EUROPEAN COMMISSION
Community research

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

# IS-ENES2 DELIVERABLE (D -N°: 5.2)

## *Assessment of impact communities' requirements*

***{File name: IS-ENES2_D5_2.pdf}***


Author(s): ***Christian Pagé, CERFACS***
***Ronald Hutjes, WUR***
***Chris Jack, CSAG***

Reviewer(s): ***Sébastien Denvil, CNRS-IPSL***
***Lars Barring, SMHI***


Reporting period: ***01/04/2016 – 31/03/2017***

Release date for review: ***09/09/2016*** Final date of issue: ***09/03/2017***

| Revision table | | | |
|---|---|---|---|
| **Version** | **Date** | **Name** | **Comments** |
| 1 | 21/01/2015 | Christian Pagé | First structure |
| 2 | 29/08/2016 | Christian Pagé | Integration of CSAG and WUR contributions |
| 3 | 01/09/2016 | Christian Pagé | Draft of abstract, summary, introduction, conclusion |
| 4 | 07/09/2016 | Ronald Hutjes/ Christian Pagé | Conclusion of WUR section. Perspectives and Executive Summary. |
| Internal Review | 09/09/2016 | Christian Pagé | Abstract and final tweaks. |
| Final | 09/03/2017 | Christian Pagé | Integration of reviewers' comments |

## Abstract


This report aims at evaluating and improving the climate4impact portal in view of users' requirements and needs. Several types of users have been considered in this analysis. Key conclusions and perspectives for further development of climate4impact and climate data portals have been identified.



| **Project co-funded by the European Commission's Seventh Framework Programme (FP7; 2007-2013) under the grant agreement n°312979** | | |
|---|---|---|
| **Dissemination Level** | | |
| **PU** | Public | X |
| **PP** | Restricted to other programme participants including the Commission Services | |
| **RE** | Restricted to a group specified by the partners of the **IS-ENES2** project | |
| **CO** | Confidential, only for partners of the **IS-ENES2** project | |

# Table of contents

# Executive Summary

This report aims at contributing to the development of guidance needed by the climate change user communities by gathering requirements. These requirements are specifically on data and general data use, but also on how to produce downscaling data and corresponding metadata for the impact communities, as well as on evaluating the downscaling methods and datasets using climate-changing-conditions informative measures (distributional similarity, ensembles, robustness, uncertainties, etc.). Several types of users have to be considered, because they have different needs in terms of guidance and requirements.

User requirements for the relatively new target audience of commercial sectorial consultants (water, agriculture, etc.) have been assessed through feedback received during a number of Climate Services Trainings in which the climate4impact portal played an important role. Also, the analysis of a large number of climate portals has given important information that will be used for further development of climate4impact. User requirements from the metrics point of view have been also highlighted as additional information able to improve the knowledge about the users' activity on going in the climate4impact portal. It is clear that there is a huge variety and diversity of portals offering different subsets of data to a diversity of end users, but key conclusions have been identified to help further improve usability and guidance.

The climate4impact portal clearly already provides ample guidance and documentation, selection filters and processing tools that often lack in many other portals. But an important design consideration could be a portal entrance that differentiates between users with varying levels of expertise and skill. It is clear from those results that the need for a portal that discloses the ESGF data universe for users, beyond the climate modelling community itself, is and will remain very high and is essential for the development of a societally relevant and economically important Climate Services sector. Only through continued efforts and close co-creation between providers and users can we hope to the exploit to the maximum the potential of the CMIP5 legacy and the yet to bloom CMIP6, along with CORDEX and with strong collaboration with Copernicus C3S.

# 1. Requirements Assessment – Focus on consultancy or boundary workers

The following assessment is based on evaluation and feedback received during the course Climate Services Training (CST), a course format developed between ISENES2, Climate KIC and Wageningen University. In the following a brief outline of the course will be given, followed by a compilation of feedback received during the following instances of this course:

- To master students of Wageningen University (March 2014-trial, 2015-full, 2016-full). Respectively 29, 46 and 52 students participated, with varying backgrounds in natural and social science aspects of climate and global change.

- To professionals of CENICAFE (November 2015), the research organisation for coffee growers in Colombia, for which the provisioning of climate services is one of their core tasks. 11 professionals participated with background in agronomy, climate and IT.

- To consultants active in the water sector in the context of a C3S project (April 2016). 10 professionals participated from SME consultancy firms active in provisioning advice in water quantity and quality issues for agriculture (irrigation), drinking water and energy and industry, in areas around Europe.

Also 'alumni' of this course occasionally contacted the teachers to further help them out with data issues, providing further feedback and illuminating user needs.
These courses and their participants represent very different user communities, especially in terms of prior knowledge of climate variability and change, in terms of technical skills and in terms of climate aspects of interest and tools/data already in use.

