

E CERFACS

ESGF F2F Workshop Washington, DC, December 2016



Server-side computing services provided by IS-ENES through the climate4impact.eu platform

CERFACS, KNMI, University of Cantabria, SMHI, Wageningen University & Research, CNRS-IPSL, CMCC, STFC, DMI, INHGA

Christian Pagé









- Platform for impact researchers to explore climate data and perform analysis
- In-depth documentation and guidance
- Use cases from impact researchers
- Perform calculations / Data processing – WPS





The IS-ENES project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration.







Online on-demand calculations

- C4I Statistical Downscaling REST API
 - Services provided by the University of Cantabria servers
 - Connected to ESGF
 - Friendly user interface on C4I
- C4I Climate Indices
 - All ETCCDI indices and simple statistics available
 - Native Python open-source ICCLIM software (fully validated against R.climdex)
 - Expandable to climate indicators as well



22914

CUMATE CHANGE

ETCCDI: The joint CCI/CLIVAR/JCOMM

Expert Team (ET) on Climate Change Detection and Indices

\bigcirc	Í
\smile	BY





Web Processing Service

Generic WPS UI

- Based on Describe Process XML file.
- Links to basket
- Selections populated dynamically
- Main WPS Processes
 - Climate indices calculation
 - Subsetting over a large time period
 - Regrid and reformat, extract regions
 - Combine climate indicators (CLIPC)
- Wizards
 - Indices calculation
 - Subsetting / file conversions (planned)

limate4impact	× Viewer × +
🕽 🔒 https://clin	matelimpact.eu/impactportal/account/wpsuseprocessor.jsp?process C 🖓 🔍 Zoeken 🕴 🏠 🛕 💆 🔌 🌸
S-CC FRASTRUCTURE FOR THE EUI R EARTH SYSTEM MODELLIN	Exploring climate model data
Home D	ata discovery Downscaling Documentation Help About us Account 1
Account Account » Proces	Basket (609) Processing Monitor jobs (5) Token API Administration page ising > Use a processor > clipc_extractnuts_execute
Overview	
Proces	SOR CLIPC Create statistics per NUTS region Execute
Identifier	clioc extractnuts execute
Abstract	The NUTS extractor calculates statistics for any NetCDF file by extracting geographical areas defined in a GeojSON file. The statistics per geographical area include minimum, maximum, mean and standard deviation. The statistics are presented in a CSV table and a NetCDF file.
Location	https://climate4impact.eu/impactportal/WPS?service=WPS&version=1.0.0&request=describeprocess& identifier=clipc_extractnuts_execute
	Processing: [Starting WCS request 2/2: data](50%)
Processing	insuite
Tocessing	inputs
FILE A (inp application/	intcof
http://open	dap.knmi.nl/knmi/thredds/dodsC/CLIPC/storyline_urbanheat/geojson/NUTS_2010_L0.geojson.nc 🦷 🛢 🛢
File B (inp	min:0 / max: 1 netcdf
http://open	dap.knmi.nl/knm/thredds/dodsC/IS-ENES/TESTSETS/tas_day_EC-EARTH_rcp26_r8i1p1_20060101-20251231 🐺 😝 🛢





Web Processing Service

Job progress can be viewed from anywhere (also in other portals using the services, e.g. CLIPC)

