

OGC® DOCUMENT: 23-020

External identifier of this OGC® document: XXXX



ENGINEERING REPORT FOR OGC CLIMATE RESILIENCE PILOT

ENGINEERING REPORT

DRAFT

Submission Date: # 2023-XX-XX

Approval Date: # 2023-XX-XX

Publication Date: # 2023-XX-XX

Editor: Guy Schumann, Albert Kettner, Nils Hempelmann

Notice: This document is not an OGC Standard. This document is an OGC Public Engineering Report created as a deliverable in an OGC Interoperability Initiative and is *not an official position* of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard.

Further, any OGC Engineering Report should not be referenced as required or mandatory technology in procurements. However, the discussions in this document could very well lead to the definition of an OGC Standard.

License Agreement

>Use of this document is subject to the license agreement at <https://www.ogc.org/license>

Copyright notice

Copyright © 2023 Open Geospatial Consortium
To obtain additional rights of use, visit <https://www.ogc.org/legal>

Note

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

CONTENTS

I.	EXECUTIVE SUMMARY	viii
II.	KEYWORDS	ix
III.	SECURITY CONSIDERATIONS	x
1.	INTRODUCTION	12
1.1.	Enhancing Interoperability for Climate Resilience Information Systems	12
1.2.	The Role of the Pilot	13
1.3.	Objectives	14
1.4.	Background	14
1.5.	Climate Indices	17
1.6.	Technical Challenges	18
1.7.	How is this Pilot Relevant to the Climate Resilience Domain Working Group?	19
2.	CONTRIBUTORS	22
2.1.	About Laubwerk	23
2.2.	About Pixalytics Ltd	23
2.3.	About Safe Software	23
2.4.	About Intact	24
2.5.	About Pelagis	24
3.	COMPONENTS	26
3.1.	Component workflow	26
4.	RAW DATA TO DATACUBES	29
4.1.	Using Datacubes for wildfire risk analysis	29
5.	RAW DATA TO ANALYSIS READY DATA (ARD)	45
5.1.	Transforming climate relevant raw data to ARD	45
5.2.	Use of API services for climate related use cases	46
5.3.	From Raw Data to ARD with the FME Platform	52
5.4.	A framework example for climate ARD generation	60
5.5.	ESRI Climate Resilience Data	69
6.	ARD TO DECISION READY INDICATOR (DRI)	77
6.1.	Wildfire hazard component	78
6.2.	The Blue Economy	81
6.3.	ECMWF – Copernicus (will be integrated with INTRODUCTION section)	84

7. FROM DATA TO VISUALIZATION	92
7.1. 5D Meta World	92
7.2. Visualizing the Impact of Climate Change and Mitigation on Vegetation	93
7.3. ESRI's Web Application	100
8. USE CASES	104
8.1. Drought Impact Use Cases (Wuhan University)	104
8.2. Analysis Ready Data (ARD) Use Case (D-100 Client instance by George Mason University)	107
8.3. Solar climate atlas for Poland – Climate Resilience Information System	111
8.4. Wildfire resilience in insurance (Intact)	113
8.5. D-100 Client (Pelagis)	116
8.6. Climate Resilience for Coastal Ecosystems	116
9. LESSONS LEARNED	120
10. FUTURE WORK	125
ANNEX A (INFORMATIVE) REVISION HISTORY	127
BIBLIOGRAPHY	129

LIST OF TABLES

Table 1	82
Table 2	83

LIST OF FIGURES

Figure 1	12
Figure 2	15
Figure 3	16
Figure 4	17
Figure 5 – CRIS overview	26
Figure 6 – High level workflow diagram that illustrates the interactions between data, models and the various components	27
Figure 7	30
Figure 8	31
Figure 9	31

Figure 10 – GNOSIS Map Server information resource for GNOSIS Global Grid zone 5-24-6E	32
Figure 11 – GNOSIS Map Server information resource for ISEA9Diamonds zone 5-24-6E	33
Figure 12 – GNOSIS Map Server information resource for ISEA9Diamonds zone 5-24-6E sections	34
Figure 13	34
Figure 14	34
Figure 15	35
Figure 16	35
Figure 17 – GNOSIS Map Server Map of tiles 3/4/17 in GNOSISGlobalGrid	35
Figure 18	35
Figure 19 – Sentinel-2 with image classification styling	36
Figure 20	36
Figure 21	36
Figure 22 – Sentinel-2 map with NDVI band arithmetic	37
Figure 23 – GeoJSON output	39
Figure 24 – GeoJSON output	41
Figure 25 – GeoJSON output	42
Figure 26	46
Figure 27 – Pixalytics drought severity workflow architecture	47
Figure 28 – Drought indicator calling code that generated the default output, which is the CDI in GeoJSON format	47
Figure 29 – Plot of the CDI for a point location in Canada (Latitude: 55.5 N Longitude: 99.1 W); generated using Copernicus Emergency Management Service information [2023]	49
Figure 30 – Global Drought Observatory Web Portal,	50
Figure 31 – Plot of the ECMWF precipitation and SPI, extended using the Canadian climate forecast data, compared to the GDO calculated SPI; generated using Copernicus Emergency Management Service information, Copernicus Climate Service and Canadian Centre for Climate Services data [2023]	51
Figure 32 – High level FME ARD workflow showing generation of climate scenario ARD and impacts from climate model, EO, IoT, infrastructure and base map inputs	54
Figure 33 – Environment Canada NetCDF GCM time series downscaled to Vancouver area. From:	55
Figure 34 – Data Cube to ARD: NetCDF to KML, Geopackage, GeoTIFF	55
Figure 35 – Extracted timestep grids: Monthly timesteps, period mean T, period max T	56
Figure 36 – Convert raster temperature grids into temperature contour areas by class	56
Figure 37 – Geopackage Vector Area Time Series: Max Yearly Temp	57
Figure 38 – FME Data Inspector: RCM NetCDF data cube for Manitoba temperature 2020-2099	57
Figure 39 – Modified ARD Worflow: NetCDF data cube to precipitation delta grids (future - historical) in Geopackage for LA	59
Figure 40 – The processing chain to produce harmonized ARD.	62

Figure 41 – Sentinel-2 RGB composite (red Band4, green Band3, blue Band2), over Hubei, acquired on October 22, 2020. (a) corresponds to the reflectance at the top of the atmosphere (L1C product); (b) corresponds to the surface reflectance after pre-processing.	63
Figure 42 – Flow of the SPEI calculation.	65
Figure 43 – The SPEI results for the date 2000_02_01.	66
Figure 44 – The UML model of WHU Data Cube.	68
Figure 45 – The coverage with the ID "2m_temperature_201602" in the Asian region.	69
Figure 46 – View API Resources	74
Figure 47 – List of buttons	75
Figure 48 – Two samples of IFC's current national wildfire hazard map	78
Figure 49 – IFC's exposure synthetic dataset, with Montreal – Ottawa corridor on the left, and close-up of Montreal on the right. Color scale represent relative risk density in each cell, while points are representative individual risks	81
Figure 50 – Architecture	83
Figure 51 – ARD Query: Monthly Max Temp Contours	85
Figure 52 – ARD Query: Max Mean Monthly Temp > 25C	85
Figure 53 – Town of Lytton - location where entire town was devastated by fire during the heat wave of July 2021 - same location highlighted in ARD query from heat risk query in previous figure	86
Figure 54 – FME Query Workflow: Geopackage precipitation delta time series to GeoJSON points	87
Figure 55 – FME Query Parameters: Geopackage precipitation delta time series to GeoJSON points	88
Figure 56 – FME Data Inspector: precipitation delta result showing potential drought risk for areas and times with significantly less precipitation than past	88
Figure 57	93
Figure 58 – Overview of the Visualized Region (Status Quo)	95
Figure 59 – Overview of the Visualized Region (Scenarios)	96
Figure 60 – Above the Corner Sunset Blvd and N Curson Ave Looking North-East (Status Quo)	96
Figure 61 – Above the Corner Sunset Blvd and N Curson Ave Looking North-East (Scenarios)	97
Figure 62 – Corner Franklin Ave And N Sierra Bonita Ave Looking East (Status Quo)	97
Figure 63 – Corner Franklin Ave And N Sierra Bonita Ave Looking East (Scenarios)	98
Figure 64 – Corner Hollywood Blvd And Camino Palmero St Looking Looking North (Status Quo)	98
Figure 65 – Corner Hollywood Blvd And Camino Palmero St Looking Looking North (Scenarios)	99
Figure 66 – US Project view	101
Figure 67 – US Project view	101
Figure 68 – Application output reports	102
Figure 69 – The technical architecture of the use-case for drought impact.	104
Figure 70 – Visualization of SPEI on a virtual globe.	105

Figure 71 – Drought risk map in part of China.	106
Figure 72 – The study area of the Poyang Lake case.	107
Figure 73 – Software Architecture	109
Figure 74 – Surface soil moisture percentile (year 2019-2022)	111
Figure 75 – Solar Climate atlas for Poland available on the IMGW website:	112
Figure 76 – FireSmart Canada's Home Ignition Zones [8]	114
Figure 77 – Wildland-Urban Interface for Canada, on the left. Extraction of the WUI using satellite-derived imagery, on the right. [5]	115

EXECUTIVE SUMMARY

In recognizing the impacts of climate change and monitoring the environment to achieve Sustainable Development Goals collaborative solutions are required. This calls for standardized data and tools. But how can we ensure an effective exchange of reliable information across disciplines without sacrificing individual users' needs? Climate services require vast volumes of data from different providers to be processed by various scientific ecosystems: Raw data needs to be transformed into analysis ready data and from there into indicators to become more useful for supporting decisions. To provide a processing infrastructure that supports collaboration, we need standards based on the principles of being findable, accessible, interoperable, and reusable. OGC standards are aligned with these principles allowing for the reuse of data refinement features across political boundaries, organizations, and administrative levels.

Policy instruments should include technological guidelines, for example mandate the use of international standards for data formats, metadata and machine to machine communication protocols, in order to foster interoperability and software reuse. This would strengthen international collaboration on software development, and in turn contribute to the deployment of effective, robust and scientifically credible climate resilience information systems. With increased data access and interoperability of data, processing tools and data infrastructure, it can reduce human and economic costs and ensure appropriate policies are implemented to benefit all.

The OGC Climate Resilience Pilot, as the initial phase of a series of long-term climate initiatives, aimed to transform geospatial data, technologies, and other capabilities into meaningful information for various stakeholders, including decision makers, scientists, policy makers, data providers, software developers, and service providers. The pilot demonstrated the establishment of data pipelines to convert vast amounts of raw data through various steps into decision-ready information. To get to decision ready information the data first needs to be organised from multiple sources into processing pipelines, towards analysis-ready formats. The importance of analysis-ready data and decision-ready indicators was emphasized through discussions on various aspects of GEODataCubes. The pilot explored scientific aspects of climate impact by examining case studies related to droughts, floods, and wildfires, highlighting assessment tools and the complexities of climate indices. Ultimately, this Climate Resilience Pilot serves as a valuable resource for making informed decisions to support and enhance climate action. It specifically assists the location community in developing powerful visualization and communication tools to effectively address ongoing climate threats such as heat, drought, floods, and wildfires.

One of the biggest gaps to date has been the challenge of translating the outputs of global climate models into specific impacts and risks at the local level. The climate modeling community has embraced standards and there is a wide array of data for modelers to exchange and compare, with numerous climate data services now available online. However, outside the weather and climate domain, planners and GIS analysts working for agencies responsible for climate change impacts have limited familiarity with and capacity to consume climate model results. Because of this, a key focus of this pilot was exploring methods for extracting essential climate variables (ECVs) from climate model output scenarios and transforming them into a form more readily consumable via GIS platforms and applied to the local level. Climate variables relevant to use case impacts were selected. Climate variable data cubes were extracted into

temporal and spatial ranges specific to the use cases. Finally the data structure was transformed from multidimensional gridded cubes into data forms more readily consumable by geospatial applications. For example open standards were employed such as 2D OGC geopackage and geojson point data and published to OGC API services — making them readily available and explorable by a much wider user community. These pilot data flows serve as useful examples of how climate model results can be translated into impacts and risks at the local level in a way that is easy to integrate into existing planning workflows.

With the target user group of non-technical decision makers, the workflow from data to visualisation is being shown at several chapters of this report. A dedicated chapter is pointing out the options and challenges of usage on artificial intelligence to establish a 5D meta world where the efficiency of climate action can be simulated. Reduction of disaster risks due to technical engineering constructions like dams are able to be simulated. Climate resilience is not only an aspect of shift of meteorological phenomena but also related to land degradation and loss of biodiversity. Therefor the vegetation is been pointed out and options of 3D vegetation simulation is shown and how different species are surviving under changing climate conditions. It could be shown that small scale urban planning is being supported by the data to visualisation application where single tree species are representing the real or simulated situation of a small scale area. The pilot is showing studies of Los Angeles.

The pilot recognizes the significant challenges associated with effectively conveying information to decision-makers, which necessitates a thorough examination of communication methods. As a result, a dedicated chapter has been incorporated into the pilot's work to address this issue. This chapter emphasizes unique approaches that facilitate effective communication with non-technical individuals who frequently hold responsibility for local climate resilience action strategies. By focusing on communication, the pilot aims to bridge the gap between technical and non-technical stakeholders, ensuring that vital information is conveyed accurately and comprehensively. The inclusion of this chapter serves as a testament to the pilot's aim to enhancing communication strategies for improved decision-making in the realm of climate resilience.

Based on the generated data information, there are several areas of focus and further exploration for the scenario tests and analysis in the context of climate data processing. Here is a breakdown of the key points that should be addressed in follow-on activities:

II

KEYWORDS

The following are keywords to be used by search engines and document catalogues.

Climate Resilience, data, ARD, component, use case, FAIR, Drought, Heat, Fire, Floods

SECURITY CONSIDERATIONS

No security considerations have been made for this document.

1

INTRODUCTION

INTRODUCTION

1.1. Enhancing Interoperability for Climate Resilience Information Systems

The OGC Climate Resilience Pilot will be the first phase of multiple long term climate activities aiming to evolve geospatial data, technologies, and other capabilities into valuable information for decision makers, scientists, policy makers, data providers, software developers, and service providers so we can make valuable, informed decisions to improve climate action. The goal is to help the location community develop more powerful visualization and communication tools to accurately address ongoing climate threats such as heat, drought, floods, fires as well as supporting the national determined contributions for greenhouse gas emission reduction. Climate resilience is often considered the use case of our lifetime, and the OGC community is uniquely positioned to accelerate solutions through collective problem solving with this initiative.



Figure 1

As illustrated, big, raw data from multiple sources requires further processing in order to be ready for analysis and climate change impact assessments. Applying data enhancement steps, such as bias adjustments, re-gridding, or calculation of climate indicators and essential variables, leads to “Decision Ready Indicators.” The spatial data infrastructures required for this integration should be designed with interoperable building blocks following FAIR data principles. Heterogeneous data from multiple sources can be enhanced, adjusted, refined, or quality controlled to provide Science Services data products for Climate Resilience. The OGC

Climate Change Services Pilots will also illustrate the graphical exploration of the Decision Ready Climate Data. It will demonstrate how to design FAIR climate services information systems. The OGC Pilot demonstrators will illustrate the necessary tools and the visualizations to address climate actions moving towards climate resilience.

1.2. The Role of the Pilot

The OGC Climate Resilience Community brings decision makers, scientists, policy makers, data providers, software developers, and service providers together. The goal is to enable everyone to take the relevant actions to address climate change and make well informed decisions for climate change adaptation. This includes scientists, decision makers, city managers, politicians, and last but not least, it includes everyone of us. So what do we need? We need data from lots of organizations, available at different scales for large and small areas to be integrated with scientific processes, analytical models, and simulation environments. We need data visualization and communication tools to shape the message in the right way for any client. Many challenges can be met through resources that adhere to FAIR principles. FAIR as in: Findable, Accessible, Interoperable, and Reusable. No single organization has all the data we need to understand the consequences of climate change. The OGC Climate Resilience Community identifies, discusses, and develops these resources. The OGC community builds the guidebooks and Best Practices, it experiments with new technologies to share data and information, and collaboratively addresses shared challenges.

The OGC Climate Resilience Community has a vision to support efforts on climate actions and enable international partnerships (SDG 17), and move towards global interoperable open digital infrastructures providing climate resilience information on users demand. This pilot will contribute to establishing an OGC climate resilience concept store for the community where all appropriate climate information to build climate resilience information systems as open infrastructures can be found in one place, be it Information about data services, tools, software, handbooks, or a place to discuss experiences and needs. The concept store covers all phases of Climate Resilience, from initial hazards identification and mapping to vulnerability and risk analysis to options assessments, prioritization, and planning, and ends with implementation planning and monitoring capabilities. These major challenges can only be met through the combined efforts of many OGC members across government, industry, and academia.

This Call for Participation solicits interests from organizations to join the upcoming Climate Resilience Pilot, an OGC Collaborative Solution and Innovation Program activity. This six-months Pilot is setting the stage for a series of follow up activities. It therefore focuses on use-case development, implementation, and exploration. It answers questions such as:

- What use-cases can be realized with the current data, services, analytical functions, and visualization capabilities that we have?
- How much effort is it to realize these use-cases?
- What is missing, or needs to be improved, in order to transfer the use-cases developed in the pilot to other areas?

1.3. Objectives

The pilot has three objectives. First, to better understand what is currently possible with the available data and technology. Second, what additional data and technology needs to be developed in future to better meet the needs of the Climate Resilience Community; and third, to capture Best Practices, and to allow the Climate Community to copy and transform as many use-cases as possible to other locations or framework conditions.

1.4. Background

With growing local communities, an increase in climate-driven disasters, and an increasing risk of future natural hazards, the demand for National Resilience Frameworks and Climate Resilience Information Systems (CRIS) cannot be overstated. Climate Resilience Information Systems (CRIS) are enabling data-search, -fetch, -fusion, -processing and -visualization. They enable access, understanding, and use of federal data, facilitate integration of federal and state data with local data, and serve as local information hubs for climate resilience knowledge sharing.

CRIS are already existing and operational, like the Copernicus Climate Change Service with the Climate Data Store. CRIS architectures can be further enhanced by providing climate scientific methods and visualization capabilities as climate building blocks. Based on FAIR principles, these building blocks enable in particular the reusability of Climate Resilience Information Systems features and capabilities. Reusability is an essential component when goals, expertises, and resources are aligned from the national to the local level. Framework conditions differ across the country, but building blocks enable as much reuse of existing Best Practices, tools, data, and services as possible.

Goals and objectives of decision makers vary at different scales. At the municipal level, municipal leaders and citizens directly face climate-related hazards. Aspects thus come into focus such as reducing vulnerability and risk, building resilience through local measures, or enhancing emergency response. At the state level, the municipal efforts can be coordinated and supported by providing funding and enacting relevant policies. The national, federal, or international level provides funding, science data, and international coordination to enable the best analysis and decisions at the lower scales.



Figure 2

Productivity and decision making are enhanced when climate building blocks are exchangeable across countries, organizations, or administrative levels (see Figure below). This OGC Climate Resilience Pilot is a contribution towards an open, multi-level infrastructure that integrates data spaces, open science, and local-to-international requirements and objectives. It contributes to the technology and governance stack that enables the integration of data including historical observations, real time sensing data, reanalyses, forecasts or future projections. It addresses data-to-decision pipelines, data analysis and representation, and bundles everything in climate resilience building blocks. These building blocks are complemented by Best Practices, guidelines, and cook-books that enable multi-stakeholder decision making for the good of society in a changing natural environment.

The OGC Innovation Program brings all groups together: The various members of the stakeholder group define use cases and requirements, the technologists and data providers experiment with new tools and data products in an agile development process. The scientific community provides results in appropriate formats and enables open science by providing applications that can be parameterized and executed on demand.

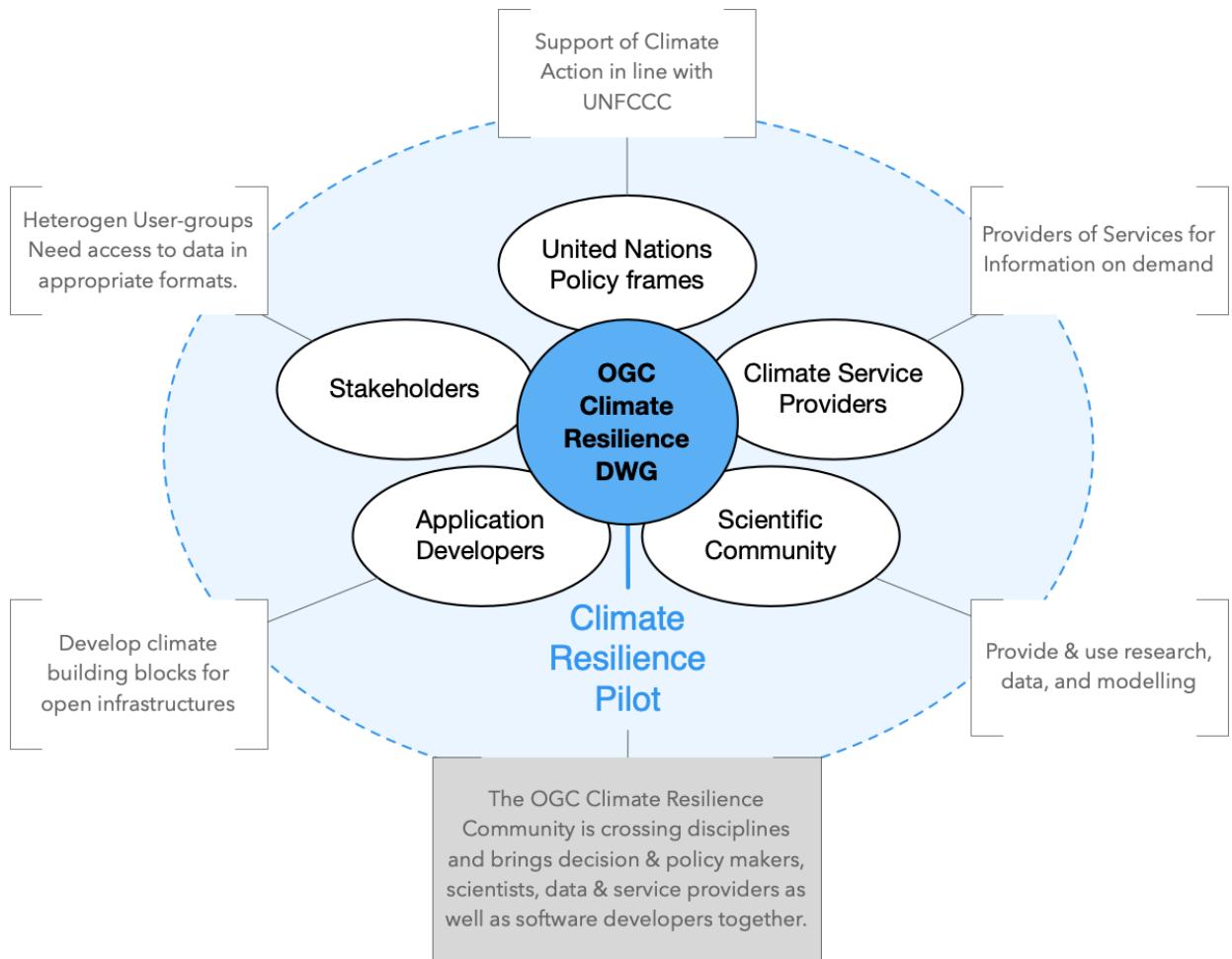


Figure 3

This OGC Climate Resilience Pilot is part of the OGC Climate Community Collaborative Solution and Innovation process, an open community process that uses the OGC as the governing body for collaborative activities among all members. A spiral approach is applied to connect technology enhancements, new data products, and scientific research with community needs and framework conditions at different scales. The spiral approach defines real world use cases, identifies gaps, produces new technology and data, and tests these against the real world use cases before entering the next iteration. Evaluation and validation cycles alternate and continuously define new work tasks. These tasks include documentation and toolbox descriptions on the consumer side, and data and service offerings, interoperability, and system architecture developments on the producer side. It is emphasized that research and development is not constrained to the data provider or infrastructure side. Many tasks need to be executed on the data consumer side in parallel and then merged with advancements on the provider side in regular intervals.

Good experiences have been made using OGC API standards in the past. For example, the remote operations on climate simulations (roocs) use OGC API Processes for subsetting data sets to reduce the data volume being transported. Other systems use OGC STAC for metadata and data handling or OGC Earth Observation Exploitation Platform Best Practices for the deployment of climate building blocks or applications into CRIS architectures. Still data handling regarding higher complex climate impact assessments within FAIR and open infrastructures

needs to be enhanced. There is no international recommendation or best practice on usage of existing API standards within individual CRIS. It is the goal of this pilot to contribute to the development of such a recommendation, respecting existing operational CRIS that are serving heterogen user groups

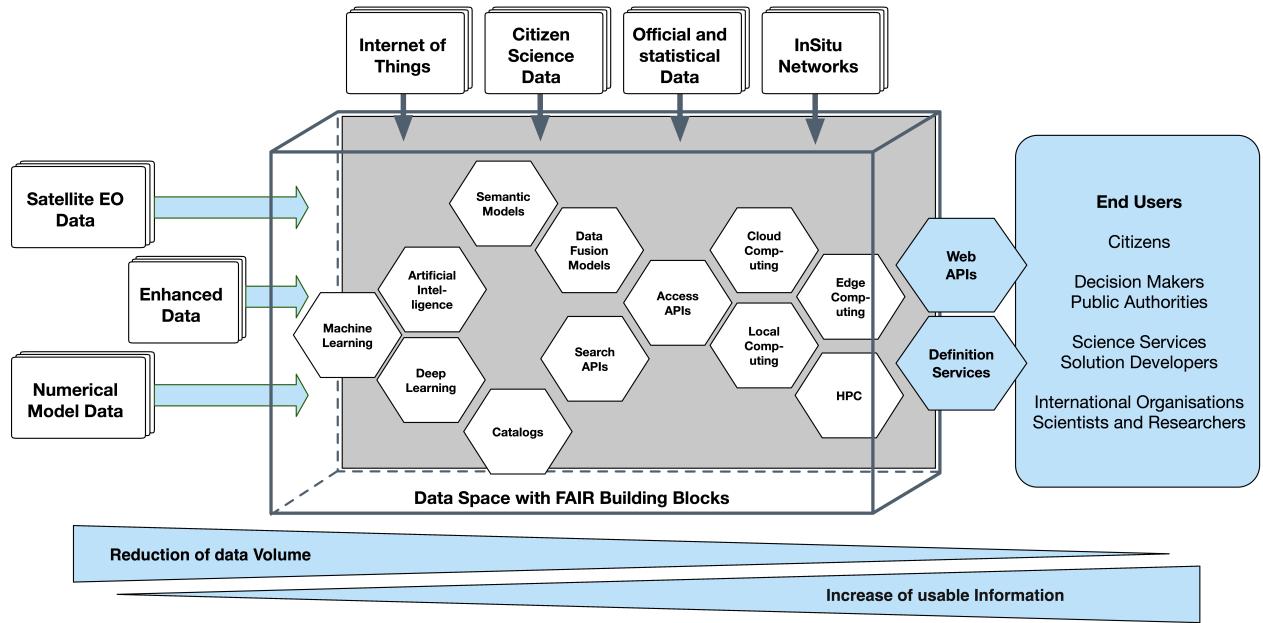


Figure 4

1.5. Climate Indices

To make planning decisions to build resilience and adapt to future climate, government officials from local to national, as well as corporate leaders and citizens need an approachable yet scientifically rigorous view of their local climate. We propose a dynamic web mapping interface and report generation tool backed by a suite of web services and downloadable data. All data and web service deliverables will be provided following FAIR principles at no cost, using the appropriate OGC standards.

The climate indices describe 47 measures of future temperature and precipitation in 3 future time periods (early, mid, late century) under 2 emission scenarios RCP 4.5 and 8.5. These indices were created to inform understanding of 5 climate hazards (Wildfire, Heat, Drought, Inland Flooding, Coastal Inundation). Wildfire and drought are the current focus of the disaster pilot and these climate indices will provide useful in those projects when considering future climate.

The project will present a pattern with reproducible workflows in an open Github repo showing the full process of transforming climate science data (CMIP model outputs) into a collection of analysis ready data layers (47 temperature and precipitation indices) and transforming those into decision ready information as climate indices summarized to local geographies such as counties and other subnational boundaries.

1.6. Technical Challenges

Realizing the delivery of Decision Ready Data on demand to achieve Climate Resilience involves a number of technical challenges that have already been identified by the community. A subset will be selected and embedded in use-cases that will be defined jointly by Pilot Sponsors and the OGC team. The goal is to ensure a clear value-enhancement pipeline as illustrated in Figure 1, above. This includes, among other elements, a baseline of standardised operators for data reduction and analytics. These need to fit into an overall workflow that provides translation services between upstream model data and downstream output – basically from raw data, to analysis-ready data, to decision-ready data. The following technical challenges have been identified and will be treated in the focus areas cycles of the Pilot accordingly:

- **Big Data Challenge:** Multiple obstacles still exist, creating big barriers for seamless information delivery starting from Data Discovery. Here the emergence of new data platforms, new processing functionalities, and thus new products, data discovery remains a challenge. In addition to existing solutions based on established metadata profiles and catalog services, new technologies such as OGC's Spatio-Temporal Asset Catalog (STAC) and open Web APIs such as OGC API Records will be explored. Furthermore, aspects of Data Access need to be solved where the new OGC API suite of Web APIs for data access, subsetting, and processing are currently utilized very successfully in several domains. Several code sprints have shown that server-side solutions can be realized within days and clients can interact very quickly with these server endpoints, thus development time is radically reduced. A promising specialized candidate for climate data and non-climate data integration has been recently published in the form of the OGC API – Environmental Data Retrieval (EDR). But which additional APIs are needed for climate data? Is the current set of OGC APIs sufficiently qualified to support the data enhancement pipeline illustrated in Figure 1? If not, what modifications and extensions need to be made available? How do OGC APIs cooperate with existing technologies such as THREDDS and OPEnDAP? For challenges of data spaces, Data Cubes have recently been explored in the OGC data cube workshop. Ad hoc creation and embedded processing functions have been identified as essential ingredients for efficient data exploration and exchange. Is it possible to transfer these concepts to all stages of the processing pipeline? How to scale both ways from local, ad hoc cubes to pan-continental cubes and vice versa. How to extend cubes as part of data fusion and data integration processes?
- **Cross-Discipline Data Integration:** Different disciplines such as Earth Observation, various social science, or climate modeling use different conceptual models in their data collection, production, and analytical processes. How can we map between these different models? What patterns have been used to transform conceptual models to logical models, and eventually physical models? The production of modern Decision-ready information needs the integration of several data sets, including census and demographics, further social science data, transportation infrastructure, hydrography, land use, topography and other data sets. This pilot cycle uses 'location' as the common denominator between these diverse data sets and works with several data providers and scientific disciplines. In terms of Data Exchange Formats the challenge is to know what data formats need to be supported at the various interfaces of the processing pipeline? What is the minimum constellation of required formats to cover the majority of use cases? What role do container formats play? Challenging on technical level is also the Data Provenance. Many

archives include data from several production cycles, such as IPCC AR 5 and AR 6 models. In this context, long term support needs to be realized and full traceability from high level data products back to the original raw data. Especially in context of reliable data based policy, clear audit trails and accountability for the data to information evolution needs to be ensured.

