

From (big)data to information visualization with birdhouse

a collection of Web Processing Services

Session A3: FROM DATA TO INFORMATION
A DISTILLATION DILEMMA

Nils Hempelmann¹, Carsten Ehbrecht², Stephan Kindermann²,
Patrick Brockmann¹, Cathy Nangini¹, Robert Vautard¹

1. Le Laboratoire des Sciences du Climat et de l'Environnement. Saclay France
2. German Climate Computing Center, Hamburg Germany



Climate Data volume grows quickly

**But on user side:
Limited storage/compute capacities**

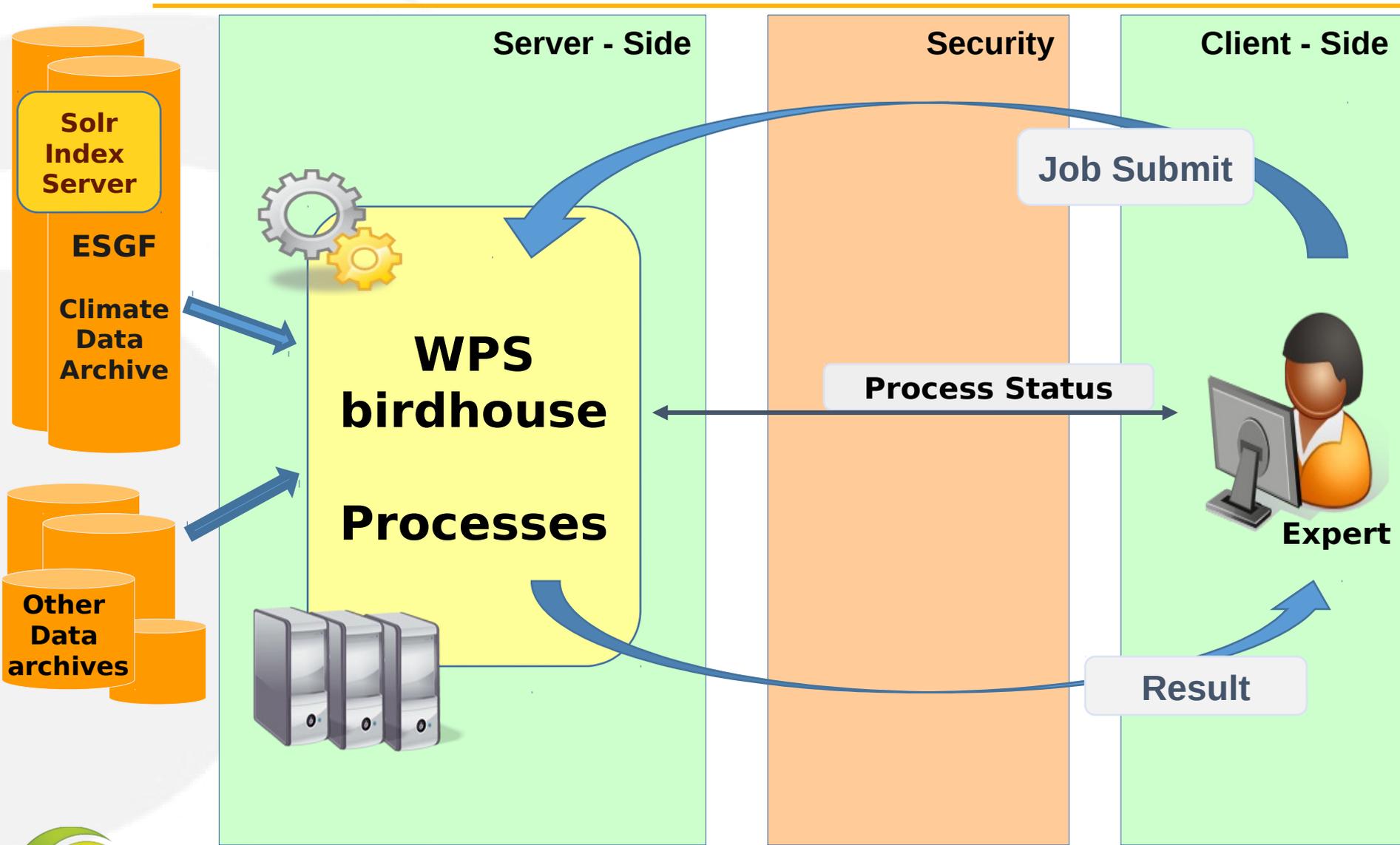
**“download and
process at home”**



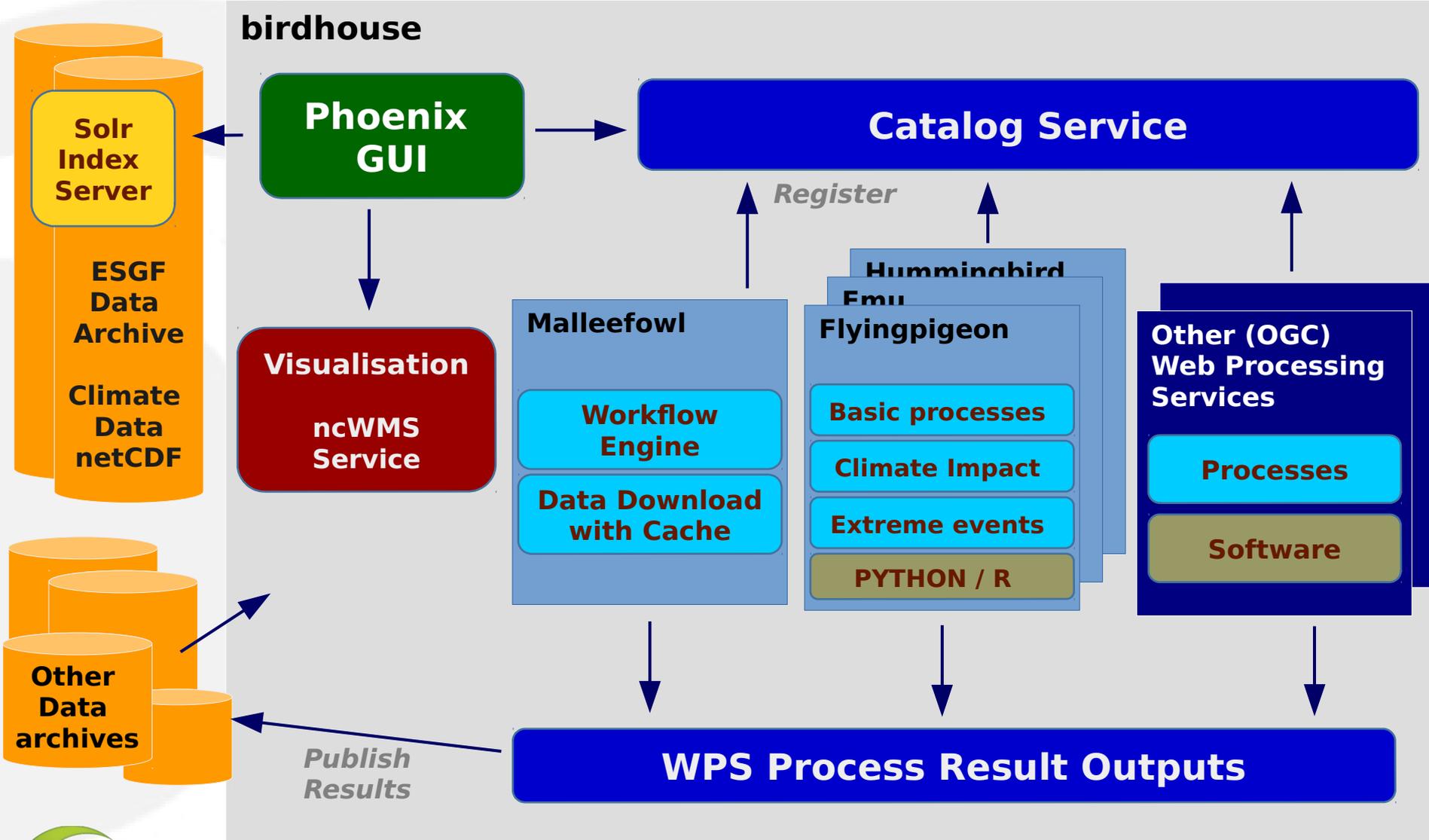
**Data processing
close to archives**

Web Processing Service
trigger computing processes remotely





Organisation



Ecosystem of WPS

WPS Compute Provider
(server side)

