

---

# **acccmip6**

***Release 2.0.3***

**May 18, 2022**



---

## Contents

---

<b>1</b>	<b>Features</b>	<b>3</b>
<b>2</b>	<b>Useful links</b>	<b>5</b>
<b>3</b>	<b>User guide</b>	<b>7</b>
3.1	Quickstart guide . . . . .	7
3.2	Installation . . . . .	8
3.3	Searching the database . . . . .	8
3.4	CMIP6DB module . . . . .	14
3.5	Downloading data . . . . .	15
3.6	Q & A . . . . .	17
3.7	Authors . . . . .	18
3.8	License . . . . .	18



acccmip6 allows user to access the 6th Coupled Model Intercomparison Project ([CMIP6](#)) database in real-time. CMIP6 is still updating its database (as of Sep. 2019) and new data for different Tier experiments are still to come. It's a hassle to go through their servers and put search terms manually everytime to find out whether your desired data is uploaded yet or not! Instead, use `acccmip6`, to get what you need in seconds. This is written in python and does not require any python programming experience. It outputs what models, experiments, variables and realizations are currently available.



# CHAPTER 1

---

## Features

---

- Real-time search and download from continuously updating CMIP6 database
- Find data for any specific items (e.g. model, experiment, variable, frequency, realm)
- Search and download any combination of the above items
- Monitor each download process
- Find the total number of available files and realizations
- Accepts inputs in multiple formats including lists
- Validate your search items
- Get suggestions if necessary
- Access definition of the experiments
- Supports python 3.5 and above
- Skips already existing files





## CHAPTER 2

---

### Useful links

---

- Source code is available on [GitHub](#).
- A [quickstart guide](#) is available with multiple demos.
- [Overview](#) of the CMIP6 experimental design and organization.



### 3.1 Quickstart guide

A basic user guide for minimal usage.

Simply type `pip install acccmip6` from your terminal to install the package.

General usage -

- Type `acccmip6 -h` for help.
- Use `acccmip6 -o S` to search the database.
- Use `acccmip6 -o D` to download from the database.

All usable arguments and their explanations -

#### Required Argument

- `-o` : Takes output type. 'S' for searching the database or 'D' for downloading from the database. Use 'M' for using the CMIP6DB module.

#### Optional Arguments

- `-m` : Model names (multiple comma separated names are allowed)
- `-e` : Experiment names
- `-f` : CMIP6 output frequency (e.g. mon, day etc.)
- `-v` : Variable names
- `-r` : Realm name (e.g. atmos, ocean etc.)
- `-rlzn` : Select a specified realization
- `-c` : 'yes' to use checker when searching or downloading. This helps to find out whether the search items are currently available. If not, it will produce suggestions that matches closely to your search.
- `-desc` : 'yes' to get the description of the experiments searched for

- `-dir` : Download directory
- `-skip` : Skip any item (model/experiment/realizations) from your download
- `-time` : 'yes' to print out all available time periods
- `-yr` : Select data for a time period (number of years)

Example usage

```
$ acccmip6 -o S -m MIROC6 -e ssp245 -v zg -f mon -r atmos
```

Output

```
TIPS: Use the check (-c) argument to check your inputs.

Currently available models based on your search:
['MIROC6']

Currently available variables based on your search:
['zg']

Currently available experiments based on your search:
['ssp245']

Number of files: 27

Available realizations: [1, 2, 3]
```

Helpful demos:

- [installation](#)
- [search](#)
- [download](#)

## 3.2 Installation

Install is as simple as typing -

```
pip install acccmip6
```

Requires python v3.5 or up and pip. Mac users can use `brew install python3` and python `get-pip.py` from terminal. Windows users can use [Windows Subsystem](#).

You may also install the package via [conda](#) -

```
conda install -c thassan acccmip6
```

## 3.3 Searching the database

For all kinds of searches use `acccmip6 -o S` and then add in the optional arguments according to your need.

### 3.3.1 Optinal arguments 1

Use these arguments with **any combination** and in **any sequence**.