- Building Blocks for processing pipelines: With a focus on Machine Learning and Artificial Intelligence which plays an increasing role in the context of data science and data integration. This focus area needs to evaluate the applicability of machine learning models in the context of the value-enhancing processing pipeline. What information needs to be provided to describe machine learning models and corresponding training data sufficiently to ensure proper usage at various steps of the pipeline? Upcoming options to deploy ML/AI within processing APIs to enhance climate services are rising challenges e.g. on how to initiate or ingest training models and the appropriate learning extensions for the production phase of ML/AI. Heterogeneity in data spaces can be bridged with Linked Data and Data Semantics. Proper and common use of shared semantics is essential to guarantee solid value-enhancement processes. At the same time, resolvable links to procedures, sampling & data process protocols, and used applications will ensure transparency and traceability of decisions and actions based on data products. What level is currently supported? What infrastructure is required to support shared semantics? What governance mechanisms need to be put in place?

1.7. How is this Pilot Relevant to the Climate Resilience Domain Working Group?

The Climate Resilience DWG will concern itself with technology and technology policy issues, focusing on geospatial information and technology interests as related to climate mitigation and adaptation as well as the means by which those issues can be appropriately factored into the OGC standards development process.

The mission of the Climate Resilience DWG is to identify geospatial interoperability issues and challenges that impede climate action, then examine ways in which those challenges can be met through application of existing OGC Standards, or through development of new geospatial interoperability standards under the auspices of OGC.

Activities to be undertaken by the Climate Resilience DWG include but are not limited to:

- Identify the OGC interface standards and encodings useful to apply FAIR concepts to climate change services platforms;
- Liaise with other OGC Working Groups (WGs) to drive standards evolution;
- Promote the usage of the aforementioned standards with climate change service providers and policy makers addressing international regional and local needs;

- Liaise with external groups working on technologies relevant to establishing ecosystems of EO Exploitation Platforms;
- Liaise with external groups working on relevant technologies;
- Publish OGC Technical Papers, Discussion Papers or Best Practices on interoperable interfaces for climate change services;
- Provide software toolkits to facilitate the deployment of climate change services platforms.

2

CONTRIBUTORS

CONTRIBUTORS

NAME	ORGANIZATION	ROLE OR SUMMARY OF CONTRIBUTION
Guy Schumann	RSS-Hydro	Lead ER Editor
Albert Kettner	RSS-Hydro/DFO	Lead ER Editor
Timm Dapper	Laubwerk GmbH	
Zhe Fang	Wuhan University	
Hanwen Xu	Wuhan University	
Peng Yue	Wuhan University	
Dean Hintz	Safe Software, Inc.	
Kailin Opaleychuk	Safe Software, Inc.	
Jérôme Jacovella-St-Louis	Ecere Corporation	
Hanna Krimm	alpS GmbH	
Andrew Lavender	Pixalytics Ltd	Development of drought indicator
Samantha Lavender	Pixalytics Ltd	Development of drought indicator
Jenny Cocks	Pixalytics Ltd	Development of drought indicator
Jakub Walawender	Walawender, Jakub P.	
Eugene Yu	GMU	
Gil Heo	GMU	
Glenn Laughlin	Pelagis Data Solutions	Coastal Resilience & Climate Adaptation
Patrick Dion	Ecere	
Tom Landry	Intact Financial Corporation	

NAME	ORGANIZATION	ROLE OR SUMMARY OF CONTRIBUTION
Nils Hempelmann	OGC	Climate resilience Pilot Coordinator

2.1. About Laubwerk

Laubwerk is a software development company whose mission is to combine accurate, broadly applicable visualizations of vegetation with deeper information and utility that goes far beyond their visual appearance. We achieve this through building a database that combines ultra-realistic 3D representation of plants with extensive metadata that represents plant properties. This unique combination makes Laubwerk a prime partner to bridge the gap from data-driven simulation to eye-catching visualizations.

2.2. About Pixalytics Ltd

Pixalytics Ltd is an independent consultancy company specializing in Earth Observation (EO). We combine cutting-edge scientific knowledge with satellite and airborne data to provide answers to questions about our planet's resources and behavior. The underlying work includes developing algorithms and software, with activities including a focus on EO quality control and end-user focused applications.

2.3. About Safe Software

Safe Software is a leader in supporting geospatial interoperability and automation for more than 25 years as creators of the FME platform. FME was created to promote FAIR principles, including data sharing across barriers and silos, with unparalleled support for a wide array of both vendor specific formats and open standards. Within this platform, Safe Software provides a range of tools to support interoperability workflows. FME Form is a graphical authoring environment that allows users to rapidly prototype transformation workflows in a no-code environment. FME Flow then allows users to publish data transforms to enterprise oriented service architectures. FME Hosted offers a low cost, easy to deploy and scalable environment for deploying transformation and integration services to the cloud.

Open standards have always been a core strategy for Safe to better support data sharing. The FME platform can be seen as a bridge between the many supported vendor protocols and open standards such as XML, JSON and OGC standards such as GML, KML, WMS, WFS and OGC APIs. Safe has collaborated extensively over the years with the open standards community. Safe actively participates in the CityGML and INSPIRE communities in Europe. We are also active within the OGC community and participated in many initiatives including test beds, pilots such

as Maritime Limits and Boundaries and IndoorGML, and most recently the 2021 Disaster Pilot and 2023 Climate Resilience Pilot. Safe also actively participates in a number of Domain and Standards working groups.

2.4. About Intact

Intact Financial Corporation (IFC) is the largest provider of Property & Casualty (P&C) insurance in Canada. Its purpose is to help people, businesses and society prosper in good times and be resilient in bad times [1]. The company has been on the front lines of climate change with our customers for more than a decade – getting them back on track and helping them adapt. As extreme weather is going to get worse over the next decade, Intact intends to double down on adapting to this changing environment and be better prepared for floods, wildfire and extreme heat [2].

With close to 500 experts in data, artificial intelligence, machine learning and pricing, the Intact Data Lab has deployed almost 300 AI models in production to-date. It is focused on improving risk selection and making operations as efficient as possible while creating outstanding interactions with customers. Within Intact's Data Lab, the Centre for Climate and Geospatial Analytics (CCGA) uses weather, climate, and geospatial data along with machine learning models and claims data to develop risk maps and other specialized products to the business.

2.5. About Pelagis

Pelagis is an OceanTech venture located in Nova Scotia, Canada. Our foundation focuses on the application of open geospatial technology and standards designed to promote the sustainable use of our ocean resources. As a member of the Open Geospatial Consortium, we co-chair the Marine Domain Working Group responsible for developing a spatially-aware federated service model of marine and coastal ecosystems.



3

COMPONENTS

COMPONENTS

The various organizations and institutes that contribute to the Climate Resilience Pilot are described below. Their input to the pilot is indicated in the figure below Figure 5.

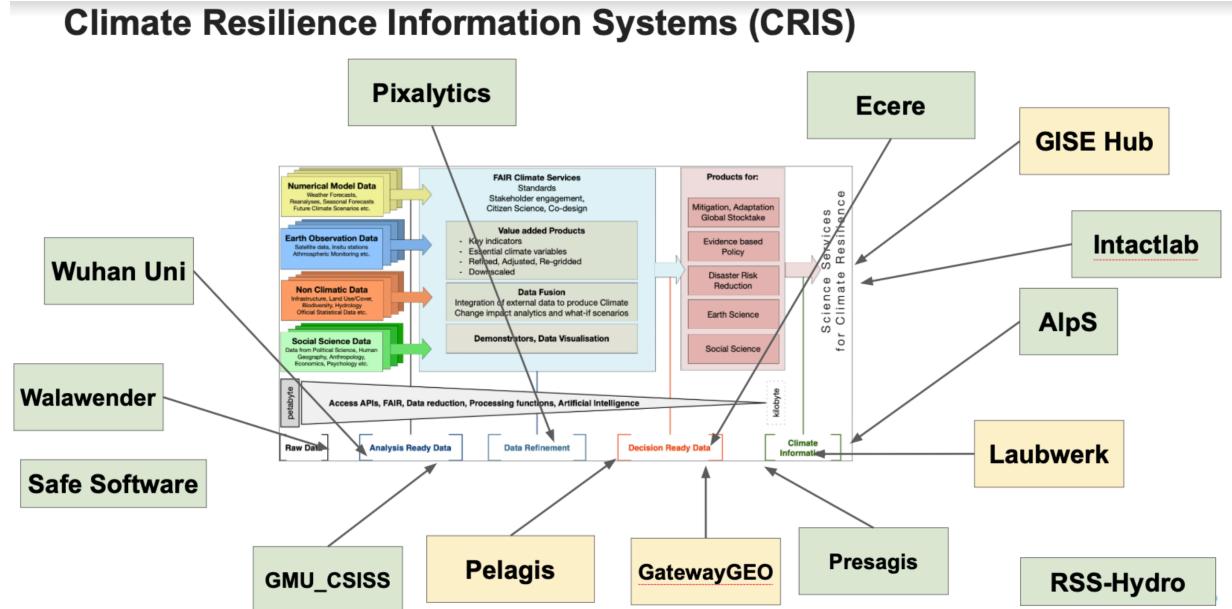


Figure 5 – CRIS overview

3.1. Component workflow

The figure below shows a high level workflow diagram that illustrates the interactions between data, models and the various components.

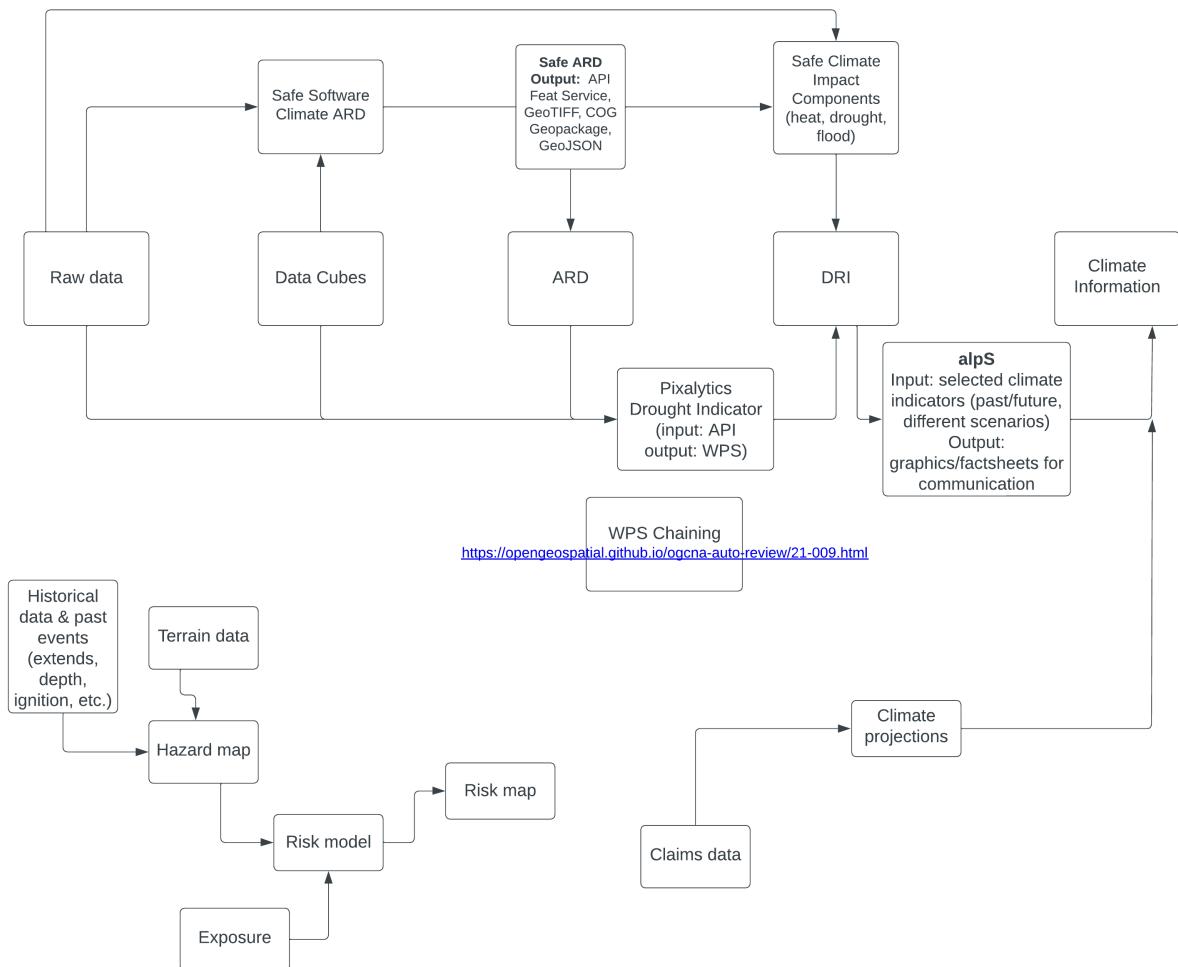


Figure 6 – High level workflow diagram that illustrates the interactions between data, models and the various components

4

RAW DATA TO DATACUBES

RAW DATA TO DATACUBES

Raw data and Datacubes are two different forms of organizing and structuring data in the context of data analysis and data warehousing.

1. Raw Data refers to the unprocessed, unorganized, and unstructured data that is collected or generated directly from various sources. It can include a variety of forms such as text, numbers, (geo) images, audio, video, or any other form of data. Raw data often lacks formatting or context and requires further processing or manipulation before it can be effectively analyzed or used for decision-making purposes. Raw data is typically stored in databases or data storage systems.
2. Datacubes, also known as multidimensional cubes are a structured form of data representation that organizes and aggregates raw data into a multi-dimensional format. Datacubes are designed to facilitate efficient and fast analysis of data from different dimensions or perspectives. They are commonly used in data warehousing.

Datacubes organize data into a multi-dimensional structure typically comprising dimensions, hierarchies, and cells. Dimensions represent various attributes or factors that define the data, such as time, geography, or products. Hierarchies represent the levels of detail within each dimension. Cells typically store the aggregated data values at the intersection of dimensions.

Datacubes enable users to perform complex analytical operations like slicing, dicing, drilling down, or rolling up data across different dimensions. They provide a summarized and pre-aggregated view of data that can significantly speed up query processing and analysis compared to working directly with raw data, something that is very valuable for the climate resilience community. Therefore, Datacubes are often used to support decision-making processes. The example below highlights an climate resilience related example of how to create and make available Datacubes for wildfire risk analysis.

4.1. Using Datacubes for wildfire risk analysis

Ecere is providing a deployment of its GNOSIS Map Server with a focus on a Sentinel-2 Level 2A data cube. OGC API - Tiles, OGC API - Coverages, OGC API - Maps, OGC API – Discrete Global Grid Systems, Common Query Language (CQL2), and OGC API – Processes – Part 3: Workflows & Chaining are the supported standards and extensions for this task.

The plan is to use machine learning process output from the Wildland Fire Fuel Indicator Workflow to identify vegetation fuel types from sentinel-2 bands, then combine with weather data to assess wildfire hazards risk in Australia. The workflow will use as input the sentinel-2 OGC API data cube from our GNOSIS Map Server.

- Component: Data Cube and Wildfire vegetation fuel map / risk analysis.

- Inputs: ESA Sentinel-2 L2A data (from AWS / Element 84), Temperature / Precipitation / Wind climate data, Reference data for training: vegetation fuel type classification, wildfire risk.

The sentinel-2 Level 2A collection is provided at <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a>

- Outputs: OGC API (Coverage, Tiles, DGGS, Maps) for Sentinel-2 data (<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a>) including full global coverage, all resolutions/scales, all bands that can be individually selected, CQL2 expressions for band arithmetics; climate data (to be added), vegetation fuel type (possibly by end of pilot, or for DP2023), wildfire risk workflow (possibly by end of pilot, or for DP2023).
- What other component(s) can interact with the component: Any OGC API client component requiring efficient access to Sentinel-2 data, clients requiring climate data once made available, clients presenting vegetation fuel type, wildfire risk (once ready, might extend into DP2023).
- What OGC standards or formats does the component use and produce:
 - OGC API (Coverage – with subsetting, scaling, range subsetting, coverage tiles; Tiles, DGGS (GNOSISGlobalGrid and ISEA9R), Maps (incl. map tiles), Styles), CQL2, OGC API – Processes with Part 3 for workflows (Nested Local/Remote Processes, Local/Remote Collection Input, Collection Output, Input/Output Field Modifiers)
 - Formats: GNOSIS Map Tiles (Gridded Coverage, Vector Features, Map imagery, and more); GeoTIFF; PNG (16-bit value single channel for coverage, RGBA for maps); JPEG.

4.1.1. Overview of standards and extensions available for outputs

4.1.1.1. OGC API – DGGS

There are two main requirements classes for this standard.

- Data Retrieval (What is here? – "give me the data for this zone"),
- Zones Query (Where is it? – "which zones match this collection and/or my query")

Example of data retrieval queries:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/dggs/GNOSISGlobalGrid/zones/3-4-11/data> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/dggs/ISEA9Diamonds/zones/E7-FAE/data>

Figure 7

Example of a zones query:

```
https://maps.gnosis.earth/ogcapi/collections/SRTM\_ViewFinderPanorama/dggs/ISEA9Diamonds/zones https://maps.gnosis.earth/ogcapi/collections/SRTM\_ViewFinderPanorama/dggs/ISEA9Diamonds/zones?f=json (as a list of compact JSON IDs)
```

Figure 8

Level, Row, Column (which encoded differently in the compact hexadecimal zone IDs) can be seen on the zone information page at:

```
https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/dggs/GNOSISGlobalGrid/zones/3-4-11 https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/dggs/ISEA9Diamonds/zones/E7-FAE
```

Figure 9

There are several different discrete global grids. Two are implemented in our service:

- Our GNOSIS Global Grid, which is geographic rather than projected, and is axis-aligned with latitudes and longitudes, but not equal area (though it tends towards equal area – maximum variation is ~48% up to a very detailed level)
- ISEA9R, which is a dual DGGS of ISEA3H even levels, using rhombuses/diamonds instead of hexagons, but much simpler to work with and can transport the hexagon area values as points on the rhombus vertices for those ISEA3H even levels. It is also axis-aligned to a CRS defined by rotating and skewing the ISEA projection.

The primary advantage of OGC API – DGGS is:

- for retrieving data from DGGS that are not axis-aligned or have geometry that cannot be represented as squares (e.g., hexagons), or
- for the zone query capability, most useful for specifying queries (e.g. using CQL2). The extent to which we implement Zones Query at this moment is still limited.

Examples of DGGS Zone information page:

Zone 5-24-6E

for GNOSISGlobalGrid DGGS (sentinel2-l2a)

(View [JSON](#), [ECON](#), [GeoJSON](#) / on [geojson.io](#) representation)

[Back to GNOSISGlobalGrid DGGS zones](#)

Hierarchy Level: 5

Latitude Row: 36

Longitude Column: 110

Zone Surface Area: **95,052.63** kilometers square (*1.82% more* than reference mean zone for level 5)

GNOSIS Hash Code: **SE**

Centroid: { lat: -12.65625, lon: 132.1875 }

Extent: { { lat: -14.0625, lon: 129.375 }, { lat: -11.25, lon: 132.1875 } }

[Download data](#)



Child	Zone ID	Code	Map
North-West	6-48-DC	SEG	
South-West	6-49-DC	SEI	
North-East	6-48-DD	SEH	
South-East	6-49-DD	SEJ	

Neighbor	Sibling	Zone ID	Code	Map
North	No	5-23-6F	SB	
West	No	5-24-6D	SE	
East	Yes	5-24-6F	SE	
South	Yes	5-25-6E	SE	

Figure 10 – GNOSIS Map Server information resource for GNOSIS Global Grid zone 5-24-6E

Zone A7-0**for ISEA9Diamonds DGGS (sentinel2-l2a)***(View [JSON](#), [ECON](#) representation)**[Back to ISEA9Diamonds DGGS zones](#)*Hierarchy Level: **0**

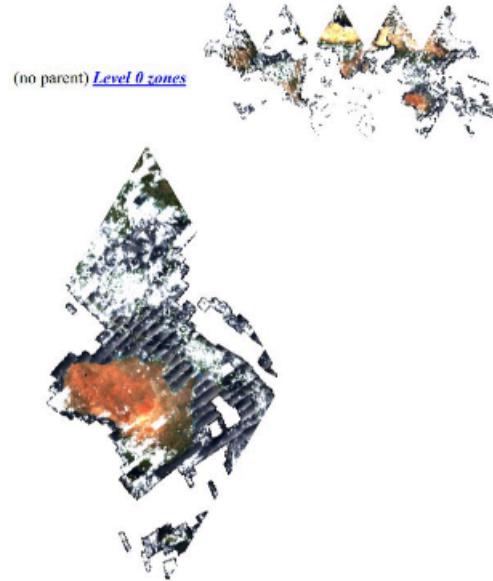
Row: 4

Column: 3

Zone Surface Area: **51,054,321.41** kilometers square (**0.09% more** than reference mean zone for level 0)

Centroid: { lat: -14.1369208534205, lon: 179.9381166703972 }

Extent: { { lat: -60.1636837367554, lon: -179.4717320723485 }, { lat: 31.8898420299144, lon: 179.9381166703972 } }

[Download data](#)**Figure 11** – GNOSIS Map Server information resource for ISEA9Diamonds zone 5-24-6E

Child	Zone ID	Map	
<i>Top-Left</i>	B7-0		
<i>Top-Center</i>	B7-1		
<i>Top-Right</i>	B7-2		
<i>Middle-Left</i>	B7-3		
<i>Middle-Center</i>	B7-4		
<i>Middle-Right</i>	B7-5		
<i>Bottom-Left</i>	B7-6		
<i>Bottom-Center</i>	B7-7		
<i>Bottom-Right</i>	B7-8		
Neighbor	Sibling	Zone ID	Map
<i>Top</i>	Yes	A6-0	
<i>Right</i>	Yes	A8-0	

Figure 12 – GNOSIS Map Server information resource for ISEA9Diamonds zone 5-24-6E sections

4.1.1.2. OGC API – Coverages with OGC API – Tiles

Because they are axis-aligned, both of these DGGS can be described as a TileMatrixSet, and therefore equivalent functionality to the OGC API – DGGS Data Retrieval requirements class can be achieved using OGC API – Tiles and the corresponding TileMatrixSets instead.

Coverage Tile queries for the same zones:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/GNOSISGlobalGrid/3/4/17> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/ISEA9Diamonds/4/373/288>

Figure 13

To request a different band than the default RGB (B04, B03, B02) bands:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/GNOSISGlobalGrid/3/4/17?properties=B08> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/ISEA9Diamonds/4/373/288?properties=B08>

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/ISEA9Diamonds/4/373/288?properties=B08>

Figure 14

To retrieve coverage tiles with band arithmetic to compute NDVI:

[https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/GNOSISGlobalGrid/3/4/17?properties=\(B08/10000-B04/10000\)/\(B08/10000+B04/10000\)](https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/GNOSISGlobalGrid/3/4/17?properties=(B08/10000-B04/10000)/(B08/10000+B04/10000)) [https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/ISEA9Diamonds/4/373/288?properties=\(B08/10000-B04/10000\)/\(B08/10000+B04/10000\)](https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/coverage/tiles/ISEA9Diamonds/4/373/288?properties=(B08/10000-B04/10000)/(B08/10000+B04/10000))

Figure 15

4.1.1.3. OGC API – Maps with OGC API – Tiles

Map Tiles queries for the same zones:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/GNOSISGlobalGrid/3/4/17> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/ISEA9Diamonds/4/373/288>

Figure 16



Figure 17 – GNOSIS Map Server Map of tiles 3/4/17 in GNOSISGlobalGrid

To retrieve a map of the Scene Classification:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/styles/scl/map/tiles/GNOSISGlobalGrid/3/4/17> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/styles/scl/map/tiles/ISEA9Diamonds/4/373/288>

Figure 18

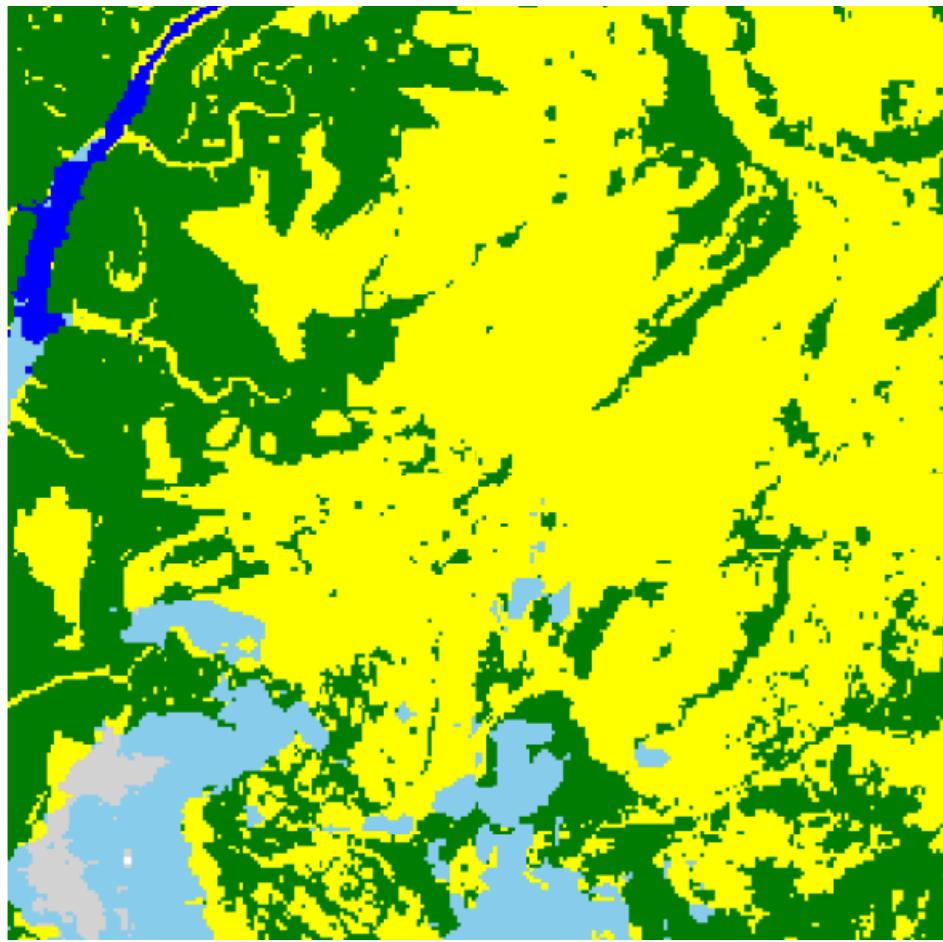


Figure 19 – Sentinel-2 with image classification styling

To filter out the clouds:

[https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/GNOSISGlobalGrid/3/4/17?filter=SCL<8 or SCL >10](https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/GNOSISGlobalGrid/3/4/17?filter=SCL<8%20or%20SCL%20>10) [https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/ISEA9Diamonds/4/373/288?filter=SCL<8 or SCL >10](https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/map/tiles/ISEA9Diamonds/4/373/288?filter=SCL<8%20or%20SCL%20>10)

Figure 20

To get an NDVI map:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/styles/ndvi/map/tiles/GNOSISGlobalGrid/3/4/17> <https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a/styles/ndvi/map/tiles/ISEA9Diamonds/4/373/288>

Figure 21

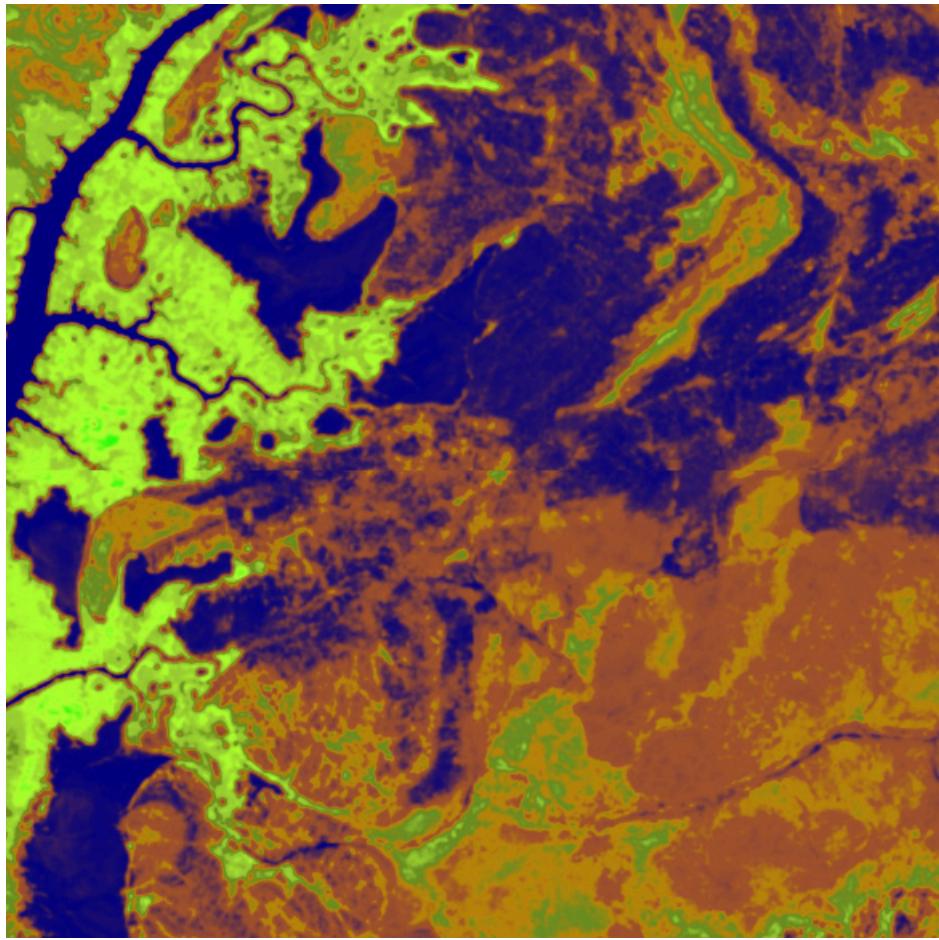


Figure 22 – Sentinel-2 map with NDVI band arithmetic

The same filter= and properties= should also work with the /coverage and /dggs end-points. The filter= also works with the /map end-points.

