DEPLOYMENT OF AI-ENHANCED SERVICES IN CLIMATE RESILIENCE INFORMATION SYSTEMS

Nils Hempelmann [*] [1], Carsten Ehbrecht[2], Etienne Plesiat[2], Gobe Hobona[1], Joana Simoes[1], David Huard[3], Trevor J. Smith[3], Ursula S. McKnight[4], Ilias G. Pechlivanidis[4], and Carmen Alvarez-Castro[5]

[1]OGC - Open Geospatial Consortium, Innovation Program, Leuven, Belgium - (nhempelmann, ghobona, jsimoes)@ogc.org

[2]DKRZ - German Climate Computing Center, Hamburg, Germany - (ehbrecht, plesiat)@dkrz.de

[3]Ouranos, Montreal, Canada - (huard.david, smith.trevorj)@ouranos.ca

[4]SMHI - Swedish Meteorological and Hydrological Institute, Norrköping, Sweden - (ilias.pechlivanidis, ursula.mcknight)@smhi.se)

[5]CMCC - Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici, Lecce, Italy - carmen.alvarez-castro@cmcc.it

Commission IV WG IV/4

KEY WORDS: Climate Resilience Information Systems, Climate Change, Climate resilience, OGC API Processes, Climate Service, Climate Indices.

ABSTRACT:

Recent advances in modelling capabilities and data processing combined with vastly improved observation tools and networks have resulted in the expansion of available weather and climate information, from historical observations to seasonal climate forecasts, as well as decadal climate predictions and multi-decadal climate change projections. However, it remains a key challenge to ensure this information reaches the intended climate-sensitive sectors (e.g. water, energy, agriculture, health), and is fit-for-purpose to guarantee the usability of climate information for these downstream users. Climate information can be produced on demand via climate resilience information systems which are existing in various forms. To optimise the efficiency and establish better information exchange between these systems, standardisation is necessary. Here, standards and deployment options are described for how scientific methods can be be deployed in climate resilience information systems, respecting the principles of being findable, accessible, interoperable and reusable. Besides the general description of OGC-API Standards and OGC-API Processes based on existing building blocks, ongoing developments in AI-enhanced services for climate services are described.

1. BACKGROUND

The need for climate services (CS) became urgent with the Paris Agreement in 2015, where adaptation came into focus as a pressing need alongside traditional mitigation measures (UNFCC, 2015). Information about future climate variability can help to inform decision-making by providing deeper insights into the potential risks, as well as supporting actions to mitigate those risks. Consequently, the field of CS has been developing rapidly, with many types of services and service providers evolving worldwide (Panenko et al., 2021).

Users of CS are expected to range from individuals or organisations with responsibilities for decisions and policies related to climate change mitigation/adaptation, to intermediary users such as consultancies, to societal actors, like the media, non-governmental organisations; other non-profit organisations (Cortekar et al., 2020). Depending on the user's needs, these data and information products may be combined with non-climate data, such as agricultural production, health indicators, population distributions in high-risk areas, road and infrastructure maps for the delivery of goods, and other socio-economic variables. The aim is to support efforts to prepare for new climate conditions and adapt to their impact on e.g. water supplies, health risks, extreme events, farm productivity, etc.

The current landscape of CS is highly diverse with an ever-growing range of programs, projects and portals involved in developing and/or providing CS at different administrative levels and spatial-temporal scales. This diversity, i.e. of producers, users and policy arenas, has furthermore generated a highly heterogeneous data- and information-oriented service landscape. Efforts will be required to harmonise and standardise the conceptualisation, operationalisation and evaluation of CS information and data (Weichselgartner and Arheimer, 2019). It is thus of paramount importance to involve the user community – especially those responsible for climate-informed decision making and climate-smart policy and planning – in the production, translation, transfer, and use of climate information and knowledge. This will lead to more effective engagement with users, increasing also the uptake of the best available climate science information and practices through improved confidence in scientific knowledge. The technical infrastructures translating raw data into useful climate information will be called Climate resilient information systems (CRIS) in this paper.

Accordingly, climate service centres emerged, translating scientific data generated by the scientific community into locally-relevant information that helps with decision making and policy setting. These translations are important across all disciplines and require various levels of external data integration and fusion. Climate service centres combine future climate projections from modelling centres, remotely sensed data from satellite instruments, and ground measurements from observation networks to create climate products and services tailored to decision makers' needs. Although the tailoring process is specific to local contexts, climate service centres face similar challenges, such as storing and processing large data sets, providing data visualization interfaces catered to various levels of users, or assembling and publishing rich data catalogues. In the ab-

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-187-2022 |

sence of explicit coordination mechanisms across climate service centres, in order to meet these complex goals, these centres must invest considerable resources solving these issues on their own. Most likely, the easiest problems will be solved many times over, while the most difficult ones are left unsolved.
This paper points out the importance of standardisation for application programming interfaces (APIs) to generate higher synergies between climate service centres, optimise the available resources and tackle more complex technical challenges. To achieve these objectives, the concept of "FAIR" data (Wilkinson et al. (2016)) should be extended to encompass FAIR climate services, with the principal components of FAIR meaning Findable, Accessible, Interoperable & Reusable. The FAIR principles should not only enrich data practices, but also the full CRIS that surround and inform their effective use. This requires agreements on metadata aspects for discovery, APIs and resource models for interaction with the climate services information systems. The use of open standards is thus essential in order to make data interoperable and facilitate access for users. This concept of FAIR climate services is only achievable when the participating organisations are agreeing on a common standard for software modules and APIs to exchange data, services, and corresponding information. The full meaning of FAIR has yet to be defined in the realm of climate services, however, we present the following; 'Findability' implies that CRIS publish public catalogues of well documented scientific analytical processes; Remote storage and computation resources should be 'Accessible' to all, including low bandwidth regions and closing digital gaps to 'Leave No One Behind'; agreeing on standards for data inputs, outputs, and processing APIs are the necessary conditions to ensure the system is 'Interoperable'; and finally, they should be built from 'Reusable' building blocks that can be realized by modular architectures with swappable components, data provenance systems and rich metadata.