↑ Link

↑ Pull

↑ Pull

Other WPS :

Climate4impact
CLIP-C
COWS
Zoo
52° North
...etc...

Phoenix

GUI Web application



Malleefowl

Basic processes



Hummingbird

Quality checks



Flyingpigeon

Climate impact
Extreme events



Local repository

Barnowl

internal processes



↑ Pull

↑ Pull / Push

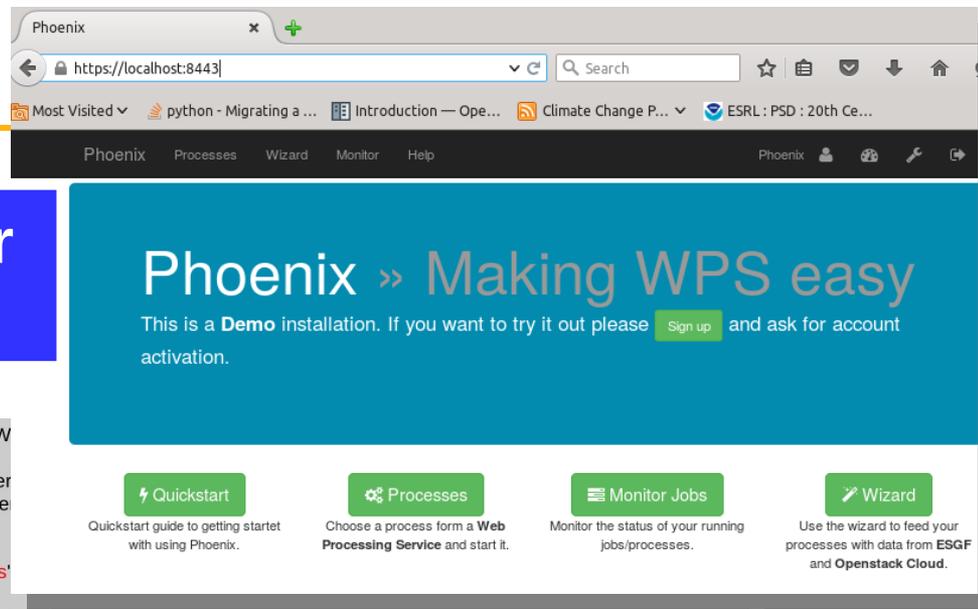
localhost users

Developer



Client Side

Web Browser GUI



Phoenix >> Making WPS easy

This is a **Demo** installation. If you want to try it out please [Sign up](#) and ask for account activation.

- Quickstart**: Quickstart guide to getting started with using Phoenix.
- Processes**: Choose a process from a **Web Processing Service** and start it.
- Monitor Jobs**: Monitor the status of your running jobs/processes.
- Wizard**: Use the wizard to feed your processes with data from **ESGF** and **Openstack Cloud**.