4.1.2. GNOSIS implementation of OGC API for climate data cube (2016-2025 CMIP5 data)

There is now a fairly complete set of variables from the CMIP5 global dataset (from the Copernicus Climate Data Store) for the 2016-2025 time period available from our GNOSIS data cube implementation at: <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5>

The variables on a single pressure level are organized as a single collection (coverage / data cube) at: <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure> (consisting of 9 fields: specific humidity, precipitation, snowfall, sea level pressure, downwelling shortwave radiation, wind speed, mean surface air temperature, maximum daily air temperature, minimum daily air temperature), while the variables on multiple pressure levels are organized into three separate collections: <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:temperature> <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:gpHeight> <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:gpDepth>

byPressureLevel:windSpeed (consisting of two separate fields for Eastward and Northward wind velocity)

The temporal resolution of this dataset is daily, while the source spatial resolution is 2.5 degrees longitude x 2 degrees of latitude, and it is for 8 different pressure levels. Currently, the API supports requesting data from this data using OGC API – Tiles (coverage tiles as well as map tiles), Coverages, Maps and DGGS. With all these APIs, a specific pressure level can be specified for the multi-pressure using e.g., `subset=pressure(500)`, while a specific time can be requested using e.g., `datetime=2022-03-01` or `subset=time("2022-03-01")`. With Coverages and Maps, a spatial area of interest can be specified using either e.g., `bbox=10,20,30,40` or `subset=Lat(20:40),Lon(10:30)`.

At the moment, the Coverages API is limited to 2D output formats (spatial trim, slicing by time and pressure): GeoTIFF and PNG (16-bit output, currently fixed scale: 2.98 and offset: 16384). There is a plan to add support for n-dimensional output formats, including netCDF, CIS JSON and eventually CoverageJSON as well. Currently, separate API requests with the above parameters are needed for different times/pressure levels.

For coverage output, the fields can be selected using `properties=` (a single field for PNG, and one or more fields for GeoTIFF) e.g., `properties=tasmin,tasmax` The fields can also be derived using CQL2 expressions that can perform arithmetic e.g., `properties=pr*1000`.

With all these APIs, it is also possible to filter fields with `filter=` also specified as a CQL2 expression e.g., `filter=tasmax>300` (unmatched cells will be replaced by NODATA values). The domains of the collections are described in the collection description (inside the `extent` property) as well as in the Coverages CIS DomainSet resource e.g., <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure?f=json> , <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage/domainset?f=json>

The ranges of the collections are described in the Coverages CIS RangeType resource as per the example below, and we are also planning to implement describing in a /schema resource that will be harmonized with the OGC API – Features schema. <https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage/rangetype?f=json>

Some sample requests: Maps

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/styles/precipitation/map?datetime=2022-09-04>

[https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:windSpeed/map?subset=pressure\(850\)&height=1024](https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:windSpeed/map?subset=pressure(850)&height=1024)

Proper symbolization here will require support for wind barbs – in the meantime the Eastward and Northward velocity are assigned to the green and blue color channels.

[https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:temperature/map?subset=pressure\(850\)](https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:byPressureLevel:temperature/map?subset=pressure(850))

Tiles

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/styles/precipitation/map/tiles/WebMercatorQuad/1/1/0?datetime=2022-09-04>

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage/tiles/WebMercatorQuad/1/1/0?f=geotiff&datetime=2022-09-04> (GeoTIFF Coverage Tile)

DGGS

Data retrieval – What is here? (equivalent to Coverage Tiles requests for DGGSs whose zone geometry can be described by a 2D Tile Matrix Set e.g., GNOSISGlobalGrid, ISEA9R, rHealPix):

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/dggs/GNOSISGlobalGrid/zones/0-0-3/data?f=geotiff&datetime=2022-09-04>

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/dggs/ISEA9Diamonds/zones/A7-0/data?f=geotiff&datetime=2022-09-04>

Zones query – Where is it?: Where is maximum daily temperature greater than 300 degrees Kelvins on September 4, 2022? (at precision level of GNOSIS Global Grid level 6)

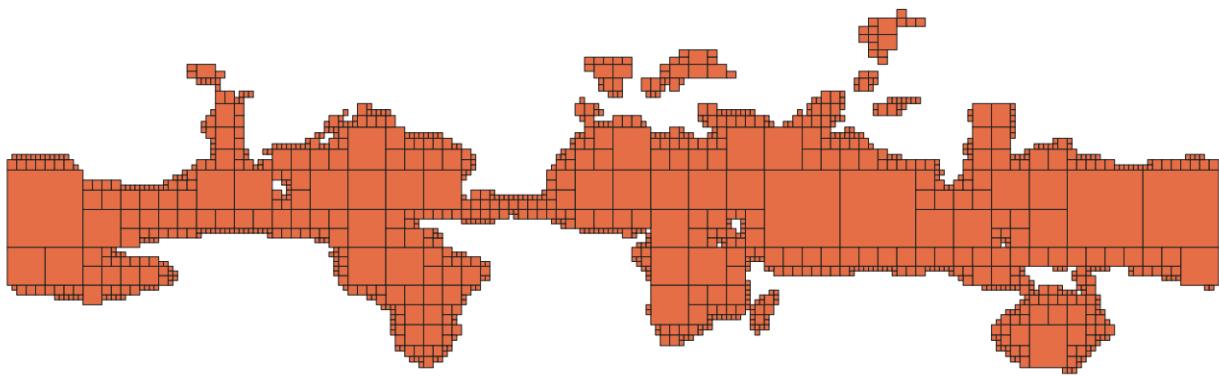


Figure 23 – GeoJSON output

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/dggs/GNOSISGlobalGrid/zones?filter=tasmax%3E300&datetime=2022-09-04&zone-level=6&f=json>

(Plain JSON Zone ID list output)

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/dggs/GNOSISGlobalGrid/zones?filter=tasmax%3E300&datetime=2022-09-04&zone-level=6&f:uint64>

(Binary 64-bit integer Zone IDs)

<https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/dggs/GNOSISGlobalGrid/zones?filter=tasmax%3E300&datetime=2022-09-04&zone-level=6&f=geotiff>

(GeoTIFF output) (using the default compact-zones=true where children zones are replaced by parent zone if all children zones are included)

By creating a kind of mask at a specifically requested resolution level, DGGS Zones Query can potentially greatly help parallelization and orchestration of spatial queries combining multiple datasets across multiple services, allowing to perform early optimizations with lazy evaluation.

Coverages

[https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage?f=png&properties=\(tasmax-250\)*400](https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage?f=png&properties=(tasmax-250)*400)

[https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage?f=geotiff&properties=tas,tasmax,tasmin,pr,psl&subset=Lat\(-90:90\),Lon\(0:180\)&height=400&datetime=2020-05-20](https://maps.gnosis.earth/ogcapi/collections/climate:cmip5:singlePressure/coverage?f=geotiff&properties=tas,tasmax,tasmin,pr,psl&subset=Lat(-90:90),Lon(0:180)&height=400&datetime=2020-05-20)

(GeoTIFF coverage with 5 bands for each field)

As a test of higher resolution data, we also loaded an hourly dataset for the ERA5 relative humidity for the April 1-6, 2023 period at: <https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity>

The spatial resolution for this one is also higher at 0.25 degrees longitude x 0.25 degrees latitude, and the data is for 37 different pressure levels. Some sample requests:

Maps

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/map?width=2048&subset=pressure\(750\)&bgcolor=0x002040](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/map?width=2048&subset=pressure(750)&bgcolor=0x002040)

Tiles

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/map/tiles/WorldCRS84Quad/0/0/0?subset=pressure\(750\)&bgcolor=0x002040](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/map/tiles/WorldCRS84Quad/0/0/0?subset=pressure(750)&bgcolor=0x002040)

[\(GeoTIFF coverage tile\)](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/coverage/tiles/WorldCRS84Quad/0/0/0?f=geotiff&subset=pressure(750))

Coverages

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/coverage?f=png&subset=pressure\(750\),Lat\(-90:90\),Lon\(0:180\),time\(%222023-04-03%22\)&properties=r*200&filter=r%3E20](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/coverage?f=png&subset=pressure(750),Lat(-90:90),Lon(0:180),time(%222023-04-03%22)&properties=r*200&filter=r%3E20)

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/coverage?f=geotiff&subset=pressure\(750\),Lat\(-90:90\),Lon\(0:180\),time\(%222023-04-03%22\)](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/coverage?f=geotiff&subset=pressure(750),Lat(-90:90),Lon(0:180),time(%222023-04-03%22))

(GeoTIFF Coverage)

DGGS

Data retrieval – What is here? (equivalent to Coverage Tiles requests for DGGSs whose zone geometry can be described by a 2D Tile Matrix Set e.g., GNOSISGlobalGrid, ISEA9R, rHealPix):

<https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones/0-0-3/data?f=geotiff&datetime=2023-04-03>

<https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/ISEA9Diamonds/zones/A7-0/data?f=geotiff&datetime=2023-04-03>

Zones query – Where is it?: Where is relative humidity at 850 hPa greater than 80% on April 3rd, 2023? (at precision level of GNOSIS Global Grid level 6) [https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure\(850\)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=json](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure(850)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=json)

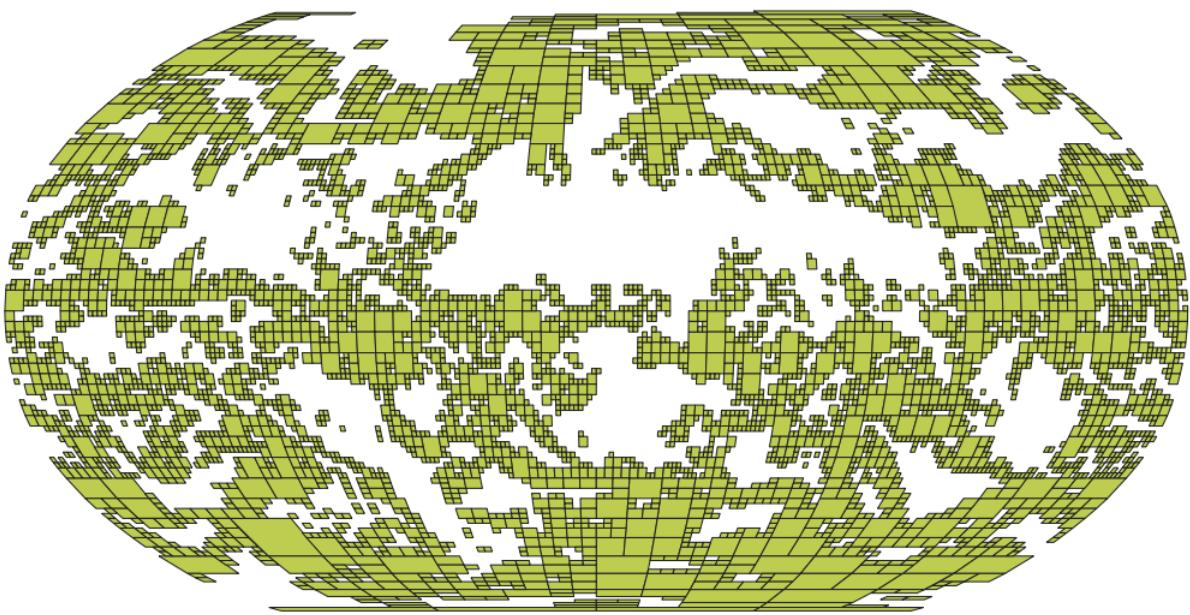


Figure 24 – GeoJSON output

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure\(850\)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=json](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure(850)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=json)

(Plain Zone ID list output)

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure\(850\)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=uint64](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure(850)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=uint64)

(Binary 64-bit integer Zone IDs)

[https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure\(850\)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=geotiff](https://maps.gnosis.earth/ogcapi/collections/climate:era5:relativeHumidity/dggs/GNOSISGlobalGrid/zones?subset=pressure(850)&datetime=2023-04-03&filter=r%3E80&zone-level=6&f=geotiff)

(GeoTIFF output) (using the default compact-zones=true where children zones are replaced by parent zone if all children zones are included)

We hope that our API and these climate datasets proves useful to other participants and can be part of Technology Integration Experiments for the pilots and/or Testbed 19 GeoDataCube.

We have also been working on our client to visualize these data sources from local netCDF files, our native GNOSIS data store, or remotely through OGC APIs, and we are working on support for EDR in order to perform integration experiments with the NOAA EDR API.

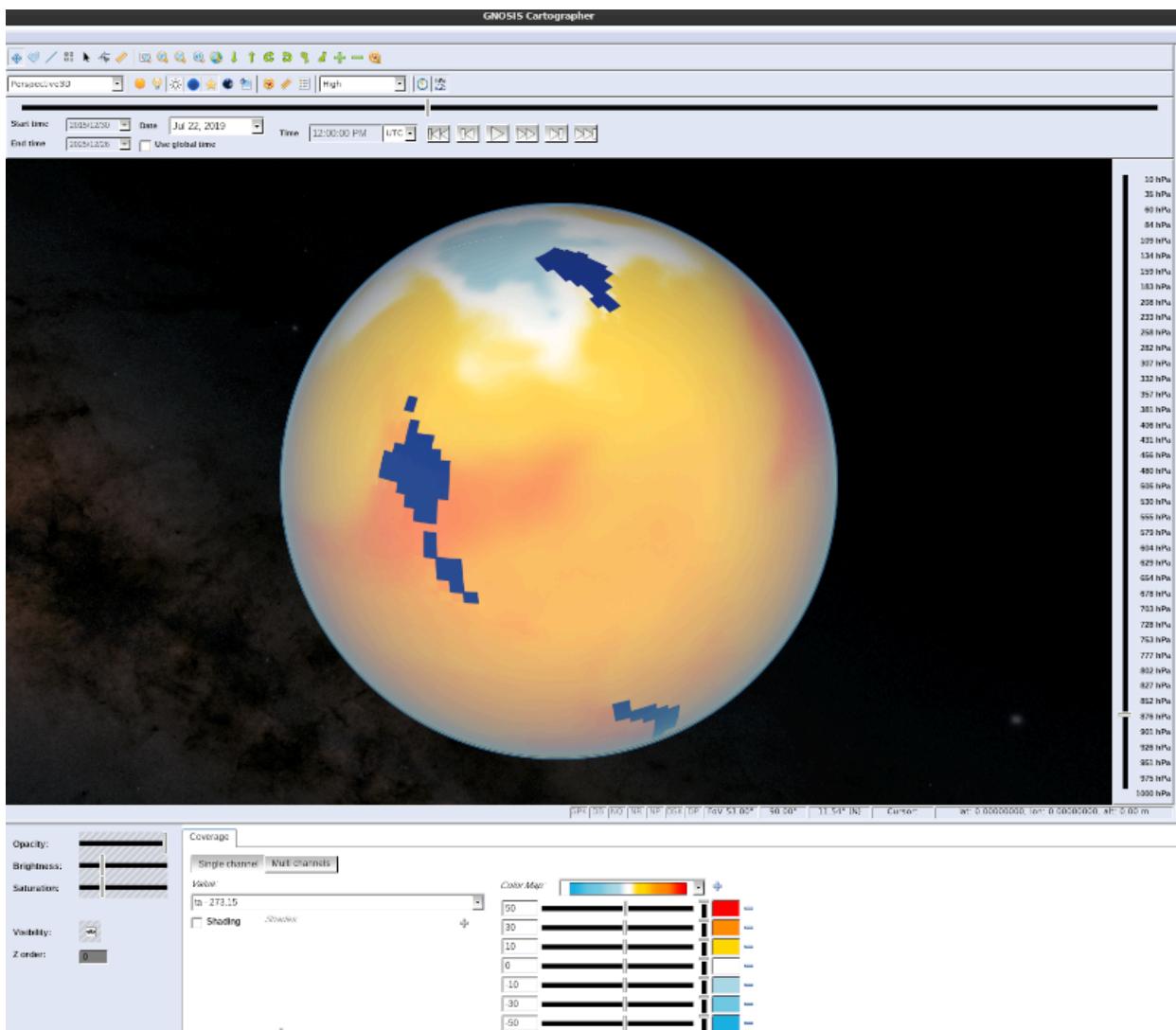


Figure 25 – GeoJSON output

We are also planning work on demonstrating the integration of these datasets as cross-collection queries and with our OGC API – Processes implementation including support for Part 3 – Workflows and Chaining.

One process we are putting together is a machine learning prediction process for classifying fuel vegetation types, based on sentinel-2 Level 2A accessed through our API at:

<https://maps.gnosis.earth/ogcapi/collections/sentinel2-l2a>

The initial training data will be using this Fuel vegetation Type coverage for the whole continental US from landfire.gov available from our API at:

<https://maps.gnosis.earth/ogcapi/collections/wildfire:USFuelVegetationTypes>

More work is being done on loading additional fire danger indices from the Copernicus Climate Data Store.

5

RAW DATA TO ANALYSIS READY DATA (ARD)

RAW DATA TO ANALYSIS READY DATA (ARD)

CEOS defines Analysis Ready Data as satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets. See <https://ceos.org/ard/>, and especially the information for data producers: https://ceos.org/ard/files/CARD4L_Info_Note_Producers_v1.0.pdf.

5.1. Transforming climate relevant raw data to ARD

Several past successful OGC testbeds, including the DP 21 to which this pilot is linked, have looked at ARD and IRD but also in terms of use cases. In this pilot, some main technical contributions have been creating digestible OGC data types and formats for specific partner use cases, so producing ARD from publically available EO and model data, including hydrological and other type of model output as well as climate projections.

These ARD will feed into all use cases for all participants, with a particular focus toward the use cases proposed for Heat, Drought and Health Impacts by participants in the pilot.

Specifically, participants provide access to the following satellite and climate projection data:

- Wildfire: Fire Radiant Power (FRP) product from Sentinel 3 (NetCDF), 5p, MODIS products (fire detection), VIIRS (NOAA); possibly biomass availability (fire fuel).
- Land Surface Temp – Sentinel 3
- Pollution – Sentinel 5p
- Climate Projection data (NetCDF, etc., daily downscaled possible): air temp (10 m above ground). Rainfall and possibly wind direction as well
- Satellite-derived Discharge Data to look at Droughts/Floods etc. by basin or other scale
- Hydrological model simulation outputs at (sub)basin scale (within reason)

The created ARD in various OGC interoperable formats created digestible dataflows for the proposed OGC Use Cases. This proposed data chain by several participants is similar to DP21, in which contributors, like RSS-Hydro, SafeSoftware, and others also participated. A generated climate indicator or EO remotely sensed data (NASA, NOAA, ESA, etc.) from various sources are “simplified” to GeoTIFF and / or vectorized geopackage per time step by other participants’ tools, such as the FME software (by SafeSoftware). Another option as an intermediate data type (IRD) can be COG – cloud optimized geotiff which would make access more efficient. The COG GeoTIFFs are optimized for cloud so data sharing can be made more efficient. ARD and IRD should become more service / cloud based wherever possible.

Besides the data format, data structures and semantics required to support the desired DRI's are important. The time series / raster, and classification to vector contour transform is an approach that worked well in DP21 and has been a good starting point also in this pilot. For example, together in the FME processing engine, time series grids can be aggregated across timesteps to mean or max values, then classify them into ranges suitable for decision making, and then write them out and expose them as time tagged vector contour tables.

In summary, the different ARD and IRD data can be created from the following data sources:

- Inputs: EO (US sources fire related: MODIS, VIIRS); Climate projections, sub catchment polygons, ESA sources; Sentinel-3, Sentinel 5-P.
- Outputs forma & instances: WCS, GeoTIFF spatial / temporal subset, Shape; NetCDF.
- Output parameters: e.g. hydrological condition of a basin (historically/current). So drought / flood etc.
- Output themes: downscaled / subset outputs, hydrologic scenarios.

Another highly relevant input are the Essential Climate Variables (ECV) Inventory (<https://climatedatainfo.info/ecvinventory/>) houses information on Climate Data Records (CDR) provided mostly by CEOS and CGMS member agencies. The inventory is a structured repository for the characteristics of two types of GCOS ECV CDRs:

Climate data records that exist and are accessible, including frequently updated interim CDRs; Climate data records that are planned to be delivered.

Figure 26

The ECV Inventory is an open resource to explore existing and planned data records from space agency sponsored activities and provides a unique source of information on CDRs available internationally. Access links to the data are provided within the inventory, alongside details of the data's provenance, integrity and application to climate monitoring.

Participants, particularly GMU CSISS have demonstrated the use of ECV record information as input with OpenSearch service endpoint (currently CMR(CWIC) and FedEO), and downloading URLs for accessing NetCDF or HDF files.

Outputs in this case include WCS service endpoint for accessing selected granule level product images (GeoTIFF, PNG, JPEG, etc.), focusing on the WCS for downloading images and WMS for showing layers on a basemap.

5.2. Use of API services for climate related use cases

Pixalytics have developed an OGC-compliant Application Programming Interface (API) service, see Figure 27, which will provide global information on droughts. The approach is to take global open data/datasets from organizations such as ESA/Copernicus, NASA/NOAA, and the WMO and combine meteorology, hydrology, and remote sensing data to produce ARD data based on

a composite of different indicators. Where globally calculated drought indicators already exist, these are being used in preference to their re-calculation, although consistency and the presence of uncertainties are also being considered.

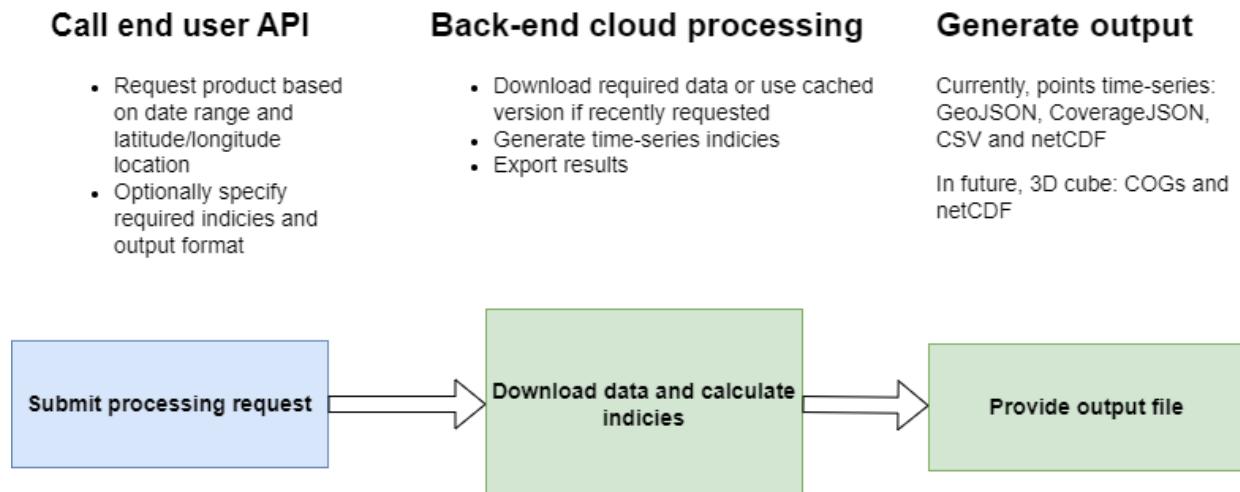


Figure 27 – Pixalytics drought severity workflow architecture

The Drought Severity Workflow (DSW) is built using individual drought indicators for precipitation (SPI), soil moisture (SMA), and vegetation drought that are together using the Combined Drought Indicator (CDI) as described by [Sepulcre-Canto 2012]. The API access has been set up following the Building Blocks for Climate Services (<https://climateintelligence.github.io/smarduck-docs/>) approach.

- Component: D100 Drought indicator.
- Inputs: Meteorological data, including Precipitation, plus Land Surface Temperature, Soil Moisture, and Vegetation Index (or optical data to calculate it from).
- Outputs: Drought Indices — default is CDI — as a time-series dataset output in a choice of download able formats: CSV, GeoJSON (default), CoverageJSON and NetCDF for point data and then COG for areas (to be developed).
- What other component(s) can interact with the component: a desire to link to visualization/DRI analysis components. A QGIS plugin called WPS Client has been updated to be able to perform a request and view the outputted JSON file, with a second QGIS plugin called Data Ploty used to visualize the time-series data. The Web Processing Service (WPS) link to access the drought indicator is: <https://api.pixalytics.com/climate/wps?request=GetCapabilities&service=wps>

An example Python query for a location in Canada (Latitude: 55.5 N Longitude: 99.1 W) for the SPI time series, with data for these dates/this location already cached, so runs quicker:

```

from owslib.wps import WebProcessingService, monitorExecution

# contact the WPS client
wps = WebProcessingService("http://api.pixalytics.com/climate/wps", skip_
caps=True)
  
```

```

# GetCapabilities
wps.getcapabilities()

# Execute
inputs = [ ("start_date", '20200101'),
           ("end_date", '20221231'),
           ("latitude", '55.5'),
           ("longitude", '-99.1')]

execution = wps.execute("drought", inputs, "output")

outfile = "temp.json"
monitorExecution(execution, download=True, filepath=outfile)

# Wait 5 seconds and check
execution.checkStatus(sleepSecs=5)

# show status
print('Percent complete {}%, generated {}'.format(execution.
percentCompleted, outfile))

# If there's an error print the error information
for error in execution.errors:
    print("Error: ",error.code, error.locator, error.text)

```

Figure 28 – Drought indicator calling code that generated the default output, which is the CDI in GeoJSON format

- What OGC standards or formats does the component use and produce: Producing data on-the-fly using the WPS, so need to pull data through preferably an API route. The speed that the input data can be made available (i.e., extracting time-series subsets) governs the speed that the drought indicator provides data. To speed this up, input data that is not changing is being cached so that it runs significantly quicker when the API is called for a second time.

Figure 29 shows an example of the output visualized within Python using Streamlit with the intermediate data (cached as NetCDF files) as input.

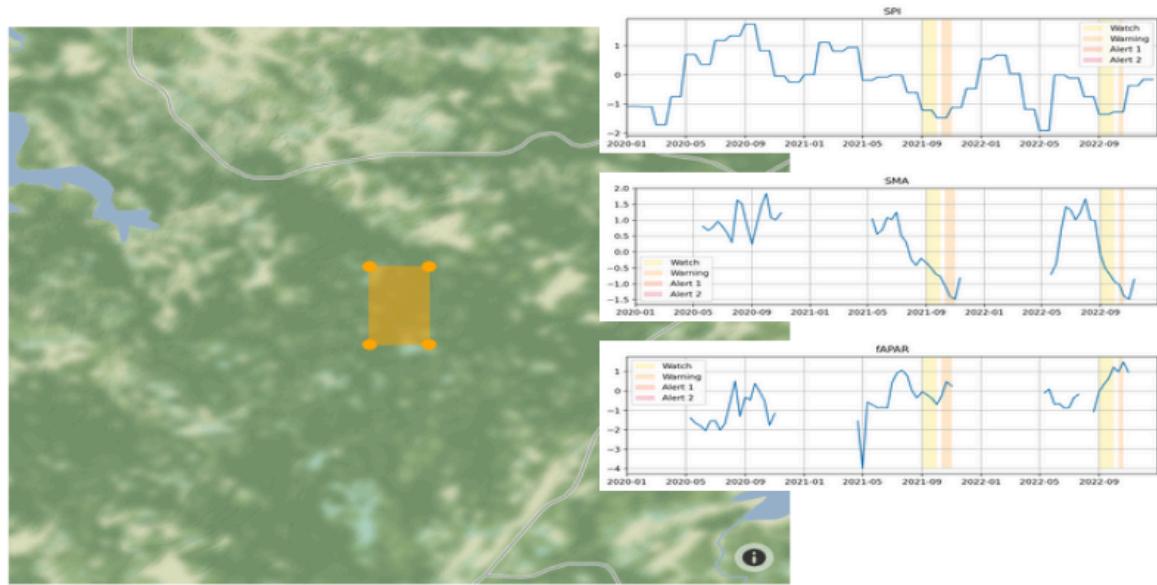


Figure 29 – Plot of the CDI for a point location in Canada (Latitude: 55.5 N Longitude: 99.1 W); generated using Copernicus Emergency Management Service information [2023]

5.2.1. Data Sources

The Global Drought Observatory

The Global Drought Observatory (GDO), owned by the Copernicus Emergency Management Services, provides a global map of coarsely-gridded agricultural drought risk, along with a breakdown of the risk for each country. The drought risk is computed using the CDI, with the variables used to compute it and other drought-related variables provided in the user portal for [download](#), but the CDI itself is not available for download and so is being calculated in the DSW.

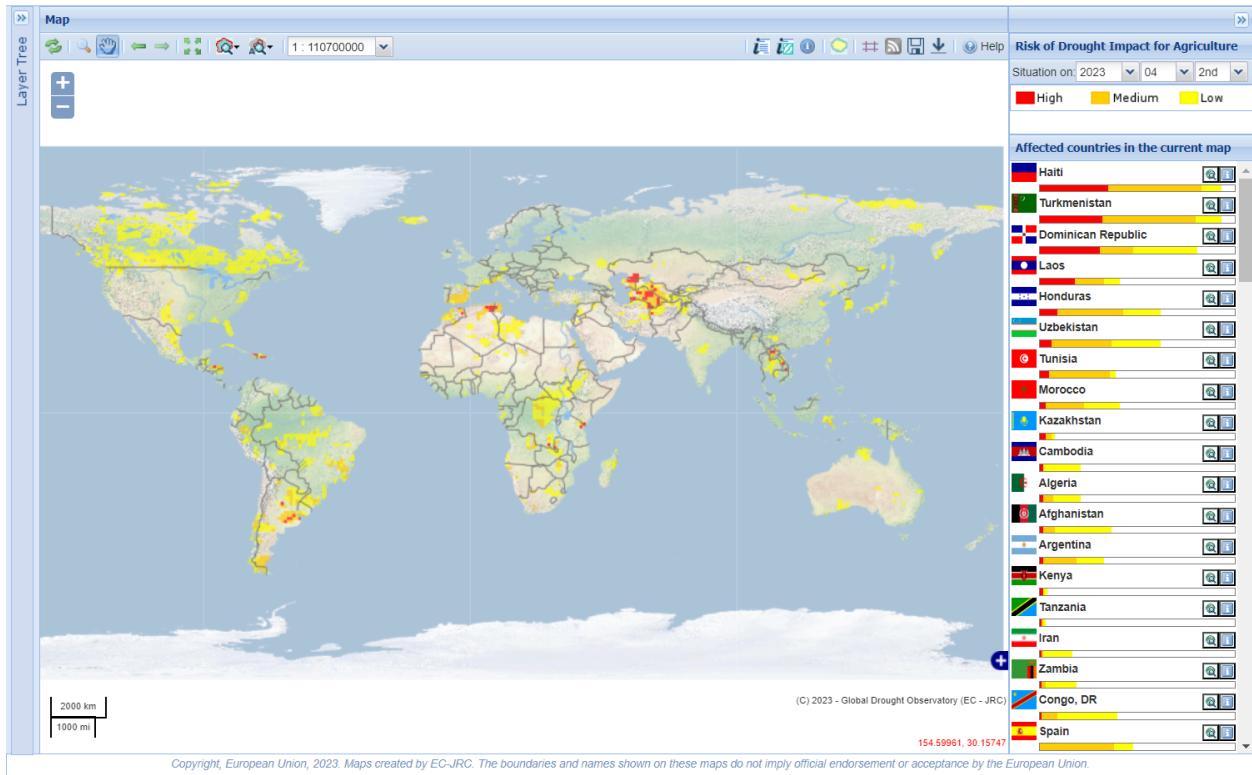


Figure 30 – Global Drought Observatory Web Portal,
<https://edo.jrc.ec.europa.eu/gdo/php/index.php?id=2001>

We obtain SMA and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) from the GDO data download service. These are provided as netCDF files and contain pre-computed anomalies, so can be assimilated directly into the back-end. The SMA uses a combination of the root soil moisture from the [LISFLOOD model](#), the MODIS land surface temperature and the ESA Climate Change Initiative (CCI) skin soil moisture [Cammalleri 2016], and the FAPAR is from NASA optical imagery.