Account	asket (621)	Processing	Monitor jobs (10)	📍 Token API	Administration	n page	
Submitted pro	ocessing jol	os					
Created at	Name		Status location		Progress	View	x
2016-12-06 20:36:36Z	clipc_extractnu	ts_execute	pywps-aec7c400-bbf3-11e6-ac70)-78e3b502c7d8.xml	ready	view	
2016-12-06 20:35:32Z	clipc_simpleinc	licator_execute	pywps-8835e09c-bbf3-11e6-887	4-78e3b502c7d8.xml	failed	view	x
2016-12-06 14:02:50Z	clipc_extractnu	ts_execute	pywps-ac876b46-bbbc-11e6-86e	c-78e3b502c7d8.xml	ready	view	x
2016-12-05 10:32:43Z	clipc_extractnu	ts_execute	pywps-27bdecbe-bad6-11e6-96c	lf-78e3b502c7d8.xml	ready	view	x
2016-12-05 10:32:00Z	clipc_extractnu	ts_execute	pywps-0ddf0c60-bad6-11e6-92d	f-78e3b502c7d8.xml	ready	view	
2016-12-03 19:25:45Z	knmi_advanced	d_combine	pywps-4984287a-b98e-11e6-8d4	b-78e3b502c7d8.xml	ready	view	x
2016-12-03 19:24:13Z	clipc_extractnu	ts_execute	pywps-1266efc6-b98e-11e6-a9a	e-78e3b502c7d8.xml	ready	view	
2016-12-02 13:16:55Z	clipc_combine_	identify	pywps-98511b00-b891-11e6-930)7-78e3b502c7d8.xml	ready	view	x
2016-12-02 12:55:03Z	knmi_advanced	d_combine	pywps-8a6e920e-b88e-11e6-b96	f-78e3b502c7d8.xml	ready	view	x
2016-12-02 10:28:06Z	clipc_extractnu	ts_execute	pywps-03058ed0-b87a-11e6-a30	f-78e3b502c7d8.xml	ready	view	x







Calculations near storage (or away from user's servers)

Dedicated high performance systems

- Replicate all needed data onto one large data service: all local data only
- Expose services through standard protocols (http REST, WPS, OpenDAP, ...)
- Give (selected) users command-line access

Remote Computing Services

- Expose computing services
- Grab remote data and perform calculations on remote high capacity servers/clusters
- Only send results to users

Central Computing Platform

- Portal/platform with standard services (WPS, ...)
- Orchestrate calculations with, when possible, delegations to external services (computations, storage)
- Grab results, and make available to users with more services (visualization, quicklook, download, ...)



















Fig. 13. CEDA's JASMIN analysis platform. JASMIN integrates cloud architecture, container technologies, and virtual machines to improve flexibility and performance and track maintenance.







Tier1 indicators using Ophidia (CLIPC)

Snow on/off – Length of snow season

REDUCTION [time]min

REDUCTION [time]max Snow off









Delegations of some calculations; integration of services



<u>c</u>





Background: Integration of European and International projects

Objective: C4I

Nodes

Delegates selected

processing to future

ESGF Computing



IS-ENES2 (FP7) 23 partners

Infrastructure for the European Network of Earth System Modelling

Access to large climate datasets for climate impact researchers C4I http://climate4impact.eu/

Objective: **C4I** backend for ondemand calculations: **icclim**



CLIPC (FP7) 22 partners

Copernicus Climate Information Platform

Developing icclim for climate indices and indicators calculation Defining metadata standards for automated processing Objective: C4I Delegates selected processing/storage to EUDAT B2 Services



EUDAT2020 (H2020) 35 partners

Collaborative Pan-European infrastructure

Developing Generic Execution Framework: Providing data processing on EGI Federated Cloud using EUDAT B2 Services



ESGF-CWT

Compute Working Team

Worldwide (mainly US-Europe) Collaboration on Data Processing of climate data near data storage

Design of API and Implementation of Processing Capabilities for ESGF





Conclusion and next steps

- ESGF + climate4impact enables impact research
 - Requires ESGF Search API, OpenDAP and THREDDS catalogs
 - Climate4impact processing, search and visualization is a layer over ESGF
 - Support of Downscaling ESGF datasets on the fly
- Climate4impact is flexible due to applied technologies and standards
 - PyWPS with ICCLIM as generic processing framework for climate indices
 - ADAGUC WMS can be used to visualize local and remote files
 - OPeNDAP can be used to access small bits of large files over the internet
 - Many APIs developed to integrate C4I services into CLIPC portal https://dev.knmi.nl/projects/impactportal/wiki/API
- Next steps
 - Implement more use cases from climate impact researchers
 - Climate indices calculation wizard
 - File abstraction: focus on physical parameters