- `-m` : takes model names
- `-e` : takes experiment names
- `-v` : takes variable names
- `-f` : takes frequency
- `-r` : takes realm name

#### General usage

```
acccmip6 -o S -m MIROC6 -v vas -f mon -r atmos
```

#### General output

```
TIPS: Use the check (-c) argument to check your inputs.

Currently available models based on your search:

['MIROC6']

Currently available variables based on your search:

['vas']

Currently available experiments based on your search:

['hist-piAer', 'ssp119', 'ssp585', 'piClim-NTCF', 'piClim-histall', 'hist-stratO3',
↪ 'piClim-OC', 'ssp534-over',
'hist-piNTCF', 'abrupt-4xCO2', 'piClim-anthro', 'piClim-2xfire', 'amip-p4K', 'amip-
↪ 4xCO2', 'histSST-piAer',
'piClim-lu', 'historical', 'piControl', 'faf-stress', 'piClim-SO2', 'faf-all',
↪ 'piClim-control', 'amip-hist',
'ssp370SST-lowBC', 'hist-GHG', 'piClim-2xdust', 'dcppA-hindcast', 'ssp245-stratO3',
↪ 'ssp245-aer', 'dcppA-assim',
'ssp460', 'faf-passiveheat', 'amip', 'amip-future4K', 'hist-CO2', 'ssp245', 'ssp370',
↪ 'ssp370SST', 'hist-aer',
'ssp434', '1pctCO2', 'ssp245-GHG', 'piClim-aer', 'piClim-histghg', 'ssp126', 'histSST-
↪ piNTCF', 'piClim-4xCO2',
'piClim-2xss', 'abrupt-0p5xCO2', 'piClim-BC', 'faf-heat', 'ssp370-lowNTCF', 'abrupt-
↪ 2xCO2', 'histSST', 'piClim-ghg',
'piClim-histaer', 'faf-water', 'ssp370SST-lowAer']

Number of files: 782

Available realizations: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]

<=====Exiting now!=====>
```

**Allows multiple inputs:** takes comma (,) separated multiple items

```
acccmip6 -o S -m MIROC6,CanESM,CNRM-CM6-1-HR -v va,ua -e ssp245-aer,hist-piAer
```

**Allows list inputs:** this is useful since the acccmip6 outputs in list format

```
acccmip6 -o S -e ['ssp245-GHG','historical'] -v ['hfls','hfss'] -f mon
```

**Recommended:** Use inputs within a ‘ ‘

```
acccmip6 -o S -m 'MIROC6, CanESM, CNRM-CM6-1-HR' -e '['ssp245-GHG', 'historical']' -f_
↪mon
```

### 3.3.2 Optional arguments 2

- `-c`: ‘yes’ to check the inputs under optional argumetns 1. Searches through the servers and checks whether the input items are available or not.

#### General usage

```
acccmip6 -o S -m MIROC6 -v vas -f mon -r atmos -c yes
```

#### General output

```
TIPS: If you are not sure about what you are looking for use CMIP6DB module
      to look for currently available models/experiments/variables and so on . . .

Checking for MIROC6 model in CMIP6 database . . .
Found: MIROC6 model.

Checking for vas variable in CMIP6 database . . .
Found: vas variable.

Checking for mon frequency in CMIP6 database . . .
Found: mon frequency.

Checking for atmos realm in CMIP6 database . . .
Found: atmos realm.

Currently available models based on your search:

['MIROC6']

Currently available variables based on your search:

['vas']

Currently available experiments based on your search:

['ssp370SST-lowBC', 'amip-future4K', 'abrupt-4xCO2', 'piClim-aer', 'piClim-histall',
↪'dcpA-hindcast', 'faf-all',
'amip-4xCO2', 'hist-CO2', 'histSST-piNTCF', 'piClim-histaer', 'piClim-2xdust',
↪'lpctCO2', 'histSST-piAer', 'ssp245',
'piClim-2xfire', 'faf-water', 'piClim-histghg', 'piClim-2xss', 'ssp245-stratO3', 'amip
↪', 'hist-aer', 'ssp245-GHG',
'piClim-OC', 'ssp370', 'faf-heat', 'piClim-NTCF', 'ssp370SST-lowAer', 'amip-hist',
↪'piClim-4xCO2', 'piClim-control',
'hst-stratO3', 'piClim-ghg', 'piClim-lu', 'histSST', 'faf-stress', 'ssp585', 'abrupt-
↪0p5xCO2', 'ssp370-lowNTCF',
'ssp119', 'piClim-BC', 'amip-p4K', 'ssp245-aer', 'ssp370SST', 'abrupt-2xCO2', 'hist-
↪GHG', 'dcpA-assim', 'hist-piNTCF',
```

(continues on next page)

(continued from previous page)

```
'faf-passiveheat', 'piControl', 'ssp534-over', 'ssp126', 'hist-piAer', 'piClim-anthro
→', 'ssp460', 'historical',
'piClim-SO2', 'ssp434']
```

Number of files: 782

Available realizations: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]

<=====Exiting now!=====>

**Get suggestions:** use the `-c` argument when in doubts whether any input item is available in the current database.

```
acccmip6 -o S -m IPSL -c yes
```

### Output with suggestions

TIPS: If you are not sure about what you are looking for use CMIP6DB module  
to look for currently available models/experiments/variables and so on . . .