ERA5 Reanalysis from ECMWF CDS

The CDS portal provides an API interface to return either hourly or monthly averages of the ERA5 variables. Requesting the hourly data is necessary to compute anything which requires a frequency greater than monthly, which is the case for most drought indicators (e.g. SMA) which are in dekads. To ensure there is no anti-aliasing, the full 24hr dataset for each day of the month must be downloaded. This is very time-consuming and requests will fail if the number of data points exceeds the limit, which will occur for a period of 2 years or more, even for a single location.

There is a separate application, which can also be accessed via API, to return daily data. The CDS employs a queue management system, which determines the priority of each request based partially on the computational demand of the request. The daily data retrieval relies upon an underlying service to compute the daily statistics from the hourly data, demanding more resources than simply extracting the hourly or monthly data which are pre-computed. This means the request is held in the queue for a long time (up to hours), so there is no time benefit over using the hourly data. However, for a longer time-period which would be rejected if

requested hourly, this provides a workaround. A further benefit of requesting daily, rather than hourly, data is that the downloaded file is smaller.

We compute SPI and SMA using variables from the CDS API. The SPI is computed from the total precipitation in monthly intervals. The SMA is computed from the soil water volume, which is available for 4 depth levels. The SMA for each depth is computed by calculating the z-score against a long term mean, using the same baseline time period as the SPI. The most relevant depth layer can then be selected by the user; for instance, a user interested in the health of crops with shallow roots may wish to access the surfacemost layer.

ERA5 Reanalysis from AWS

Input precipitation data was also tested using ERA5 data held within the Registry of Open Data on AWS versus the CDS API and found the Amazon Web Service (AWS) Simple Storage Service (S3) stored data could be accessed faster once virtual Zarrs has been setup, but there is a question over provenance as the data on AWS was put there by an organization other than the data originator and the Zarr approach didn't work for more recent years as the S3 stored NetCDFs have chunking that is inconsistent. An issue was raised for the Python kerchunk library, to be able to cope with variable chunking, as this isn't current supported. The issue has also been raised with the organisation storing the data on S3, and they are investigating.

NOAA API

The NOAA Climate Environmental Data Retrieval (EDR)

Safe Software extraction of climate forecast data

We reviewed the GeoJSON Feature point data extracted from the Climate scenario RPC4.5 downscaled for Canada provided by Safe Software. As a preliminary test of combining the reanalysis and forecast data the SPI was calculated using reanalysis data up until the end of 2022 and then forecast data for 2023 and 2024; see comparisons in Figure 31.



Figure 31 – Plot of the ECMWF precipitation and SPI, extended using the Canadian climate forecast data, compared to the GDO calculated SPI; generated using Copernicus Emergency Management Service information, Copernicus Climate Service and Canadian Centre for Climate Services data [2023]

5.2.2. Further work

The work in this Pilot has focused on building this initial version of the workflow, deploying it via WPS and pulling data from different sources to understand the advantages and disadvantages of the different sources, including straightforwardness and speed of accessibility. For future Pilot activities we plan to continue to build the robustness of the approach, including testing and improving on the robustness of the interfaces to the input data sources and output provided to other Pilot components.

Other additions to the workflow which could support future collaboration with other OGC contributors include:

- The current work has focused on the extraction and generation of a point time-series, and so there are plans to expand the code to the extraction and generation of a 3D data cube. This might involve changing the output API interface to the OGC EDR API standard.
- The sources used to calculate the combined indicator currently includes only historical data. If Soil Moisture or FAPAR projections can be obtained, these could be combined with future predictions of precipitation from Safe Software and NOAA to project the CDI into the future.
- We currently focus on the CDI, a combination of precipitation, soil moisture and vegetation health data, which is most useful in agricultural contexts. Other combined indicators could be developed which are more applicable to other challenges; for instance, combining rainfall and temperature indicators may produce a drought warning more applicable to public health.
- The ECMWF Soil Moisture data contains information for multiple depths of soil. The service currently returns the moisture of just one of these layers, however, the most applicable layer will vary with location and crop type. For some use-cases, providing the option to choose the soil layer and providing guidance on how this can be done would be beneficial to the end user.

5.3. From Raw Data to ARD with the FME Platform

5.3.1. Component Descriptions

D100 – Client instance: Analysis Ready Data Component

Our Analysis Ready Data component (ARD) uses the FME platform to consume regional climate model and EO data and generate FAIR datasets for downstream analysis and decision support.

The challenge to manage and mitigate the effects of climate change poses difficulties for spatial and temporal data integration. One of the biggest gaps to date has been the challenge of

translating the outputs of global climate models into specific impacts at the local level. FME is ideally suited to help explore options for bridging this gap given its ability to read datasets produced by climate models such as NetCDF or OGC WCS and then filter, aggregate, interpolate and restructure it as needed. FME can inter-relate it with higher resolution local data, and then output it to whatever format or service is most appropriate for a given application domain or user community.

Our ARD component supports the consumption of climate model outputs such as NetCDF. It also has the capacity to consume earth observation (EO) data, and the base map datasets necessary for downstream workflows, though given time and resource constraints during this phase we did not pursue consumption of other data types besides climate data.

5.3.1.1. ARD Workflow

The basic workflow for generating output from the FME ARD component is as follows. The component extracts, filters, interrelates and refines these datasets according to indicator requirements. After extraction, datasets are filtered by location and transformed to an appropriate resolution and CRS. Then the workflow resamples, simplifies and reprojects the data, and then defines record level feature identifiers, ECV values, metadata and other properties to satisfy the target ARD requirements. This workflow is somewhat similar to what was needed to evaluate disaster impacts in DP21. Time ranges for climate scenarios are significantly longer – years rather than weeks for floods.

Once the climate model, and other supporting datasets have been adequately extracted, prepared and integrated, the final step is to generate the data streams and datasets required by downstream components and clients. The FME platform is well suited to deliver data in formats as needed. This includes Geopackage format for offline use. For online access, other open standards data streams are available, such as GeoJSON, KML or GML, via WFS and OGC Features APIs and other open APIs. For this pilot we generated OGC Geopackage, GeoJSON, CSV and OGC Features API services.

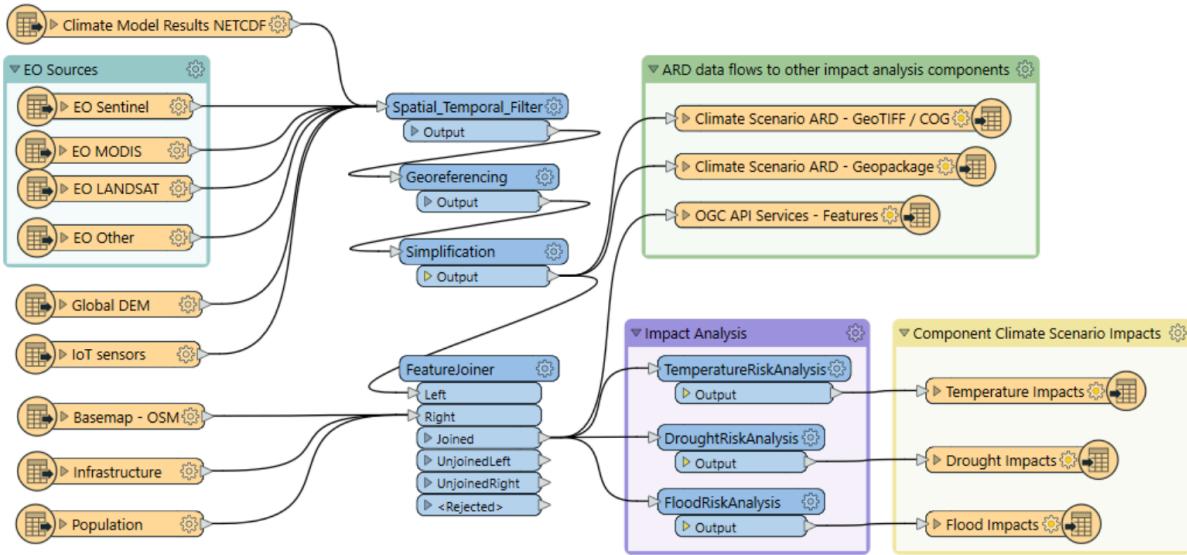


Figure 32 – High level FME ARD workflow showing generation of climate scenario ARD and impacts from climate model, EO, IoT, infrastructure and base map inputs

As our understanding of end user requirements continues to evolve, this will necessitate changes in which data sources are selected and how they are refined, using a model based rapid prototyping approach. We anticipate that any operational system will need to support a growing range of climate change impacts and related domains. Tools and processes must be able to absorb and integrate new datasets into existing workflows with relative ease. As the pilot develops, data volumes increase, requiring scalability methods to maintain performance and avoid overloading downstream components. Cloud based processing near cloud data sources using OGC API web services supports data scaling. Regarding the FME platform, this involves deployment of FME workflows to FME Cloud. Note that in future phases, we are likely to test how cloud native datasets (COG, STAC, ZARR) and caching can be used to scale performance as data transactions and volume requirements increase.

It is worth underlining that our ARD component depends on the appropriate data sources in order to produce the appropriate decision ready data (DRI) for downstream components. Risk factors include being able to locate and access suitable climate models of sufficient quality, resolution and timeliness to support indicators as the requirements and business rules associated with them evolve. Any data gaps encountered are documented under this section under Challenges and Opportunities and in the common Lessons Learned chapter and the end of the ER.

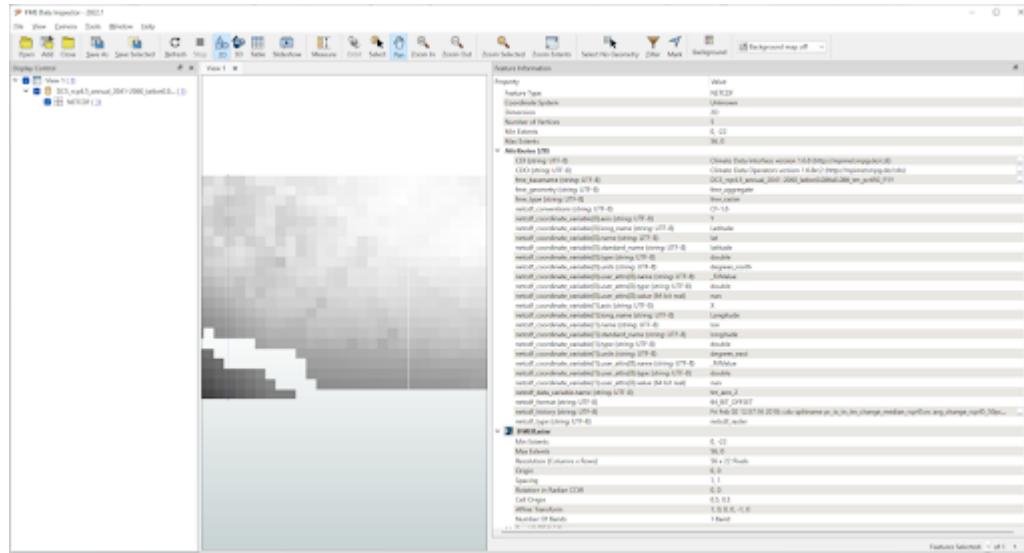


Figure 33 – Environment Canada NetCDF GCM time series downscaled to Vancouver area. From: <https://climate-change.canada.ca/climate-data/#/downscaled-data>

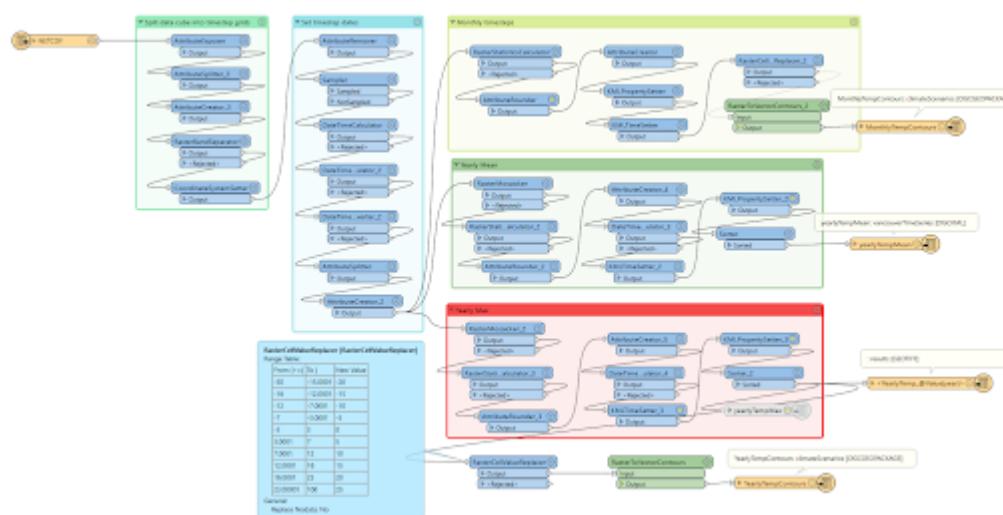


Figure 34 – Data Cube to ARD: NetCDF to KML, Geopackage, GeoTIFF

Original Data workflow: - Split data cube - Set timestep parameters - Compute timestep stats by band - Compute time range stats by cell - Classify by cell value range - Convert grids to vector contours

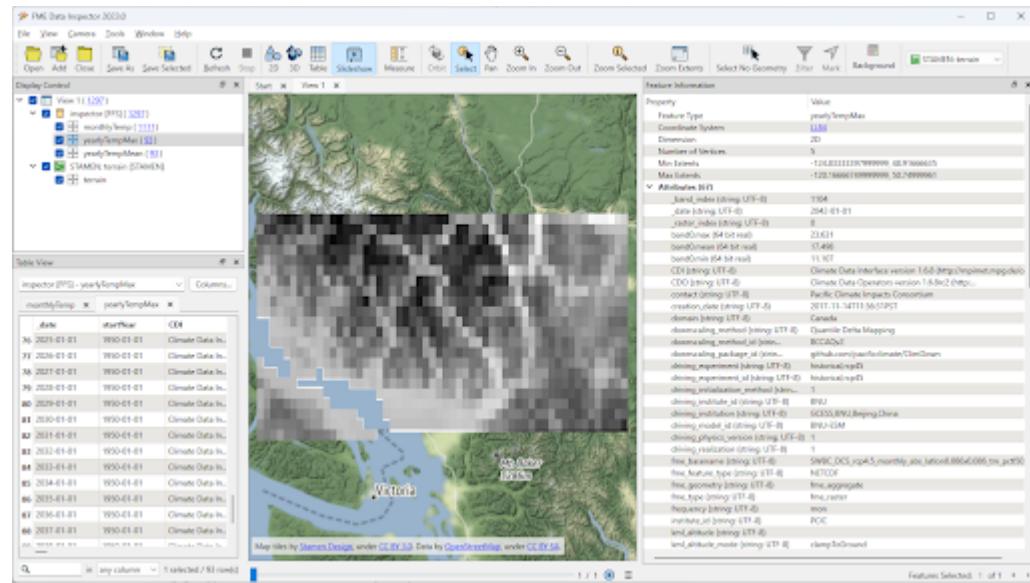


Figure 35 – Extracted timestep grids: Monthly timesteps, period mean T, period max T

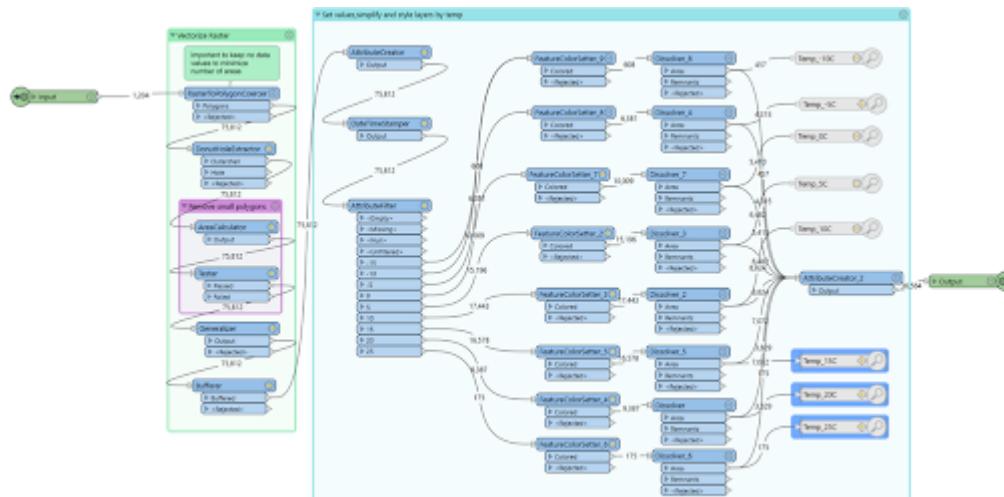


Figure 36 – Convert raster temperature grids into temperature contour areas by class

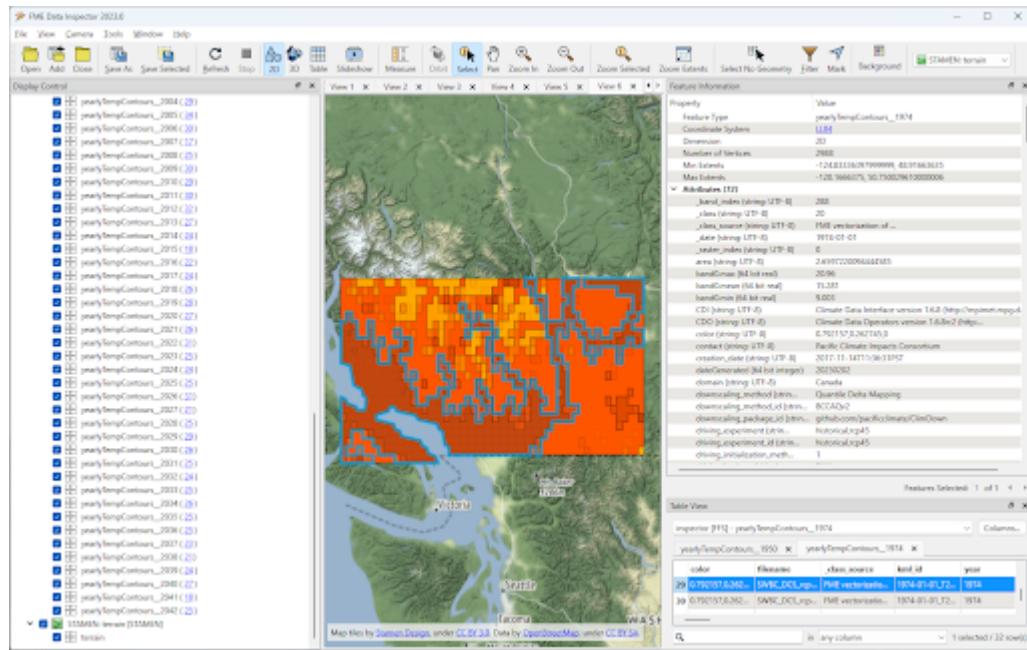


Figure 37 – Geopackage Vector Area Time Series: Max Yearly Temp

5.3.1.2. ARD Development Observations

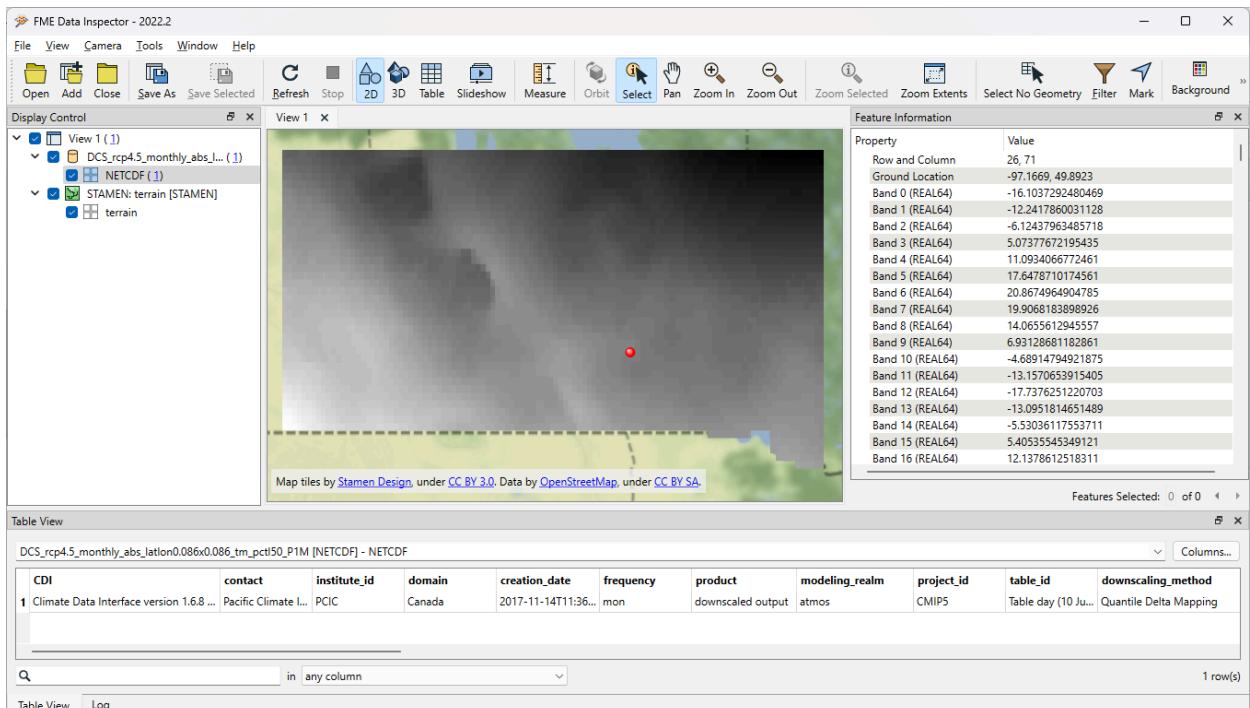


Figure 38 – FME Data Inspector: RCM NetCDF data cube for Manitoba temperature 2020-2099

Disaster Pilot 2021 laid a good foundation for exploring data cube extraction and conversion to ARD with using the FME data integration platform. A variety of approaches were explored

for extraction, simplification and transformation including approaches to select, split, aggregate, and summarize time series. However, more experimentation was needed to generate ARD that can be queried to answer questions about climate trends. This evolution of ARD was one of the goals for this CRP. This goal includes better support for both basic queries, and analytics, statistical methods, simplification & publication methods, including cloud native – NetCDF to Geopackage, GeoJSON and OGC, APIs.

In consultation with other participants, we learned fairly early on in the pilot that our approach to temperature and precipitation contours or polygons inherited from our work in DP21 on flood contours involved too much data simplification to be useful. For example, contouring required grid classification into segments, such as 5 degree C or 10mm of precipitation etc. However, this effective loss of detail oversimplified the data to the point where it no longer held enough variation over local areas to be useful. In discussion with other participants, it was determined that simply converting multidimensional data cubes to vector time series point data served the purpose of simplifying the data structure for ease of access, but retained the ECV precision needed to support a wider range of data interpretations for indicator derivation. It also meant that as a data provider we did not need to anticipate or encode interpretation of indicator business rules into our data simplification process. By simply providing ECV data points, the end user was free to run queries to find locations and time steps where temp > or precipitation < some threshold of interest.

Initially it was thought that classification rules need to more closely model impacts of interest. For example, the business rules for a heat wave might use a temperature range and stat type as part of the classification process before conversion to vector. However, this imposes the burden of domain knowledge on the data provider rather than on the climate service end user who is much more likely to understand the domain they wish to apply the data to and how best to interpret it.

Modified ARD Data workflow: - Split data cube - Set timestep parameters - Compute timestep stats by band - Compute time range stats by cell - Convert grids to vector contours

Further scenario tests were explored, including comparison with historical norms. Calculations were made using the difference between projected climate variables and historical climate variables. These climate variable deltas may well serve as a useful starting point for climate change risk indicator development. They also serve as an approach for normalizing climate impacts when the absolute units are not the main focus. Interesting patterns emerged for the LA area that we ran this process on deltas between projected and historical precipitation. While summers are typically dry and winters are wet and prone to flash floods. Initial data exploration seemed to show an increase in drought patterns in the spring and fall. More analysis needs to be done to see if this is a general pattern or simply one that emerged from the climate scenario we ran. However, this is the type of trend that local planners and managers may benefit from having the ability to explore once they have better access to climate model scenario outputs along with the ability to query and analyze them.

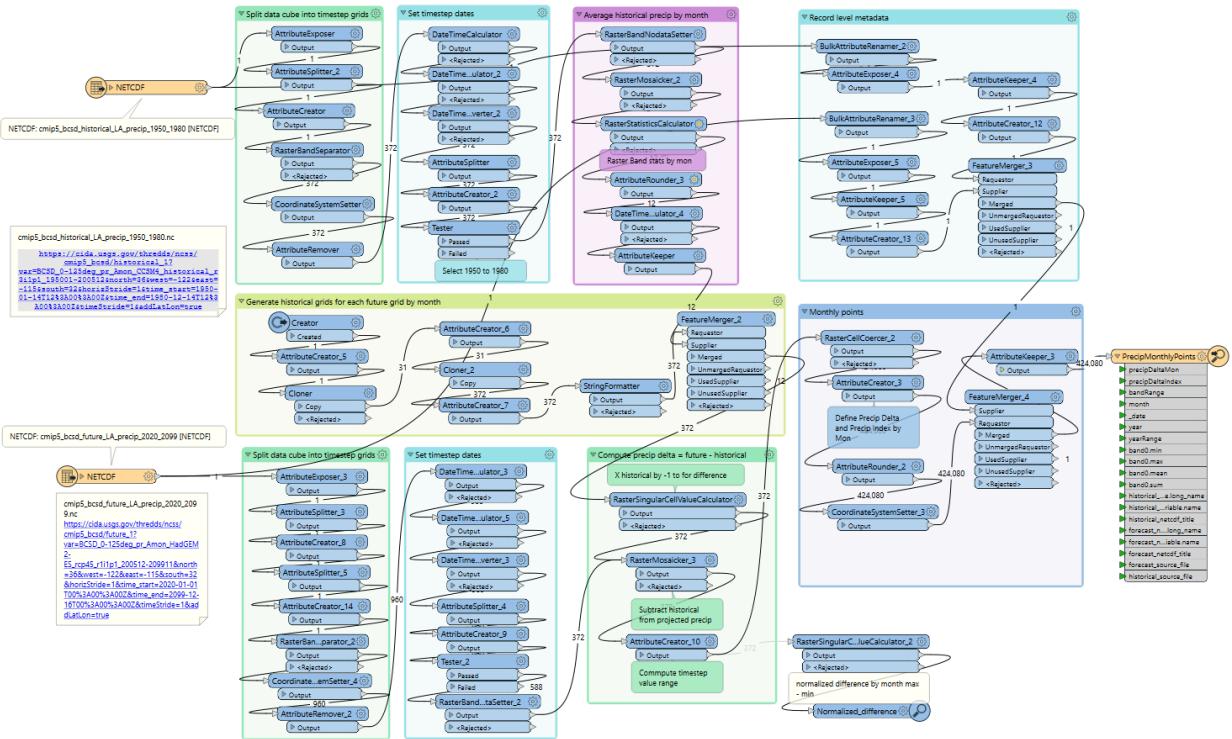


Figure 39 – Modified ARD Worflow: NetCDF data cube to precipitation delta grids (future - historical) in Geopackage for LA

ARD Climate Variable Delta Data workflow: - Split data cubes from historic and future netcdf inputs - Set timestep parameters - Compute historic mean for 1950 – 1980 per month based on historic time series input - Multiply historic mean by -1 - Use RasterMosaiker to sum all future grids with -1 * historic mean grid for that month - Normalize environmental variable difference by dividing by historic range for that month delta / (max – min) - Convert grids to vector contours - Define monthly environment variables from band and range values

More analysis needs to be done with higher resolution time steps – weekly and daily. At the outset monthly time steps were used to make it easier to prototype workflows. Daily time step computations will take significantly more processing time. Future pilots should explore ways of better supporting scalability of processing through automation and cloud computing approaches such as the use of cloud native formats (STAC, COG, ZARR etc).

5.3.1.3. OGC API Features Service

Compared to OGC WFS2, OGC APIs are a simpler and more modern standard based on a REST and JSON / openAPI approach. However we found implementation of OGC API services somewhat challenging. There seems to be more complexity in terms of number of ways for requesting features, and too many options for representing service descriptions. As every client tends to interpret and use the standard a bit differently – it becomes a challenge to derive how to configure service for a wide range of clients. In particular, QGIS / ArcPro were a challenge to debug given limited logging. For QGIS, we had to examine cache files in the operating system temp directories to look for and resolve errors.

Once correctly configured, OGC API feature services seemed to perform well and likely are more efficient than the equivalent WFS2 / GML feature services. A key aspect of performance improvement was achieving query parameter continuity by passing query settings from the client all the way to the database reader configuration. For example, it was important to make sure the spatial extent and feature limits from the end user client were implemented in the database SQL extraction query and not just at an intermediate stage. We will need to explore better use of caching to further optimize performance. There may also be opportunities for pyramiding time series vector data at a lower resolution for wide area requests. This may better serve those interested in observing large area patterns who don't necessarily need full resolution at the local level.

It should also be noted that while OGC API services should be a priority for standards support, for a climate and disaster management context, given the relative recent nature of these standards many users may be less than familiar with or prepared to use these standards. As such, there should also be provision to access data directly in well accepted open standards such as GeoJSON, CSV, Geotiff, Geopackage or Shape. In this project, some users preferred direct access to GeoJSON or CSV over OGC API access.

5.4. A framework example for climate ARD generation

Wuhan University (WHU) is a university that plays a significant role in researching and teaching all aspects of surveying and mapping, remote sensing, photogrammetry, and geospatial information sciences in China. In this Climate Resilience Pilot, WHU will contribute three components (ARD, Drought Indicator, and Data Cube) and one use-case (Drought Impact Use-cases).

5.4.1. Component: ARD

- Inputs: Gaofen L1A data and Sentinel-2 L1C data
- Outputs: Surface Reflectance ARD
- What other component(s) can interact with the component: Any components requiring access to surface reflectance data

Surface Reflectance (SR) is the fraction of incoming solar radiation reflected from the Earth's surface for specific incidents or viewing cases. It can be used to detect the distribution and change of ground objects by leveraging the derived spectral, geometric, and textural features. Since a large amount of optical EO data has been released to the public, ARD can facilitate interoperability through time and multi-source datasets. As the probably most widely applied ARD product type, the SR ARD can contribute to climate resilience research. For example, the SR-derived NDVI series can be applied to monitor wildfire recovery by analyzing vegetation index increases. Several SR datasets have been assessed as ARD by CEOS, like the prestigious Landsat Collection 2 Level 2, and Sentinel-2 L2A, while many other datasets are still provided at a low processing level.

