Tutorial

This tutorial refers to the monthly report in July 2021 in South Africa regarding the cloud coverage which was created using the CM SAF R-Toolbox. All steps on how to create the maps and plots with the R-Toolbox are documented in here. The order is based on the order of the report.

- To draw up this kind of report, you first need to download daily and monthly data for your chosen area. The exact data names are written down in the report under "Accessed data".
- All data is available on: https://wui.cmsaf.eu/
- The extract for South Africa used in the application example:
  - Latitude: 20-40° S
  - Longitude: 0-40° E
- Plotted locations in South Africa:

Location	Latitude	Longitude
Pretoria	-25.731340	28.218370
Port Elizabeth	-33.958252	25.619022
Bloemfontein	-29.087217	26.154898
Cape Town (Kapstadt)	-33.918861	18.423300

How to create a…

✓ …general map of cloud coverage in a selected country in a selected month?

1. Get monthly mean data (.tar-file) for your selected country and year
2. Select the data range you want to analyze
3. Choose "cfc" as the variable and reduce the spatial coverage if necessary
4. Select "monthly statistics" as a group of operators as well as "monthly means"
5. Choose your month of interest in the visualizer options and start brighten up your map!

➔ **Result**: „Average Cloud Fractional Cover (CFC) in July 2021 in South Africa" (see fig. 1)

- ❖ To provide a better overview, plot several locations in your country
- ❖ In addition, you can create particular maps of each month of the year by selecting every time step given in the visualizer options. Disposing them side by side afterwards, makes a comparison between them easy.

➔ **Result**: „Comparison of the monthly maps (January-July 2021)" (see fig. 3)

✓ …global map of cloud coverage in the selected month?

1. Get global monthly mean data for the current year

2. The process is now the same as before

➔ **Result**: „Map of the cloud coverage in July 2021 illustrated as a globe, centered and zoomed in on Africa" (see fig. 4)

✓ …monthly anomaly of cloud coverage in a selected month?

1. Get monthly mean data for the following time span for your country of interest: 1991-2020, the period of time must cover 30 years
   - 1991-2019 are available as TCDR
   - 2019-2020 must be downloaded as ICDR

   → the downloaded orders can be combined by giving them same names in ascending order

   For example:

   - *ORD44122.tar → ORD44122_1.tar*
   - *ORD44123.tar → ORD44122_2.tar*
   - *ORD44124.tar → ORD44122_3.tar*

2. Get monthly mean data for the current year

> Please note that you always have to examine the exactly same spatial extract for all data packages when downloading or processing the data in the R-Toolbox!

3. **Prepare:** Untar and unzip the whole time series and select the "cfc" variable, the resulting file will be: cfc_1991-01-01-2020-12-01.nc

4. **Analyze:**
   a) Create "multi-year monthly means" by choosing the group of operators "monthly statistics" and hook the box saying: "do you want to apply another operator afterwards?"
   b) Choose the operators "selection" and "select list of months" to be able to pick out the month to be analyzed and create your long-term mean. The created output-file is called "cfc_selmon…".

   ❖ The generated long-term-mean-file can be used as a reference. It is now possible to compare the map of the cloud coverage that represents the current month and year to the reference!

➔ **Result**: „Current map of July 2021 compared to the long-term mean map of July 1991 to 2020)" (see fig. 5)

   c) Untar and unzip the selected month for the current year (f.e. 2021-07-01 to 2021-07-01)
   d) Then compute the difference between the selected month (July 2021) and the mulit-year mean that has been generated before by choosing "Mathematical operators" and "Substract values from another file". The difference is calculated as follows:
   July 2021 minus mean (July 1991-2020)

   → The result will be a monthly anomaly map. Illustrating the map in the color "blue-red 3" is recommended. The R-Toolbox creates the overall analysis automatically. It can be downloaded easily afterwards.

➔ **Result:** „Monthly anomaly map of cloud coverage in July 2021 in South Africa" (see fig. 6)

✓ …time-series of monthly anomalies for a specific location (1991-2020)?

First option

1. **Starting with Analyze**: the already created .nc-file (cfc_1991-01-01-2020-12-01) can now be used to depict the monthly anomalies as a time-series

2. Choose "monthly statistics" as well as "monthly anomalies" and hook the box saying: "do you want to apply another operator afterwards?" again

3. Next, select the operators “Selection” and “Select data at given point”. In the following fields, you can type in the location (longitude and latitude) of any location within your chosen country!

4. You can also export the data (for example as .csv) and design your plot individually in any other program

   ❖ If you want to get a better overview about your plot, you can insert trendlines. They show the overall change in the cloud coverage throughout the time span!

   ➔ **Result:** „Time series of monthly anomalies of fractional cloud cover in the city Pretoria (South Africa) from 1991 to 2020” (see fig. 8)

Second option

1. **Analyze:** choose again the already created .nc-file (cfc_1991-01-01-2020-12-01)

2. this time, select the operators “Climate Analysis” and “Time Series Plot” as well as your country

   ❖ Advantage: using “climate analysis” and “Time Series Plot” as the operators makes it possible to plot the time-series in red and blue colors to visualize the anomalies better

   ➔ **Result:** „Time series of monthly anomalies of fractional cloud cover all over the country South Africa from 1991 to 2020” (see fig. 9)

✓ …time-series of monthly averages for a specific location (1991-2020)?