Python Script

```
from owslib.wps import WebProcessingService

wps = WebProcessingService('http://localhost:8443')

output = wps.execute(
    identifier="niceprocess",
    outputs=[
        ("parameter_1", "argument"),
        ("parameter_2", "42"),
        # ("parameter_3", "0.987"), # use the default value
        ("file_identifier", "https://thredds/fileServer1/test/file1.nc"),
        ("file_identifier", "https://thredds/fileServer1/test/file2.nc"),
        ("file_identifier", "https://thredds/fileServer2/test/file3.nc")],
    output=("output", True))

# time for a coffee

for o in execute.processOutputs:
    print o.reference
```

```
https://mouflon.dkrz.de:8090/wpsoutputs/flyingpigeon/output_graphic-697dee76-d722-93ae-9789bf75cf44.png
https://mouflon.dkrz.de:8090/wpsoutputs/flyingpigeon/output_netCDF-697dee76-d722-93ae-9789bf75cf44.nc
https://mouflon.dkrz.de:8090/wpsoutputs/flyingpigeon/output_text-697dee76-d722-93ae-9789bf75cf44.txt
```

```
[nhempel@lsce3199 ~]$ export WPS_SERVER=http://localhost:8443
[nhempel@lsce3199 ~]$ birdy -h

usage: birdy [<options>] <command> [<args>]

Options:
  -h, --help            show this help message and exit
  -v, --verbose          enable debug mode
```

```
command:
  List of available commands (wps processes)

(visualisation,sdm,segetalflora,indices_single,subset_countries,eobs_to_cordex,ensembleRobustness,analog,fetch)
Run "birdy <command> -h" to get additional help.

visualisation  Visualisation of netcdf files:          Just testing a nice script to visualise some variables
sdm             Species distribution model:           Species distribution model
segetalflora   Segetal Flora:                        Species biodiversity of segetal flora. Input files: variable:tas , domain: EUR-11 or EUR-44

indices_single Calculation of climate indice (single variable): This process calculates climate indices based on one single variable.
subset_countries Subset netCDF files:           This process returns only the given polygon from input netCDF files.
eobs_to_cordex EObs to CORDEX:                        downloads EObs data in adapted CORDEX format
ensembleRobustness Calculation of the robustness of an ensemble: Calculates the robustness as the ratio of noise to signal in an ensemble of timeseries

analog         Days with analog pressure pattern:    Search for day with analog pressure pattern
fetch         Download Resources:                    This process downloads resources (limited to 50GB) to the local file system and returns a textfile with appropriate paths
```

Terminal Call



ICRC-C

DKRZ

Eucleia



birdhouse

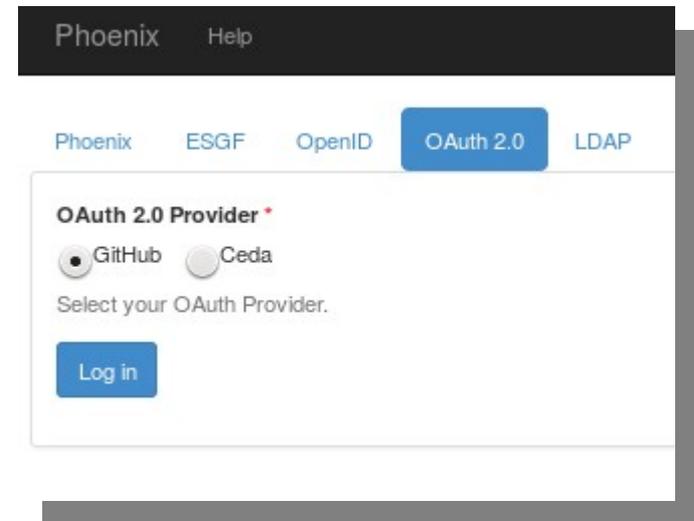
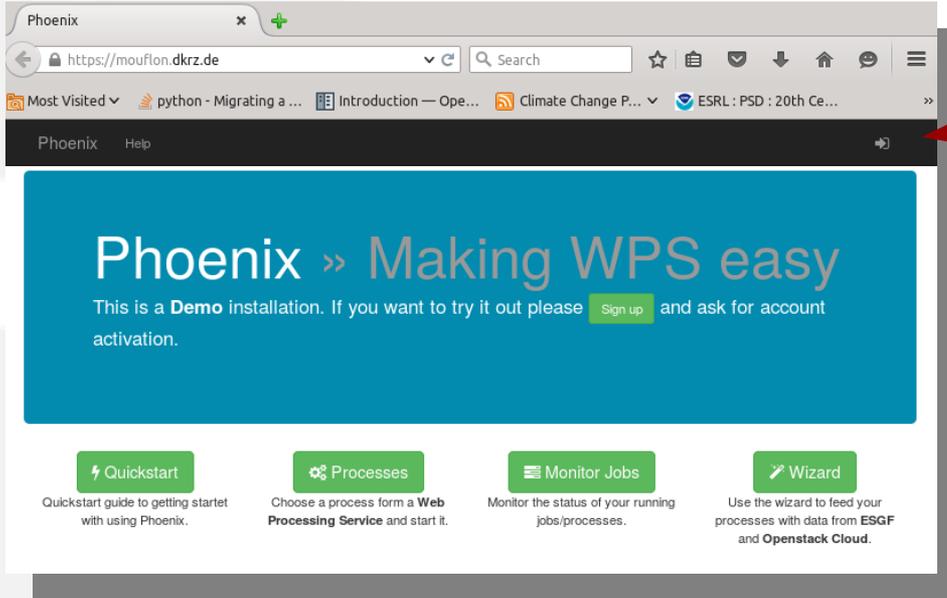
- Based on Open Source
- Open Geospatial Consortium (OGC) Standards
- Climate Data processing