Checking for IPSL model in CMIP6 database . . .

Cannot find model.

Looking for other options . . .

Option 1 IPSL-CM6A-ATM-HR

Option 2 IPSL-CM6A-LR

Did you mean any of the above?

### 3.3.3 Optional arguments 3

- `-desc`: 'yes' to get the description of the experiments searched with `-e` argument. The descriptions comes at the end of the general search results.

#### general usage

```
acccmip6 -o S -e hist-piNTCF,hist-piAer -v vas -m MIROC6 -desc yes
```

#### General output

TIPS: Use the check (`-c`) argument to check your inputs.

Currently available models based on your search:

```
['MIROC6']
```

(continues on next page)

(continued from previous page)

Currently available variables based on your search:

```
['vas']
```

Currently available experiments based on your search:

```
['hist-piNTCF', 'hist-piAer']
```

Number of files: 12

Available realizations: [1, 2, 3]

< < < Here are the experiment descriptions > > >

hist-piNTCF:

Impose historical WMGHG **and** halocarbon concentrations. Near Term Climate Forcers (NTCFs: methane, tropospheric ozone **and** aerosols, **and** their precursors), to be fixed at 1850 emission levels. These simulations parallel the "CMIP6 historical", **and** differ only by fixing the anthropogenic emissions **or** concentrations of a specified **class of** species. All other forcing agents must evolve **as in** "CMIP6\_↪historical".

hist-piAer:

Historical WMGHG **and** Halocarbon concentrations. Historical ozone precursor emissions (e.g. NOx). Aerosols **and** aerosol precursors fixed at 1850 emission levels.

### 3.3.4 Extra arguments

Use these optional arguments with in addition to **Optional arguments**.

- `-rlzn`: select a realization
- `-skip`: skip items during download
- `-time`: 'yes' print out all available time periods

#### general usage

```
acccmip6.py -o S -v pr -e historical -f mon -m NorESM2-LM -time yes
```

**general output:** Available time periods are 1850-1859, 1860-1869 and so on until 2010-2014.

TIPS: Use the check (`-c`) argument to check your inputs.

Currently available models based on your search:

```
['NorESM2-LM']
```

Currently available variables based on your search:

(continues on next page)



(continued from previous page)

```
['pr']

Currently available experiments based on your search:

['historical']

Number of files: 51

Available realizations: [1, 2, 3]

< < < Data available for these time periods > > >

['1850', '1860', '1870', '1880', '1890', '1900', '1910', '1920', '1930', '1940', '1950',
↪', '1960', '1970', '1980', '1990', '2000', '2010']
```

- `-yr` : select data for a specific time period

**general usage:** To download only the first 10 years (e.g. 1850-1859), use `-yr 10` or the last 5 years (e.g. 2010-2014), use `-yr -5`.

```
python acccmip6.py -o S -v pr -e historical -f mon -m NorESM2-LM -yr -5 -time yes
```

**general output:** Note that the number of files is reduced from 51 to 3! This is very useful if the the data are as frequently chunked as this particular example. `-yr 11` will download the 1850-1859 and 1860-1869 chunks. So, even though the user needs only the first 11 years of data, the package is limited to download whatever chunk is available to fulfill that 11 years of time period.

TIPS: Use the check (`-c`) argument to check your inputs.

```
Currently available models based on your search:

['NorESM2-LM']

Currently available variables based on your search:

['pr']

Currently available experiments based on your search:

['historical']

Number of files: 3

Available realizations: [1, 2, 3]
```

(continues on next page)

(continued from previous page)

```
< < < Data available for these time periods > > >
['2010']
```

**acccmip6** package accesses all publicly available CMIP6 data servers. Currently available servers -

- USA, PCMDI/LLNL (California)
- France, IPSL
- Germany, DKRZ
- UK, CEDA

### 3.4 CMIP6DB module

To activate **CMIP6DB** module use `acccmip6 -o M`. This module allows all **Optinal arguments 1**. However, the input is always show.

This module can be used to search all currently available models, experiemnts, variables, frequecies and realms in the CMIP6 database.

