

Rook

A Web Processing Service for the Copernicus
Climate Data Store

Ag Stephens, CEDA

Climate Projection Workshop, 6 May 2021



Rook





Remote Operations On Klimadaten

(The K is not a typo)

<https://rook-wps.readthedocs.io/en/latest/>

Climate Data Store

CMIP6 is now live in CDS ... using rook



Home Search Datasets Applications Toolbox FAQ Live

CMIP6 climate projections

Overview

Download data

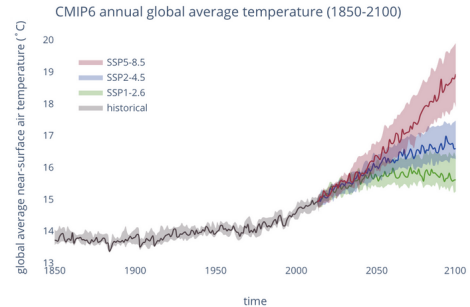
Documentation

This catalogue entry provides daily and monthly global climate projections data from a large number of experiments, models and time periods computed in the framework of the sixth phase of the Coupled Model Intercomparison Project (CMIP6).

CMIP6 data underpins the Intergovernmental Panel on Climate Change 6th Assessment Report. The use of these data is mostly aimed at:

- addressing outstanding scientific questions that arose as part of the IPCC reporting process;
- improving the understanding of the climate system;
- providing estimates of future climate change and related uncertainties;

CMIP6 annual global average temperature (1850-2100)



Year	Historical (°C)	SSP5-8.5 (°C)	SSP2-4.5 (°C)	SSP1-2.6 (°C)
1850	13.8	-	-	-
1900	13.9	-	-	-
1950	14.0	-	-	-
2000	14.1	14.1	14.1	14.1
2050	-	16.5	15.5	15.0
2100	-	19.5	17.5	16.5

Climate Data Store - Download data

CMIP6 climate projections

Overview

Download data

Documentation

Temporal resolution

☐ Monthly

☒ Daily

☐ Fixed (no temporal resolution)

Experiment ?

☒ Historical

☐ SSP1-1.9

☐ SSP1-2.6

☐ SSP4-3.4



☐ SSP5-3.4OS

☐ SSP4-6.0

☐ SSP3-7.0

☐ SSP5-8.5

Climate Data Store - Toolbox



Climate Change
Service

James Varndell [Logout](#)

Your feedback helps us to improve the service

[Home](#) [Search](#) [Datasets](#) [Applications](#) [Your requests](#) [Toolbox](#) [Help&Support](#) [Live](#)

cmip6-tests

Console

History

Your queue

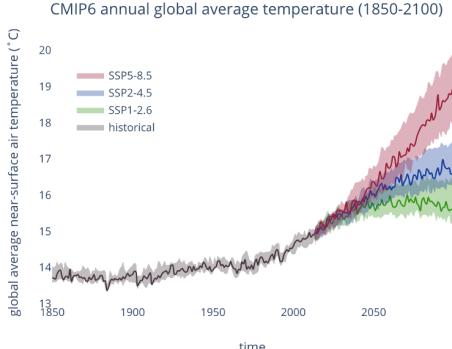
Runtime profile

Layout ▾

Copy Save Run

```
1
2 import itertools
3 import cdstoolbox as ct
4
5
6 MODELS = [
7     'awi_cm_1_1_mr',
8     'cams_csml_0',
9     'cmcc_cm2_sr5',
10    'hadgem3_gc31_ll',
11    'iitm_esm',
12    'kace_1_0_g',
13    'mpi_esml_2_lr',
14    'miroc6',
15    'cnrm_cm6_1_hr',
16 ]
17
18 EXPERIMENTS = [
19     'historical',
20     'ssp1_2_6',
21     'ssp2_4_5',
22     'ssp5_8_5',
23 ]
24
25
26 @ct.application()
27 @ct.output.livefigure()
28 def application():
29     """
30     Simple application to retrieve CMIP6 monthly temperature for three
31     scenarios, calculate the global average and plot the result.
32     """
33     fig = None
34     all_data = {}
35     for experiment, model in itertools.product(EXPERIMENTS, MODELS):
```