- **<https://github.com/bird-house>**
- **<http://birdhouse.readthedocs.org/en/latest/>**
- **<https://lists.dkrz.de/mailman/listinfo/wps>**
- **<https://lists.dkrz.de/mailman/listinfo/wps-dev>**

- **DEMO GUI: <https://mouflon.dkrz.de>**



Login



http://pyramid-phoenix.readthedocs.io/en/latest/user_guide.html#login



ICRC-CORDEX 2016



DKRZ

Eucleia



Choose a Favorite

Favorite

ensembleRobustness DEMO

Next

Cancel

Choose a Web Processing Service

Web Processing Service

- Flyingpigeon - Processes for climate data, indices and extrem events
- Hummingbird - WPS processes for general tools used in the climate science
- Emu - WPS processes for testing and demos.
- British Antarctic Survey - Web Processing Service - Meteorological Data - British Antarctic Survey, Cambridge (implementation: 52 North WPS 3.2.0)

Previous

Next

Cancel



Process

- Visualisation of netcdf files - Just testing a nice
- Species distribution model - Species distributic
- Weather Regimes - Weather Regimes based
- Extract Coordinate Points - Extract Timeseries
- Segetal Flora - Species biodiversity of segetal
- Calculation of climate indice (single variable) -
- Calculation of percentile based climate indices
percentils of a referece period.
- Subset netCDF files - This process returns onl
- EOBS to CORDEX - downloads EOBS data ir
- Calculation of the robustness of an ensemble -
- Days with analog pressure pattern - Search fo
- Download Resources - This process downloac

Previous

Next

Cancel



ICRC-CORDEX 2016



DKRZ

Eucleia



Process

- Visualisation of netcdf files - Ju
- Species distribution model - Sp
- Weather Regimes - Weather F
- Extract Coordinate Points - Ext
- Segetal Flora - Species biodive
- Calculation of climate indice (si
- Calculation of percentile based
- percentils of a referece period.
- Subset netCDF files - This proc
- EOBS to CORDEX - downloa
- Calculation of the robustness o
- Days with analog pressure patt
- Download Resources - This pro

[Previous](#)[Next](#)[Cancel](#)[Home](#) / [Wizard](#) / Literal Inputs

Literal inputs of Calculation of the robustness of an ensemble

Method of robustness calculation

Detailed information about the methodes can be found in the documentation

Start Year

Beginn of the analysed period (e.g 1971; if not set, the first consistend year of the ensemble will be taken)

End Year

End of the analysed period (e.g. 2050 if not set, the last consistend year of the ensemble will be taken)

Time slice

Time slice (in years) for robustness reference (default=10))

Variable

Variable to be expected in the input files (Variable will be detected if not set,)

[Previous](#)[Next](#)[Cancel](#)

Choose Data Source

Source

- Earth System Grid (ESGF)
- Swift Cloud
- Thredds Catalog Service
- Local Storage
- Birdhouse Solr Search