**Usage:** without any arguments

```
acccmip6 -o M
```

#### output

```
<<You are using the CMIP6DB module now>>

Use 'show' as the -m/-e/-f/-v/-r argument option instead of the names

Example: acccmip -o M -m show >>> will generate a list of all currently available_
↪models in the CMIP6 database.
```

**Usage:** with argument

```
acccmip6 -o M -m show
```

#### output

```
<<You are using the CMIP6DB module now>>

Use 'show' as the -m/-e/-f/-v/-r argument option instead of the names

Example: acccmip -o M -m show >>> will generate a list of all currently available_
↪models in the CMIP6 database.

Currently 49 models has outputs!

Available models:

['AWI-CM-1-1-MR', 'BCC-CSM2-MR', 'BCC-ESM1', 'CAMS-CSM1-0', 'CESM2', 'CESM2-WACCM',
↪'CMCC-CM2-HR4', 'CMCC-CM2-VHR4', 'CNRM-CM6-1', 'CNRM-CM6-1-HR',
'CNRM-ESM2-1', 'CanESM5', 'E3SM-1-0', 'EC-Earth3', 'EC-Earth3-Veg', 'ECMWF-IFS-HR',
↪'ECMWF-IFS-LR', 'FGOALS-f3-L', 'FGOALS-g3', 'FIO-ESM-2-0',
```

(continues on next page)

(continued from previous page)

```
'GFDL-AM4', 'GFDL-CM4', 'GFDL-CM4C192', 'GFDL-ESM4', 'GFDL-OM4p5B', 'GISS-E2-1-G',
↪ 'GISS-E2-1-G-CC', 'GISS-E2-1-H', 'HadGEM3-GC31-HM', 'HadGEM3-GC31-LL',
'HadGEM3-GC31-LM', 'HadGEM3-GC31-MM', 'IITM-ESM', 'IPSL-CM6A-ATM-HR', 'IPSL-CM6A-LR',
↪ 'MCM-UA-1-0', 'MIROC-ES2L', 'MIROC6', 'MPI-ESM1-2-HR',
'MRI-AGCM3-2-H', 'MRI-AGCM3-2-S', 'MRI-ESM2-0', 'NESM3', 'NICAM16-7S', 'NICAM16-8S',
↪ 'NICAM16-9S', 'NorESM2-LM', 'SAM0-UNICON', 'UKESM1-0-LL']
```

## 3.5 Downloading data

For all kinds of download use `acccmip6 -o D` and then add in the optional arguments.

All arguments listed in **Optional arguments 1** and **Optional arguments 2** are allowed for downloading the data.

### 3.5.1 Extra arguments

Use these optional arguments with in addition to **Optional arguments 1** and **Optional arguments 2**.

- `-dir`: select directory. If kept blank, `acccmip6` will ask for a directory. With no inputs, download will continue in CMIP6 directory.
- `-rlzn`: select realization
- `-skip`: skip items during download
- `-yr`: select data for a specific time period

#### General usage

```
acccmip6 -o D -e hist-piNTCF -v vas -m MIROC6
```

#### General output

```
Finding server . . .

Current directory: /mnt/d/Work/acccmip6_download

Default directory: /mnt/d/Work/acccmip6_download/CMIP6
Please specify a directory here:
```

**with no input:** creates directory/folder if not available

```
Finding server . . .

Current directory: /mnt/d/Work/acccmip6_download

Default directory: /mnt/d/Work/acccmip6_download/CMIP6
Please specify a directory here:

Selected directory: /mnt/d/Work/acccmip6_download/CMIP6

/mnt/d/Work/acccmip6_download/CMIP6 doesn't exist. Creating one...

vas_Amon_MIROC6_hist-piNTCF_rlilplfl_gn_185001-194912.nc is available!
```

(continues on next page)

(continued from previous page)

```
Downloading 36% || 44/120MB 25.62 MB/s
```

**with input directory:** creates directory/folder if not available

```
Finding server . . .

Current directory: /mnt/d/Work/acccmip6_download

Default directory: /mnt/d/Work/acccmip6_download/CMIP6
Please specify a directory here:
/mnt/d/Work/acccmip6_download/download_here
Selected directory: /mnt/d/Work/acccmip6_download/download_here

/mnt/d/Work/acccmip6_download/download_here doesn't exist. Creating one...

vas_Amon_MIROC6_hist-piNTCF_r1i1plf1_gn_185001-194912.nc is available!

Downloading 41% || 50/120MB 29.80 MB/s
```

**with -dir argument**

```
acccmip6 -o D -e hist-piNTCF -v vas -m MIROC6 -dir /mnt/d/Work/acccmip6_download/
↳download_here
```