- The MSc students have an adequate to good understanding generally of the current concepts and knowledge of climate variability and change, uncertainties with respect to climate sensitivity, local/regional vs global uncertainties and how these are addressed in multi model ensemble projections and analysed in statistical terms (e.g. pdf's). This often in addition to considerable training in more sectorial subjects (hydrology, agronomy, ecology, planning, etc.). Part of them are well acquainted with modern data standards (NetCDF) and processing tools (scripting, python, R). They have some initial knowledge of using such data in research impact models operating at larger spatial scales (hydrology, crops) and long time horizons. They will enter the job market within 1-2 years and will be the professionals of tomorrow in research organisations, consultancy and policy.

- The present professionals had their education between 5 years and one or more decades ago. Their expertise on climate change and especially the associated uncertainties and approaches to deal with those is much less. Their educational background is more mono-disciplinary. The have often a good understanding of local climate and its impacts on the sector they work in. They work with sometimes highly specific operational tools/models e.g. for a single catchment/irrigation system, sometimes with very specific data requirements (e.g. standard time series, specific indices), but that are in terms of IT complexity fairly simple (spreadsheet type models, text-based data formats, etc.), but the same time often linked to GIS systems.

### 1.1 Overview of the Climate Services Training (CST) course

A previous full description and evaluation of the course was presented in D6.2. Here a brief summary of its main elements. The course started with 3 lectures:

- Climate change and adaptation consultancy: asking the right questions, selecting the right questions. The lecture basically follows the structured approach to climate change adaptation problems developed for the MEDIATION/PROVIA Adaptation Platform (http://www.mediation-project.eu/platform/)

- The Climate data processing chain: dealing with uncertainties in climate projections. The lecture focusses on guidelines for selecting appropriate scenarios, models and datasets from the CMIP5 and CORDEX archives, that ensure proper accounting for climate uncertainties and on bias correction and downscaling approaches as prerequisites for using climate data to force impact models.

- Disclosing Climate Data: visualisation and communication of climate knowledge. The lecture focusses on co-design of climate risk and adaptation projects and state-of-art results visualisation in e.g. (interactive) climate atlases that combine climate and non-climate aspects relevant for stakeholders at regional (sub national) and (very) local scale.

The main element of the course consists of practical exercises structured around Use-case examples (from real world projects around the globe) to be elaborated by the students, by selecting and obtaining climate data through climate4impact portal, skill/bias assessment against observed data, trend analysis of mean and extremes for various future time slices.

Finally, the results of this 'mini project' had to be presented by the students to their peer, receiving recommendations for improvement of the analysis from each other and the teacher alike. In the professionals course the participants were allowed to present their own use-cases and customer experiences shortly after beginning the course.

### 1.2 Evaluation of climate4impact.eu portal data access and processing functionalities

• **Registration and group subscription**: Registration was generally straightforward and fast, although the availability of different registration nodes varied and the reason for having multiple registration nodes was totally unclear. More problematic was the group registration process (CMIP5 / CORDEX, academic / commercial): if this was forgotten at first registration, it is very hard to find where/how to do this later on. More fundamental issue: it is unclear why there is a distinction between academic use and commercial use, what the limitations of the latter are and why that is (to be better explained at registration). Many non-academic users were unhappy with this situation and argued for true and cost-free open access, e.g. like the current open data policies in the (US) Earth Observation.

• The more than **6 months ESGF downtime** (or at least much limited functionality) is difficult to understand to any time-constrained end-users. For the climate modellers community there may have been acceptable reasons for downtime prior to the CMIP6/DECK

main upload phase, but for the user community that uses such data continuously this heavily interferes with any project planning and negatively affects customer satisfaction, and consequently their confidence in the ESGF and climate4impact facilities. But users also need to understand that ESGF is still build and run with a best effort approach.

- **File restructuring and ease of access**. After the ESGF upgrade many long-term data sets (all historical and RCP runs) have been cut-up in 5-year time slices. By many end users this is perceived as a step back in user friendliness of the data access system. As teachers we try to convince end-users to use 30-year time slices shorter for reasons of statistical robustness of any projection trends. The new situation is perceived as not supporting the importance of the recommendation to use 30-year time slices whenever possible. A way to improve the situation for users is a tool to aggregate such data again, prior to download.

- **Native projection vs regular lat/lon for RCMs**. Likewise, after the ESGF upgrade all CORDEX data are only available in their native projection, whereas before also copies were available in regular lat/lon, which is conceived very difficult to use by many end-users. Re-projection tools prior to download or even online visualisation is recommended. This is already available in climate4impact, but not always found, so an easier and better access to re-projection tools is needed. This requires better guidance also because not all interpolation methods are adequate for every quantity. For example, conservative interpolation is correct for precipitation, but not for surface wind. This shows that significant guidance is also needed along with those tools.

- **Processing tools**. A number of processing tools are already offered on the climate4impact portal, but both their functionality and how-to are not easily grasped (except perhaps for the sub-setting tool). Better guidance on which to use for what and especially detailed how-to's need to be provided. In addition to the tools already available high priority should be given to fool-proof bias correction and downscaling tools (and associated documentation and guidance) as support for these processing steps appear most on the end user wish lists. This also requires that observed data (CRU, ECA-D, etc.) should be searchable and accessible through the climate4impact portal (including an observation filter/facet in the data search page).