Previous

Next

Cancel

ESGF Search *

Datasets found: 4

Search Options

Freetext Search

Your keyword selections

project:CORDEX x

time_frequency:sem x

variable:tas x

domain:EUR-11 x

Categories

data_node

ensemble

experiment_family

institute

model

Keywords: project

CORDEX

Date

Previous

Next

Cancel



My Jobs Public Private All ProcessSucceeded

Process Status

Job List 

<input type="checkbox"/>	Status	Job	Userid	Process	Service	Duration	Finished	Public	Progress
<input type="checkbox"/>	✓	e6b42aaa-183a-11e6-98af-a5ed03284ad0	phoenix@localhost	ioos_cchecker	Hummingbird	0:00:18	4 days ago	<input type="checkbox"/>	<input type="button" value="100%"/>
<input type="checkbox"/>	✓	e495fd88-16c9-11e6-8b3b-13a2c6ede0e4	phoenix@localhost	ultimatequestionprocess	Emu	0:00:12	6 days ago	<input type="checkbox"/>	<input type="button" value="100%"/>
<input type="checkbox"/>	✓	07c2e446-1211-11e6-b10b-955ee593571a	cehbrecht@github.com	sdm	Flyingpigeon	0:13:51	12 days ago	<input checked="" type="checkbox"/>	<input type="button" value="100%"/>
<input type="checkbox"/>	✓	817bdc12-1210-11e6-9355-83b8fa870647	cehbrecht@github.com	sdm	Flyingpigeon	0:06:10	12 days ago	<input type="checkbox"/>	<input type="button" value="100%"/>



Status

ProcessSucceeded

Duration

0:00:48

Finished

less than 1 minute ago

Progress

100%

Status Location

XML

Outputs

Log

```
1 0%: Process workflow accepted
2 0%: processtarted workflow wizard_esgf_search prepared.
3 0%: processtarted esgsearch: status_location=http://local
4 0%: processtarted esgsearch: Process esgsearch accepted
5 10%: processtarted download: status_location=http://local
6 50%: processtarted ensembleRobustness: status_location=htt
2c678fe.xml
7 50%: processtarted ensembleRobustness: Process ensembleRot
8 52%: processtarted ensembleRobustness: processtarted argu
9 100%: PyWPS Process workflow successfully calculated
```



Output files

Outputs

[Log](#)

Output

Value

Sourcefiles

test file with a list of the used input data sets

text/plain



Signal

netCDF file containing calculated change of mean over the timeperiod and ensemble members

application/x-netcdf



Mask for areas with low agreement

netCDF file containing calculated robustness mask

application/x-netcdf



Mask for areas with high agreement

netCDF file containing calculated robustness mask

application/x-netcdf



Graphic

PNG graphic file showing the signal difference with high and low ensemble agreement marked out

image/png



ICRC-CORDEX 2016



DKRZ

Eucleia



Output files

Output	Value
Sourcefiles test file with a list of the used input data sets	text/plain
Signal netCDF file containing calculated change of mean over the timeperiod and ensemble members	application/x-netcdf
Mask for areas with low agreement netCDF file containing calculated robustness mask	application/x-netcdf
Mask for areas with high agreement netCDF file containing calculated robustness mask	application/x-netcdf
Graphic PNG graphic file showing the signal difference with high and low ensemble agreement marked out	image/png