1.1 Key actors and IT landscape

In Europe, interoperable building blocks for climate information are being developed and deployed on behalf of the Copernicus Climate Change Service (C3S), comprising the subproject "Climate Projections for the Climate Data Store" (CP4CDS (Ehbrecht et al., 2018b)). CP4CDS provides the required data and services for global climate projections (CMIP5) to the Copernicus Climate Data Store (CDS). The CDS catalogue contains the geophysical information needed to perform processing in a consistent and harmonized way, made available by open communication services standards (Web Processing Services; WPS) deployed in the back-end. A geographically-distributed network of CP4CDS data nodes provide a highly available set of environmental data and compute services. They are load balanced across three sites hosted from the leading European climate compute centres: CEDA, IPSL and DKRZ. This infrastructure model represents an initial take on the FAIR principles as applied to climate services, and further development and deployments are expected in the near future. Built upon the CDS is the European Copernicus Climate Change Service (C3S (Street, 2016; Su et al., 2018)), an example of an operational CRIS. Similar systems such as this can be found in the U.S.-based and NOAA-developed Climate.gov, the Australian Biodiversity and Climate Change Virtual Laboratory, and the World Bank Climate Knowledge Portal.
Detailed in the United Nations Framework Convention on Climate Change (UNFCCC) policy frameworks, research and systematic observation plays an important role to foster climate information production and provision. International initiatives like the European Space Agency (ESA) Climate Change Initiative (CCI) or the Copernicus Atmospheric Monitoring Services (CAMS) revolve around the production of satellite data. For example, a critical need for mitigating climate change involves enhancing the monitoring certainty of atmospheric trace gas emissions like carbon dioxide, CO_2, or methane, CH_4. The World Meteorological Organization (WMO) has furthermore set up the Global Framework for Climate Services (GFCS), a global partnership of governments and national institutional structures that produce and use climate information and services, to expand access to the best available climate data and information (Lúcio and Head, 2016; Giuliani et al., 2017). GFCS thus accelerates and coordinates the technically- and scientifically-sound implementation of measures to improve climate-related outcomes at national, regional and global levels, as well as build capacity in countries for managing the risks and opportunities of climate variability and change. National institutional structures (GFCS, 2014; Cullmann et al., 2019; Timofeyeva-Livezey et al., 2017) can further support science delivery, particularly for the special case of climate information for policy-makers, through their own CRISes (WMO, 2017).
Following the recommendations of the United Nations Committee of Experts on Global Geospatial Information Management (UNGGIM), open standards should be taken into account at multiple facets of CRIS as well. Indeed, the implementation plan of the GFCS (GFCS, 2014) clearly underlines the importance of avoiding duplication, sharing data at the global scale, working with common standards, and seamlessly linking the global, regional, and national entities encompassing the CRIS infrastructure. The modern development of CRIS has thus largely focused on efforts to improve climate data access and automate components of climate services delivery (Déandreis et al., 2014). These systems facilitate access to either raw or processed data from climate projections and Earth observation platforms through web-based portals; they can be used to quantify the impacts of climate change, for efficient reporting mechanisms required by international agreements (Akhtar-Schuster et al., 2017); or can be used as tools for the implementation of policy measures (Muraya, 2018).
The Open Geospatial Consortium (OGC) is the organization responsible for developing and maintaining standards for geospatial information. The structure of OGC comprises a large network of experts and technical innovations, and accordingly, also the appropriate standardisations are being developed, enhanced and distributed. It starts with discovery, which is facilitated by the use of metadata and ensures that search engines are aware of the data, or, at least, of the metadata. With the data portals of the European Environmental Agency (EEA) and the European Commission (EU) General Multilingual Environmental Thesaurus (GEMET) and the INSPIRE registry, promising starting points exist for achieving semantic interoperability, though linked data principles need to be further explored in terms of how data were captured, produced, processed, and fused to make sure that we stay on top of data complexity and integration.

1.2 Climate intelligence CLINT

This paper is tailored to the deployment of AI-enhanced services which are (or will be) developed in the project 'Climate Intelligence (CLINT)', funded by Horizon 2020 of the European Commission (EC). Figure 1 shows the principal(/principle) scientific method workflow concerning the development of AI-enhanced Science to the deployment of services in CRIS. In

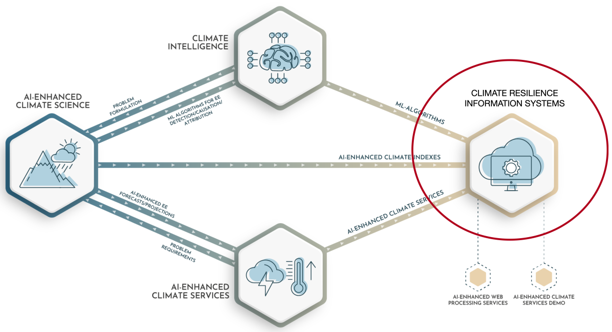

Figure 1. Scheme of the CLINT project. Developed AI enhanced services are being prepared for deployment in CRIS. Source (modified): https://climateintelligence.eu/

particular, the principles and approaches detailing how the scientific method is wrapped into deployable and FAIR climate service building blocks will be highlighted in this paper. We describe the suggested OGC-API standards, introduce example scientific methodologies undertaken in CLINT to develop the blueprints within the project, and point to existing building blocks under production and running in CRIS.