### 1.3 Evaluation of climate4impact.eu portal guidance

As in previous user-requirements assessments guidance is at the top of user needs. However, with a shift to users yet not connected to the climate impacts research community (consultants) a more directive guidance would be appreciated. The present guidance is conceived as rather complete, but as a result also very extensive and practical guidelines are not easily found. Also they are often formulated as a bit academic discussions, not as straightforward *DO's and DON'T's*. In this context the following requests for guidance keep emerging:

- **Guidance on scenario selection**. Selecting an RCP in relation to the time horizon targeted in their projects is not trivial for many end users. More *prescriptive guidance* like the

following would help: to use any (or one specific one) of the RCPs for impact evaluation up to 2050, stating that one RCP suffices, put effort in analysis of more than one RCM/GCM (and vice versa for time horizons towards the end of the century).

- **Guidance model selection GCM (and RCM)**. Most end-users are overwhelmed by the multitude of GCMs available and have no idea which one to choose. For RCMs combined with their driving GCM the situation is perceived even more confusing, though many end users feel they should use RCMs rather than GCMs (see also next bullet on downscaling). More *prescriptive guidance* again would help. The climate community is challenged to *provide a kind of decision tree for model selection*, based on model quality/skill, climate sensitivity (though still an abstract concept for end users) or change signal for certain time slices and certain (broad) regions, availability of specific RCPs, etc. Such a decision tree should be supported by online explanations/documentation and appropriate tools for the creation e.g. dP/dT plots

Less related to choosing in the face of many possibilities, further guidance / explanation is still needed on:

- **Guidance CMIP/AMIP difference and implications for use**. Many data end-users have difficulty understanding that CMIP historical runs should not be compared to historical observations on a year-to-year or day-to-day basis, but only in statistical terms. Although the present climate4impact documentation gives an explanation of this (somewhere) in the context of AMIP/CMIP differences, this is easily overlooked and many end users are confused when comparing historical model and observed data. More prominent highlighting of this issue is recommended.

- **Downscaling (statistical) guidance**. Most end-users have the notion that some form of downscaling must be done before they can use the GCM or even RCM data in their impact assessment. In-fact ‘downscaling’ always ranked among the highest on the list in enquiries about things-they-want-to-learn, prior to the courses we gave. Online processing tools for statistical downscaling and observational datasets that can serve as reference should get high priority in improving the functionality of the data portals. This will also require information and guidance on uncertainties that come with such downscaling.

- **Impact indicators vs Climate indicators**. Also a processing tool for calculation of climate indices has recently been developed for the climate4impact/CLIPC portals, though it has not been used/evaluated yet in any of our courses. Nevertheless, flexibility in these is warranted. Flexibility in terms of user specified parameters for certain indices, e.g. user specified thresholds for threshold based indices, user specified percentiles for percentile based indices. Other periods instead of monthly statistics, e.g. weekly, 10-day, seasonal (period user specified). For various impact sectors other, but often very similar, indicators are commonly used. For example Flow Duration Curves or Flood Recurrence Intervals used to quantify hydrological extremes are very similar to percentiles and cumulative PDFs already available

in the indices toolbox. Such similarities in concept but differences in vocabulary should be made clear across sectors where relevant in the climate4impact portal.

In several C3S, ERA4CS and others projects, sector specific climate based indicators are being developed for e.g. wind/solar energy, hydrology and hydropower, agriculture. Further harmonisation in data and processing standards between such sectors, where possible, is recommended to facilitate online provisioning of sector specific indices.

- **Climate4impact not found**. On search engines climate4impact is not easily found when searching for 'climate data' (not on at least first 5 pages), whereas once found its functionality is highly appreciated compared to others. We recommend to improve the HTML Metadata in order to improve indexing and ranking in common search engines.

### 1.4 Evaluation of Climate Services Training for consultants

- **Lecture elements**. By nature of the current format of this short course Climate Services Training, the lectures are very dense in information and provided at relatively high speed. Better annotated PowerPoint material and more extensive guidance for further-reading could be provided as reference for use by participants after the course. Offering follow up activities like refresher webinars or online question time *after* the participants had ample time to ingest the course material at home may further ease this.

- **Exercise elements**. Working directly on teacher provided or participants own use cases is highly appreciated and provides strong motivation for engagement in the data search and processing exercises. Processing is mostly done offline (in spreadsheet software) as the how-to of online processing tools are yet not easily grasped. More time and better help documentation would be needed to also instruct on these. A lot is learned from presentation of exercise results and feedback from teachers and peer students and is definitely worth the relatively large time investment (depending on the student numbers).

- **Post data handling and Mediation Toolbox**. For many consultancy end-users a climate impact assessment is only one part of an adaptation study and there-in only one future trend affecting their clients' operations and assets. Choosing and using the appropriate tools for any impact, vulnerability and adaptation problems has been well structured in the *Mediation/Provia Adaptation Pathfinder* (MAP) and Toolbox. In the course more extensive treatment of it possibilities as well as some exercises on this is recommended, possibly in parallel to the more technical data handling exercises. Also on the climate4impact portal reference to the MAP portal could be made more prominently. If needed the course duration could be extended by half or one day.