WHU is developing a pre-processing framework for SR ARD generation. The framework supports radiometric calibration, geometric ratification, atmospheric correction, and cloud mask. To address the inconsistencies in observations from different platforms, including variations in band settings and viewing angles, we proposed a processing chain to produce harmonized ARD. This will enable us to generate SR ARD with consistent radiometric and geometric characteristics from multi-sensor data, resulting in improved temporal coverage. In the first stage of our mission, we are focusing on the harmonization of Chinese Gaofen data and Sentinel-2 data, as shown in Figure 1, the harmonization involves spatial co-registration, band conversion, and bidirectional reflectance distribution function (BRDF) correction. Figure 2 shows the Sentinel-2 data before and after pre-processing. Furthermore, we wish to seek the assessment of CEOS-ARD in our long-term plan.

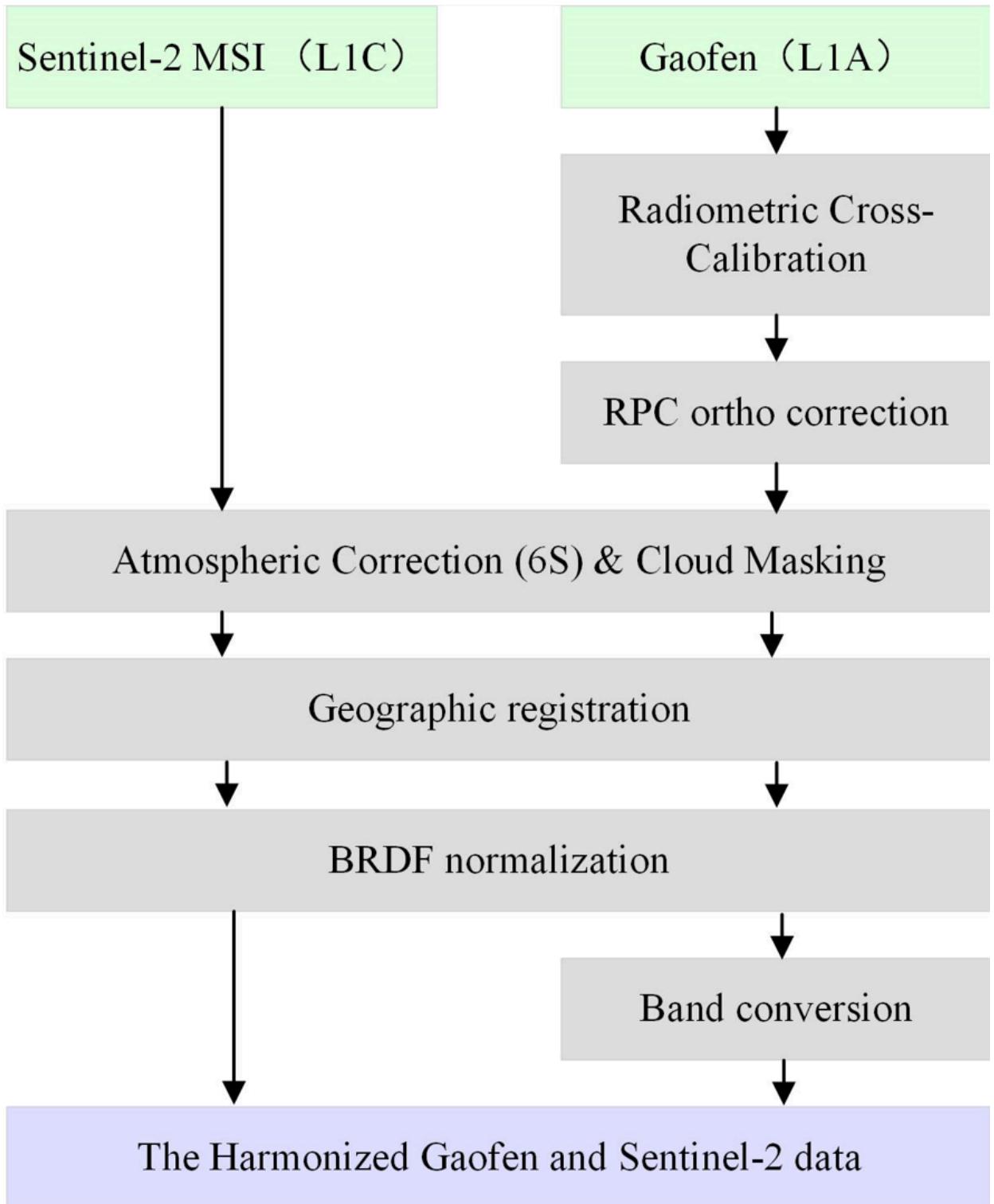


Figure 40 – The processing chain to produce harmonized ARD.

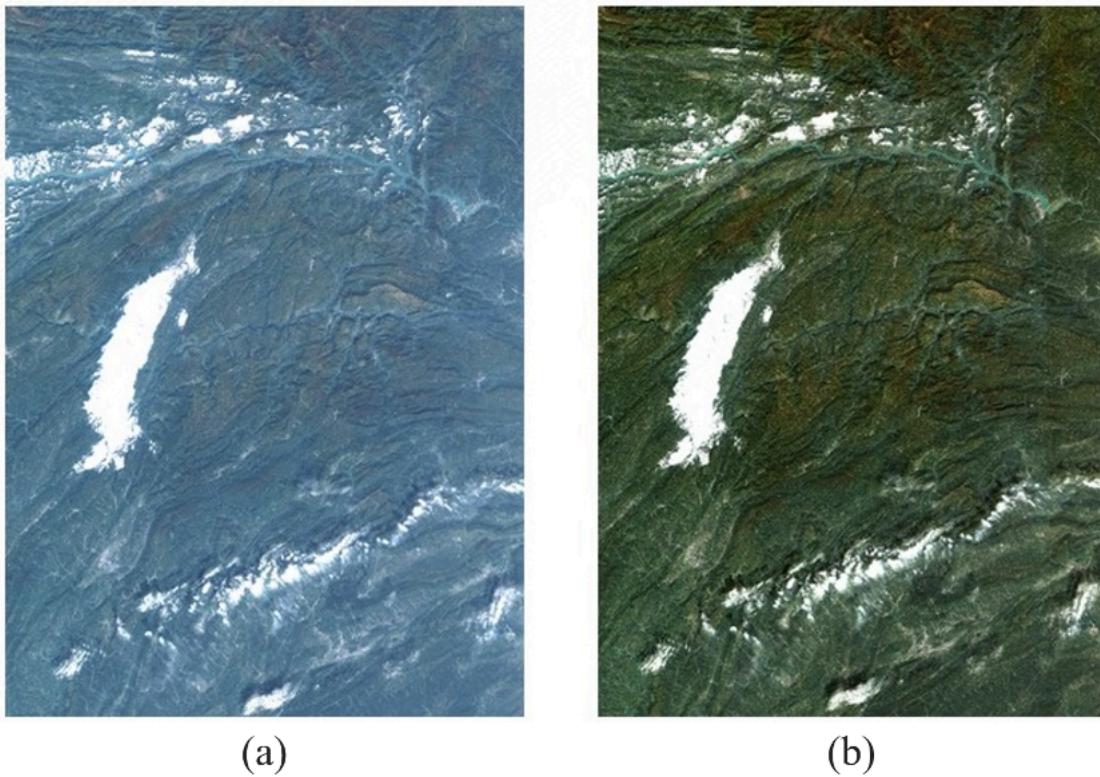


Figure 41 – Sentinel-2 RGB composite (red Band4, green Band3, blue Band2), over Hubei, acquired on October 22, 2020. (a) corresponds to the reflectance at the top of the atmosphere (L1C product); (b) corresponds to the surface reflectance after pre-processing.

5.4.2. Component: Drought Indicator

- Inputs: Climate data, including precipitation and temperature
- Outputs: Drought risk map derived from drought indicator
- What other component(s) can interact with the component: Any components requiring access to drought risk map through OGC API
- What OGC standards or formats does the component use and produce: OGC API – Processes

Drought is a disaster whose onset, end, and extent are difficult to detect. Original meteorological data, such as precipitation, can be obtained through satellites and radar, which can be used for drought monitoring. However, the accuracy is easily affected by detection instruments and terrain occlusion, and the ability to retrieve special shapes, such as solid precipitation, is limited. In addition, many meteorological monitoring stations on the ground can provide local raw meteorological observation data. The SPEI is a model to monitor, quantitatively analyze, and determine the spatiotemporal range of the occurrence of drought

using meteorological observation data from various regions. It should supplement the result of drought monitoring with satellite and radar.

SPEI has two main characteristics: 1) it considers the deficits between precipitation and evapotranspiration comprehensively, that is, the balance of water; 2) multi-time scale characteristics. For 1) drought is caused by insufficient water resources. Precipitation can increase water, while evapotranspiration can reduce water. The differences between the two variables simultaneously and in space can characterize the balance of water. For 2), the deficits value of different usable water sources is distinct at different time scales due to the different evolution cycles of different types, resulting in various representations in temporal. By accumulating the difference between precipitation and evapotranspiration at different time scales, agricultural (soil moisture) droughts, hydrological (groundwater, streamflow, and reservoir) droughts, and other droughts can be distinguished by SPEI.

In our project, the dataset for SPEI calculation is ERA5-Land monthly averaged data from 1950 to the present. We selected years of data about partial areas of East Asia for experiments.

Through the following flow of the SPEI calculation, we obtain the SPEI value for assessments of drought impact. The flow of the SPEI calculation is shown in Figure 3.

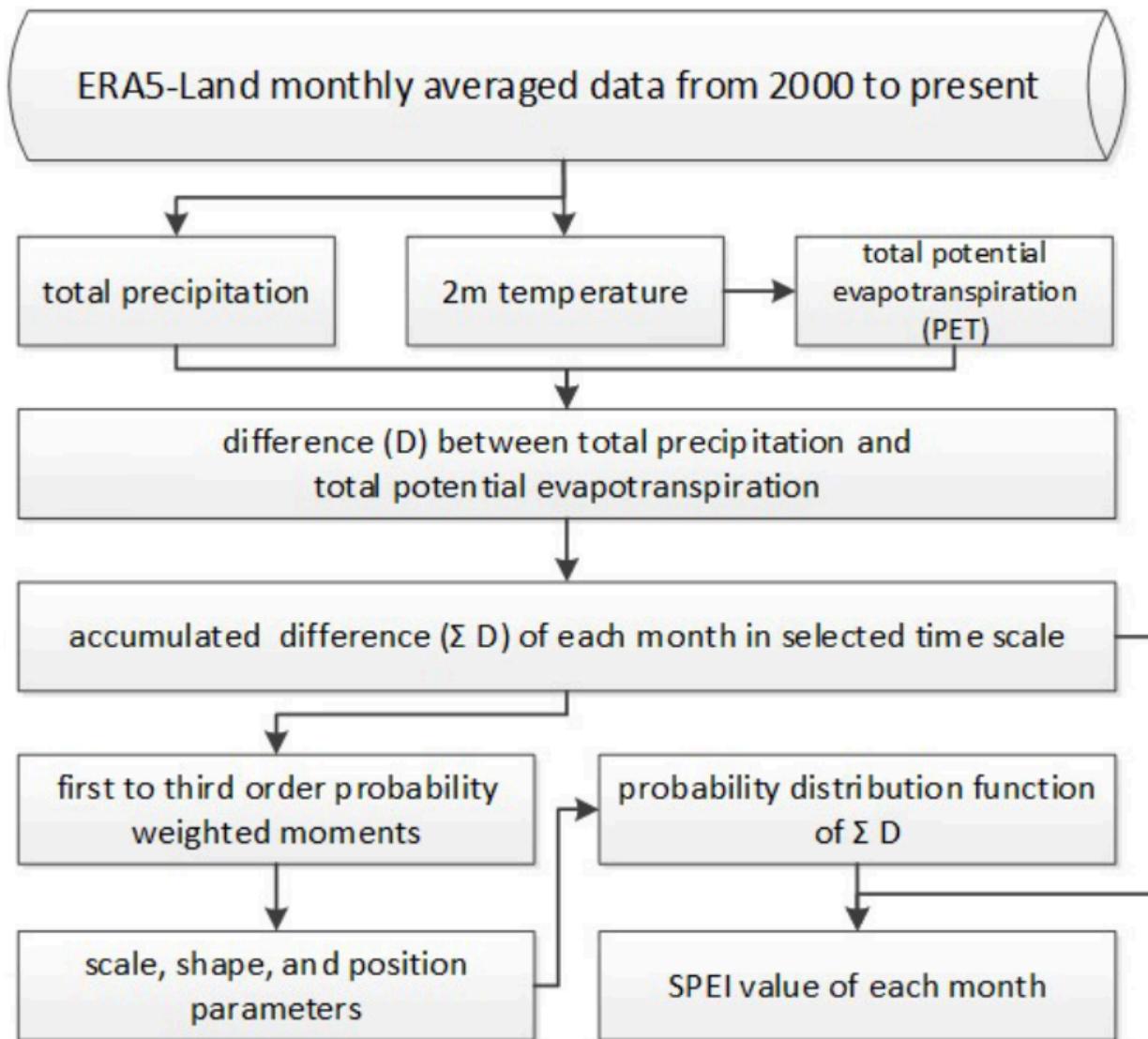


Figure 42 – Flow of the SPEI calculation.

WHU has provided the SPEI drought index calculation services through the OGC API – Processes, enabling interaction with other components. The current endpoint for OGC API – Processes is http://oge.whu.edu.cn/ogcapi/processes_api. This section will explain how to use this API for calculating the drought index.

- Example : /processes http://oge.whu.edu.cn/ogcapi/processes_api/processes The API endpoint for retrieving the processes list.
- Example : /processes/{processId} http://oge.whu.edu.cn/ogcapi/processes_api/processes/spei The API endpoint for retrieving a process description (e.g. spei). This returns the description of “spei” process, which contains the inputs and outputs information.
- Example : /processes/{processId}/execution http://oge.whu.edu.cn/ogcapi/processes_api/processes/spei/execution The API endpoint for executing the process. The spei

process exclusively supports asynchronous execution, resulting in the creation of a job for processing. The request body:

```
{ "inputs": { "startTime": "2010-01-01", "endTime": "2020-01-01", "timeScale": 5, "extent": { "bbox": [73.95, 17.95, 135.05, 54.05], "crs": "http://www.opengis.net/def/crs/OGC/1.3/CRS84" } } }
```

- Example : /processes/{processId}/jobs/{jobId} http://oge.whu.edu.cn/ogcapi/processes_api/processes/spei/jobs/{jobId} The API endpoint for retrieving status of a job.
- Example : /processes/{processId}/jobs/{jobId}/results http://oge.whu.edu.cn/ogcapi/processes_api/processes_spei/jobs/{jobId}/results The API endpoint for retrieving the results of a job, which are encoded as : [{ "value": { "time": "2000_02_01", "url": "http://oge.whu.edu.cn/api/oge-python/data/temp/9BC500C1B0E3438C090AF5C6F8602045/8d0357fb-8ffb-4e62-9c3a-55ad17a5831a/SPEI_2000_02_01.png" } },]

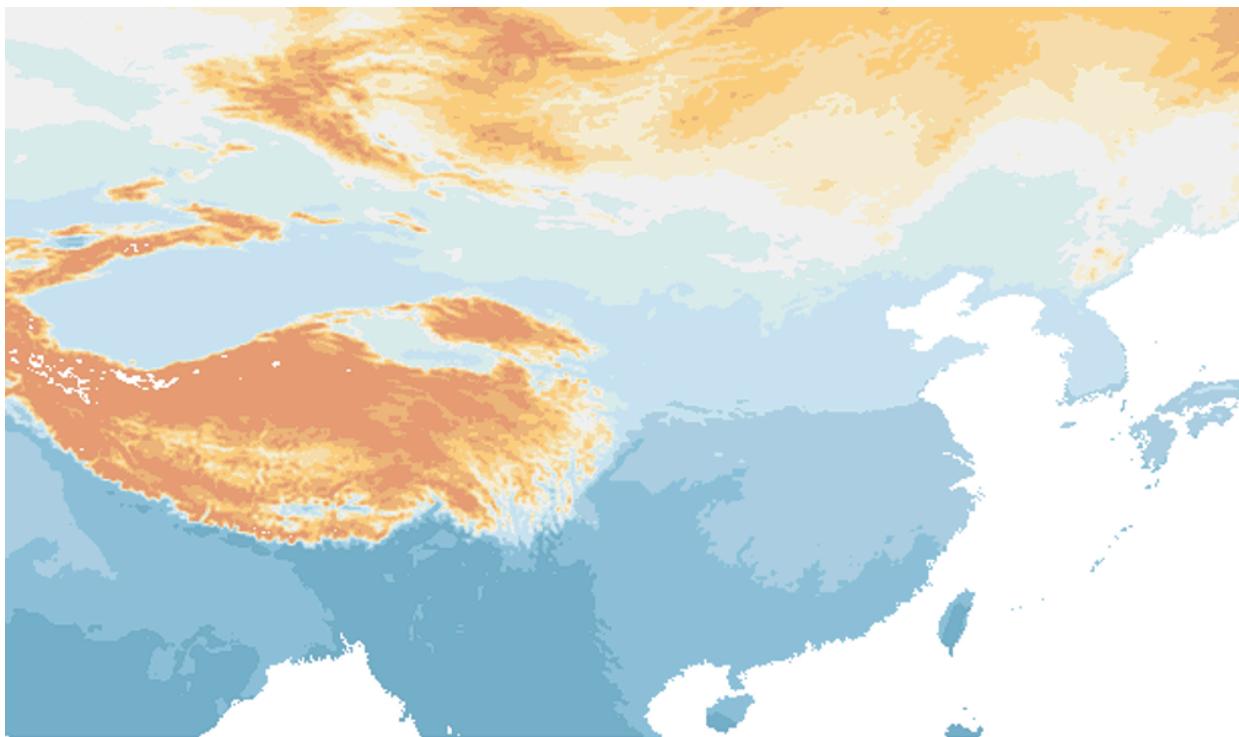


Figure 43 – The SPEI results for the date 2000_02_01.

5.4.3. Component: Data Cube

- Inputs: ERA5 temperature and precipitation data
- Outputs: Results in the form of GeoTIFF after processing in Data Cubes
- What other component(s) can interact with the component: Any components requiring access to temperature and precipitation data in part of Asia through OGC API

- What OGC standards or formats does the component use and produce: OGC API-Coverages

WHU has introduced GeoCube as a cube infrastructure for the management and large-scale analysis of multi-source data. GeoCube leverages the latest generation of OGC standard service interfaces, including OGC API-Coverages, OGC API-Features, and OGC API-Processes, to offer services encompassing data discovery, access, and processing of diverse data sources. The UML model of the GeoCube is given in Figure 5, and it has four dimensions: product, spatial, temporal, and band. Product dimension specifies the thematic axis for the geospatial data cube using the product name (e.g. ERA5_Precipitation or OSM_Water), type (e.g. raster, vector, or tabular), processes, and instrument name. For example, the product dimension can describe optical image products by recording information on the instrument and band. Spatial dimension specifies the spatial axis for the geospatial data cube using the grid code, grid type, city name, and province name. The cube uses a spatial grid for tiling to enable data readiness in a high-performance form. Temporal dimension specifies the temporal axis for the geospatial data using the phenomenon time and result time. Band dimension describes the band attribute of the raster products according to the band name, polarization mode that is reserved for SAR images, and product-level band. The product-level band is the information that is extracted from the original bands. For example, the Standardized Precipitation Evapotranspiration Index (SPEI) band is a product-level band that takes into account the hydrological process and evaluates the degree of drought by calculating the balance of precipitation and evaporation.

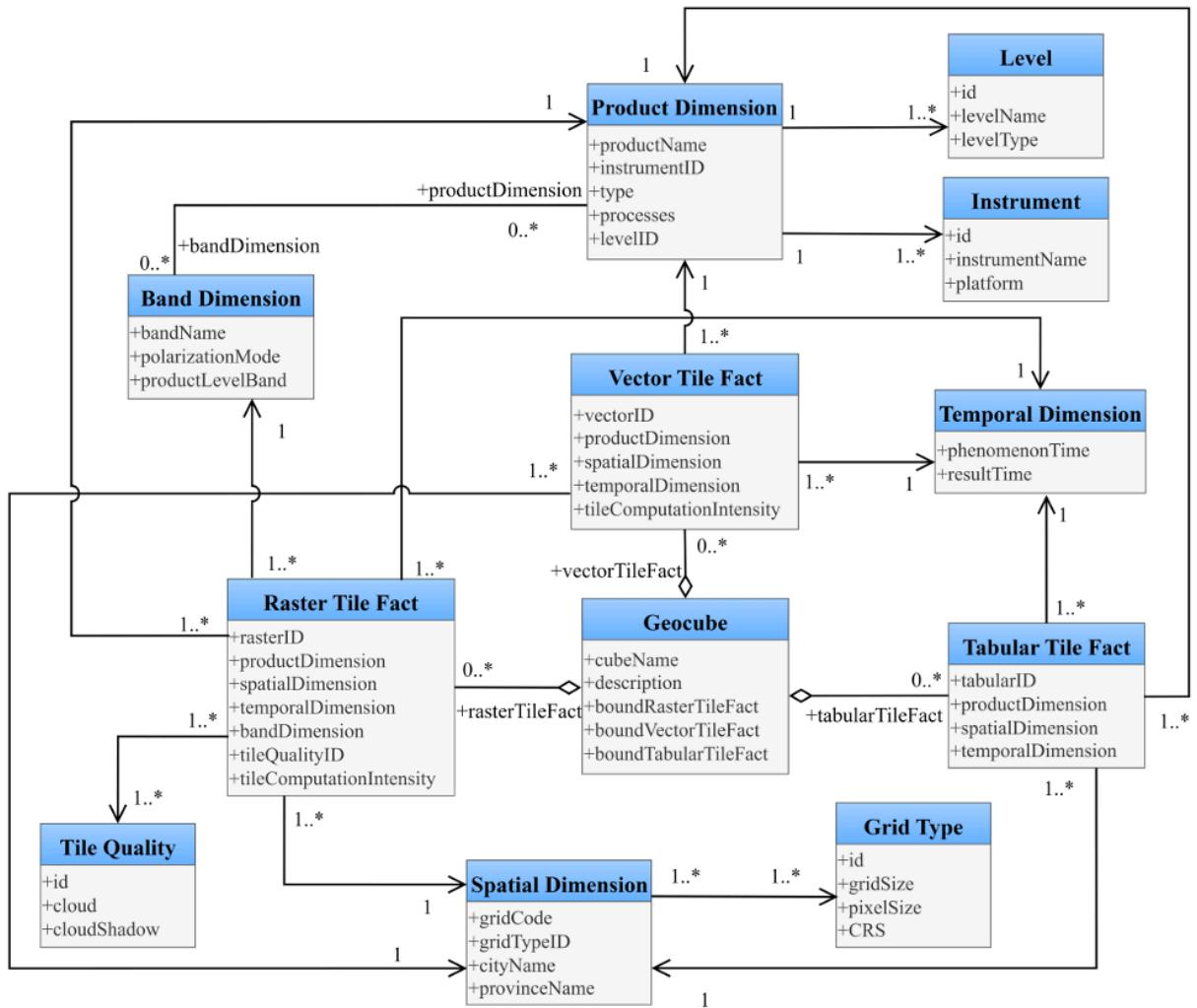


Figure 44 – The UML model of WHU Data Cube.

WHU has organized ERA5 temperature and precipitation data into a cube and offers climate data services through the OGC API – Coverages, supporting the computation of various climate indices. The API endpoint is http://oge.whu.edu.cn/ogcapi/coverages_api, allowing users to query and retrieve the desired data from the cube. This section provides examples demonstrating how to access the data from the cube using OGC API – Coverages.

- Example : /collections http://oge.whu.edu.cn/ogcapi/coverages_api/collections?bbox=112.65942,29.23223,115.06959,31.36234&limit=10&time=2016-01-01T02:55:50Z/2018-01-01T02:55:50Z The API endpoint for querying datasets from the cube, and the query parameters including limit, bbox, and time.
- Example : /collections/{collectionId} http://oge.whu.edu.cn/ogcapi/coverages_api/collections/2m_temperature_201602 The API endpoint for retrieving the description of the coverage with the specified ID from the cube.

- Example : /collections/{collectionId}/coverage http://oge.whu.edu.cn/ogcapi/coverages_api/collections/2m_temperature_201602/coverage The API endpoint for retrieving the coverage in GeoTIFF format for the specified ID. Here is an example of the response:

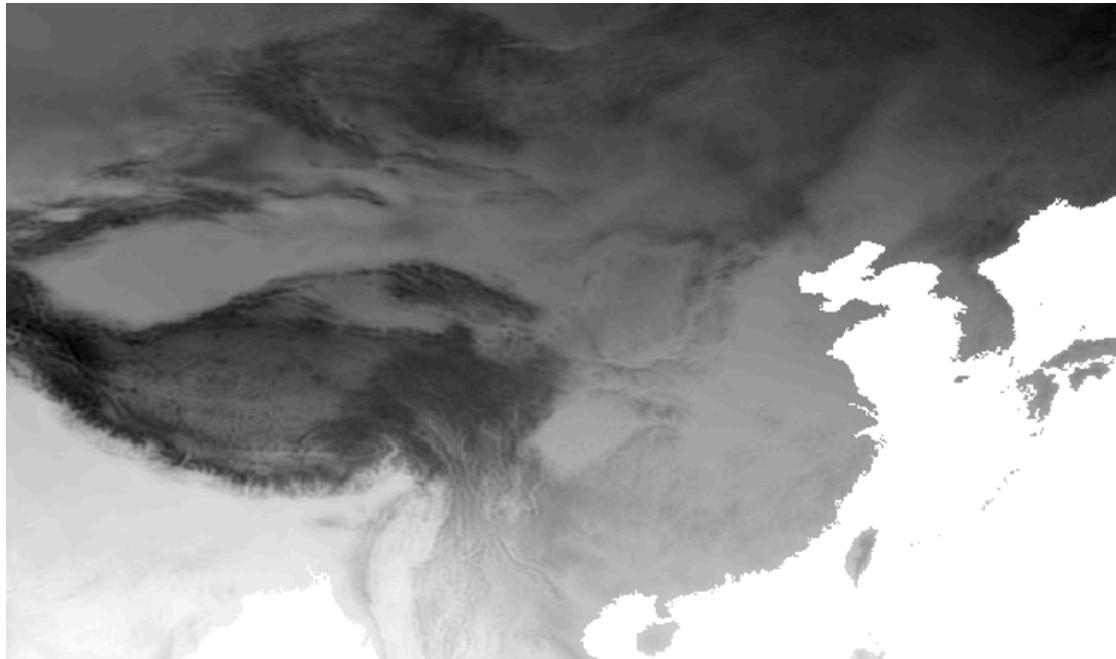


Figure 45 – The coverage with the ID "2m_temperature_201602" in the Asian region.

- Example : /collections/{collectionId}/coverage/rangetype http://oge.whu.edu.cn/ogcapi/coverages_api/collections/2m_temperature_201602/coverage/rangetype The API endpoint for accessing the range type of the coverage, which is part of the band dimension members in the cube. In this example, the coverage consists of only one band dimension member.
- Example : /collections/{collectionId}/coverage/domainset http://oge.whu.edu.cn/ogcapi/coverages_api/collections/2m_temperature_201602/coverage/domainset The API endpoint for the domain set of the coverage, which is also the domain set of the cube.

5.5. ESRI Climate Resilience Data

5.5.1. Climate Projection Data

To make climate projection data more easily usable we transformed CMIP5 data (version 1 of our project), now working on CMIP6, into an Analysis Ready Data collection of indices of future temperature and precipitation. Climate summaries for the contiguous 48 states were derived from data generated for the 4th National Climate Assessment. These data were accessed from the Scenarios for the National Climate Assessment website. The 30-year mean values

for 4 time periods (historic, early-, mid-, and late-century) and two climate scenarios (RCP 4.5 and 8.5) were derived from the Localized Constructed Analogs (LOCA) downscaled climate model ensembles, processed by the Technical Support Unit at NOAA's National Center for Environmental Information.

- Historical: 1976-2005
- Early-Century: 2016-2045
- Mid-Century: 2036-2065
- Late-Century: 2070-2099

In order to display the full range of projections from individual climate models for each period, data originally obtained from USGS THREDDS servers were accessed via the Regional Climate Center's Applied Climate Information System (ACIS). This webservice facilitated processing of the raw data values to obtain the climate hazard metrics available in CMRA.

As LOCA was only generated for the contiguous 48 states (and the District of Columbia), alternatives were used for Alaska and Hawaii. In Alaska, the Bias Corrected Spatially Downscaled (BCSD) method was used. Data were accessed from USGS THREDDS servers. The same variables provided for LOCA were calculated from BCSD ensemble means. However, only RCP 8.5 was available. Minimum, maximum, and mean values for county and census tracts were calculated in the same way as above. For Hawaii, statistics for two summary geographies were accessed from the U.S. Climate Resilience Toolkit's Climate Explorer: Northern Islands (Honolulu County, Kaua'i County) and Southern Islands (Maui County, Hawai'i County).

This data is being updated to CMIP6 and will be available in the latter half of 2023. The system is being expanded globally using NASA NEX CMIP6 data using the same time periods and climate scenarios.

5.5.2. Climate Indices

To provide a more approachable context to future climate, a collection of 47 indices of future temperature and precipitation are computed. These indices build upon prior work on Climdex indices and additional indices developed for National Climate Assessment 4 (NCA4).