2. OGC APIS FOR PROCESSES

Producing and providing useful information for climate services requires vast volumes of data to come together. This data typically comes from different institutions and different disciplines, underlining the importance of technical standards enabling interoperability. One such example of these standards, the NetCDF data format, has been adopted as the de facto community standard for storing and sharing climate datasets. Metadata attributes embedded in NetCDF files are themselves expected to follow the Climate and Forecast (CF) Conventions, ensuring that variables and their units are uniquely defined, that coordinate systems and metadata information are clearly described, and that results are reusable.
To draw useful and actionable information from raw climate data, considerable processing is required, typically starting with the extraction of data over a region and time of interest. From there a wide range of analytical or machine learning algorithms can be applied to extract information relevant to climate adaptation, mitigation, or risk analysis. From a user perspective, the proliferation of different and independent CRIS could either be seen as a boon or as a plague. While more data and services are available than ever before, there is a risk of developing a landscape of fragmented and incompatible systems, with users having to master multiple different APIs and technologies, converting data from one format to another, in order to combine data and processes from different sources to effectively perform climate analyses.

2.1 Importance

Despite broad community agreement on the need for easier access to climate data and associated server-side processing, and encouraging ongoing initiatives, there is no broad consensus on the technology to use to achieve this. Public, private, and academic initiatives propose multiple competing or complementary solutions, creating a technological maze that individual scientists are forced to master and navigate to pursue their research.
Amongst a set of challenges affecting spatially-enabling climate services, Giuliani et al. (2017) identify data volume as a challenge needing to be addressed. This challenge is caused, in part, by both the complexity and quantity of climate data, as such data requires efficient processing workflows and powerful computing infrastructure to process the data. For an appropriate data processing to generate climate information Giuliani et al. (2017) suggest to tailor the CRIS to the WPS standard respective its emerging successor, OGC API - Processes, offer a potential solution to this challenge.
There are multiple reasons why OGC API - Processes is well suited to supporting climate science. First and foremost, the standard enables execution of distributed processing on a network, which greatly facilitates collaboration between multiple scientific institutions. Second, the standard enables scientists to invoke processing on high performance computing clusters from their desktop computers. Third, when implemented according to the Best Practice for Earth Observation Application Packages, the standard enables analytical code to be deployed closer to the data, thereby reducing bandwidth, bandwidth costs for data providers and CRIS providers that impede climate science.
The ability to uniquely identify inputs and types of outputs enables implementations of OGC API - Processes to support requests that specify Essential Climate Variables (ECVs) as inputs or outputs. ”An ECV is a physical, chemical or biological variable or a group of linked variables that critically contributes to the characterization of Earth’s climate.” (WMO). It means that two or more separate scientific institutions can chain their applications together through their APIs, such that the outputs of one application can unambiguously serve as inputs into the application of another. To achieve this, however, the identifiers of the ECVs have to be consistently named by the participating applications. The OGC API - Processes standard provides a schema that supports feature collections. This means that ECVs that are represented according to the Observations and Measurements (O&M) standard can be exchanged between implementations of the OGC API - Processes standard whilst also maintaining the semantics of the information being exchanged. The O&M standard leverages the Simple Knowledge Organisation System (SKOS), an application profile of the Resource Description Framework (RDF), to support the representation of semantics. The formal definition of the ’observation type’ vocabulary can be found in the OGC Definitions Server.

2.2 Description

The OGC API - Processes standard supports the wrapping of computational tasks into executable processes that can be offered by a server through a Web API and/or be invoked by a client-side application. The standard specifies a processing interface to communicate over a RESTful protocol using JavaScript Object Notation (JSON) to encode messages in client-server communication. Typically, these processes execute well-defined algorithms that ingest vector and/or coverage data to produce new datasets. Part 1 of the standard, referred to as OGC API - Processes - Part 1: Core, specifies the core capabilities that all implementations of the API standard must support.
The standard enables an application to describe its interface through the use of the OpenAPI Specification. The OpenAPI Specification is a specification for machine-readable definition documents that describe the interfaces of RESTful services. The OpenAPI definition documents identify the resource paths, parameters, types of errors, and schemas supported by an API. By formally specifying the capabilities supported by an API, the API provider enables more climate scientists to implement applications and analytical notebooks (e.g. Jupyter notebooks)

that make use of the API.
At the core of the OGC API - Processes standard is the ”Process” model. This is a data model that enables a client-based application to furnish the server with an identifier of the process to execute, a mode to determine whether the process should be executed asynchronously, the inputs to provide to the process, and identifiers of the outputs to expect from the process. For long-running processes, an additional field identifying the subscriber makes it possible to return the results of a job asynchronously.
The OGC API - Processes standard addresses the data volume challenge identified by Giuliani et al. (2017) in a number of ways. First, the standard enables the implementation of RESTful services that primarily communicate through JSON, which is easier to parse than XML. Second, the standard supports both push-based and callback mechanisms meaning that, for long-running processes, a client application can poll to check whether a job has completed or alternatively a client application may subscribe to a job and be notified by the server when the job is completed. Third, a candidate extension to the standard will enable deployment of code close to the data that it processes.