CMIP6 annual global average temperature (1850-2100)

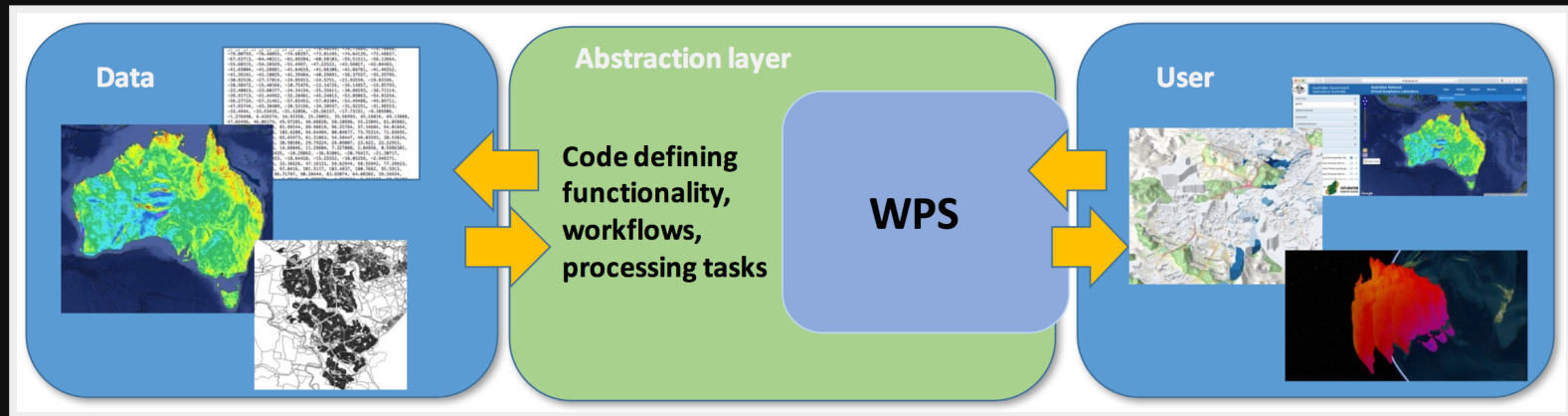


Year	Historical (°C)	SSP1-2.6 (°C)	SSP2-4.5 (°C)	SSP5-8.5 (°C)
1850	13.8	-	-	-
1900	13.9	-	-	-
1950	14.0	-	-	-
2000	14.1	14.1	14.1	14.1
2050	-	15.5	16.5	17.5
2100	-	16.5	18.0	19.5

Climate Data Store - Rook

- The climate data is accessed remotely
- Using rook: download only a subset of the data
- Example: Temperature, 1990, Africa

Web Processing Service



Call a function remotely

Rook - WPS

- An OGC Web Processing Service
- Using PyWPS - GeoPython
- Providing climate data operators as a service
- Used for data reduction: Temperature, 1990, Africa

Rook - Operators

- Subsetting - time, area, level
- Averaging - over dimensions (time, ...)
- Regridding (a pain!)
- ??? - can be extended


Rook - Clisops

- The Python library implementing these operators
- Using xarray - low level library
- Joint effort together with Ouranos, Canada
<https://clisops.readthedocs.io/en/latest/>

Rooki

- Python WPS client - interactive or as library
- Using OWSLib - GeoPython
- Joint effort with Ouranos, Canada
- <https://rooki.readthedocs.io/en/latest/>