- **Room for exchanging experiences**. In both professionals courses the opportunity to share experiences, problems and solutions for approaching climate problems for their clients was evaluated as an important element of the course (it should be noted that in neither course direct competitors participated, as e.g. their respective home-markets were geographically separated). In a future course more room for such exchanges should be made. This also

requires a minimum number of participants (>10) to make sure all will find potential 'matches' with colleagues. On/around the climate4impact portal also a user community should be facilitated, where experiences with the portal can be shared, perhaps as well as more general experiences, problems and solutions for approaching (local) climate problems.

### 1.5 Conclusions Climate Services Training for consultants

Addressing a relatively new target audience of commercial sectorial consultants (water, agriculture, etc.) with a Climate Services Training and the functionalities of the climate4impact portal requires new levels of guidance and easy-to-use data access and processing tools.

The Climate Services Training is very information-dense for professionals with limited climate expertise. Concise "further reading" guidance with the lecture material and some form of aftercare in the form of webinars or remotely organised question time might ease this. The opportunity to exercise with each participant own use cases and to present and discuss results with peers is a strong course element. Quantitative climate impact assessments is generally not the only task consultants do for their clients, guidance on vulnerability and adaptation strategies is requested and can be easily provided in a structured context using the *Mediation/Provia Adaptation Pathfinder* and associated Toolbox.

Though not easily found, the *climate4impact* portal offers functionalities that are highly appreciated. For this target audience the guidance can be even more directive (with more straightforward *do's and don'ts*'). For a certain minimal quantification of uncertainties preselected ensembles of RCPs and GCM/RCM combinations can be offered and easy-to-use, well documented processing tools, working on ensembles rather than individual datasets would be highly beneficial for non-climate specialists working with climate data.

## 2. CSAG Climate Portal – Survey and Evaluation

Researchers at the Climate System Analysis Group, part funded through IS-ENES2 undertook a comprehensive review of climate data/information portals/websites. This was undertaken in response to the growing number of such websites available to users and the increasing concern over the diversity of information available to users and decision makers. The review has been documented formally in a paper (WIRES under review). This report describes some of the key findings of potential relevance to IS-ENES2 and in particular the climate4impact portal.

The review attempts to strike a balance between an objective analysis of key characteristics as well as a more subjective "user experience". The review was performed by a range of researchers extending from those with highly technical and long experience through to relatively junior, low technical and minimal experience. This allowed the review to capture a diversity of user experiences without burdening external users with extensive surveys or formalised user experience evaluations.

The review notes the expected high diversity in data/information provided to users, the common difficult users experience in accessing both the data as well as, more importantly, descriptions of the data (meta-data) and information on limitations of the data for particular use cases. The defensibility of the data provided is often questionable and it is easy for users to feel that the data provided is appropriate or informative for their contexts when in reality it may not be.

### 2.1 Approach

The review considered 42 climate information portals:

**Table 1: Climate Information Portals reviewed**

| Climate Information Website | Link |
|---|---|
| Arctic Climate Research at the University of Illionis | http://arctic.atmos.uiuc.edu/ |
| aWhere | http://www.awhere.com/ |
| Canada Centre for Climate Modelling and Analysis | http://www.ec.gc.ca/ccmac-cccma/ |
| CCAFS Downscaled GCM Data Portal* | http://ccafs.cgiar.org/ |
| Centre for Climate Change Research (CCCR) | http://cccr.tropmet.res.in/ |
| Climate Change in Australia | http://www.climatechangeinaustralia.gov.au/ |
| Climate Change Knowledge Portal * | http://sdwebx.worldbank.org/ |
| Climate CHIP | http://www.climatechip.org/ |
| Climate Data Online | http://reg.bom.gov.au/climate/data/ |
| Climate Information Portal (CIP) | http://cip.csag.uct.ac.za/ |
| Climate Wizard | http://www.climatewizard.org/ |
| Climatic Research Unit | http://www.cru.uea.ac.uk/ |
| CliMond | https://www.climond.org/ |
| CLIMsystems* | http://www.climsystems.com/ |
| CORDEX East Asia | https://cordex-ea.climate.go.kr/ |
| Downscaled CMIP3 and CMIP5 Climate & Hydrological Projections archive | http://gdo-dcp.ucllnl.org/ |