2.3 Transfer scientific methods to CRIS services

To transfer a scientific method performed individually by a scientist into an interoperable climate building block, the appropriate algorithms needs to be wrapped up into interoperable APIs. For deployment in CRIS the climate building blocks needs to be tailored to FAIR climate services principles to achieve a resilient interoperable global CRIS. It needs to be adaptable to local circumstances and flexible enough to quickly update to stay functional when science, policy and technology is changing over time. A software framework designed as stand-alone components is like a federated political union with member countries where common standards are agreed while a national independence is kept. Software developer communities contributing to a framework independently of each other and developed software components stays in their ownership and responsibility. Here we suggest OGC API as common standards.
An appropriate utility to facilitate the development is the Cookiecutter, which is a command-line utility to create projects from templates. The cookiecutter-birdhouse template creates a bare bone PyWPS server, adhering to Birdhouse conventions. It comes complete with a framework for installation, configuration, deployment, documentation and tests. With the template, it is possible to create a project then get started writing new WPS processes in only few minutes. Also, the project can be later updated with the latest cookiecutter template using Cruft.

Birdy is a Python library to work with Web Processing Services (WPS). It is using OWSLib from the GeoPython project to communicate with WPS services. Birdy is suitable to be used in Jupyter notebooks to interact with WPS services. It translates automatically each offered process by the WPS service into a pythonic function call. A remote function call can be used like a simple Python function.

3. EXISTING BUILDING BLOCKS

In recent years, scientific methods have already been deployed in CRIS, tailored to OGC Standards enabling interoperability (Hempelmann et al. (2018); Landry et al. (2019); Ehbrecht et al. (2018a)). Three example building blocks are described below. Further descriptions of the collection of OGC Standard-based

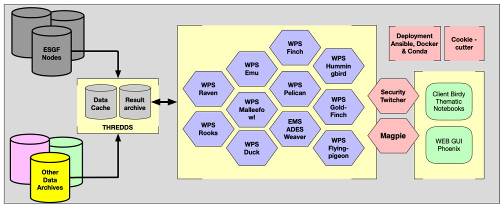

Figure 2. Example repositories of the birdhouse GitHub organisation: OGC Standard based building blocks for CRIS.

climate building blocks following the FOSS4G principles can be found in the birdhouse GitHub organisation. The transparent visibility of scientific algorithms and multiple usage moves FOSS4G solutions up to high quality, which is essential to enhance communication, accountability, and support efficiently the decision making processes (Street, 2016). Furthermore, a facilitation to reuse components avoids doubling of work and enables synergies of scientific collaboration.

3.1 Spatio-temporal subsetting with Rooks

Implemented in the COPERNICUS Climate change C3S are the Remote Operations on Climate Simulations ’roocs’. It is deployed there with is a set of tools and services to provide ”data-aware” processing of ESGF (Earth System Grid Federation) and other standards-compliant climate datasets from modelling initiatives such as CMIP6 and CORDEX. One example is ‘Rook’ an application using the OGC Web Processing service (WPS) standard, that enables remote operations, such as spatio-temporal subsetting, on climate model data. It exposes all the operations available in the ‘daops’ library based on Xarray.

3.2 Climate indices computation with Finch

For climate information to be useful to decision makers, it often has to be embedded in standard workplace practices. This typically implies converting large ensembles of climate projections to a few meaningful numbers relevant to operations or for planning purposes. For example, a public health professional could be interested in the frequency, intensity and duration of heatwaves, while an energy utility might want to see projections for heating and cooling degree days. Both measures are based on temperature time series, however, the value of the relevant indices is what gives meaning to the raw data.
To ensure the quality, reproducibility, and efficiency of those indices, the Ouranos Consortium, a Quebec-based climate service provider, developed xclim, a Python library for calculating numerous climate indices from a wide range of domains. The library has two layers, an inner computational layer handling unit conversions and implementing indice algorithms using xarray, and an outer compliance and metadata layer handling missing values, CF-Convention attributes and metadata language translations. xclim also provides data-inspection and quality control tools, and the project development has invoked changes to its base components, further empowering researchers whose workflows rely on tools such as xarray and pint. Building on xclim, another library called Finch uses PyWPS to wrap the outer layer as a Web Processing Service, letting third party applications compute climate indices remotely over CF-compliant NetCDF datasets.
Anecdotal experience suggests that climate analysts prefer working with xclim’s outer layer rather than WPS services based on them. Having computations mediated by a client-server interface creates barriers to code inspection and debugging. Al-

though indice algorithms are thoroughly tested, input datasets may contain corrupted data or attributes that are far easier to identify when the code runs locally. On the other hand, when inputs datasets are well known and quality-controlled, WPS services are appreciated by web developers, who are able to run on-the-fly computations via HTTP request.

