

Wikiprint Book

Title: 1. Data

Subject: Igcmg_doc - Doc/DataAnalyse

Version: 26

Date: 06/29/24 09:50:03

Table of Content

Data and Analyse	3
1. Data	3
1.1. CMIP archive	3
1.1.1. Start with CMIP archives	3
1.1.2. Access CMIP archives	3
1.1.3. Dive into CMIP archives at IPSL	3
1.1.4. What's a dataset?	3
1.2. THREDDS access (aka DODS service)	3
2. Analyse	4

Data and Analyse

1. Data

1.1. CMIP archive

1.1.1. Start with CMIP archives

We highly recommend to read the user's guide of CMIP projects available at: <https://pcmdi.llnl.gov/mips/>

1.1.2. Access CMIP archives

CMIP archives are available through the ESPRI platform including CICALAD and ClimServ machines. You can create an account on ESPRI mesocentre to transparently log to CICALAD or ClimServ at <https://meso-account.ipsl.fr/>.

1.1.3. Dive into CMIP archives at IPSL

The IPSL-CM model produces CMIP-compliant files. The CMIP convention include the CF convention with additional rules from the WIP panel (ex. a CMIP file must describe only one variable). Once the IPSL-CM data are quality checked, the CMIP-compliant files are migrated to a dedicated filesystem at TGCC following the CMIP directory structure. This part of the TGCC filesystem is read-only mounted on the ESPRI mesocentre. On CICALAD and ClimServ, the CMIP data archives are then available under :

```
/bdd/CMIP6
```

```
/bdd/CMIP5
```

The CMIP data from the other GCMs are downloaded at the IDRIS and IPSL. They are organized in the same CMIP directory structure and transparently available in the same `/bdd` root.

Example: `/bdd/CMIP5/output/IPSL` provides access to the **whole** IPSL-CM data production for CMIP5 exercice. `/bdd/CMIP5/output/BCC` provides access to BCC data requested by the IPSL community.

We adopted an "on-demand" process to add CMIP data from other GCM. Please send your request at glipsl@ipsl.fr with the following template:

```
cat < EOF > my_template.txt
project=CMIP5
experiments=historical amip
models=IPSL-CM5A-LR CNRM-CM5
ensembles=all
variables[atmos][3hr]=cltc tas
variables[land][fx]=sftgif
variables[seaIce][mon]=sic evap
EOF
```

1.1.4. What's a dataset?

A "dataset" (as defined by ESGF) is **one** version of a data set resulting from a single simulation (i.e., characterized by a unique option of each CMIP facet before the version such as the institute, the model, the domain, the experiment, the frequency, the ensemble, etc.).

Examples: CMIP5 dataset: `CMIP5/output1/IPSL/IPSL-CM5A-LR/ 1pctCO2/mon/atmos/Amon/r1i1p1/v201110427`

A dataset is the finest granularity for ESGF publication.

1.2. THREDDS access (aka DODS service)

The IPSL ESGF node includes a THREDDS data server which replaces the deprecated DODS server. This service automatically provides free access to parts of filesystems or remotely-mounted filesystems. From TGCC, the shared space `/ccc/work/cont003/thredds/$LOGIN` is expose at <https://vesg.ipsl.upmc.fr/thredds/catalog/work/catalog.html>.

Be careful that the THREDDS spaces at TGCC should only includes hard-links of data from your project spaces on WORK. The THREDDS folders were not designed to host deep copies of your files because of high constraints on storage and inodes.

2. Analyse