Text File

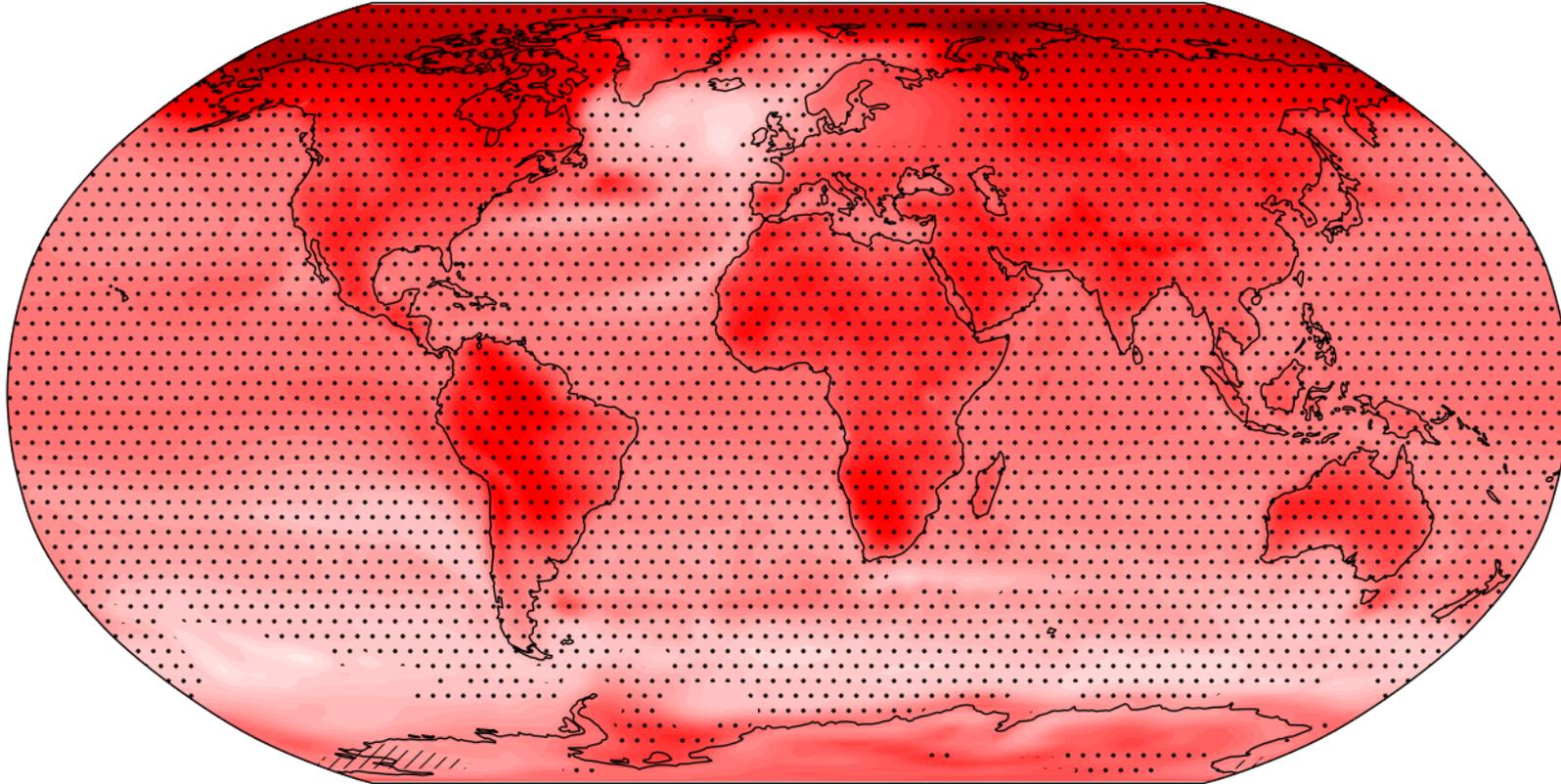
NetCDF Files

png Graphic

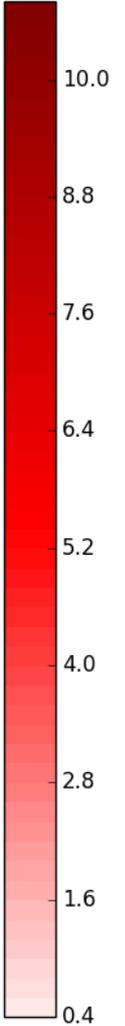


png Graphic

Change of tas (difference of mean 2081-2100 to 2006-2025)