3.3 Hydrological modelling with Raven

At the same time Finch was created, another WPS server called RavenWPS was set up to host hydrological modelling services. The idea explored in this project was to see if the complex workflows required to set up and run hydrological models under different climate scenarios could be streamlined using chained web processes. The motivation for this came from the realization that climate impact studies are very often multidisciplinary endeavours, with outputs from one discipline serving as inputs to another. In large academic projects, this serial dependence can create logistical complexity and anguish, where delays accrue and amplify as data cascades down the chain. The hope was that if each discipline packaged its expertise into dedicated WPS, using commonly-adopted standard data formats for inputs and outputs, each team could work in parallel, and run other teams' services, thereby alleviating serial dependency issues.
In practice, this approach worked reasonably well for simple processes running on known data sets, for example processes delineating watersheds or extracting physiographic information. For hydrological modelling itself, we implemented global hydrological models using WPS, but again the *distance* with the code imposed by the client-server interaction created additional friction in the development cycle. Also, hydrological models can get fairly complex, and creating a WPS-compatible interface for every model option is labour-intensive. It is very difficult to diagnose issues stemming from data-model-configuration interactions when the process is run remotely. Indeed, due to security concerns, error messages from servers do not include a full stack trace, leaving users to guess where errors are stemming from. Ongoing work deals with semi-distributed hydrological modelling, magnifying the complexity of model configuration. We are reevaluating which steps can be outsourced to remote web processes, and which are more easily done locally. Hopefully a balance can be found between the simplicity offered by remote services, and scientists' needs for interactive code inspection.

4. EXAMPLES OF AI ENHANCED CLIMATE SERVICES

The following examples are ongoing developments in the running CLINT project (sec. 1.2).

4.1 Processes to infill missing values

The WPS focusing on the spatial infilling of missing values will be the first service by the CLINT project. The process will have an interface to read NetCDF files, fill in the gaps of the given datasets by using a pre-trained neural network, produce an infilled NetCDF file and an illustrative visualization of the infilling (see Fig. 3).
To infill the missing climate data, we are using an inpainting method employing a neural network with a U-Net architecture and partial convolutional layers (Liu et al., 2018). Compared to standard convolutional layers, the partial convolution approach

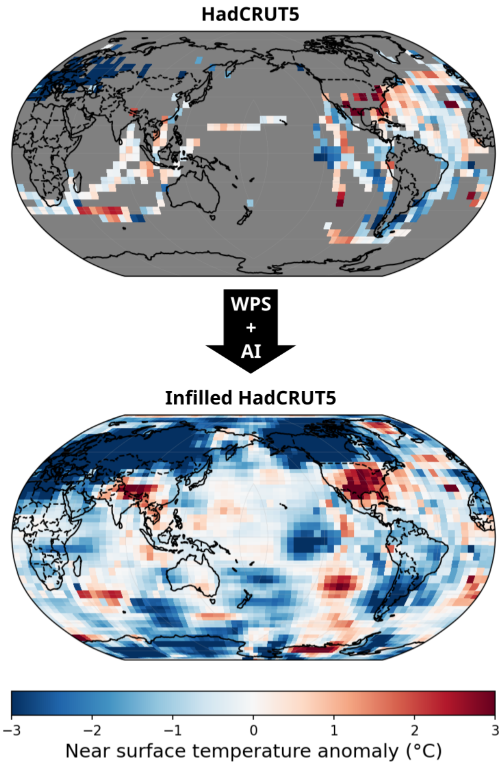

Figure 3. Illustration of the infilling process of the HadCRUT5 dataset. Top figure: original HadCRUT5 data for January 1850. Bottom figure: same as top figure where the missing data have been infilled using our deep learning based inpainting technology.

has shown to produce accurate results when applied to spatial data with large and irregular regions of missing data. As shown in (Kadow et al., 2020), this technique is particularly adapted to infill observational climate datasets which often contain many missing values, due to failures of the measuring instruments or to the impossibility to perform systematic measurements in earlier times.
The first model to be integrated to the WPS will be trained on near-surface air temperature anomalies from the Twentieth-Century Reanalysis (20CR) dataset and will support the infilling of HadCRUT4 and HadCRUT5 datasets.

In a later stage, the technology will be further developed and applied to extreme variables to determine their climate trends. The models resulting of this work will be integrated to the WPS in order to infill additional climate variables relevant to extreme events as well as extreme indices from e.g. the HadEX3 dataset.

4.2 Processes for bias-adjustment of seasonal climate predictions and hydro-climatic prediction skill assessment

Seasonal predictions of daily mean precipitation and temperature were taken from the CMCC seasonal prediction system (CMCC-SPS3.5) (Sanna et al., 2017). The CMCC-SPS3.5 hindcasts consist of 40 ensemble members available at a 0.5 degree grid resolution and for the period 1993–2016. The system provides 6-month predictions initialized at the beginning of each month.
The hindcast data were bias adjusted using the Distribution Based Scaling (DBS) method, which belongs to the family of quantile-mapping methods. The DBS parameters are conditioned on the lead month and the issue date, and the adjustment is done on