Sustainability Issues

- CLIPC and IS-ENES2 ending
 - Currently 2-3 Coding Sprints per year for development and implementation: needed to improve but not only.
 - Very good core of dev people within IS-ENES2/CLIPC
 - Long to-do list based on user requirements
- What will happen after?
 - New project proposals to further develop and use the C4I platform
 - Limited staff, but we want to continue working together (own institutes funding, related projects)
 - Will need to adapt the platform and portal to future ESGF changes
 - Involvement in ESGF CWT will continue
 - Will need to tackle new datasets, new projects, CMIP6 and +
 - More dissemination needed
 - Support to users
 - Improvements
 - Delegate calculations to ESGF Computing Nodes, External Computing Resources (EGI, Ophidia, etc.)









http://icclim.readthedocs.org/







		Samenvatting			
Rapportageperiode Eerste bezoek Laatste bezoek	Maand jan. 2015 01 jan. 2015 - 01.16 31 jan. 2015 - 23.24				
	Unieke bezoekers	Aantal bezoeken	Pagina's	Hits	Bytes
Bekeken verkeer *	1002	1644 (1.64 bezoeker/bezoeker)	43149 (26.24 Pagina's/bezoek)	142263 (86.53 Hits/bezoek)	48.08 GB (30668.11 KB/bezoek)
Niet-bekeken verkeer *			80506	82125	3.66 GB
a set of the term of the term of the	1 A A A A A A A A A A A A A A A A A A A				

* "Niet bekeken" is verkeer dat gegenereerd werd door robots of wormen, of respons met een speciale HTTP-statuscode.



Maand	Unieke bezoekers	Aantal bezoeken	Pagina's	Hits	Bytes
jan. 2015	1002	1644	43149	142263	48.08 GB
febr. 2015	1059	1855	32931	153179	1.90 GB
mrt. 2015	2186	3923	64718	247721	29.51 GB
april 2015	2255	4430	41533	187479	2.46 GB
mei 2015	2490	5920	429910	525534	7.25 GB
juni 2015	2843	8358	383313	505838	12.60 GB
juli 2015	1806	2819	15752	81211	2.01 GB
aug. 2015	1907	3027	91892	145005	4.20 GB
sept. 2015	1688	3099	42509	123899	2.82 GB
okt. 2015	2074	3924	43705	139625	2.08 GB
nov. 2015	2049	4019	39164	115474	1.85 GB
dec. 2015	1497	3317	122402	191622	2.50 GB
Totaal	22856	46335	1350978	2558850	117.26 GB





		Samenvatting			
Rapportageperiode Eerste bezoek Laatste bezoek	Maand jan. 2016 01 jan. 2016 - 00.00 31 jan. 2016 - 23.54				
	Unieke bezoekers	Aantal bezoeken	Pagina's	Hits	Bytes
Bekeken verkeer *	1556	3063 (1.96 bezoeken/bezoeker)	193052 (63.02 Pagina's/bezoek)	319625 (104.35 Hits/bezoek)	4.21 GB (1441.05 KB/bezoek)
Niet-bekeken verkeer *			159418	161593	3.08 GB

* "Niet bekeken" is verkeer dat gegenereerd werd door robots of wormen, of respons met een speciale HTTP-statuscode.

Maandelijkse historie



Maand	Unieke bezoekers	Aantal bezoeken	Pagina's	Hits	Bytes
jan. 2016	1556	3063	193052	319625	4.21 GB
febr. 2016	955	1684	73453	154556	2.22 GB
mrt. 2016	0	0	0	0	0
april 2016	0	0	0	0	0
mei 2016	0	0	0	0	0
juni 2016	0	0	0	0	0
juli 2016	0	0	0	0	0
aug. 2016	0	0	0	0	0
sept. 2016	0	0	0	0	0
okt. 2016	0	0	0	0	0
nov. 2016	0	0	0	0	0
dec. 2016	0	0	0	0	0
Totaal	2511	4747	266505	474181	6.43 GB





Use of open standards and open source software

- Data access over OPeNDAP \rightarrow THREDDS
- Online analysis using WPS
- Online visualization using WMS \rightarrow KNMI ADAGUC
- → PyWPS and CERFACS ICCLIM
- Single Sign On using OpenID, and OAuth2, delegation using MyProxy X509







