15_bis`. There are already included changes made by Laura Rontu to introduce new 11 CAMS aerosol masses from ASCII file. Many routines were touched, among others `eincli9.F90`. Original routine `eincli9.F90` is reading four Tegen aerosol data stored in the ASCII format, interpolating and writing to ALADIN climate file. We decided to not change the original `eincli9.F90` but instead to make a new analogue routine `eincli12.F90`. To make the e923 procedure able to process new type and content of aerosol data (binary containing 11 types instead of ASCII containing 4 Tegen types), two routines had to be created/changed:

```
src/local/aladin/c9xx/eincli12.F90

src/local/arpifs/c9xx/incli0.F90
```

The important changes in the new `eincli12.F90` routine with respect to the original `eincli9.F90` are:

- changed number and names of aerosols
- opening the binary file with record length 4 x number of longitudes, `NDATX`
- reading the data first to an auxiliary single precision (`JPRS`, 4 bytes) variable `ZAUX(NDATX)` and then rewriting it to the actual precision (`JPRB`, 4 or 8 bytes) field `ZFLD(NDATX+4,NDATY+4)`

Subroutine is called from the `incli0.F90` routine. There are added:

- a block with defined values of `LIEEE=.TRUE.` (input is unformatted), `NDATX` and `NDATY`; all this can be changed in the `&NAMCLI` namelist
- protection if value of `N923` is smaller than 0 or bigger than 12 to abort; instead of previous protection in setup `sumcc.F90` which was putting all unintroduced options to 11
- if `N923==12` call `eincli12.F90`

6 Running the e923 procedure

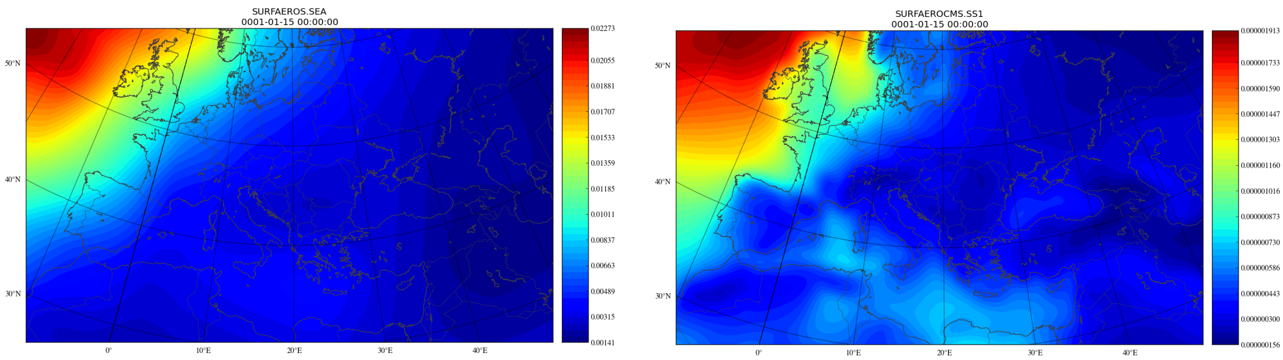
Climate files containing the new 11 CAMS aerosol masses can be obtained now by running step 12 of the e923 procedure. Inputs necessary to run it are:

1. climate file in which aerosols will be added, named `Const.Clim`
2. namelist named `fort.4`
3. binary file containing the data, named `aero_cams` (preparation described in Section 3)

Modifications in the namelist fort.4 include mainly:

- setting up the geometry according to `Const.Clim` file in the namelists: `&NAMDIM, &NEMDIM, &NEMGEO`
- setting the value of `N923=12` in `&NAMCC`

Comparison between the old Tegen sea salt aerosol optical depth data and the new CAMS EAC4 data for sea salt mass (small size) are shown in the Figure 1.



(a) (b)

Figure 1: Sea salt Tegen climatology (a) and new CAMS EAC4 global reanalisys sea salt, small size (b)