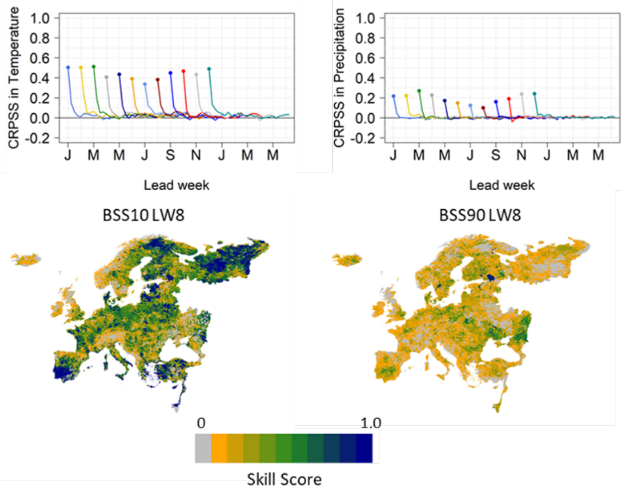

Figure 4. Median prediction skill over the pan-European domain in terms of CRPSS after bias-adjustment for temperature (top left) and precipitation (top right) for all initialization months and lead months. The bottom row shows the prediction skill in terms of BSS for low (BSS10) and high (BSS90) river discharges and for lead week 8.

all monthly initialized predictions using the HydroGFD data set (Berg et al., 2018) as reference. After bias adjustment, the cumulative distribution of daily precipitation and temperature predictions follows closely the one of the HydroGFD data. The bias-adjusted seasonal predictions forced the pan-European E-HYPE hydrological model (Hundecha et al., 2016) in order to simulate river discharge.
We finally quantified the skill for precipitation and temperature based on the Continuous Ranked Probability Skill Score (CRPSS). We also quantified the skill for low and high hydrological extremes (based on the 10th and 90th percentile, respectively) using the Brier Skill Score (BSS). For both metrics, positive skill values indicate higher performance from the CMCC-SPS3.5 system than climatology, which is used here as benchmark. Fig. 4 presents the results indicating skill for all initialisation months and variables. The skill varies in space, initialization month and lead time, which is considered as a success since the predictions could add value for long-term decision-making.

4.3 Processes for heatwave assessments

Heatwaves have a great impact on society, both from a health and economic point of view. In response, warning systems dedicated to heat stress are developed, requiring meteorological forecasts from which relevant indicators are developed expressing the impact of high and low temperatures on humans. Due to that, we describe the development of a service that performs a prediction of heatwaves based on the algorithms developed using machine learning. This service include three processes: 1) Detection of heatwaves through heatwave indices, 2) Drivers leading heatwaves, 3) Future projections of heatwaves using storylines (Fig. 5).
The detection of heatwaves from the input data (i.e. ERA5,C3S, CMIP) at certain domain (cases studies in Europe) will be focus in the computation of the indices such us the HWMI (Russo et al. (2014); Prodhomme et al. (2022)) in order to take into account different components (Magnitude, Intensity, duration, extension). These indices will be computed using machine learning. The output of the process will be a heatwave index.

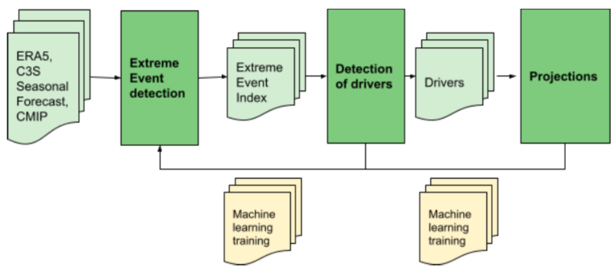

Figure 5. Scheme of the development of the service based on machine learning that performs a) detection of heatwaves, b) detection of drivers leading heatwaves, and c) projections of heatwaves using different datasets and timescales.

Once the heatwave index is computed, the drivers leading to heatwaves are detected through different variables (atmospheric circulation, soil moisture, albedo, etc) identified by the machine learning techniques developed in CLINT project. The drivers identified in this phase include large-scale fields as well as local land surface ones (e.g. land use, vegetation indices, albedo and soil moisture), and in this way two complementary setups of input drivers are used to train the statistical model.
The third process related to heatwaves will be focused on future changes of heatwaves, as an example of AI-enhanced Climate Science integrating efforts across CLINT project. The developed service will allow the user to get different storylines of plausible future changes in European heatwaves and their associated large-scale fields, including the best and the worst-case outcomes for different levels of global warming.

5. DISCUSSION

It is safe to say that due to the significant structural changes in our world as a result of climate change, "climate science is indeed under closer societal scrutiny than other scientific disciplines". In other words, climate science is held to higher standards than other fields, especially in terms of reproducibility and transparency. This is reflected for example in the rigorous review process IPCC assessment reports go through, or the meticulous curating of data and code underlying the US National Climate Assessments.
Climate services are responsible for providing customized tools, products and services to assess and monitor the range of climate change impacts and their long-range evolution. They should deliver on-demand climate information tailored to support sustainable development and achieve climate resilience (Commission, 2015; Dolman et al., 2016). In the year 2019, WMO reported 137 countries having climate services in place to support the agriculture and food security sector, especially in Africa and Europe (Cullmann et al., 2019). The various political instruments targeting climate resilience often strive for the so-called "policy-interoperability", in the sense that the objectives of one policy framework contribute to the success of other policies. However, policies remain vague on how to foster interoperability at the technical level.
At the moment, these CRIS are developed without international coordination or technological guidelines, and the result is a bazaar of heterogeneous, independent platforms that do not interoperate easily with each other. Beyond the wasted resources to reinvent the wheel, the current situation also creates confusion with users who are faced with multiple portals that do roughly the

same thing.