- Cooling Degree Days: Cooling degree days (annual cumulative number of degrees by which the daily average temperature is greater than 65°F) [degree days (degF)]
- Consecutive Dry Days: Annual maximum number of consecutive dry days (days with total precipitation less than 0.01 inches)
- Consecutive Dry Days Jan Jul Aug: Summer maximum number of consecutive dry days (days with total precipitation less than 0.01 inches in June, July, and August)
- Consecutive Wet Days: Annual maximum number of consecutive wet days (days with total precipitation greater than or equal to 0.01 inches)

- First Freeze Day: Date of the first fall freeze (annual first occurrence of a minimum temperature at or below 32degF in the fall)
- Growing Degree Days: Growing degree days, base 50 (annual cumulative number of degrees by which the daily average temperature is greater than 50°F) [degree days (degF)]
- Growing Degree Days Modified: Modified growing degree days, base 50 (annual cumulative number of degrees by which the daily average temperature is greater than 50°F; before calculating the daily average temperatures, daily maximum temperatures above 86°F and daily minimum temperatures below 50°F are set to those values) [degree days (degF)]
- Growing-season: Length of the growing (frost-free) season (the number of days between the last occurrence of a minimum temperature at or below 32degF in the spring and the first occurrence of a minimum temperature at or below 32degF in the fall)
- Growing Season 28F: Length of the growing season, 28°F threshold (the number of days between the last occurrence of a minimum temperature at or below 28°F in the spring and the first occurrence of a minimum temperature at or below 28°F in the fall)
- Growing Season 41F: Length of the growing season, 41°F threshold (the number of days between the last occurrence of a minimum temperature at or below 41°F in the spring and the first occurrence of a minimum temperature at or below 41°F in the fall)
- Heating Degree Days: Heating degree days (annual cumulative number of degrees by which the daily average temperature is less than 65°F) [degree days (degF)]
- Last Freeze Day: Date of the last spring freeze (annual last occurrence of a minimum temperature at or below 32degF in the spring)
- Precip Above 99th pctl: Annual total precipitation for all days exceeding the 99th percentile, calculated with reference to 1976-2005 [inches]
- Precip Annual Total: Annual total precipitation [inches]
- Precip Days Above 99th pctl: Annual number of days with precipitation exceeding the 99th percentile, calculated with reference to 1976-2005 [inches]
- Precip 1in: Annual number of days with total precipitation greater than 1 inch
- Precip 2in: Annual number of days with total precipitation greater than 2 inches
- Precip 3in: Annual number of days with total precipitation greater than 3 inches
- Precip 4in: Annual number of days with total precipitation greater than 4 inches
- Precip Max 1 Day: Annual highest precipitation total for a single day [inches]
- Precip Max 5 Day: Annual highest precipitation total over a 5-day period [inches]
- Daily Avg Temperature: Daily average temperature [degF]
- Daily Max Temperature: Daily maximum temperature [degF]

- Temp Max Days Above 99th pctl: Annual number of days with maximum temperature greater than the 99th percentile, calculated with reference to 1976-2005
- Temp Max Days Below 1st pctl: Annual number of days with maximum temperature lower than the 1st percentile, calculated with reference to 1976-2005
- Days Above 100F: Annual number of days with a maximum temperature greater than 100degF
- Days Above 105F: Annual number of days with a maximum temperature greater than 105degF
- Days Above 110F: Annual number of days with a maximum temperature greater than 110degF
- Days Above 115F: Annual number of days with a maximum temperature greater than 115degF
- Temp Max 1 Day: Annual single highest maximum temperature [degF]
- Days Above 32F: Annual number of icing days (days with a maximum temperature less than 32degF)
- Temp Max 5 Day: Annual highest maximum temperature averaged over a 5-day period [degF]
- Days Above 86F: Annual number of days with a maximum temperature greater than 86degF
- Days Above 90F: Annual number of days with a maximum temperature greater than 90degF
- Days Above 95F: Annual number of days with a maximum temperature greater than 95degF
- Temp Min: Daily minimum temperature [degF]
- Temp Min Days Above 75F: Annual number of days with a minimum temperature greater than 75degF
- Temp Min Days Above 80F: Annual number of days with a minimum temperature greater than 80degF
- Temp Min Days Above 85F: Annual number of days with a minimum temperature greater than 85degF
- Temp Min Days Above 90F: Annual number of days with a minimum temperature greater than 90degF
- Temp Min Days Above 99th pctl: Annual number of days with minimum temperature greater than the 99th percentile, calculated with reference to 1976-2005
- Temp Min Days Below 1st pctl: Annual number of days with minimum temperature lower than the 1st percentile, calculated with reference to 1976-2005

- Temp Min Days Below 28F: Annual number of days with a minimum temperature less than 28degF
- Temp Min Max 5 Day: Annual highest minimum temperature averaged over a 5-day period [degF]
- Temp Min 1 Day: Annual single lowest minimum temperature [degF]
- Temp Min 32F: Annual number of frost days (days with a minimum temperature less than 32degF)
- Temp Min 5 Day: Annual lowest minimum temperature averaged over a 5-day period [degF]

The individual web services of climate indices and raster data for download can be accessed at:
<https://resilience.climate.gov/pages/climate-model-content-gallery>

Or for each scenario:

- Historical: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-historical/about>
- RCP 4.5 Early Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-4-5-early-century/about>
- RCP 4.5 Mid Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-4-5-mid-century/explore?location=34.597533%2C-95.830000%2C5.00>
- RCP 4.5 Late Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-4-5-late-century/about>
- RCP 8.5 Early Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-8-5-early-century/about>
- RCP 8.5 Mid Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-8-5-mid-century/about>
- RCP 8.5 Late Century: <https://resilience.climate.gov/maps/nationalclimate::u-s-climate-thresholds-loca-rcp-8-5-late-century/explore?location=34.561983%2C-95.830000%2C5.00>

The data can be viewed directly in the online map viewer or opened in ArcGIS Online, ArcGIS Desktop, or a StoryMap. To view in other softwares GeoService and KMZ URLs are on the right side of the page under View API Resources.

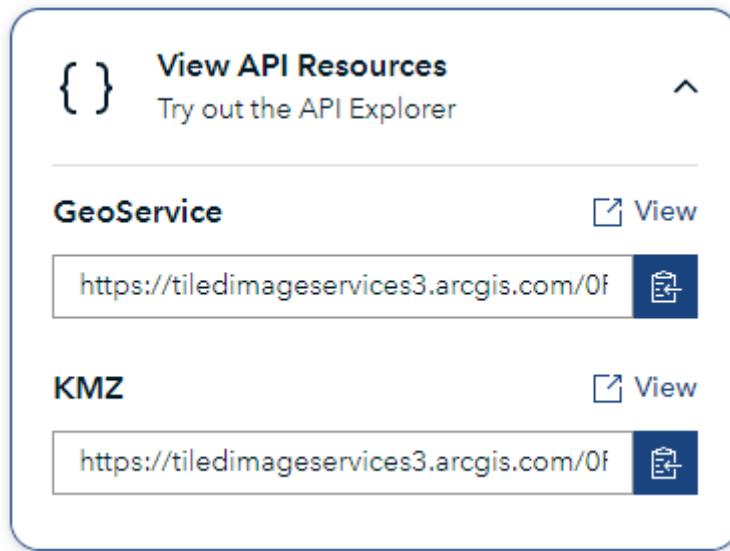


Figure 46 – View API Resources

5.5.3. Summarized Indices for Locations

To support easier interpretation and local decision making, the above indices were summarized by county, census tract, and tribal areas using the Zonal Statistics as Table utility in ArcGIS Pro. The results were joined into the corresponding geography polygons. A minimum, maximum, and mean value for each variable was calculated. This process was repeated for each time range and scenario. Precomputing enables quick map and graph response in the web application, and also provides as easily reusable download for someone who wants to utilize the data elsewhere.

To reuse the summarized services outside of the CRMA application or to download the processed data visit the links below for the geography of interest.

- Counties: <https://resilience.climate.gov/datasets/nationalclimate::climate-mapping-resilience-and-adaptation-cmra-climate-assessment-data/explore?layer=0&location=0.000000%2C0.000000%2C2.74>
- Census Tracts: <https://resilience.climate.gov/datasets/nationalclimate::climate-mapping-resilience-and-adaptation-cmra-climate-assessment-data/explore?layer=1&location=-0.000000%2C0.000000%2C2.76>
- American Indian/Alaska Native/Native Hawaiian Areas: <https://resilience.climate.gov/datasets/nationalclimate::climate-mapping-resilience-and-adaptation-cmra-climate-assessment-data/explore?layer=2&location=-0.000000%2C0.000000%2C2.71>

On these pages, a list of buttons allow you to filter the selection to a subset by attribute or geography, download into a variety of formats, and translate the descriptive documentation for viewing in other languages.



Figure 47 – List of buttons

6

ARD TO DECISION READY INDICATOR (DRI)

ARD TO DECISION READY INDICATOR (DRI)

A decision Ready Indicator (DRI) is information and knowledge that is in such a format that it provides specific support for actions and decisions that have to be made. These indicators are pre-determined, using a set recipe which pulls together one or more ARDs to create an indicator of action and/or decision. DRIs hold significant importance as they serve as benchmarks or criteria to determine when a decision-making process is adequately prepared and can proceed efficiently. Their importance lies in several aspects. Firstly, DRIs facilitate efficient decision-making by signaling that all necessary information, analysis, and resources are available, minimizing delays and preventing hasty or uninformed decisions. Secondly, they ensure quality assurance by setting standards for the decision-making process, ensuring thorough consideration of relevant factors, accurate analysis, and reliable information. DRIs also promote accountability and transparency by defining expectations and providing a framework for evaluation, enabling stakeholders to understand the reasoning behind decisions and hold decision-makers accountable. Additionally, DRIs aid in effective resource allocation by identifying the point at which resources can be allocated, preventing wastage on underprepared decisions. They also assist in managing risks associated with decision-making by encouraging thorough analysis and consideration of potential risks. Furthermore, DRIs promote consistency and standardization, reducing subjectivity and increasing fairness across different decisions. In summary, DRIs play a crucial role in ensuring well-prepared, informed, and accountable decision-making processes, enhancing efficiency, quality, transparency, and resource management.

Analyze Ready Data (ARD), data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets, form the building blocks for DRIs. The transition from ARDs to DRIs encompasses a series of steps designed to extract meaningful insights and facilitate informed decision-making. It commences with the collection and preparation of data, where relevant information is gathered from diverse sources and formatted appropriately for analysis. This involves data cleaning, standardization, and transformation to ensure consistency and reliability. Following data preparation, the integration stage merges multiple data sources, aligning them based on common variables or identifiers, thereby creating a comprehensive dataset.

Subsequently, data exploration and analysis techniques are employed to delve into the dataset's intricacies. Through statistical analysis, data visualization, and data mining, analysts uncover patterns, relationships, and trends that enable a deeper understanding of the underlying information. Feature engineering plays a crucial role in enhancing the analytical model's performance. By selecting pertinent features, transforming existing variables, handling missing data, and encoding categorical variables, analysts optimize the model's ability to extract insights from the data.

Once the data is prepared and features are engineered, model development ensues. Depending on the nature of the problem and the data at hand, analysts choose appropriate algorithms, such as regression, classification, clustering, or machine learning, to build predictive or analytical models. These models are then trained using a portion of the data, often referred to as the training set. Validation is performed using a separate portion of the data, the validation set, to assess the model's performance and fine-tune it for optimal results.

With the validated model in place, the focus shifts to generating Decision Ready Indicators (DRIs). These indicators are specific metrics, scores, or predictions derived from the model's outputs, providing actionable insights relevant to the decision-making process. The DRIs serve as valuable tools that support decision-makers in interpreting the analyzed data and guide them in making well-informed choices.

The generated DRIs become pivotal components in the decision-making process. Decision-makers leverage these indicators to assess different scenarios, evaluate risks, and identify opportunities. By incorporating the insights gained from the analyzed data and model outputs, decision-makers can make more informed and data-driven decisions, enhancing their ability to achieve desired outcomes.

It is worth noting that while the outlined steps provide a general framework, the specific implementation of the process may vary based on the unique context, data characteristics, and analytical techniques employed. Nonetheless, the overarching objective remains constant: to transform Analyze Ready Data into Decision Ready Indicators that facilitate effective decision-making. Below we provide examples on what DRIs can be developed in relation to Climate Resilience.

6.1. Wildfire hazard component

To develop its component, Intact migrated its previous proprietary wildfire hazard model to a private on-premise data science environment. For key inputs to the model, external connections to several open data repositories were established. To facilitate these access tests, several public open source datasets, such as climate model outputs, Earth observations, weather and geospatial, were vetted by the appropriate cybersecurity boards. The tests also informed experts of changes in platforms offerings, of new data products specifications, applicable licenses, and of current authoritative scientific references.

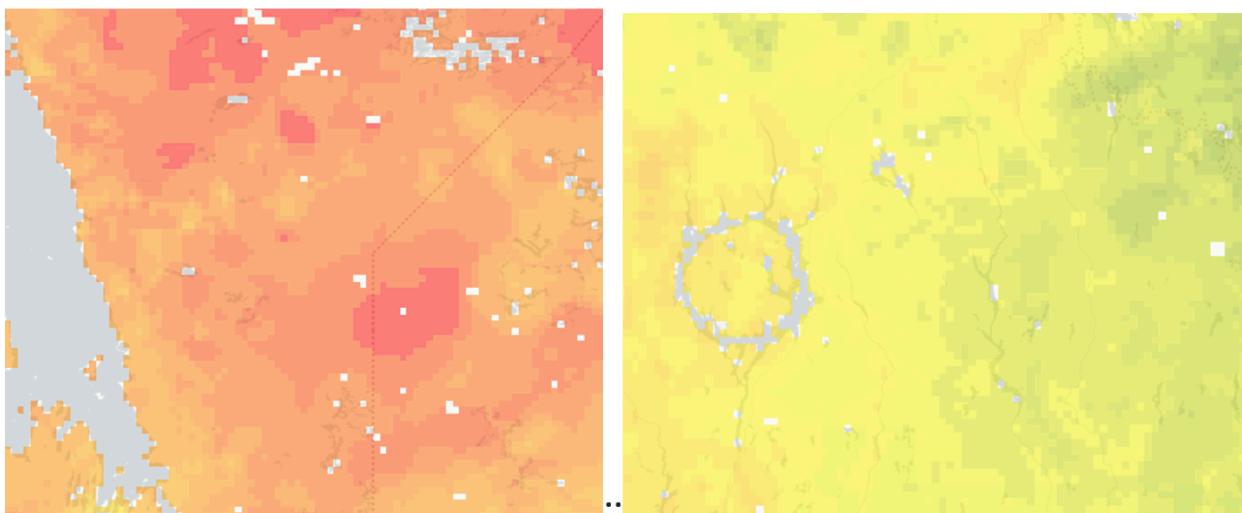


Figure 48 – Two samples of IFC's current national wildfire hazard map

The table below shows the datasets accessed by Intact during the pilot.

6.1.1. Technical Interoperability Experiments (TIE) Table

DATASET	SOURCE	URL	NOTES
National Fire Database fire polygon data	NRCan	https://cwfis.cfs.nrcan.gc.ca/datamart/download/nfdbpoly	Unable to establish SSL connection into private network
Fire Weather Index and its components	NRCan	https://cwfis.cfs.nrcan.gc.ca/downloads/fwi_obs/	Unable to establish SSL connection into private network
Forest Fuels	NRCan	ftp://ftp.nofc.cfs.nrcan.gc.ca/pub/fire/cwfis/data/fuels/	
Vegetation concentration and mass	NRCan	http://tree.pfc.forestry.ca/	503 Service Unavailable from private network
Daily reanalysis composites	NOAA	https://psl.noaa.gov/data/composites/day/	
Monthly reanalysis composites	NOAA	https://psl.noaa.gov/cgi-bin/data/composites/printpage.pl	
Global temperature anomalies/trends	NASA	https://data.giss.nasa.gov/gistemp/maps/	
Elevation at 30 meters	NASA	https://lpdaac.usgs.gov/products/nasadem_hgtv001/	
Canadian Drought Monitor	AAFC	https://agriculture.canada.ca/atlas/data_donnees/canadianDroughtMonitor/data_donnees/shp/	
Canadian Lightning Detection Network	NRCan	ftp://ftp.nofc.cfs.nrcan.gc.ca/pub/fire/CLDN/	Connection timed out, can't find alternate source
Topography	USGS	https://topotools.cr.usgs.gov/gmted_viewer/viewer.htm	Interactive map, not layers
Road segments	NRCan	ftp://ftp.nofc.cfs.nrcan.gc.ca/pub/fire/cwfis/data/base_data	Connection timed out, can't find alternate source
Population of the world	Columbia U.	https://beta.sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density/data-download	

DATASET	SOURCE	URL	NOTES
CanVec Manmade Structures	NRCAN	http://ftp.geogratis.gc.ca/pub/nrcan_rncan/vector/canvec/shp/ManMade/	503 Service Unavailable from private network

Below is a summarized list of the key datasets required to produce or update a wildfire hazard map.

- National fire database polygon data
- Fire Weather Index (FWI) daily maps
- Land cover maps
- Drought conditions
- Digital Elevation Model (DEM)
- Population density
- Fuel and vegetation data

Intact's wildfire hazard map is developed exclusively for internal use. Aside from intellectual property terms, it is meant to be deployed in highly secured data environments, and as such it cannot readily interact with other components of the pilot at this point of time. The intent is to develop geospatial infrastructures and legal terms that would allow a closer collaboration with pilot's participants.

Very early in the project, Intact also developed an H3 synthetic exposure dataset (see next figure) composed of 14M points spread out across the country in a statically representative pattern. The purpose of this dataset was to facilitate visualization and analysis of the exposure. It was also supposed to allow pilot participants to have a common exposure reference on which to develop decision-ready use cases for insurance, thus advancing towards standardization. Unfortunately, time constraints prevented update and sharing of this dataset.

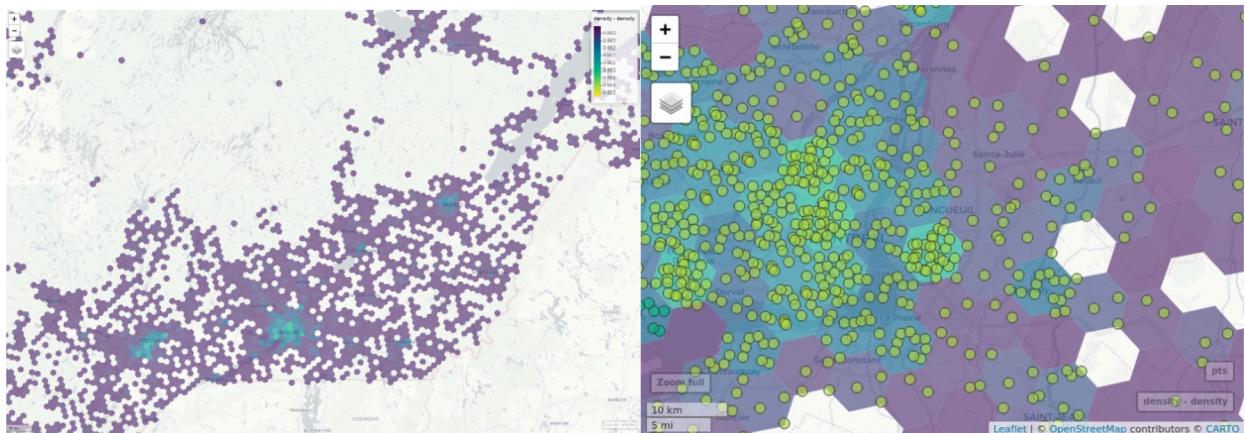


Figure 49 – IFC’s exposure synthetic dataset, with Montreal – Ottawa corridor on the left, and close-up of Montreal on the right. Color scale represent relative risk density in each cell, while points are representative individual risks

6.2. The Blue Economy

Pelagis’ participation in the Climate Resilience pilot focuses on enhancing our view of a global oceans observation system by combining real-world ground observations with analysis ready datasets. Monitoring aspects of our oceans through both a temporal and spatial continuum while providing traceability through the observations process allows stakeholders to better understand the stressors affecting the health of our oceans and investigate opportunities to mitigate the longer term implications related to climate change.

The approach to address the needs for a sustainable ocean economy is to make Marine Spatial Planning a core foundation on which to build out vertical applications. Pelagis’ platform is based on a federated information model represented as a unified social graph. This provides a decentralized approach towards designing various data streams each represented by their well-known and/or standardized model. To date, service layers based on the OGC standards for Feature, Observations & Measurements, and Sensors APIs have been developed and extended for adoption within the marine domain model. Previous work provides for data discovery and processing of features based on the IHO S-100 standard (Marine Protected Areas, Marine Traffic Management, ...); NOAA open data pipelines for major weather events (Hurricane Tracking, Ocean Drifters, Saildrones ...); as well as connected observation systems as provided by IOOS and its Canadian variant, CIOOS.

6.2.1. From Raw Data to ARD and Decision Ready Indicators

The United Nations Framework Convention on Climate Change (UNFCCC) is supported through a number of organizations providing key observation data related to climate change. Of primary interest to this project scenario is the Global Climate Observing System (GCOS) and Global Ocean Observing System (GOOS), and the Joint Working Group on Climate (WG Climate) of the

Committee on Earth Observation Satellites (CEOS). In-situ data sources are provided through a number of program initiatives sponsored through NOAA and provide key indicators for climate change that cannot be directly inferred from raw satellite information.

GCOS defines 54 Essential Climate Variables of which 18 ECVs apply to the oceans domain. Of these, only 6 ECVs may be inferred from satellite based earth observations while the remainder must be inferred through in-situ site observations and/or sampling programs.

The following table identifies the ocean-specific ECVs and associated providers.

Table 1

VARIABLE	DESCRIPTION	SOURCE OF INDICATOR
Ocean Colour	Provides indication of phytoplankton based on Ocean Colour Radiance (OCR)	ESA CEOS
Carbon Dioxide Partial Pressure	Primary indicator of the exchange of CO ₂ at the ocean surface	NOAA
Ocean Acidity	pH of ocean water as measured at varying depths and locations	NOAA PMEL
Phytoplankton	Indicator of the health of the ecosystem associated with the food web and directly a result of increased CO ₂ and eutrophication	NOAA
Sea Ice	Sea ice coverage associated with the ocean surface and a concern reflected in warming surface temperatures and sea level rise	
Sea Level	Sea level global mean and variability leading to sea level rise	
Sea State	Wave height, direction, wavelength as indicators of energy at the ocean surface	
Sea-surface Salinity	The proportion of ocean water comprised of salt and indicator of mortality rates in shellfish	
Sea-surface Temperature	Directly affects major weather patterns and ecosystems	ESA CEOS; NOAA Monitoring Stations; NOAA Saildrone program
Surface Current	Transports heat, salt and passive tracers and has a large impact on seaborne commerce and fishing	

As well, social and economic key indicators related to the area of interest are ingested to identify relationships between the immediate effects of climate change on the associated human activity.

Table 2

VARIABLE	DESCRIPTION	SOURCE OF INDICATOR
AQ Landings	Annual yields associated with Aquaculture sites within a region of interest	MaineAQ
GDP	Gross Domestic Product (\$USD) associated with dependent human activities within the region of interest	US Census
Employment	Number of individuals dependent on the targeted ecosystem	US Census
Population	Number of people inhabiting the area of interest associated with the ecosystem	US Census

6.2.2. Approach

Each ECV applicable to the use case is resolved as a service endpoint representing the area of interest, associated samplings and observations, and where possible inferred from earth observation datasets transformed as 'analysis ready'. Earth observation datasets are sourced through the ESA GCOS service endpoint; Ocean related samplings and in-situ observations are sourced through NOAA; socio-economic data is sourced from various open data portals available through government agencies.

The project effort centers around 3 key challenges * the ability to collect data relevant to Climate Resilience; * the ability to apply the data in a coherent and standardized manner in which to draw out context; * and the ability to impart insight to community members and stakeholders so as to identify, anticipate and mitigate the effects of climate change across both local and international boundaries.

Each of these activities aligns with the best practices and standards of the OGC and are used as input to the MarineDWG MSDI reference model.

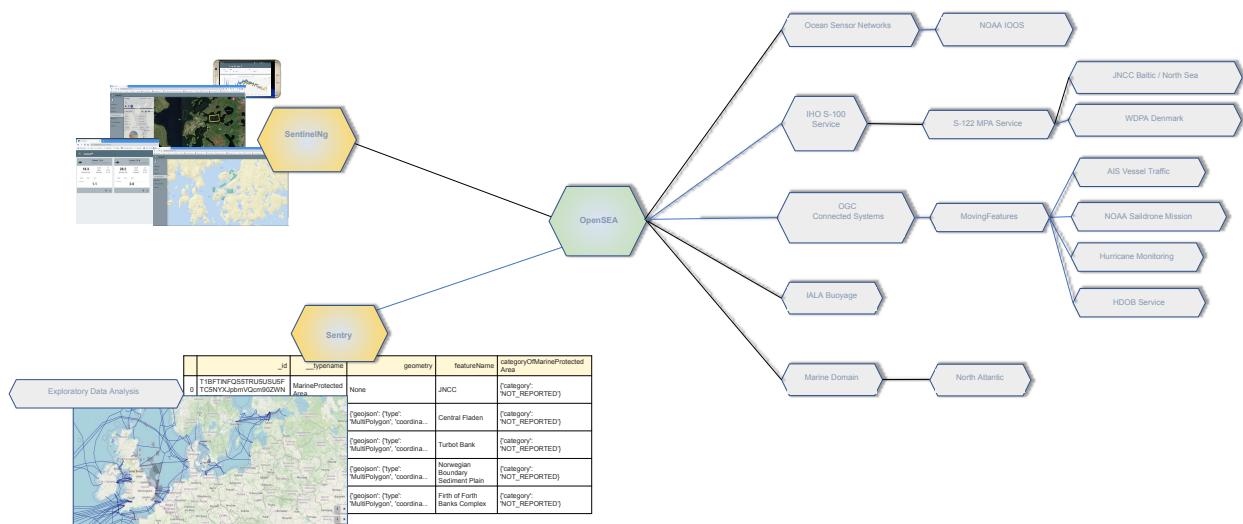


Figure 50 – Architecture

6.3. ECMWF – Copernicus (will be integrated with INTRODUCTION section)

- Component: Copernicus services.
- Outputs: Copernicus Services, including Climate Data Store (CDS) <https://cds.climate.copernicus.eu/> and Atmosphere Data Store (ADS) <https://ads.atmosphere.copernicus.eu/>.
- What other component(s) can interact with the component: CDS and ADS provide access to data via different interfaces: UI and API. It also offers a toolbox with a set of expert libraries to perform advanced operations on the available data. CDS and ADS catalogue metadata is also accessible via standard CSW. <https://cds.climate.copernicus.eu/geonetwork/srv/eng/csw?SERVICE=CSW&VERSION=2.0.2&REQUEST=GetCapabilities>
- What OGC standards or formats does the component use and produce:
 - CDS and ADS catalogues exposed via CSW.
 - Access to ESGF datasets via WPS.
 - WMS is offered in some published applications.
 - CADS 2.0 (under construction) will implement OGC APIs.

6.3.1. DRI: Heat Impact and Drought Impact Components – Safe Software

6.3.1.1. Heat Impact DRI Component

This component takes the climate scenario summary ARD results from the ARD component and analyzes them to derive estimated heat impacts over time, based on selected climate scenarios. Central to this is the identification of key heat impact indicators required by decision makers and the business rules needed to drive them. Process steps include data aggregation and statistical analysis of maximum temperature spikes, taking into account the cumulative impacts of multiple high temperature days. Heat exhaustion effects are likely dependent on duration of heat spells, in addition to high maximum temperatures on certain days.

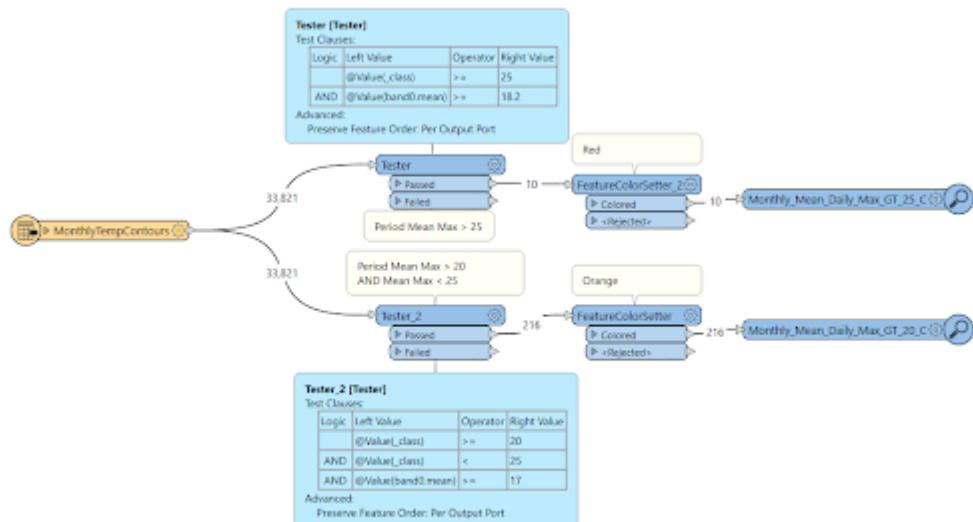


Figure 51 – ARD Query: Monthly Max Temp Contours

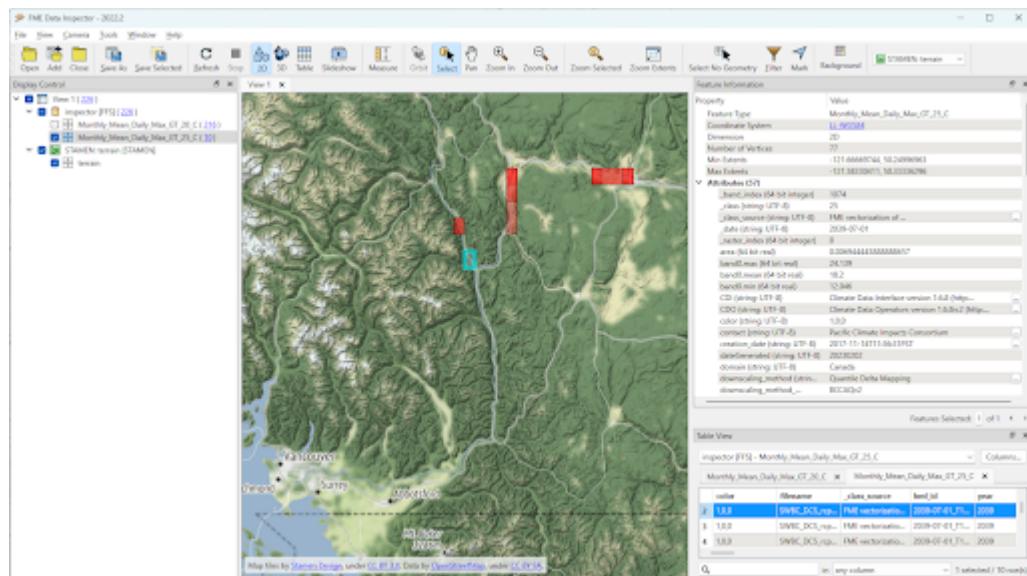


Figure 52 – ARD Query: Max Mean Monthly Temp > 25C

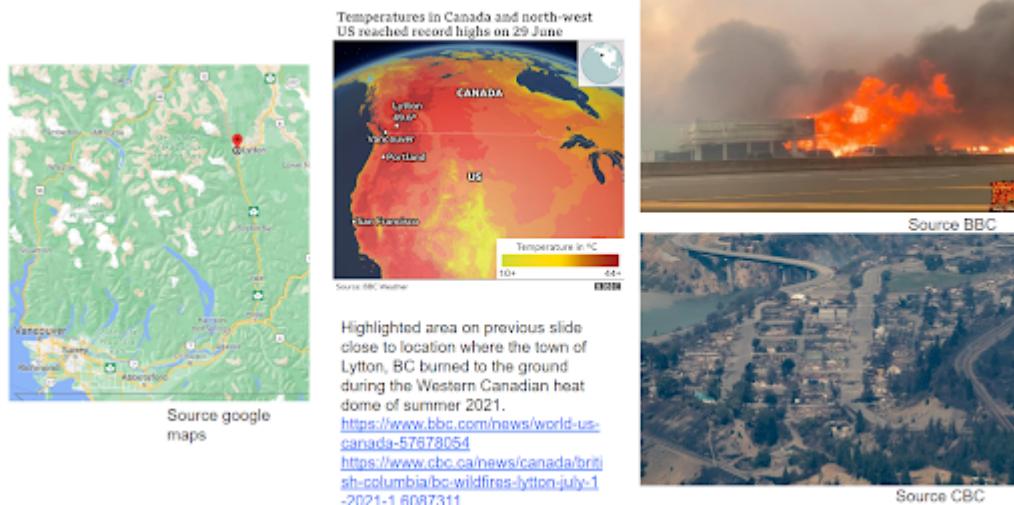


Figure 53 – Town of Lytton - location where entire town was devastated by fire during the heat wave of July 2021 - same location highlighted in ARD query from heat risk query in previous figure

6.3.1.2. Drought Impact DRI Component

This component takes the climate scenario summary ARD results from the ARD component and analyzes them to derive estimated drought risk impacts over time based on selected climate scenarios. It also feeds drought related environmental factors to other pilot DRI components for more refined drought risk analysis. For the purposes of this pilot, it was recognised that more complex indicators such as drought are likely driven by multiple environmental and physical factors. As such, our initial goal was to select and provide primary climate variable data that would be useful for deriving drought risks in combination with other inputs. Given that the primary input to drought models is precipitation, or lack thereof, we developed a data flow that extracted total precipitation per month and made this available both as a time series CSV and GeoJSON datasets, as well as OGC API features time series points. This climate scenario primary drought data was provided for the province of Manitoba and for Los Angeles. These two regions were chosen since we had pilot participants interested in each of these regions and in the case of Manitoba there is also a tie in to future work as this is an area of interest for the subsequent Disaster Pilot 2023.