| | |
|---|---|
| Earth System Grid Federation | http://esgf.llnl.gov/ |
| EDENext Data Portal | http://www.edenextdata.com |
| European Climate Assessment & Dataset (ECA&D) | http://www.ecad.eu/ |
| European Space Agency Climate Change Initiative (esa cci) open data portal | http://cci.esa.int/ |
| Giovanni | http://giovanni.sci.gsfc.nasa.gov/ |
| IPCC Data Distribution Centre (DDC) | http://www.ipcc-data.org/ |
| IRI/LDEO Climate Data Library | http://iridl.ldeo.columbia.edu/ |
| IS-ENES Climate4impact portal | https://climate4impact.eu/impactportal/ |
| KlimafolgenOnline | http://www.klimafolgenonline.com/ |
| KMNI Data Centre | https://data.knmi.nl/ |
| KNMI Climate Explorer | http://climexp.knmi.nl/ |
| Med CORDEX | https://www.medcordex.eu/ |
| NCAR's GIS Program Climate Change Scenarios GIS data portal | https://gisclimatechange.ucar.edu/ |
| Nevada Climate Change Portal | http://sensor.nevada.edu/NCCP/ |
| NOAA Climate.gov | https://www.climate.gov/ |
| NOAA Geophysical fluid dynamics laboratory | http://www.gfdl.noaa.gov/ |
| Ontario Climate Change Data Portal | http://www.ontarioccdp.ca/ |
| Pacific Climate Futures | http://www.pacificclimatefutures.net/ |
| Pacific Climate Impacts Consortium Data Portal | https://www.pacificclimate.org/ |
| Regional Clearinghouse Database* | http://clearinghouse.caribbeanclimate.bz |
| South African Risk & Vulnerability Atlas* | http://sarva.dirisa.org/ |
| Global and Regional Adaptation Support Platform (ci-grasp)* | http://pik-potsdam.de/cigrasp-2/ |
| The Satellite Application Facility on Climate Monitoring (CM SAF) | http://www.cmsaf.eu/ |
| USGS Geo Data Portal | http://cida.usgs.gov/gdp/ |
| Wisconsin Initiative on Climate Change Impacts | http://www.wicci.wisc.edu/ |
| WoodForTrees.org | http://woodfortrees.org/ |

## 2.2 Quantitative approach

The review approach identified a number of quantitative and categorical metrics that were evaluated for each portal as best as was possible. In some cases, it was not possible to determine the value of the metric because of a lack of documentation or some other obscurity. Some metrics involve a diversity of understandings. For example, identifying the target audience or "next user" was not always easy as different portals use different terminology. Some may talk about the IAV community while others will mention development practitioners.

**Table 2: Quantitative and categorical metrics**

| Criterion | Type of answer |
|---|---|
| Country in which climate service provider headquarter/ main office(s) is based | Country name (s) |
| CIW provider types | Category options |
| Access control (Log-in or registration required) | Category options |
| Target group (explicit next-user) | Category options |
| Geographic domain of data/information focus | Region/country name |
| How data/information are presented/provided | Category options |
| The number of GCMS presented | Number |
| SRES scenarios applied | Yes/No |
| RCP scenarios applied | Yes/No |
| The number of dynamical downscaling products presented | Number |

| | |
|---|---|
| The number of statistical downscaling products presented | Number |
| GCM and downscaling results separately accessible | Yes/No |
| Personal judgement: Do they point explicitly to the fact that downscaled data is based on GCM output? | Yes/No |
| Personal judgement: Can a decision maker get the information he/she needs and is he/she able to manoeuvre this portal? | Yes/No/With some difficulty |

To provide the most robust analysis possible given the resources, the indices were evaluated for all portals by two different researchers and then again by the researchers working together to try to resolve differences. It is acknowledged that the results may not perfectly reflect each portal.

### 2.3 Narrative approach

The second component of the review involved a narrative approach where researchers engaged with each portal with a hypothetical case study question to answer. Three different use cases were identified and four researchers spent 1.5 to 2 hours per portal attempting to obtain the required information. The researchers spanned a diversity of technical knowledge and experience and so the narrative reports approximate the real world diversity of users.

**Table 3: Narrative use case scenarios**

| **Use-case scenarios**<br>*Perspectives taken by narrators* | | |
|---|---|---|
| In my role as a Peace Corps member I would like to work with local residents to develop climate change adaptation responses, and therefore need to understand current climate vulnerability and vulnerability to future climate change. | Water demand in the city is expected to exceed supply by 2019. In order to conduct long term planning in my role as water management official, I need to understand how climate change may impact future water availability. | In my role as an agricultural extension officer I am tasked with understanding possible changes in the climate and to use this information to develop more resilient strains of maize seed for the climatic circumstances of the future. |

### 2.4 Results

The results showed that the overwhelming number of portals are hosted in first world developed nations, Europe (13), USA and Canada (17), and Australia (6), with only a few hosted by less developed nations. While this likely reflects the dominance of developed nation research institutes it is also an interesting result in that much critical need for climate information to inform adaptation lies within less developed nations.