Rooki - Notebook

 jupyter
nbviewer

JUPYTERFAQ</>

rooki / notebooks / demo

Run subset by area operation

Rooki calls climate data operations on the **rook** processing service.

```
In [ ]: import os
os.environ['ROOK_URL'] = 'http://rook.dkrz.de/wps'

from rooki import rooki
```

parameters of subset operation

```
In [ ]: rooki.subset?
```

run subset by area

```
In [ ]: resp = rooki.subset(
    collection='c3s-cmip6.CMIP.IPSL.IPSL-CM6A-LR.historical.r1i1p1f1.Amon.rlds.gr.v20180803',
    time='1860-01-01/1980-12-30',
    area='0.,49.,10.,65'
)
resp.ok
```

Deployment - Birdhouse Tools

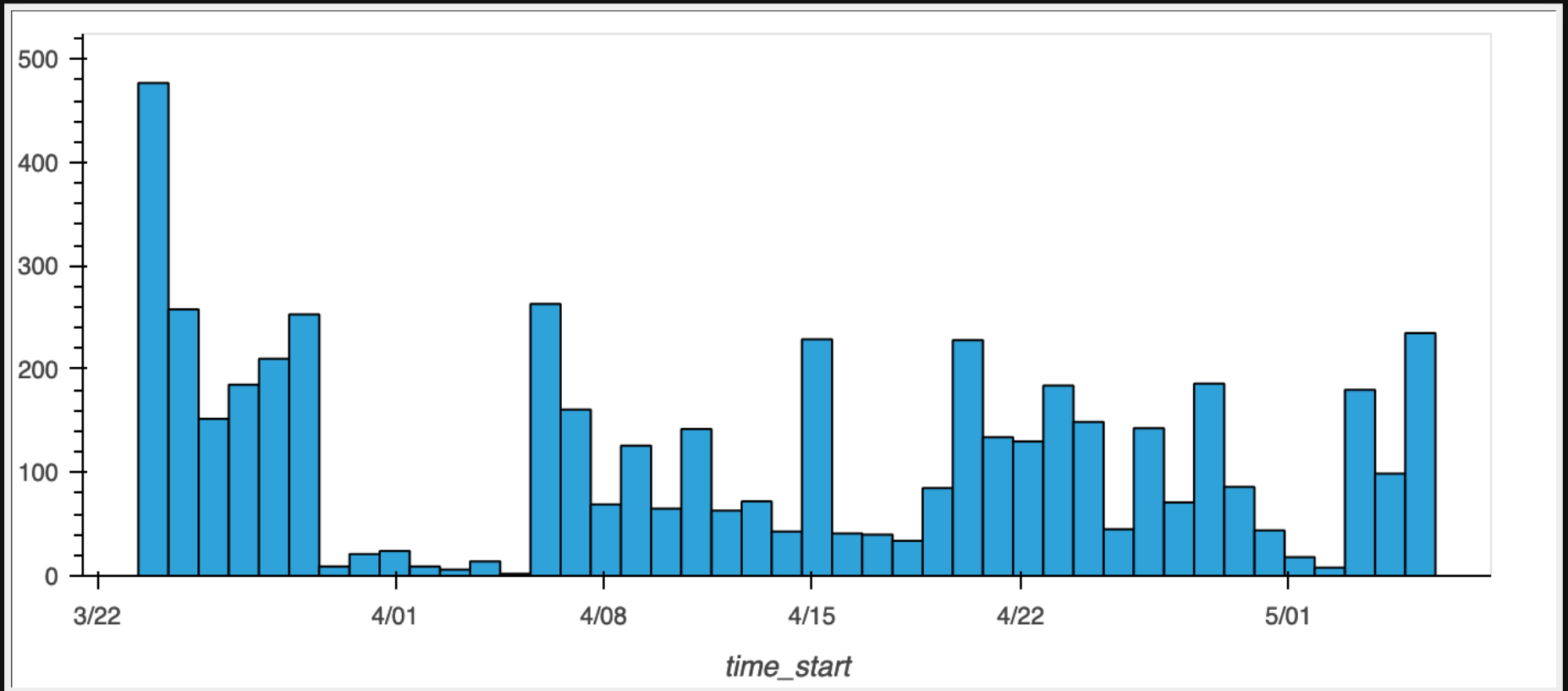
- Rook generated from a Cookiecutter template
- Ansible playbook to roll out on cluster with Slurm scheduler
- Joint effort with Ouranos, Canada