For the LA use case, we worked with Laubwerk to provide them with climate change impact data that could help drive a drought impact that could affect their future landscape visualization model. The idea is that based on changes to climatic variables, certain areas may be more or less suited to different vegetation types, causing the distribution of vegetation to change over time. For more on their component, please refer to section 7: From Data To Visualization.

In the case of this visualization component, simply providing precipitation totals per month were not sufficient to drive the needs of their vegetation model. In this case we did not have an intermediate drought model to feed climate variables to. In the absence of a more comprehensive drought model, we decided to develop a proxy drought risk indicator by normalizing the difference between future precipitation and past.

Calculations were made using the difference between time series grids of projected precipitation and historical grids of mean precipitation per month. These precipitation deltas were then divided by the historical max – mean per month to derive a precipitation index. The goal was to provide a value between -1 and +1 where 1 = 100% of past mean precipitation for that month. Naturally this approach can generate values that exceed the range of -1 to 1 if the projected precipitation values exceed the historic max or min. The goal was not so much to predict future absolute precipitation values but rather generate an estimated for precipitation trends given the influence of climate change. For example, this approach can help answer the question – in 30 years for a given location, compare to historical norms, by what percentage do we expect precipitation to increase or decrease. Laubwerk can then take these results and decide what degree of drought stress will cause a specific vegetation species to die out for a particular location.

Interesting patterns emerged for the LA area that we ran this process on deltas between projected and historical precipitation. While summers are typically dry and winters are wet and prone to flash floods. Initial data exploration seemed to show an increase in drought patterns in the spring and fall. More analysis needs to be done to see if this is a general pattern or simply one that emerged from the climate scenario we ran. However, this is the type of trend that local planners and managers may benefit from having the ability to explore once they have better access to climate model scenario outputs along with the ability to query and analyze them.

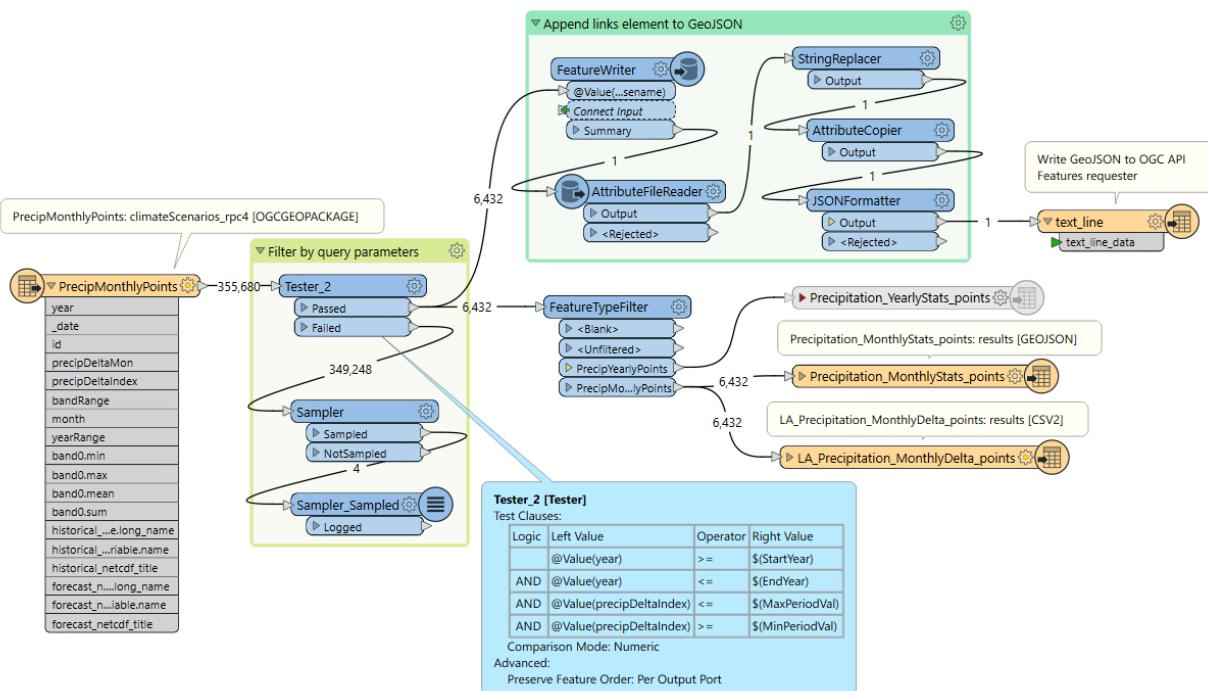


Figure 54 – FME Query Workflow: Geopackage precipitation delta time series to GeoJSON points

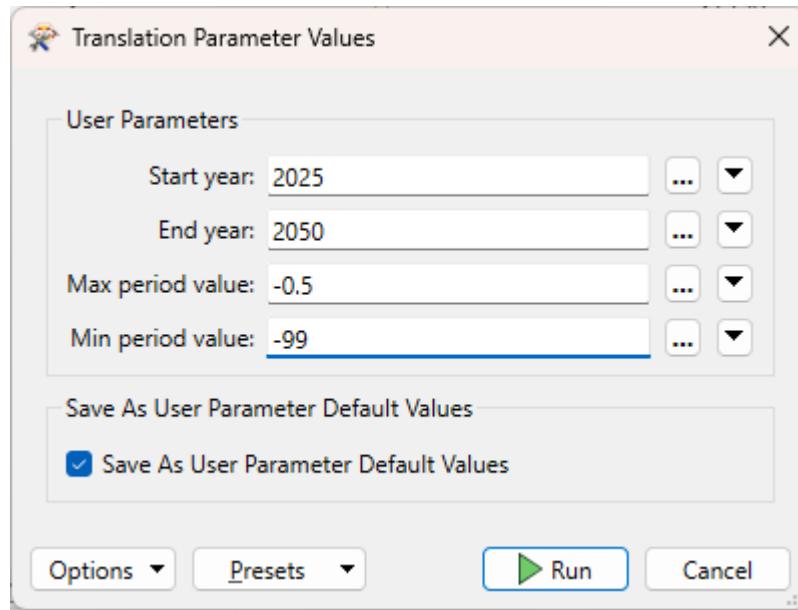


Figure 55 – FME Query Parameters: Geopackage precipitation delta time series to GeoJSON points

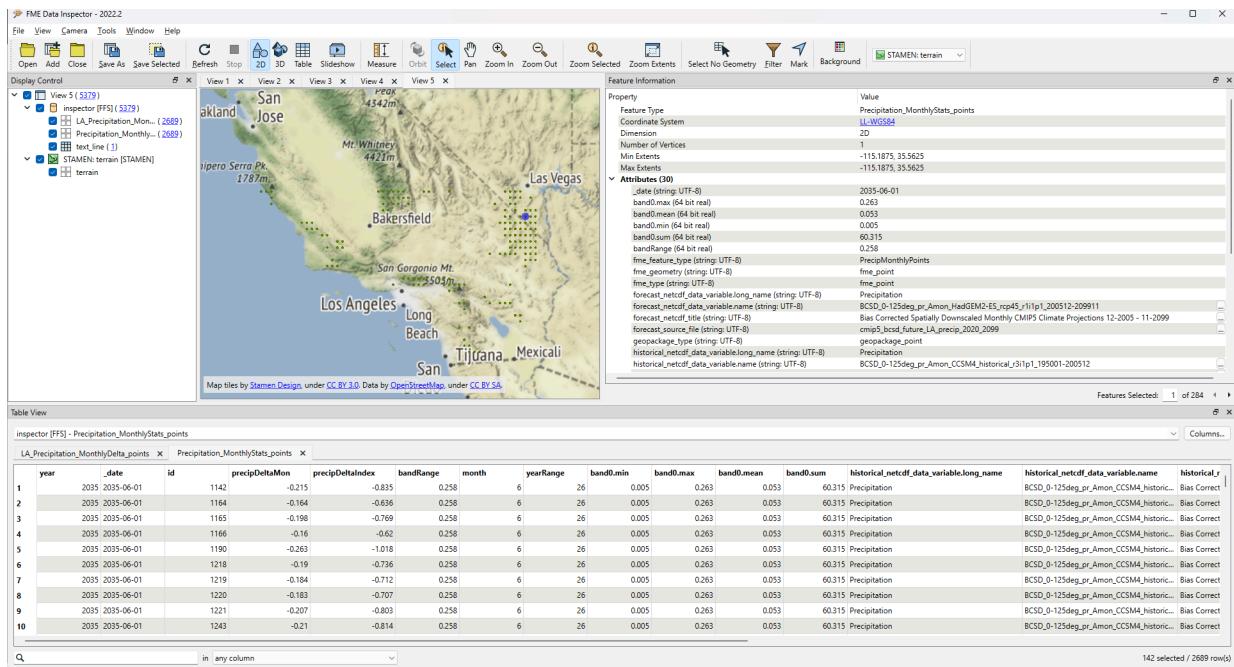


Figure 56 – FME Data Inspector: precipitation delta result showing potential drought risk for areas and times with significantly less precipitation than past

This approach is only a start and just scratches the surface in terms of what is possible for future drought projection based on climate model scenario ECVs. The specific business rules used to assess drought risk are still under development. FME provides a flexible data and business rule modeling framework. This means that as indicators and drought threshold rules are refined, it's relatively straightforward to adjust the business rules in this component to refine our risk projections. Also, business rule parameters can be externalized as execution parameters so

that end users can control key aspects of the scenario drought risk assessment without having to modify the published FME workflow. However one of the main goals of this pilot was not so much to produce highly refined forecast models for drought but rather to demonstrate the data value chain whereby raw climate model data cube outputs can feed a data pipeline that filters, refines, simplifies the data and ultimately can be used to drive indicators that help planners model visualize and understand the effects of climate change on the landscapes and environments within their communities.

To support future drought risk estimates for Manitoba, we also provided a precipitation forecast time series to Pixalytics as an input to their drought analytics and DRI component. Their component provides a much more sophisticated indicator of drought probability since in addition to precipitation it also takes into account soil moisture and vegetation. The goal was to extract precipitation totals per time step from the downscaled RCM – regional climate model ECV outputs for Manitoba based on CMIP5 (Coupled Model Intercomparison Project Phase 5) model results obtained from Environment Canada. For this use case the grids have a spatial resolution of roughly 10km and a temporal resolution a monthly time step. Pixalytics then ran their drought model based on these precipitation estimates in order to assess potential future drought risk in southern Manitoba. The data was provided to Pixalytics initially as a GeoJSON feed of 2d points derived from the data cube cells with precipitation totals per cell. We later also provided this same data feed as a OGC API Feature service.

For future phases of the climate or disaster pilots, it may be useful to explore additional approaches for both precipitation data analysis and combination with other related datasets and external models. It may be useful to segment cumulative rainfall below a certain threshold P_t within a certain time window (days, weeks or months), since cumulative rainfall over time will be crucial for computing water budgets by watershed or catch basin. To do this we would like to test the use of a higher resolution time step such as daily, to see if the increased resolution reveals patterns of interest that the coarser monthly time step does not. There are also other statistical RCM results that might be useful to make available (mean, min, max). Besides precipitation, climate models also generate soil moisture predictions which could be used by this component to assess drought risk. This component would also benefit from integration with topography, DEMs and hydrology related data such as river networks, water bodies, aquifers and watershed boundaries. Therefore rather than just computing precipitation deltas at the cell level, it would likely be useful to sum precipitation by catch basin and compute future trends that may indicate potential drought or flood.

The specific business rules used to assess drought risk are still under development. FME provides a flexible data and business rule modeling framework. This means that as indicators and drought threshold rules are refined, it's relatively straightforward to adjust the business rules in this component to refine our risk projections. Also, business rule parameters can be externalized as execution parameters so that end users can control key aspects of the scenario drought risk assessment without having to modify the published FME workflow.

It should be stressed that the field of drought modelling is not new and there are many drought modelling tools available that are far more sophisticated than anything described above. As such, subsequent Climate and Disaster pilots should explore how future climate projections can be funneled into these more mature climate models in an automated fashion to produce more refined estimates of projected drought risk. That said, we need to start somewhere, and it is hoped that this basic demonstration of the raw data to ARD to DRI value chain for drought

can provide some insights into what type of indicators we may want to generate to help better understand future drought risks, and where we may want to improve on this process.

7

FROM DATA TO VISUALIZATION

FROM DATA TO VISUALIZATION

Advances in data representation and visualization have revolutionized the way we understand and analyze information. The ability to transform raw data into meaningful visual representations has become increasingly important across various fields, including climate change. The exponential growth of data generated by various sources such as in-situ sensors, EO sensors, and social media has all led to the emergence of big data. Data visualization techniques help in extracting insights, identifying patterns, and making data-driven decisions in the face of vast and complex datasets. Visualization plays a crucial role in exploring, summarizing, and communicating the results of data analysis, making it easier for decision-makers to comprehend complex information. Data visualization enhances storytelling by presenting information in a visually engaging and intuitive manner. It helps convey complex ideas more effectively, enabling clearer communication of data-driven narratives to both technical and non-technical audiences.

Above all, general need for data visualization arises from the complexity and volume of data that is involved with climate change adaptation. Data visualizations are stimulated by the desire for actionable insights, and the importance of clear communication in various domains.

Below we provide some examples of how big data can be visualized in such a way that it captures the impact of climate change on for example vegetation in urban areas, or the impact of climate change on climate hazards. And how to overcome challenges to realize these visualizations.

7.1. 5D Meta World

Presagis offered the V5D rapid 3D (trial) Digital Twin generation capability to Laubwerk Presagis gathered open source GIS dataset for the Hollywood region in order to match the location of the tree dataset from Laubwerk Using V5D, Presagis created a representative 3D digital twin of the building and terrain. Presagis imported Laubwerk tree point dataset providing vegetation type information inside V5D Presagis provided V5D Unreal plugin to Laubwerk in order to allow the insertion of the Laubwerk 3D tree (as Unreal assets) into the scene. Using V5D, Laubwerk is capable of adapting the tree model in order to demonstrate the impact of climate change on the city vegetation

Presagis also provided to Laubwerk its V5D AI extracted vegetation dataset in order to complement the existing tree dataset as needed.



Figure 57

7.2. Visualizing the Impact of Climate Change and Mitigation on Vegetation

One of the biggest challenges in communicating climate change is to tie global changes to the local impact they will have. Photorealistic visualization is a critical component for assessing and communicating the impact of environmental changes, and possibilities for mitigation. For this to work, it is crucial for visualizations to reflect the underlying data accurately and allow for quick iteration. In this regard, manual visualization processes are inferior. As much as possible, visualizations of real-life scenarios should be driven directly by available data of present states and simulations of possible scenarios. Our contribution is a first attempt at doing just that, determining what already works and what doesn't with existing data and technology.

As our contribution to the Climate Resilience Pilot we explored such data-driven high-quality visualizations, focusing on the impact on vegetation. Due to the nature of this being a pilot, we constrained ourselves in terms of coverage area, to account for limited time and to cope with potentially limited data availability. This ensured that we were able to make the full connection from input data to final visualization, drawing valuable conclusions for broader application in the future. This size limitation will allow us to produce meaningful results if data transfer and processing is slow or even if it must be processed in manual or half-automated ways due to

inconsistent formatting. It also lets us visualize a high level of detail without having to account too much for the sheer amount of data we could face with very large areas.

We selected a relatively small section of Los Angeles for actual visualization. The rationale behind this choice of location had several components:

- The given area that will (and already does) see considerable direct impact of climate change through heat, drought, wildfires, etc.
- It contains different areas of land use (from deeply urban and sub-urban to unmanaged areas).
- Since it is part of a major metro area, the results will be relevant to a large population base
- Some known mitigation measures that can be considered for visualization are in place.
- Other external (non climate change) known influences on vegetation, such as pests, irrigation limitations, known life spans of relevant plant species, etc.) are in play that could be considered.

7.2.1. Source Data

Our visualization ties data that is very global together with data that is hyper-local. That means we need to draw on data from a wide variety of sources that are not usually combined. Examples of data sources used for our visualization are:

- Satellite Imagery
- Building Footprints and Heights
- Plant Inventory from Bureau of Street Services and Department of Recreation and Parks
- Results from Climate Models, particularly RPC 4.5 data that was pre-processed for this purpose by Safe Software as part of their work for this pilot (see the Safe Software ARD component in this document for more details)
- 3D Plant Models from the Laubwerk database
- Plant Metadata to Judge Climate Change Impact on Specific Species through given Environmental factors, also from the Laubwerk database
- Information on local mitigation measures from various sources

7.2.2. Results

The aforementioned data sources were combined to create a detailed visualization of the area in question. The pairs of images below show a visualization of the status quo as first image and then a composite of the four scenarios we visualized. The scenarios are projections of a possible climate scenario for 2045 without any 2070 without any mitigation measures (plants that

were likely to die off due to adverse climate events were just removed based on a probability measure) as well as 2045 and 2070 in which plants that were removed have been replaced by plants that are more resilient and are part of the aforementioned initiatives for more climate resilience.

It should be stressed that this is a visualization of a possible outcome, but there are many factors that make hard to make exact predictions! This contribution is merely meant as an example of how data could be used to drive scenario-based hyper-local visualization.



Figure 58 – Overview of the Visualized Region (Status Quo)



Figure 59 – Overview of the Visualized Region (Scenarios)



Figure 60 – Above the Corner Sunset Blvd and N Curson Ave Looking North-East (Status Quo)



Figure 61 – Above the Corner Sunset Blvd and N Curson Ave Looking North-East (Scenarios)



Figure 62 – Corner Franklin Ave And N Sierra Bonita Ave Looking East (Status Quo)



Figure 63 – Corner Franklin Ave And N Sierra Bonita Ave Looking East (Scenarios)



Figure 64 – Corner Hollywood Blvd And Camino Palmero St Looking North (Status Quo)



Figure 65 – Corner Hollywood Blvd And Camino Palmero St Looking North (Scenarios)

7.2.3. Challenges and Learnings

The goal of a visualization like we did is to make data and its implications visible on a hyper-local level. The hope behind this is to turn a large amount of abstract data into something the general public can better judge the very local impact of global changes.

This hyper-locality brings to light a number of problems with the granularity, availability, and machine readability of existing data. Relating to our specific inputs, this means:

- Producing a high fidelity photorealistic 3D model of a specific area is still not easy. Even in an urban area of an industrialized country like we picked (which usually have better data availability), we had to resort to relatively simple elevation data and building footprints. There are solutions for this on the horizon, but general availability is not a given, yet. 3D models based on photogrammetry seem like a promising approach to reach higher fidelity where available, but that generally available datasets like these currently lack classification, so we would not be able to remove and replace vegetation elements. This will probably improve and become more widely available in the near future.
- Information about existing vegetation is of varying quality and completeness. Detailed data is sometimes maintained by different authorities with different scopes. In our case we used data from the [Bureau of Street Services](#) as well as the Department of Recreation and Parks. Those datasets have different data layout, different depth and quality of data. OpenStreetMap also sometimes has vegetation data, but coverage and data quality is also problematic. None of the aforementioned really cover individual plants on private property or unmanaged land, which we had to fill in from photogrammetry, satellite imagery, and aerial photography.

- Climate projection data is pretty widely available and generally easy to process in terms of data volume, because the areas a visualization will typically cover is pretty small compared to the resolution of most climate models. What is still a challenge is to turn climate scenario data into properties that are needed to easily model the impact on vegetation, like the probability of extreme drought, heat, or fire events. This was partially addressed by other contributions to this pilot and we expect it to see further improvements.
- Exact data on average plant behavior in the context of relevant climate indicators is extremely patchy. Most data is only qualitatively in nature. Data gathering is complex because of the large number of factors at play when judging health of plants. This is a complex research topic that will need more work, both to produce more reliable projections based on existing research, but also on how to gather data about or predict plant health more reliably on a large scale.
- Information about climate change mitigation is often not present in a machine readable format. In our specific case, we gathered information manually from publicly available material, mostly websites. Part of the problem here is that several stakeholders are working on mitigation measures, from different local government organizations, over non-profit organizations, to private companies. Examples relevant to our specific example are [City Plants](#) (a non-profit supported by Los Angeles Department of Water and Power) and the [County of Los Angeles Parkway Trees Program](#). This manual way of data gathering obviously will not scale, is prone to data being missed, and has no unified format. All of this makes automated processing next to impossible at the moment.
- There may be further factors that need to be considered, which are not part of any of the existing data sources. In this specific case we have the pretty high average age and also various pests and diseases that the Mexican fan palm (*Washingtonia robusta*), which has become such a distinctive feature of Southern California, especially Los Angeles, is suffering from. While this isn't directly related to climate change, it still needs to be considered for any visualization to be accurate.

As was expected, the data-driven visualization of very local phenomena and changes is a challenging problem which surfaces lots of issues in terms of data availability as well as standardization and compatibility of storage formats.

7.3. ESRI's Web Application

Decision makers, public authorities, and citizens will primarily access the data via a custom web application, providing a simple dashboard interface for viewing interactive maps and graphs of the indices, and output formatted reports. The indices are grouped by 5 climate hazard types (Wildfire, Heat, Drought, Inland Flooding, Coastal Inundation). The current US project (<https://livingatlas.arcgis.com/assessment-tool/explore/details>) can be explored to gain context of what the global project will be.

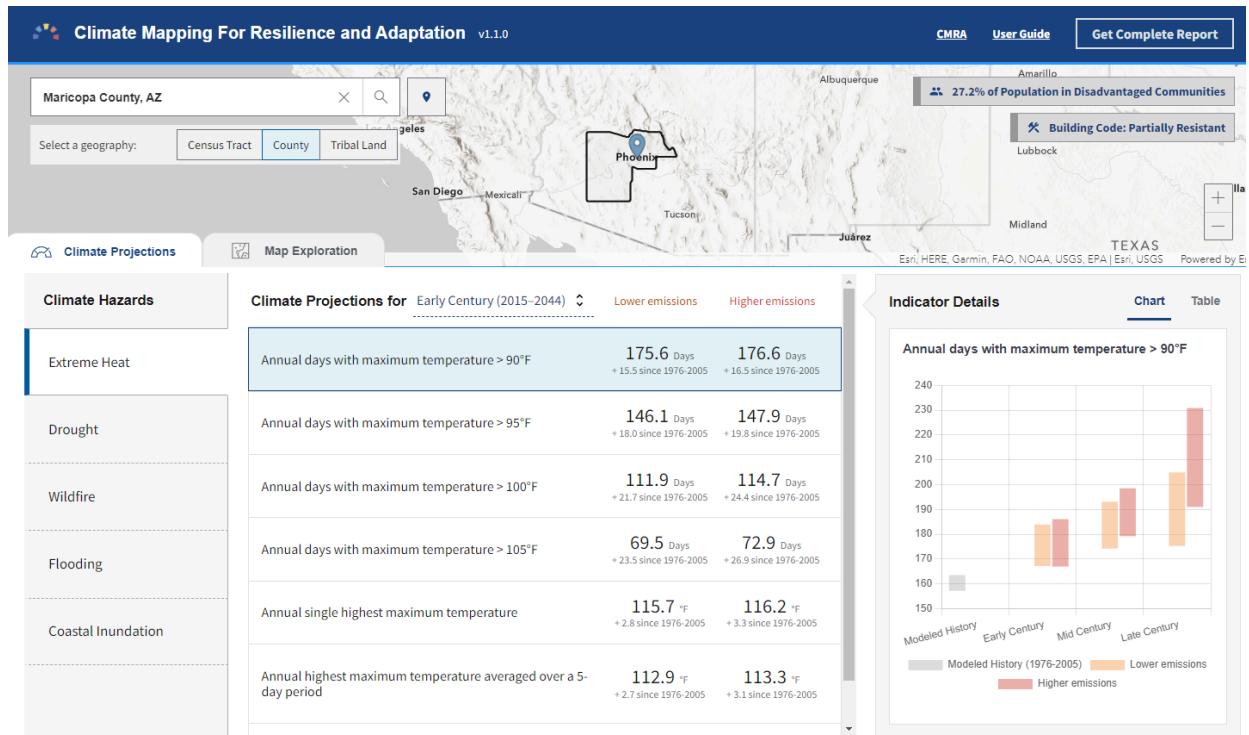


Figure 66 – US Project view

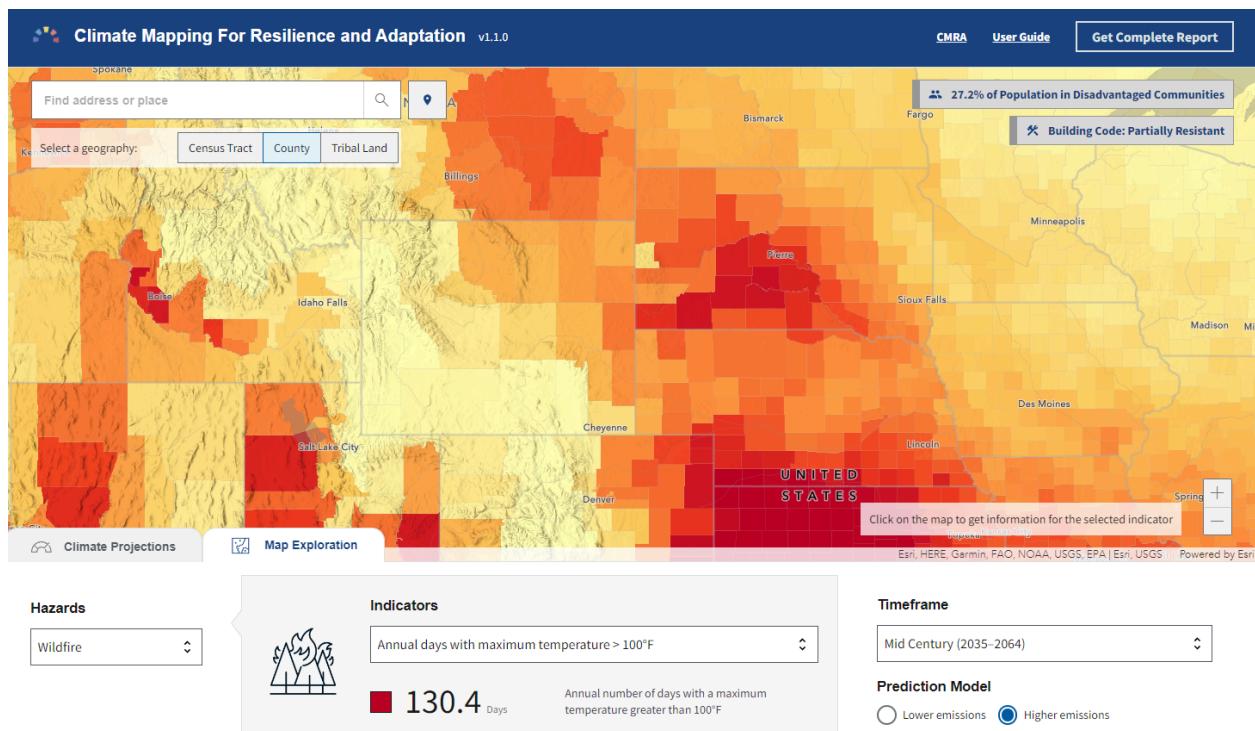


Figure 67 – US Project view

The application also outputs formatted reports by county or census tract summarizing the data in a format easy to share with others.



Figure 68 – Application output reports

For each of those 5 climate hazards there is a corresponding StoryMap to further explain that hazard type, visualize the current and future hazard, and provide links to additional relevant resources.

- Extreme Heat: <https://storymaps.arcgis.com/stories/5e482f11d2514191bb89c20638d98b3c>
- Drought: <https://storymaps.arcgis.com/stories/634ee231bb6743b88d23bda96fb838e9>
- Wildfire: <https://storymaps.arcgis.com/stories/ae2a8072429643f395f8f509df955ae6>
- Flooding: <https://storymaps.arcgis.com/stories/4ea811276aa641018f3a8d4e28585244>
- Coastal Inundation: <https://storymaps.arcgis.com/stories/f3ce292c0211400699b6e36985e561a6>

8

USE CASES

8.1. Drought Impact Use Cases (Wuhan University)

Based on the ARD, drought indicator, and data cube components, WHU develops three use-cases based on self-developed Open Geospatial Engine (OGE) for drought impact for rapid response to drought occurrences. Figure 69 shows the technical architecture of the OGE. It has the following features: 1) For data discovery, a catalogue service from OGE data center following OGC API is provided, allowing users to search geospatial data both available from WHU data stores and remote data stores. 2) For data integration, data can be integrated into the WHU software in the form of data cubes with three efforts: formalizing cube dimensions for multi-source geospatial data, processing geospatial data query along cube dimensions, and organizing cube data for high-performance geoprocessing. 3) For data processing, a processing chain is enabled in OGE using a code editor and modelbuilder. 4) For data visualization, a Web-based client for visualization of spatial data and statistics is provided using a virtual globe and charts.

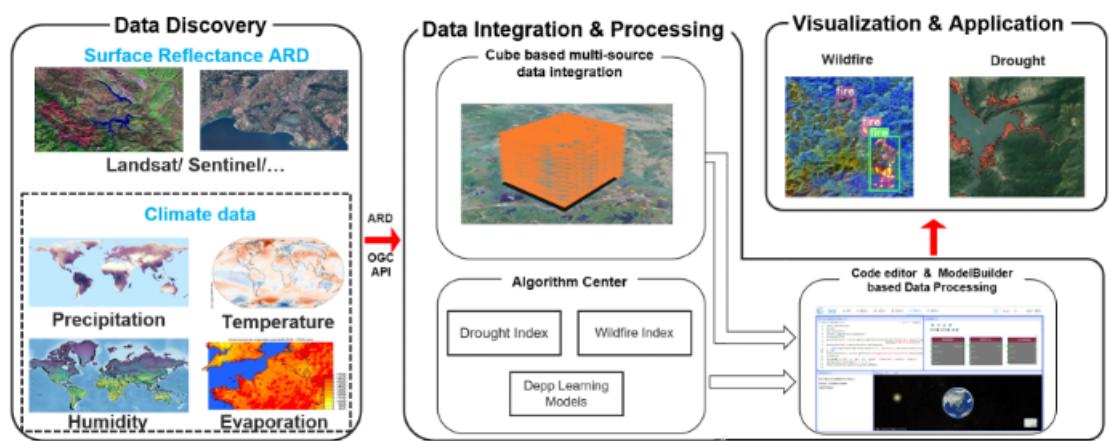


Figure 69 – The technical architecture of the use-case for drought impact.