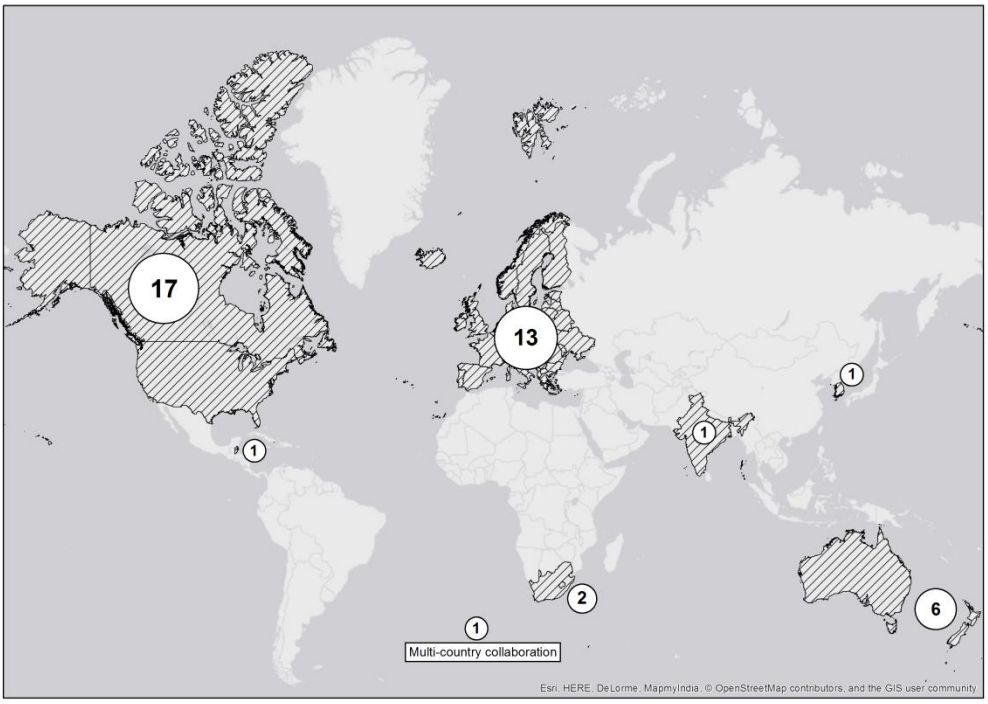



**Figure 1:** Geographic distribution of the 42 CIWs assessed in the paper

### a) Targeted user groups and commercial portals

The results also that university based portals and national research institution based portals dominate the landscape with relatively few commercial, government, and not for profit. It is encouraging that a number of portals reflect multi-national collaboration.

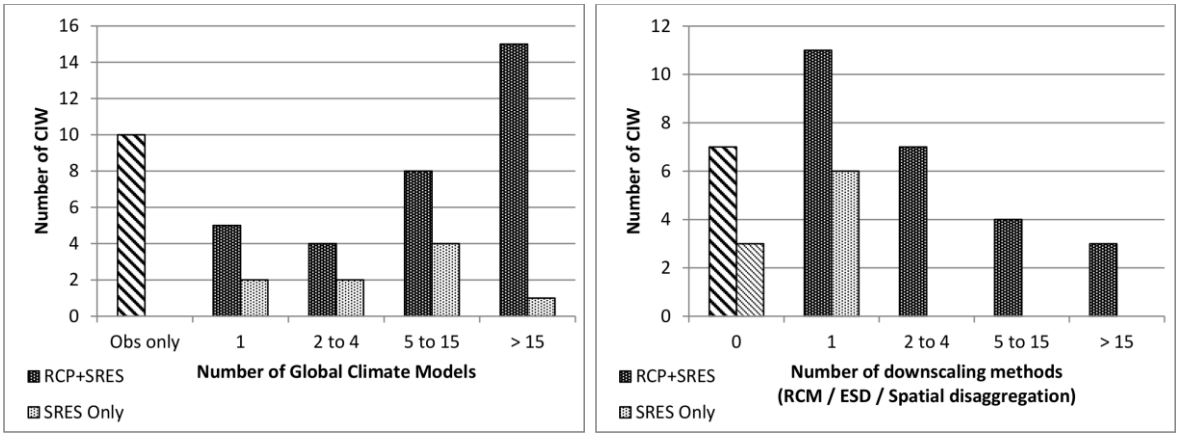
While it was often not possible to determine the target user group (17 cases), where this was possible researchers (14) and government officials (10) dominated the user groups targeted. The fact that in many cases it was impossible to determine the target user group is concerning as it would be expected that portals would be designed with a target user group in mind and that this would be clearly articulated and described. As it stands it is difficult for an end user to determine if a particular portal is providing information targeting that particular users level of expertise or field of interest.

Only two portals required payment with the majority offering extensive datasets freely to the public or private sector. This is remarkable and reflects a strong investment in open data and free access to information.

### b) Types of data provided

Around 25% of portals only provided historical climate data which is quite appropriate for certain end users. Around 50% of the remaining portals provided information from a wide selection of GCMs (> 15 models) and thus represented a spread of projections data. Five portals offered only a single GCM representing no spread or uncertainty in the projections.

However, less portals provided multiple downscaled datasets. This likely reflects the lack of such datasets and the tendency for institutions to implement single methods. Most portals only provided a single downscaled product.




a) b)

**Figure 2:** The number of CIWs in relation to the number data/information products offered. In each pair of bars, the right bar is the number of sites with only SRES-based products. The Global Climate Models are shown in panel (a), and include the number of CIWs that offer only historical information (left-most bar). Downscaled (statistical or dynamical) methods are shown in panel (b), where the bars for zero downscaling methods indicate instead the number of CIWs with only GCM products.