We argue here that there is potential for informal synergies through standard-prescriptive policy. That is, the policy drivers mentioned above should go beyond stating that information sharing is desired, and prescribe sets of standards through which information should be shared, or at least mechanisms through which standards should be developed. This would focus development efforts on a subset of technologies, and hopefully bring them to maturity faster.

There is of course a balance to achieve between prescription and innovation. If policies constrain technological options too strongly, innovators will work outside policy-bounded technological subset, defeating the purpose of mandating standards. If policies are too vague, as is the case currently, development efforts are spread thin over multiple technologies competing for limited funding and expertise.

Another approach has been for data providers to offer analytical services through backend applications. For example, the ESGF is coordinating efforts to develop and offer subsetting, regridding and averaging services that would let users pre-process datasets before downloading them locally. Someone interested in projections of average annual precipitations over Kenya would then only have to download one annual time series per simulation, instead of daily values over the entire globe. The Earth observations community has been embracing this approach, leveraging public and private cloud computing services running on analysis-ready "data cubes". In fact, the cloud computing model is becoming the dominant mode of work for most medium and large-scale Earth observation applications.

Over the last years, a number of projects have proposed virtual labs, where analytical services over curated datasets allow users to conduct part of their analysis remotely without having to download and host the original data. Another approach targeted at less technically-savvy scientists has been the development of user-friendly web portals, where pre-computed data products are offered via a graphical web interface. Such portals are however rarely cited by researchers (Sanderson et al., 2016) and not always designed for the broad array of potential users it targets (Swart et al., 2017).

Climate science and climate services draw upon a large variety of data from station observations, space and aerial remote sensing, model simulations, paleo-climate proxies, etc. For example, just within Model Intercomparison Projects, dozens of model outputs are combined to compute ensemble statistics. To make this possible, scientists had to agree on a common data format, conventions for meta-data, and a common vocabulary for variable names and units, among others.

Model intercomparison projects are examples where scientific communities have achieved a high degree of interoperability, allowing researchers to work relatively painlessly with data from different climate models. By comparison, the climate services world is still relatively far from this situation, but this is partially by design. Indeed, climate services strive to generate products according to end-users specifications. Nonetheless, there is certainly potential for increased interoperability in the realm of climate services, once agreements are reached on common vocabulary and metadata standards for typical algorithms and climate indicators.

The challenges and preliminary prototypes for services which are based on OGC API standards for processing and implementation of Artificial Intelligence (AI) solutions are going to be the future technology to enhance CRIS. The here presented blueprints on how AI-based scientific workflows can be ingested into CRIS to enhance climate services related to extreme weather and impact events can be seen as the workflow on how to transfer scientific methodes to technical climate services.

References

Akhtar-Schuster, M., Stringer, L. C., Erlewein, A., Metternicht, G., Minelli, S., Safriel, U., Sommer, S., 2017. Unpacking the concept of land degradation neutrality and addressing its operation through the Rio Conventions. *Journal of Environmental Management*, 195, 4–15.

Berg, P., Donnelly, C., Gustafsson, D., 2018. Near-real-time adjusted reanalysis forcing data for hydrology. *HESS*, 22(2), 989–1000.

Commission, E., 2015. A European Research and Innovation Roadmap for Climate Services.

Cortekar, J., Themessl, M., Lamich, K., 2020. Systematic analysis of EU-based climate service providers. *Climate Services*, 17, 100125.

Cullmann, J., Dilley, M., Fowler, J., Grasso, V. F., Kabat, P., Lúcio, F., Nullis, C., Repnik., M., 2019. Climate Services Information System. *2019 STATE OF CLIMATE SERVICES*, 21–33.

Déandreis, C., Pagé, C., Braconnot, P., Bärring, L., Bucchignani, E., de Cerff, W. S., Hutjes, R., Joussaume, S., Mares, C., Planton, S., Plieger, M., 2014. Towards a dedicated impact portal to bridge the gap between the impact and climate communities : Lessons from use cases. *Climatic Change*, 125(3), 333–347. http://dx.doi.org/10.1007/s10584-014-1139-7.

Dolman, A. J., Belward, A., Briggs, S., Dowell, M., Eggleston, S., Hill, K., Richter, C., Simmons, A., 2016. A post-Paris look at climate observations.

Ehbrecht, C., Kindermann, S., Stephens, A., Denvil, S., 2018a. Web Processing Services for Copernicus Climate Change Service. *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, 6491.

Ehbrecht, C., Landry, T., Hempelmann, N., Huard, D., Kindermann, S., 2018b. PROJECTS BASED ON THE WEB PROCESSING SERVICE FRAMEWORK BIRDHOUSE. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W8, 43–47. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-4-W8/43/2018/.

GFCS, 2014. Annex to the Implementation Plan of the Global Framework for Climate Services – Climate Services Information System Component. Technical report, World Meteorological Organization.

Giuliani, G., Nativi, S., Obregon, A., Beniston, M., Lehmann, A., 2017. Spatially enabling the Global Framework for Climate Services: Reviewing geospatial solutions to efficiently share and integrate climate data & information. *Climate Services*, 8(August), 44 - 58. http://dx.doi.org/10.1016/j.cliser.2017.08.003 http://www.sciencedirect.com/science/article/pii/S2405880716300772.

Hempelmann, N., Ehbrecht, C., Alvarez-Castro, C., Brockmann, P., Falk, W., Hoffmann, J., Kindermann, S., Koziol, B., Nangini, C., Radanovics, S., Vautard, R., Yiou, P., 2018.