1. **Analyze:** choose again the already created .nc-file (cfc_1991-01-01-2020-12-01)

2. Then, select the operators “monthly statistics” and “monthly averages” and hook the box again to apply another operator afterwards

3. Next, select the operators “Selection” and “Select data at given point”. In the following fields, you can type in the location (longitude and latitude) of any city within your chosen country!

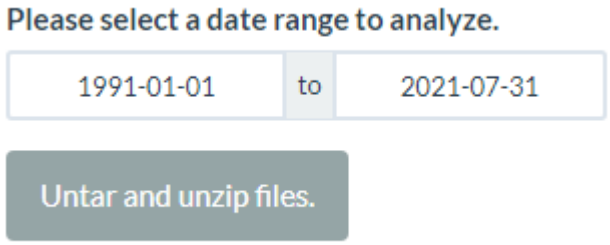
   ➔ **Result**: „Time series of the monthly average fractional cloud cover from 1991 to 2020 in Pretoria” (see fig. 11)

✓ …current year in comparison to the climatology?

1. Get daily mean data for the following time span for your country of interest: 1991-2021 and generate a .nc-file (Prepare) by combining them to one file

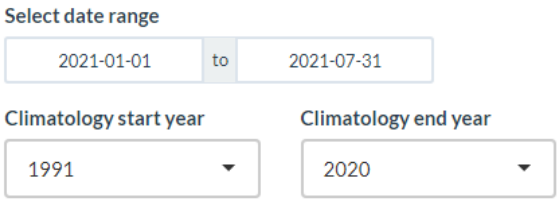
It is important to note that you select the last day of the month (30. /31.) of the current year.

→ for example:



→ cfc_1991-01-01-2021-07-31.nc

2. **Analyze**:
   - Select the operators "Climate Analysis" as well as "Fieldmean and anomaly map"
   - Then select the date range and the start and end year of the climatology, for example:



❖ Take a look into RStudio. In the console, you can read off the minimum and maximum values of cloud coverage within the specified time span!

➔ **Result**: „Plot of the current year (2021) in comparison to the climatology (on the left) as well as a map of the fractional cloud cover anomaly in South Africa in 2021 (on the right)" (see fig. 12)

✓ …map of the number of time steps above, below or equal to an uniquely defined threshold

→ New operators are required to generate this kind of map!

1. Make sure that the latest version of the “cmsaf” and “cmsafops” R-packages are installed on your machine!

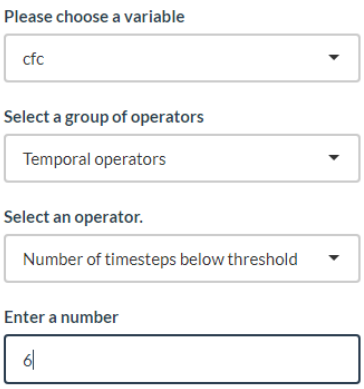
2. Get daily mean data for the current or observed month (in this example July 2021) and generate the .nc-file (cfc_2021-07-01-2021-07-31)

3. **Analyze:**
   - Select the operators “Temporal operators” and “Number of time steps above/below/equal to threshold”
   - Choose appropriate thresholds, for example:
     - ≤ 5% = „cloudfree“
     - ≥ 95% = „cloudy“

   Therefore, enter the chosen threshold value into the field “Enter a number”

   > It might be important to know, that right now, it is only possible to choose between “above” and “below”, which means that you might enter the number “6” to display cloudfree (≤ 5%) conditions/number of days.

   Example:

   - If you select the operator “below threshold”, the map shows the number of days where the cloud coverage was below this threshold. In that context, “zero” means that there were no days during the selected month with a cloud coverage below the threshold.

➔ **Result:** „ Map of the number of cloudfree-days in South Africa in July 2021” (see fig. 13)

✓ **…map showing the mean number of cloudfree-days (1991-2020)?**

1. Create (or select the already generated) .nc-file for daily mean data from 1991 to 2020 (cfc_1991-01-01-2020-12-31.nc)

2. Always choose “cfc” as a variable

3. **Analyze:**
    a) First, choose “monthly statistics” as well as “monthly number of timesteps below/above/equal to threshold” and enter an appropriate number as threshold
    b) Check the box “Do you want to apply an operator afterwards?” and apply the operator
    c) Next, select again “monthly statistics”, but now “multi-year monthly means”
    d) Uncheck the box and apply the operator

4. **Visualize:** Select the Time Step 1991-07-01, which represents the mean of July 1991 to 2020, and modify your map! Now you can compare the number of cloud free days of the current month to the mean number of cloud free days in July 1991 to 2020.

➔ **Result:** “Map showing the mean number of cloudfree-days in South Africa within the time period 1991 to 2020” (see fig. 15)

Further Tips & Tricks:

- Sometimes it happens that the window to select the .tar or .nc-file opens behind all other open tabs. If that is the case, just minimize all open windows until you are able to choose the data file!
- If you choose the operator “multi-year monthly means” and select only one month for visualization, you can only select the start year, for example: 1991. This might be confusing because this year represents the whole time range (1991-2020) which you selected beforehand!

Created by: Janine Besch