

# Accessing CMIP5/6 data at NCI

Aurel Moise (BOM), Paola Petrelli (CLEX), Kate Snow (NCI), Tim Erwin (CSIRO), Claire Trenham (CSIRO), Louise Wilson (BOM), Francois Delage (BOM)

During 2017/8 there have been ongoing discussions about the re-structure of directories and projects at NCI as well as how to best prepare ourselves for the onslaught of CMIP6 data. Thanks to a small technical group with representatives from CSIRO, Bureau, CLEX and NCI, we have progressed work on facilitating better access to the data. Additionally, the DEVL project (Data Enhanced Virtual Laboratory – an ARDC funded project which is led by NCI but also involves CSIRO, Bureau and CLEX contributions) has actual funding to support these (and more) activities.

We would like to start communicating some of the outcomes of these activities to the NESP ESCC Hub researchers. We start with this article on accessing CMIP data. A lot of the information here is also available from the new [CMIP Community site](#). As part of NCI's preparation for CMIP6 they have improved the way climate reference datasets are managed at NCI, and have new ways for you to find and access the Earth System Grid Federation (ESGF) data. The new data collection project details and the schedule for **transitioning from the old data in ua6** is outlined below.

## Changes to how you can access ESGF published data at NCI:

A new official replica of CMIP5 data has been downloaded to the *project al33*. This provides a more complete and robust data collection to work from and we will thus begin the process of decommissioning some of the data in ua6. As previously advised, the Australian published data is available from rr3. A *new project, oi10*, has been set up in preparation for the new CMIP6 replicated data.

These projects are available for use, and data can be searched and accessed using CleF: <https://clef.readthedocs.io/en/latest>. In order to use the data, you will also need to be a member of the project. Simply click on the link provided in the following table for project registration.

| Data SubCollection Name                                | Current project location    | Join new project to access |
|--|-----------------------------|----------------------------|
| <a href="#">ESGF CMIP5 Replication Data</a>            | ua6/ unofficial-ESG-replica | <a href="#">al33</a>       |
|  | ua6/DRSv2                   |                            |
|  | ua6/NCI_replica_tmp         |                            |
| <a href="#">ESGF CMIP5 Australian Data Publication</a> | ua6/authoritative           | <a href="#">rr3</a>        |
| <a href="#">ESGF CMIP6 Replication Data</a>            | N/A                         | <a href="#">oi10</a>       |

**What is CleF?** To search for available variables under CMIP5 and CMIP6, we recommend using the CleF Climate Finder tool developed by ARCCSS and CLEX – a python module designed for ease of searching and accessing data at NCI. Documentation for CleF is available [here](#) and membership of any of the NCI CMIP projects oi10, al33, rr3 or ua6 is required for use (if you are a new member it may take up to 1 day for access to the NCI database used by CleF to be granted).

**Accessing CMIP data:** Access to the CMIP and related data is available via three key methods (summarised below). Method 1 is the preferred method for you to use. Information on which CMIP5 and CMIP6 data is on NCI (and what will be downloaded in case of CMIP6) can be found [here](#). Please always check the holdings on NCI first before searching the global Earth System Grid (ESG).

- (1) Access on the NCI filesystem
  - a) Users with an NCI login may obtain direct access to the data replicated locally on the NCI filesystem and available with both the HPC facilities of Raijin and the Virtual Desktop Infrastructure (VDI).
  - b) Users are encouraged to use **CleF** - a python module developed for ease of searching and accessing CMIP data at NCI.
  
- (2) The ESGF data portal
  - a) [https://esgf.nci.org.au/projects/esgf\\_nci/](https://esgf.nci.org.au/projects/esgf_nci/)
  - b) The Earth System Grid Federation (ESGF) data portal provides access to CMIP datasets published by NCI and hosted across the international data nodes. Note, replica data is not currently published on NCI ESGF site. If you cannot use the data on Raijin or the VDI, you may search for the CMIP data from the ESGF site. Note that a login account with an ESGF node is required to download data.
  
- (3) The Geonetwork catalog
  - a) You may search for the CMIP data published at NCI via the Geonetwork metadata catalog, which also provides links to data services on the THREDDS Data Service.
  - b) Australian published CMIP5 data is available on Geonetwork under: [https://geonetwork.nci.org.au/geonetwork/srv/eng/catalog.search#/metadata/f3525\\_9322\\_8600\\_7716](https://geonetwork.nci.org.au/geonetwork/srv/eng/catalog.search#/metadata/f3525_9322_8600_7716).

**Requesting new CMIP data:** Information on how to request CMIP data can be found [here](#) on the [CMIP Community site](#).

There will be further emails about the changes sent by the NESP ESCC Hub Data Management Group closer to implementation, but remember to join the projects to continue to access the data. If you have any questions regarding data access then please contact the Data Manager within your project, or alternatively [Kate.Snow@anu.edu.au](mailto:Kate.Snow@anu.edu.au) from the NCI Data Collections Team or via [help@nci.org.au](mailto:help@nci.org.au).

**Data for evaluation of climate models:** We are in the process of re-structuring the *project rr7* where a lot of our reanalysis and observational data sets are located. Most of you will already have access to this area. We will send more updated information about this restructure soon.