Web processing service for climate impact and extreme weather event analyses. Flyingpigeon (Version 1.0). *Computers and Geosciences*, 110(Supplement C), 65 - 72.

Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I., 2016. A regional parameter estimation scheme for a pan-European multi-basin model. *Journal of Hydrology: Regional Studies*, 6, 90–111.

Kadow, C., Hall, D. M., Ulbrich, U., 2020. Artificial intelligence reconstructs missing climate information. *Nature Geoscience*, 13(6), 408–413.

Landry, T., Byrns, D., Caron, D., Charette-Migneault, F., 2019. OGC Earth System Grid Federation (ESGF) Compute Challenge Engineering Report. https://www.osti.gov/biblio/1570389.

Liu, G., Reda, F. A., Shih, K. J., Wang, T. C., Tao, A., Catanzaro, B., 2018. Image Inpainting for Irregular Holes Using Partial Convolutions. *Computer Vision—ECCV 2018 Lecture Notes in Computer Science*, 11215, 19–35.

Lúcio, F. D. F., Head, G., 2016. The global framework for climate services (GFCS). *Climate Services, 2*, 3, 52–53.

Muraya, C., 2018. Knowledge Management as an Enabler of the Paris Agreement Implementation in Africa. W. Leal Filho, E. Manolas, A. M. Azul, U. M. Azeiteiro, H. McGhie (eds), *Handbook of Climate Change Communication: Vol. 1: Theory of Climate Change Communication*, Springer International Publishing, Cham, 309–328.

Panenko, A., George, E., Lutoff, C., 2021. Towards the development of climate adaptation knowledge-action systems in the European Union: An institutional approach to climate service analysis. *Climate Services*, 24, 100265. https://doi.org/10.1016/j.cliser.2021.100265.

Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., Fargkoulidis, G., García-Serrano, J., 2022. Seasonal prediction of European summer heatwaves. *Climate Dynamics*, 58(7), 2149-2166. https://doi.org/10.1007/s00382-021-05828-3.

Russo, S., Dosio, A., Graversen, R., Sillmann, J., Saiote Carrao, H., Dunbar, M., A, S., P, M., P, B., Vogt, J., 2014. Magnitude of extreme heat waves in present climate and their projection in a warming world. *JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES*, 119(22), 12500-12512.

Sanderson, H., Hilden, M., Russel, D., Dessai, S., 2016. Database support for adaptation to climate change: An assessment of web-based portals across scales. *Integrated Environmental Assessment and Management*, 12(4), 627–631. http://doi.wiley.com/10.1002/ieam.1755.

Sanna, A., Borrelli, A., Athanasiadis, P., Materia, S., Storto, A., Tibaldi, S., Gualdi, S., 2017. CMCC-SPS3: The CMCC Seasonal Prediction System 3. *Centro Euro-Mediterraneo sui Cambiamenti Climatici*, CMCC Tech. Rep.(RP0285), 61. https://www.cmcc.it/it/publications/rp0285-cmcc-sps3-the-cmcc-seasonal-predictionsystem-3/.

Street, R. B., 2016. Towards a leading role on climate services in Europe: A research and innovation roadmap. *Climate Services*, 1, 2–5. http://dx.doi.org/10.1016/j.cliser.2015.12.001 http://www.sciencedirect.com/science/article/pii/S2405880715300157.

Su, Z., Timmermans, W., Zeng, Y., Schulz, J., John, V. O., Roebeling, R. A., Poli, P., Tan, D., Kaspar, F., Kaiser-Weiss, A. K., Swinnen, E., Toté, C., Gregow, H., Manninen, T., Riihelä, A., Calvet, J.-C., Ma, Y., Wen, J., 2018. An Overview of European Efforts in Generating Climate Data Records. *Bulletin of the American Meteorological Society*, 99(2), 349–359. https://doi.org/10.1175/BAMS-D-16-0074.1.

Swart, R. J., de Bruin, K., Dhenain, S., Dubois, G., Groot, A., von der Forst, E., 2017. Developing climate information portals with users: Promises and pitfalls. *Climate Services*. http://dx.doi.org/10.1016/j.cliser.2017.06.008.

Timofeyeva-Livezey, M. M., Horsfall, F. M. C., Pulwarty, R. S., Klein-Tank, A., Kolli, R. K., Hechler, P., Dilley, M., Ceron, J. P., Goodess, C., 2017. Climate Services Information System Activities in Support of The Global Framework for Climate Services Implementation. *AGU Fall Meeting Abstracts*, 2017, IN51A–0007.

UNFCC, 2015. Adoption of the Paris Agreement. Technical Report December, Paris, France.

Weichselgartner, J., Arheimer, B., 2019. Evolving Climate Services into Knowledge–Action Systems. *Weather, Climate, and Society*, 11(2), 385–399. https://doi.org/10.1175/WCAS-D-18-0087.1.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. https://doi.org/10.1038/sdata.2016.18.

WMO, 2017. White Paper on the Contribution of the Global Framework for Climate Services to Transforming our World : the 2030 Agenda for Sustainable Development (Agenda 2030).

ACKNOWLEDGMENT

The project 'Climate Intelligence' leading to the examples of AI-enhanced climate services showing in this paper has received funding from the European Union's Horizon 2020 Framework Programme in the call for 'Building a low-carbon, climate resilient future: climate action in support of the Paris Agreement (H2020-LC-CLA-2018-2019-2020)',
project number: 101003876.