**output**

```
Finding server . . .
creating /mnt/d/Work/acccmip6_download/download_here

vas_Amon_MIROC6_hist-piNTCF_r1i1plf1_gn_185001-194912.nc is available!

Downloading 41% || 50/120MB 28.79 MB/s
```

**selecting specific realization**

```
acccmip6 -o D -e hist-piNTCF -v vas -m MIROC6 -dir /mnt/d/Work/acccmip6_download/
↳download_here -rlzn 2
```

**output**

```
Finding server . . .

vas_Amon_MIROC6_hist-piNTCF_r2i1plf1_gn_185001-194912.nc is available!

Downloading 33% || 40/120MB 26.54 MB/s
```

**skipping items:** skip any item (models/experiments/variables) you don't want to download

```
acccmip6 -o D -e ssp245-aer,ssp245,hist-aer -MIROC6,CanESM5 -f mon -r atmos -skip_
↳hist-aer,ua,va,zg
```

## 3.6 Q & A

### What is the recommended use for the package?

- Be more specific in your search for better result. Try to include -v, -m, -e, -f options in your search.
- If the user is interested to download 30 models, for 4 different variables, with 5 different experiments; it is suggested to chunk them out into... let's say, 30 models for 1 variable, and 1 experiment.
- Users can always run a separate chunk (for a different variable/experiment) on different terminal tabs (maximum 4 connections are allowed through ESGF).

### Can I use a bash/loop script to automate the process?

- Yes! This is why the CLI system works really well.

A simple bash script may look like this

```
#!/bin/bash

for i in CNRM-ESM2-1 HadGEM3-GC31-LL CESM2 GFDL-ESM4 INM-CM5-0 CESM2-WACCM-FV2 BCC-
↳ESM1 INM-CM4-8 MRI-ESM2-0 NorESM2-LM NorESM2-MM GFDL-CM4 MIROC-ES2L MIROC

do
    acccmip6 -o D -v mmrso4,mmrbc,mmroa,mmrdust,mmrss,mmrpm2p5 -f mon -e_
↳piControl -m $i -yr 50 -dir /download_dir/
done
```

### What can I do when my download is stuck or slow?

- The package has a built-in method that would skip to the next file if the current download speed is below 0.08 MB/s or the download is stuck (no response from the server) for more than 5 minutes.
- Impatient users can use `ctrl+c` to end the process and re-run the same command. It will skip over the already downloaded files.
- If there's one model or variable that is giving you a hard time, you can always skip that item using the `-skip` argument.

Download stuck or slow

```
vas_Amon_CNRM-ESM2-1_historical_r1i1p1f2_gr_185001-201412.nc is available!

Downloading 5% || 10/195MB 0.08 MB/s
download too slow! retrying...
Removing file . . .
```

### During download, some files were skipped because of slow downloading speed. How can I retrieve those files?

- You will get a message about how many files were downloaded (e.g. n out of m files). You can download those (m-n) files just by re-running the command.

### Is it possible to select a specific variant label? For instance, can I select only 'r1i1p1f1' out of 'r1i1p1f1', 'r1i1p1f2' and 'r1i1p2f1'?

- Yes, the package is designed to download all available realizations with `-rlzn` option regardless of the variant. If you do not want to download a certain variant you can always use the `'skip'` option.
- For instance, if you do not want the latter 2 variants in the example question, add `-skip 1p2, 1f2` in your command line.

- This is also usable for any other cases, such as, choosing the grid option. For instance, using `-skip _gr_` will only download available `gn` (native grid) data.

**Can I download or search data within a specific time period?**

- Yes, using the `-yr` option. Checkout the extra arguments.

**Some files are not downloading after showing a 401 Unauthorized error! How can I download those files?**

- This is a common server issue found in CMIPs. `acccmip6` produces an error and skips over these files by producing

```
<<401 Unauthorized: restricted access!!>>

From ESGF: Before you can download this data, you have to join a data access_
↪control group
since acknowledgement of a policy is a condition for this data download.

Requires registration/manual download . . . :(
```

- It stores all unresolved files to a `wget` script in the same download directory once the download finishes.
- You can then use that script with your openid and password with `-H` option (e.g. `./wget_script -H`).

## 3.7 Authors

**Maintainer**

- Taufiq Hassan Mozumder <[taufiq.hassanmozumder@email.ucr.edu](mailto:taufiq.hassanmozumder@email.ucr.edu)>

**Contributors**

None yet.

## 3.8 License

The MIT License (MIT)

Copyright (c) 2019 Taufiq Hassan Mozumder

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.