Availability

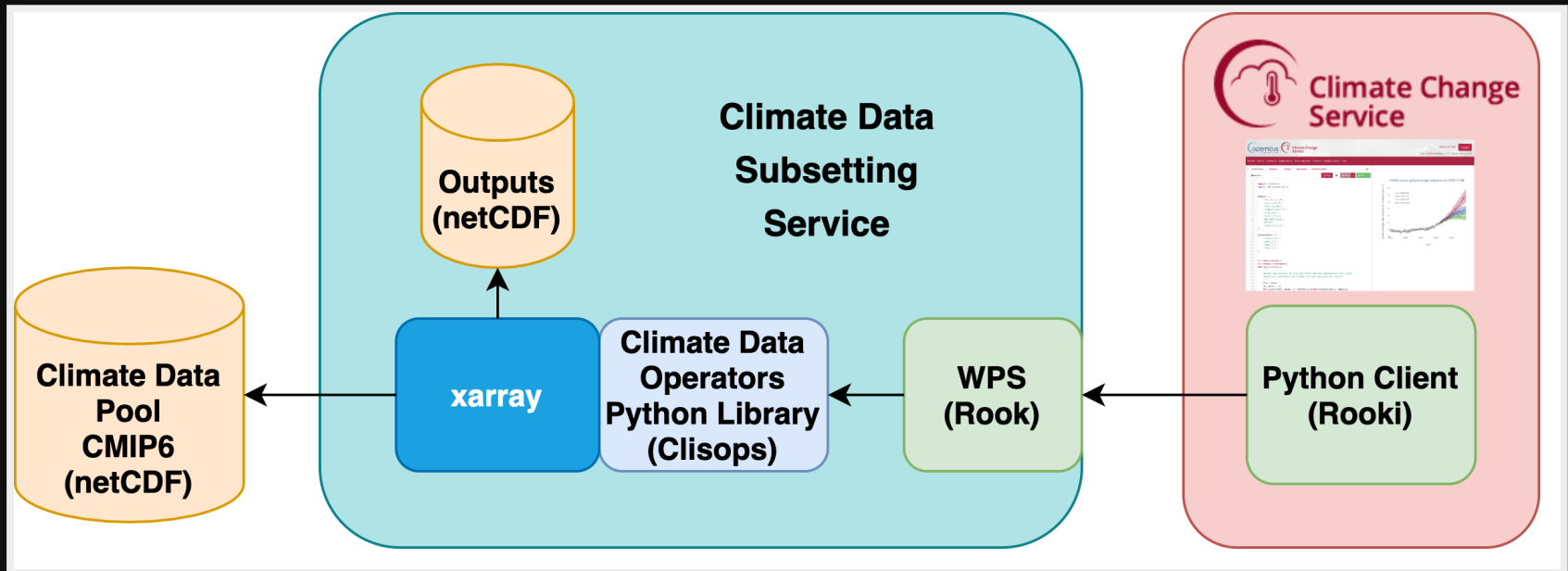
- Data pool is replicated to three sites
- Load-balanced access to rook WPS

Rook - Requests

200 per day ... can serve much more



All together



Status

- Deployed at CEDA and DKRZ
- Used for CMIP6
- Subset (time, area, level) operator
- Original CMIP6 files are downloaded from data nodes

Next steps?

- CMIP6 + Decadal
- CORDEX?
- CMIP5?
- Averaging and Regridding
- Other operations?

Projects

- Copernicus C3S: <https://climate.copernicus.eu/>
- Roocs: <https://roocs.github.io/>
- Birdhouse: <http://bird-house.github.io/>
- GeoPython: <https://geopython.github.io/>

Thanks

Questions?