Search using ESGF Search API – New faceted search

)# C Q	Zoeken
INFLATING CIEF FOR THE ENDOPEIN NETWORK	IS-ENES Contact Account 🍹 🖩
Home Data discovery Downscaling Documentation Help About us Account 1	2
Filters Project (23) Parameter (1721) Frequency (16) Experiment (177) Domain (30) Model (142) Date Geobox Fr > show all filters clear all filters	? Help
Quick select Parameter All Parameter properties (1721)	^
Image: Stremperature (tas) Image: Precipitation Image: Precipitation	Wind sfcWind) /ind (sfcWindmax) d (uas) d (vas)
Radiation Pressure Revaporation SW Radiation Dn (rsds) Pressure (ps) Act. Evap. (evspsbl) SW Radiation Up (rsus) SL Pressure (ps) Pot. Evap. (evspsblopt) LW Radiation Dn (rsds) Pressure (pfull) Soil Evap. (evspsblopt) Diff. Radiation Dn (rsdsdiff) Pressure (pfull) Soil Evap. (evspsblopt) Clouds (clt) Clouds (clt) Pressure (pfull)	g)
Selected filters	
none	
Found 672402 datasets. Displaying page 1 of 26897.	
« Previous 1 2 3 4 5 6 7 8 9 10 11 26897 Next »	Export to CSV
mip5.CSIRO-BOM.ACCESS1-0.historicalExt.mon.sealce.Olmon.r2i1p1.v1	es-doc
mip5.CSIRO-BOM.ACCESS1-0.historicalExt.mon.ocean.Omon.r2i1p1.v1	es-doc Eath System Documentation







Search using ESGF Search API – **New** faceted search

- Search interface has been improved based on feedback from impact users
- Search interface speed has greatly been improved by using short lived (1 minute) caches
- Improved error handling in C4I for ESGF data nodes, catalogs are checked in advance for availability. Catalog status is clearly indicated to the user.
- Started with handling search queries as an aggregated dataset, e.g. use a search query as input for your processing.







Download

- By default the basket contains:
 - "Remote data" for links
 - "My data" for your own data
- Script based download: select ۲ and batch download multiple files
- The basket allows for uploading your own files
 - Can be used in processing or visualization
 - A per-user OpenDAP server on files in user's basket
- In development: abstraction of the file concept. Packaging results.

		★ 11	☆自	1 * - C
Exploring climate model data			IS-ENES	5 Contact Acco
ome Data discovery Downscaling Documentation Help About us	Account	•		
Account Basket (11) Processing Monitor jobs (8)				
ount »Basket				
asket				
ile	DAP	HTTP	Filesize	Date
Remote data				
- 📄 0.50 deg. regular grid				2015-01-22
🖿 0.44 deg. rotated grid			-	2015-01-22
🚡 tx_0.44deg_rot_v10.0.nc	true		691.9M	2015-01-22
🚡 tn_0.44deg_rot_v10.0.nc	true		691.9M	2015-01-22
- b tg_0.44deg_rot_v10.0.nc	true		691.9M	2015-01-22
tasmax day IPSL-CM5A-LR historical r1i1p1 18500101-18991231.nc	true	true	673.2M	2015-03-19
tasmax day IPSL-CM5A-LR historical r11p1 18500101-19491231.nc	true	true	1.346G	2015-03-19
tasmax day IPSL-CM5A-LR historical r111p1 19000101-19491231.nc	true	true	673.2M	2015-03-19
tasmax_day_IPSI-CM5A-I_R_historical_r1i1p1_19500101-19991231.nc	true	true	673 2M	2015-03-19
tasmax day IPSI -CM5A-I R historical r1101 19500101-20051231 nc	true	true	754 OM	2015-03-19
tasmax AER-44 CNRM-CEREACS-CNRM-CM5 rcn45 r11n1 CI Mcom-CCI M4-8-17 v1 day 2096	true			2015-04-01
				2010 01 01
polar stereo m.nc	true	true	906.824K	2015-01-23
tas WAS-44 ECMWE-ERAINT evaluation r11p1 IITM-RegCM4-1 v411 mon 198901-199012 nc	true	true	2.314M	2015-01-23
tas WAS-44 ECMWE-ERAINT evaluation r1i1p1 IITM-RepCM4-1 v411 day 19890101-19901231	true	true	70 463M	2015-01-23
View/Browse file Download file Script download	Upload file	🗱 Dele	te file(s)	Reload basket