### c) Transparency

Table 5 shows that in cases where downscaled data are provided. 50% of portals do not explain the downscaling process. In cases where downscaling is not provided around, the modelling chain is more likely to explained (60%).

| | | | |
|---|---|---|---|
| Are downscaled data available separately to the GCM data | Yes | 5 | 5 |
| | No | 8 | 6 |
| | | Yes | No |
| | | Is the modelling chain explained explicitly | |

**Table 5:** The number of CIWs in relation to the transparency about the included downscaled data.

### d) Narratives

The narratives reflected the experience of junior researchers in their attempts to access information to satisfy one of the three case studies described above. Some key quotes from the narratives are informative (climate portals not named):

*"There is a straight forward and welcoming climate data interface, but a lack of obvious pathways to the desired information, and a multitude of interesting yet potentially distracting information. The lack of clear guidance on how to manoeuvre among the choices provided, and the inconsistent presentation of data from different models makes confidence around the extraction of robust messages hard to come by."*

*"The multitude of data entry points and data displays breed confusion, yet the easily accessible multi-model comparison and averaging is encouraging, despite not fully grasping why such comparison and averaging is important. Despite clear messages regarding data being non applicable at a local scale, the site tempts one to extract local scale messages with relative confidence simply because it is possible."*

*"Having located the non-apparent climate information section, one is overwhelmed by options, and without substantial reading time the extensive supportive material does not make for easy manoeuvring the jargon-laden landscape of options. Technical challenges and unfamiliar data file formats further alienate the user"*

Some key conclusions from the paper:

a) Much is assumed of the user's familiarity with terminology.
b) Navigation is complicated with complementary data often spread across multiple sections of the CIW, testing the user's patience and raising frustration.
c) There is often a lack of clarity about what is being displayed, either in terms of explanations of how it was generated, or how robust the information may be taken to be.
d) The choices presented are multiple and often confusing.
e) Guidance is commonly minimal, unclear, or hard to find.
f) Different avenues through the CIW structure can lead to different outcomes.
g) There is a presumed degree of (significant) technical skills by the user.
h) In many cases the level of effort required by the users is in contrast with the expectation that the CIW will simply deliver information.

### 2.5 Conclusions

This above represents a brief summary of a much larger work (WIRES Climate Change, in review) but highlights some key information about climate information portals. It is clear that there is huge variety and diversity of portals offering different subsets of data to a diversity of end users. Possibly the greatest challenges for end users are: (a) Identifying the appropriate portal to use, and related to that, (b) determining the appropriateness, caveats, and limitations of the data available.

Furthermore, challenges with actually accessing complex and technical data remain, though this is improving in recent years with more sophisticated web interfaces.

# 3. Improvements and metrics requirements

A relevant contribution to the continuous improvement of the climate4impact portal could come from an accurate analysis of the system usage by the actual final users. Significant requirements could be inferred keeping track of the users' activities and preferences, leading the service to better fit its final purpose.

A monitoring system used in the context of IS-ENES2 and, most widely, within the Earth System Grid Federation (ESGF) is the FASM-N (Federate Archive System Monitoring - Notification) module.

It was originally established as a distributed architecture where each ESGF data node collects local metrics and one or few collector nodes are responsible for gathering global statistics related to the whole community. However, the extremely modular nature of the system could enable a single-node approach and allow to choose which metrics to track and, consequently, to instruct the system on the new configuration.

Currently, the monitoring module installed on ESGF data nodes collects information about the data usage, the status of the services and the clients' distribution both at local and global level. More specifically, it is able to extract relevant and reliable information about the most downloaded datasets, the variables analysed, the models and the experiments most used as well as the availability of the node and the registered users. Moreover, it provides an intuitive user interface that shows time series, tables and maps as a dashboard with the most important information at a glance.

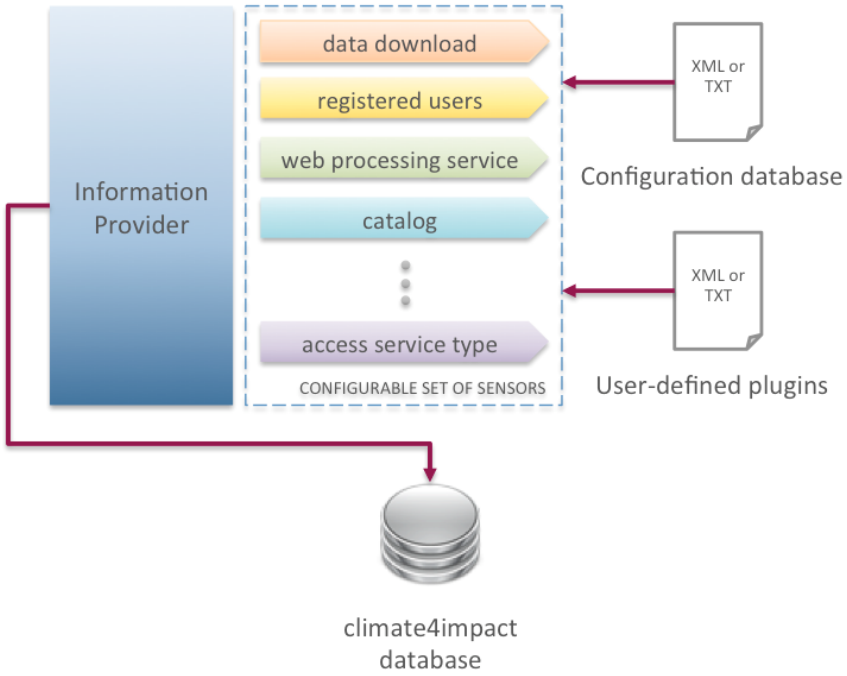
Such a tool could be properly configured to support the climate4impact portal in its continuing evolution on the basis of the user requirements. For instance, the monitoring system could provide different types of feedback useful both for the final user and for the administrator user of the portal.