// = low model ensemble agreement
.. = high model ensemble agreement



Web Mapping Server



Dynamic service from outputs/flyingpigeon/output_signal-0b69f1e0-1bba-11e6-9494-1d41b2c678fe.nc

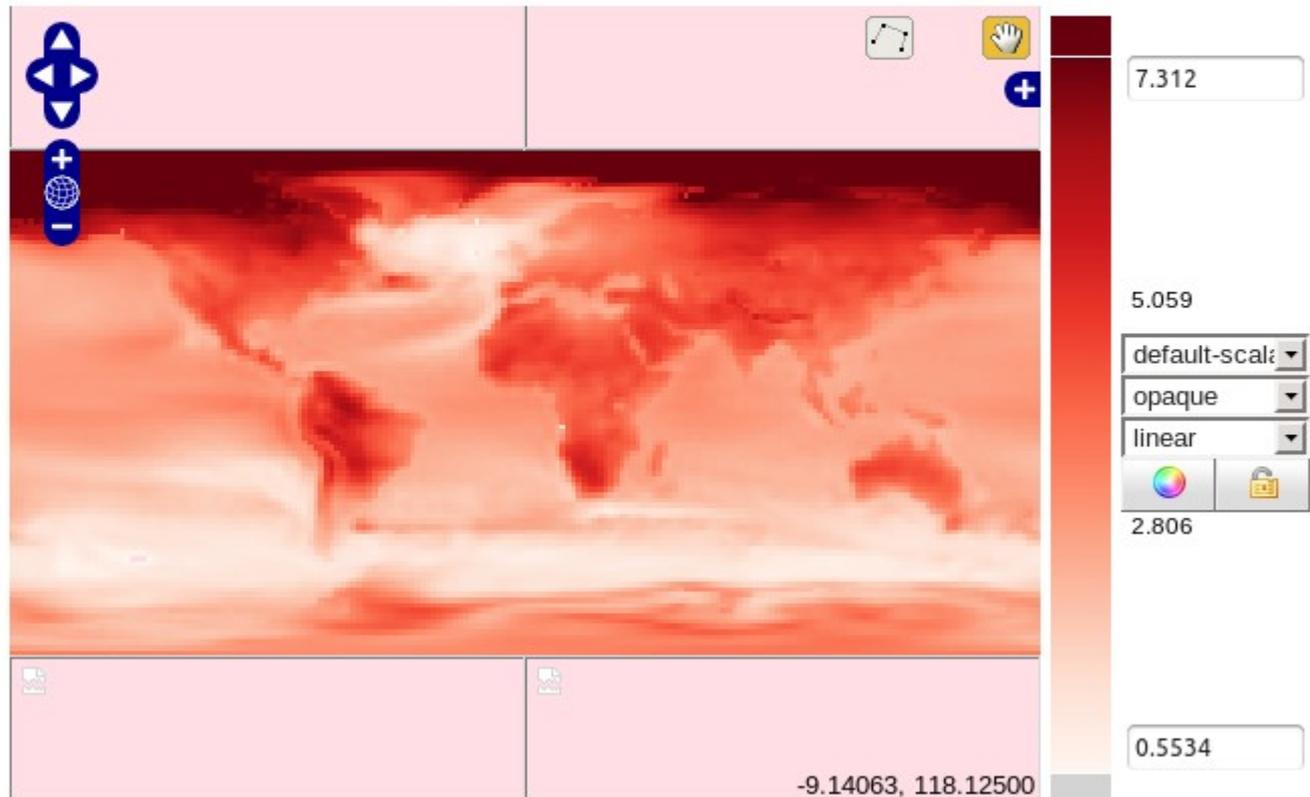
> tas

Units: K

Time: 2091-01-01 00:00:00.000Z

Elevation:

le0-1bba-11e6-9494-1d41b2c678fe.nc



[Open in Google Earth](#)

[Permalink](#)

[Email Link](#)

[Export to PNG](#)



rs



Thanks

Contact :
Nils.Hempelmann@Isce.ipsl.fr



ICRC-CORDEX 2016



DKRZ

*Eucleia*α