Distinguishing between these two roles, lots of findings can derive from the study of the statistics provided by the dashboard.

Concerning the final user, she will have the possibility to monitor all the activities executed on the portal and obtain an idea, for instance, on which data she mostly worked on or which processes she mainly submitted. This information could lead the user to provide useful feedbacks and suggestions to the improvements of the portal from a final user point of view.

Regarding the administrator user, instead, a wide number of global information can be inferred from the statistics collected by the monitoring tool and, most specifically, a very large set of these information can provide a clear idea of the actual use of the portal by the user and help the deciders and the developers to improve the user experience and the usability of the portal.

The figure below shows a schematic view of the system, underlying its configurability and modularity.




A configurable set of sensors could allow the monitoring of different kinds of metrics strictly related to the climate4impact portal.
At the moment, the dashboard could support the collection of statistics for registered users and data downloads:

- downloaded files
- downloaded datasets
- volume of data downloaded
- number of users who performed a download operation
- users' geographic distribution
- most analysed variables, models, data sources, experiments.

Moreover, opportunely tuned, it could extract information related to the usage of the different web processing services (conversion, subsetting, indices calculation, spatial selection, etc.) as well as catalogues and different access services offered by the portal (API or web interface).
The collected data could be stored and properly aggregated in the system database to be visualized through a user-friendly web interface.
An easy access to such information by the decision makers will definitely help to enhance the user experience in accessing and processing climate data by investing in the most used services and providing innovative and more suitable tools.

## 4. Perspectives

The climate4impact portal has been developed for many years, with periodic user assessments through workshops and sessions. All the main functionalities are now also available as services that can be reused to built targeted portals: this has been used notably to built the

CLIPC portal. This is a quite strong feature and it enables the easier setup and building of portals targeted to different user communities using the same common tools. It is nonetheless very important to get user requirements periodically to adapt the tools and also improve the user experience and needs when they use portals to get proper climate related data and products.

Both the user requirements assessment through feedback from participants in various Climate Services Trainings that heavily invoked the climate4impact portal, and the one through user-centred evaluation of more than 40 climate data portals world-wide, point at a continued need for improving or developing portals that serve the better the user's needs. This should probably start with much more explicit identification of the target audience and a recognition of its assumed level of climate expertise and technical skills. Then, more intense forms of co-creation/co-design, of contents, functionalities, guidance/help/documentation and UI/navigability, between users and portal developers may improve on the current situation. This applies to the development of comprehensive portals that try to serve 'all' by disclosing 'all' and consequentially will probably always require a considerable level of expertise and skills of the user. It applies even more perhaps to portals that target a smaller, more specific audience, that offer limited functionalities and reduced ensemble data (but *not* just single model/RCP data) and are consequentially easy to navigate and use.

The climate4impact portal clearly falls in the first category and already provides ample guidance and documentation, selection filters and processing tools that often lack in many other portals. A few more good examples exist like the CLIPC portal, and perhaps future efforts should be more concentrated on continued development of one (combination) of the two. An important design consideration could be a portal entrance that differentiates between users with varying levels of expertise and skill, combined with developing/providing a high level API that portals of the second category might exploit for the most specific (sectorial) audiences. At the same time, as reported in Section 3, a metrics support from the dashboard component (FASM-N) would help in better understanding the activity performed by the user in the climate4impact, the data usage, the most used processing services, etc.

The need for a portal that discloses the ESGF data universe for users beyond the climate modelling community itself, is and will remain very high and is essential for the development of a societally relevant and economically important Climate Services sector. Only through continued efforts and close co-creation between providers and users can we hope to the exploit to the maximum the potential of the CMIP5 legacy and the yet to bloom CMIP6, as well as other datasets such as observations and reanalysis, downscaling, and national/regional specific datasets, along with CORDEX and with strong collaboration with Copernicus C3S. We also must realise that Copernicus C3S will be an operational service, while the ESGF is still targeting mainly research users, on a best effort, and is not a full 24/7 operational service like the future C3S.