8.1.1. Case study 1: Visualization for drought indicator

On the fundamant of SPEI and OGE, we visualize the drought risk map on a virtual globe, as given in Figure 70 (a). The color matching of the visualization result is referred to the classification standard for the drought grade of SPEI illustrated in Table 1. The red and orange area in the visualization result represents a trend of drought ($SPEI \leq -0.5$), while the green and blue represent wetness. The SPEI is calculated for each month of the input dataset, and users can visualize the SPEI of any month on the virtual globe for flexible drought analysis. Meanwhile, the use case also supports cubes-based SPEI visualization for time series drought analysis as given in Figure 70 (b), where the height of the cube is a range of time arranged in order of month, and each layer in the cube represents drought impact of one month.

GRADE	TYPE	SPEI VALUE
1	Normal	-0.5 < SPEI
2	Light drought	-1.0 < SPEI ≤ -0.5
3	Moderate drought	-1.5 < SPEI ≤ -1.0
4	Severe drought	-2.0 < SPEI ≤ -1.5
5	Extreme drought	SPEI ≤ -2.0

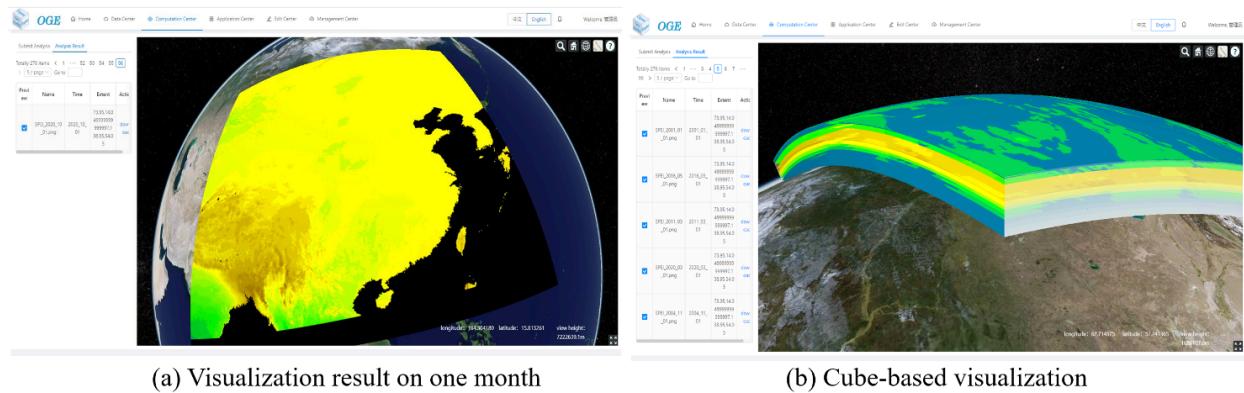


Figure 70 – Visualization of SPEI on a virtual globe.

8.1.2. Case study 2: Drought Risk analysis of Yangtze River basin

In the summer of 2022, an extreme drought hit the Yangtze River basin, posing huge impacts on agriculture, the ecosystem, and human livelihoods. It developed rapidly in the upper, middle, and lower reaches of the Yangtze River, intensifying on a large scale in 10 provinces (municipalities) in the basin (<https://doi.org/10.1002/rvr.2.23>). The water area of Poyang Lake has been reduced by 90%, threatening the habitat for fish and migratory birds, etc. To analyze drought trends in the Yangtze River Basin, we visualized monthly SPEI for 2022, as shown in Figure 71. From the Figure, it can be seen that the drought index in the Yangtze River Basin has been rising since March. In July, the drought risk map turned light yellow, indicating a moderate drought. In August and September, the drought further intensified and reached an extreme drought situation. In October, the drought eased somewhat, and it had subsided mainly by November.

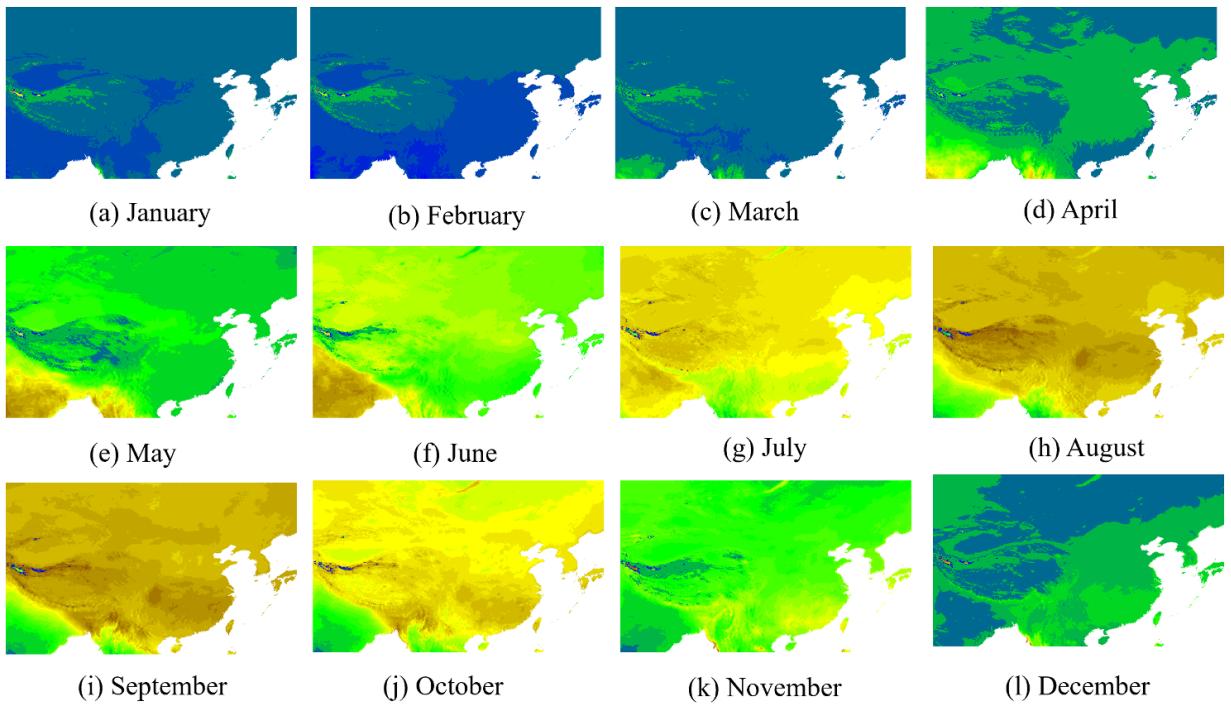


Figure 71 – Drought risk map in part of China.

8.1.3. Case study 3: Drought risk analysis of Poyang Lake

Due to the extreme drought in the Yangtze River Basin, the water inflow into Poyang Lake, the largest freshwater lake in China, declined dramatically due to continuous hot weather with little rain since early summer. Hence, we developed a use case of drought analysis applying multi-source SR ARD.

In this use case, we collected Sentinel-2 SR and Landsat-8 SR, and produced Gaofen-1 WVF SR data in the center area of Poyang Lake (As shown in Figure 72) before and during the drought period. NDWI indices were calculated to monitor water area changes in Poyang Lake. Water bodies typically exhibit positive NDWI values, making it an effortless method to extract water areas. As illustrated in [WHU_image10], the first column represents Poyang Lake before the drought, while the last three columns represent Poyang Lake which was currently experiencing the drought. It is evident from the RGB composite that the water body of Poyang Lake has significantly decreased due to the drought weather. The water body extraction results by NDWI indicate that from May to October, the water area in the study area decreased from ~1800 square kilometers to ~350 square kilometers, representing a reduction of ~ 80% in water area.

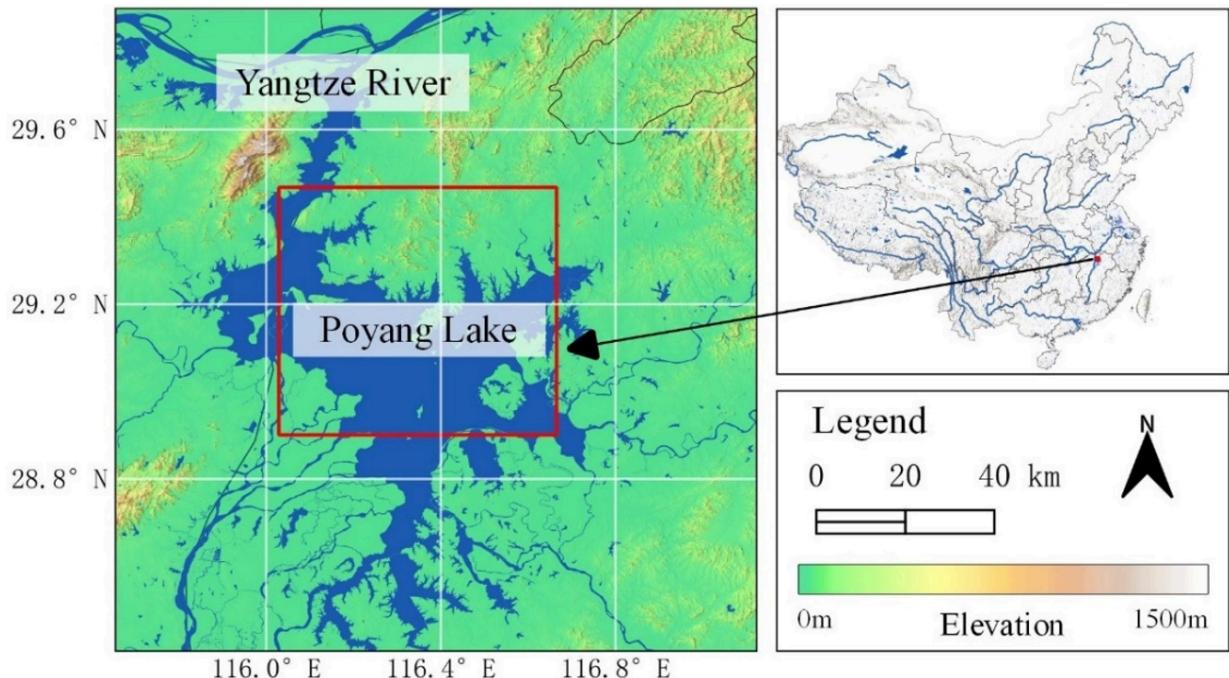


Figure 72 – The study area of the Poyang Lake case.

8.2. Analysis Ready Data (ARD) Use Case (D-100 Client instance by George Mason University)

8.2.1. Background

Definition of Analysis Ready Data (ARD) (defined by CEOS):

Analysis Ready Data (ARD) is remote sensing data and products that have been pre-processed and organized to allow immediate analysis with little additional user effort and interoperability both through time and with other datasets.

Major steps in preparing satellite data into ARD include conversion of raw reading into radiometric quantity, quality assessment, quantity normalization, and temporal integration. The ARD should follow the FAIR (Findable, Accessible, Interoperable, and Reusable) Data Principles.

Immediate analysis requires that data obtained by the data users exactly matches users' specification in the format, projection, spatial/temporal coverage and resolution, and parameters so that it can be ingested into user's analysis system immediately without further efforts. Since individual data users and projects have different requirements personalized services for customizing the data must be provided in order to meet the requirement of immediate analysis, which we call ARD services.

Essential Climate Variables (ECV) are key data sets for climate change studies. [ECV Inventory](#) houses information on Climate Data Records (CDR) provided mostly by CEOS and CGMS member agencies. The inventory is a structured repository for the characteristics of two types of GCOS ECV CDRs:

- Climate data records that exist and are accessible, including frequently updated interim CDRs
- Climate data records that are planned to be delivered.

The ECV Inventory is an open resource to explore existing and planned data records from space agency sponsored activities and provides a unique source of information on CDRs available internationally. Access links to the data are provided within the inventory, alongside details of the data's provenance, integrity and application to climate monitoring.

The client is used the existing [CEOS WGISS Community Portal](#). The portal is capable of providing automated discovery and customization services of ECV and satellite data. The client will be able to discover and access ECV and other remote sensing data and customize them into ARD for anywhere in the world to support various climate change resilience analysis.

8.2.2. Approach

The client instance is implemented as a Web application to support the creation and delivery of ARD for climate change impact assessment.

The Carbon Portal conducted data discovery and access in two steps:

- step 1: Data collection search
- step 2: Granule search to search granules in the collection

ARD services are enabled on results of granule search if the collection is an ECV. If the ECV data provider has implemented the WCS service for the dataset, the portal will directly communicate with ECV provider's WCS server for ARD service. If the ECV data provider does not have the WCS service, the portal's server will download entire granule and stage it on the portal server to provide ARD service.

Most of ECV data providers don't provide such service.

The following figure is a software architecture of the CEOS WGISS Carbon Community Portal.

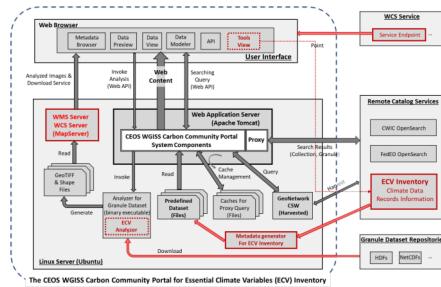


Figure 73 – Software Architecture

ECV Inventory v4.1 records are converted as a unified form of the portal predefined metadata format by a converting tool. Retrieve collection metadata for ECV entries from CWIC/FedEO OpenSearch referred by Data Record Information. There is 1251 ECV inventory records (Same as WGClimate, 870 for Existing, 381 for Planned). The portal supports totally 1910 predefined ECV relative collection datasets from ECV Records.

ARD service for ECVs in case that providers have no WCS services:

- Support when user select one granule entry
- Download granule dataset file from given repository, and manipulate it for serving WCS
- Stage the data in portal backend server and generate a list of all coverages in the granule
- User specifies the specifications of data to download
- User obtains the customized data by downloading via WCS GetCoverage request

ARD service for ECVs with data providers' WCS:

- Directly talk to provider's WCS
- Without granule downloading and stage steps in the portal's backend server.

8.2.3. Use Case: The climate change impact on crop production in Turkmenistan

The use case of the climate change impact on crop production in Turkmenistan. However, the portal can switch to another use case or support multiple use cases if this pilot requests us to do so.

Drought is one of the major climate-related natural hazards that cause significant crop production loss in Turkmenistan. Climate change increases the risk of drought in Turkmenistan. Crop models (such as WOFOST) are often used to support the decision-making in long-term adaptation and mitigation. The client will be used to prepare data to be readily used as parameters and drivers in such modeling processes. Drought impact analysis data may include long time series of precipitation, temperature, or indices for crop conditions, water content, or evapotranspiration. Many of these climate data and products from satellite sensors are served

at NASA's Goddard Earth Sciences Data and Information Services Center, such as GPM data products, MERRA assimilated climate data. These will be used in the case of drought impact assessment in Turkmenistan.

The drought impact ARD case will demonstrate:

1. Applicability of open standards and specifications in support of data discovery, data integration, data transformation, data processing, data dissemination and data visualization
2. Transparency of metadata, data quality and provenance
3. Efficiency of using ARD in modeling and analysis
4. Interoperable dissemination of ARD abiding by FAIR principles

The searching is starting with the following information:

- Keyword: surface soil moisture
- Filter: daily
- Date: 10/1/2021, 10/1/2020, 10/1/2019, 10/1/2018
- Area: Turkmenistan (Bbox: 52.264(Left), 35.129(Bottom), 66.69(Right), 42.8(Top))

Choose a collection dataset:

Groundwater and Soil Moisture Conditions from GRACE and GRACE-F0 Data Assimilation L4 7-days 0.25 x 0.25 degree Global V3.0 (GRACEDADM_CLSM025GL_7D) at GES DISC

Choose the following granule data file:

GRACEDADM_CLSM025GL_7D.3.0:GRACEDADM_CLSM025GL_7D.A20220926.030.nc4 (for year 2022)
GRACEDADM_CLSM025GL_7D.3.0:GRACEDADM_CLSM025GL_7D.A20210927.030.nc4 (for year 2021)
GRACEDADM_CLSM025GL_7D.3.0:GRACEDADM_CLSM025GL_7D.A20200928.030.nc4 (for year 2020)
GRACEDADM_CLSM025GL_7D.3.0:GRACEDADM_CLSM025GL_7D.A20190930.030.nc4 (for year 2019)

Retreve the file and choose a variable:

sfsm_inst (Surface soil moisture percentile)

Adjust legend color (0 is the least soil moisture), and get the following results:

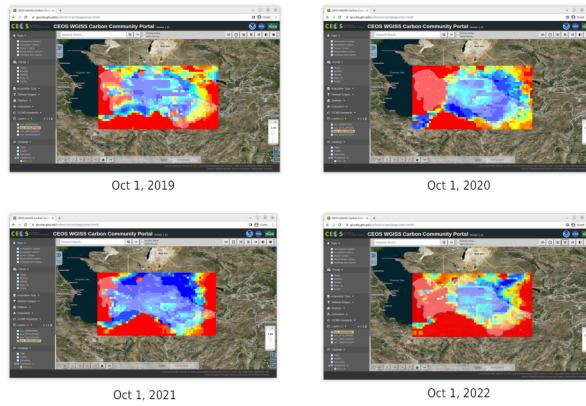


Figure 74 – Surface soil moisture percentile (year 2019-2022)

8.3. Solar climate atlas for Poland – Climate Resilience Information System

Jakub P. Walawender (Freelance climate scientist and EO/GIS expert)
email:contact@jakubwalawender.eu

The project aims at updating previously created solar climate atlas for Poland by:

- increasing spatial and temporal resolution of the datasets;
- extending time span
- replacing static maps with a dynamic and interactive interface;
- using practical solar radiation parameters instead of physical variables;
- making datasets (+ metadata) available for downloaded in interoperable file formats for further use
- sharing a solar climate knowledge base and data/service user guide

in order to:

- advance development of the solar-smart society and economy in PL
- provide know-how and tools, which are easily reusable in other geographical regions

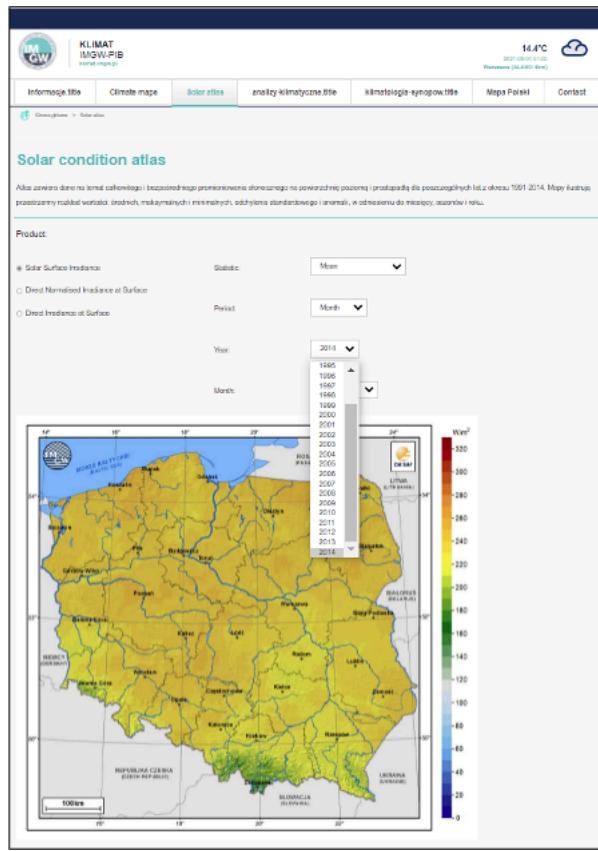


Figure 75 – Solar Climate atlas for Poland available on the IMGW website: <https://klimat.imgw.pl/en/solar-atlas>

Newly created **solar climate data cube and web map service** will be more **FAIR** as they will be made available online, possibly on the official website of the Polish Hydrometeorological Service (IMGW) for an increased findability, upon future agreement (to be discussed) to make them more **Findable** by the general public. The whole process of data access (including authentication) will be transparent and accompanied by appropriate instructions so that the **Accessibility** could be much higher. The format of the datasets in the data cube will be an OGC netCDF standard compliant with the CF (Climate and Forecast) convention, which is suitable for encoding gridded data for space/time-varying phenomena and commonly known in the climate science community but also easily readable with other common spatial data processing and visualization software including most of the GIS software to keep fully **Interoperable**. Finally, even though the proposed solar climate information system (maps+ dataset) are limited to the area of Poland, all processing scripts will be made available on github along with a well-described processing steps (both Jupyter notebooks and instructional videos will be considered) to provide **Reusability** for other countries or geographical regions.

Two objectives for the pilot OGC Climate Resilience Pilot are:

- to document existing solar radiation datasets (satellite, model and reanalysis data) and services (both freely accessible and commercial)
- to verify the accuracy of the in situ measurements and satellite climate data records for the selected solar radiation parameters using proper statistical methods

8.4. Wildfire resilience in insurance (Intact)

The main focus of IFC's participation to this project is to better understand end-to-end hazard and risk modelling workflows, in turn supporting the climate services required for decision-making in the business. This participation is also intended to further open up Intact Lab to the outside world, by exchanging information on wildfire risks and climate resiliency in the context of the insurance industry.

The project centered the efforts around these challenges:

- Identify current usages of wildfire maps at Intact by interviewing various business units;
- Revisit and update previous wildfire hazard map, using external open data sources;
- Identify and seek collaboration opportunities with pilot participants;
- Inform internal architectural, infrastructure and procurement processes of new geospatial standards and trends;
- Identify and develop insurance wildfires risk use cases to help build resilient communities.

These activities should align with the best practices and standards of the OGC and current and proposed themes in OGC's climate resilience Domain Working Group (DWG).

Wildfire risk in Canada is prominent and even though major events do not occur every year, they can cause unprecedented damage. Costs from the wildfire events of summer 2021 in British Columbia reached \$77 and \$78 millions in insured damage at White Rock Lake and Lytton, respectively [6]. Wildfire activity is expected to go up due to an increase in fire-prone conditions across the country [7].

In an insurance company, wildfire risk impacts the work of a wide array of users, such as claim adjusters, insurance brokers, engineers, data scientists, actuaries, portfolio managers, and executives. IFC's stakeholders were invited to provide information about current and potential uses of wildfire risk products within their operations. This information was used to identify use cases supporting this pilot project, as well as prospective proof-of-concepts for wildfire resiliency. It was determined that wildfires can impact numerous activities in the business, including but not limited to restoration, claims, portfolio management, CAT modelling, risk management and loss prevention. A resiliency and adaptation use case relevant to the topic of climate resilience is presented below.

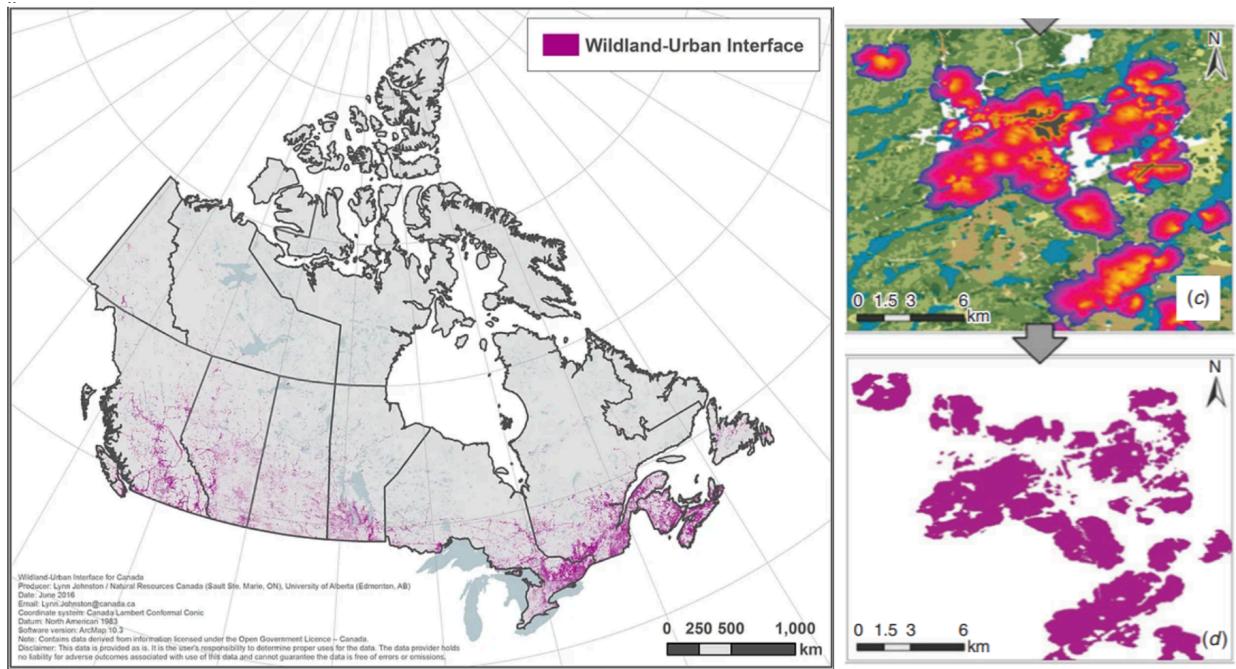
Through granting programs, Intact is investing in communities across Canada to protect people from the effects of climate change and build more resilient communities [9]. The Regional Municipality of Wood Buffalo and the community of Lac La Biche are both at an increased risk of being affected by wildfires. Their respective programs provide rebates and other incentives to residents to participate in home FireSmart assessments, and to upgrade their homes.



Figure 76 – FireSmart Canada’s Home Ignition Zones [8]

Homeowners are informed of building materials options in the immediate zone to reduce their risk of serious property damage. Residents and communities are also presented with landscaping practices for the intermediate zone, further helping reduce the risk of wildfires in the area. The Acadia First Nation’s member communities are acting in the extended zone, creating 10 to 30 meters fire breaks to increase time for emergency response in case of fire and decrease the risk of fire spread.

Ignition zones can be seen as interfaces between individual homes or structures, and the surrounding area. In the scientific literature, the area where wildland meets or mixes with human-built structures is called the Wildland-Urban Interface (WUI). As the WUI is the area that is the most at risk of wildfire, it is important to closely consider it when modelling risk. The first WUI dataset for Canada was generated in 2018, and it was identified that 3.8% of the national land area is located in the WUI [5].



**Figure 77 – Wildland-Urban Interface for Canada, on the left.
Extraction of the WUI using satellite-derived imagery, on the right. [5]**

A more comprehensive view of the WUI considers industrial areas as well public infrastructures, such as power lines and railroads. This area is called the Wildland-Human Interface (WHI) and covers 13.0% of the national land. It is estimated that within the WHI, 19.4% of the area is in zone of wildfire recurrence ≤ 250 years [4]. By the end of the century, this number could increase to 28.8% under Representative Concentration Pathway (RCP) 2.6 low emissions scenario, and to 43.3% under RCP 8.5 high emissions scenario. Integrating WUI in climate scenarios can help conduct portfolio stress testing and evaluate future risk.

As cities will keep sprawling as population increases, the WUI is also expected to grow. This is an issue since increased fire activity due to climate change is to be expected. Furthermore, this increased exposure will reach more vulnerable communities. It was shown that WUI is significantly related to socioeconomic variables such as GDP per capita, population density, road density and proportion of population above 65 years old [3].

The Canadian WUI dataset [5] is unfortunately not available for download but could be replicated with open data sources, for instance through Natural Resources Canada (NRCan) spatial infrastructures. When developing a WUI dataset, an important parameter for users to finetune is the ember transport distance. Values can vary between the median value of maximum travel distances, which is 600m (Storey et al, 2020), and the maximum travel distance of 2400m which is the official standard in the United States. Novel wildfire risk models can also dynamically adapt fuel classes within the WUI to represent propagation more accurately [10]. Producing, hosting and integrating WUI datasets can therefore support creation of better risk indices, but also help identify vulnerable areas to support further adaptation.

8.5. D-100 Client (Pelagis)

8.6. Climate Resilience for Coastal Ecosystems

The following use case(s) examine various scenarios designed to qualify the risks and pending impacts of climate change to coastal ecosystems. The scenarios are designed to leverage Analysis Ready Datasets combined with in-situ observations to draw direct relationships between a changing environment and dependent human activities.

The core of this exercise is focused on the application of OGC standards & specifications as adapters to accessing various datasets supporting key ocean and coastal climate indicators.

8.6.1. Ocean Acidification & Food Security

The ocean is responsible for upwards of 30% of the absorption of carbon dioxide from the atmosphere. As CO₂ is taken in, it combines with the water to form carbonic acid causing the pH to lower. As concentrations of CO₂ in the atmosphere continue to increase, the pH of the ocean has fallen by as much as 0.1 pH units – representing a 30% increase in ocean acidity. As acidity rises, available carbonate ions bond with excess hydrogen ions, impeding the development of calcifying organisms such as oysters and shellfish. Of critical importance is the recognition that, as ocean acidity increases, the ability of the ocean to effectively act as a carbon sink for atmospheric CO₂ is directly reduced further spiralling the future impact of anthropogenic activities and CO₂ emissions.

This use case attempts to relate the trends in changing climate variables to the ocean's ability to support the shellfish aquaculture industry along the North-East coastline of the United States. Of particular importance is the direct relationship between essential climate variables and the carrying capacity of coastal environments to support dependent socio-economic activities. Indirectly this use case attempts to identify the role of coastal ecosystems within a nature-based climate resilience strategy.

8.6.2. Background

The study combines publicly available socio-economic data with climate change indicators relevant to an area of interest off the coast of Maine USA. This area is supported through a number of observation platforms to measure ocean surface temperatures, salinity, wave heights and other important characteristics related to the ocean's state. Raw data processed to ARD provide additional metrics of the ocean's regional climate indicators.

The framework takes advantage of previous efforts made through the OGC MarineDWG implementing a 'federated marine spatial data infrastructure' (FMSDI). In this case, the framework is designed to incorporate each data source as an independent service endpoint

encoded as an OGC-compliant implementation of Feature, Coverage and/or Observation Collection. The service endpoints are developed and aligned with the OGC Features API, OGC EDR API and the OGC Observations, Measurements & Sampling (OMSv3) standards respectively. The goal being to take advantage of these standards and specifications as an adapter to the custom encoding of each raw data source allowing for a predictable semantic relationship and a loosely-coupled distributed feature schema.

This use case extends the concept of Analysis Ready Data to include processed data pipelines sourced from in-situ observation collections and sampling programs. Raw data, such as NetCDF datasets provided through the NOAA Saildrone program for monitoring ocean conditions, are processed into an 'ARD' encoded using the OGC Moving Features specification, mf-json. Extending the concept of ARD to include datasets sourced from non-satellite based observing platforms allows for a consistent view of important datasets independent of their originating platforms and associated processes and procedures. Where possible, this use case applies the OGC OMSv3 concepts of Host, Observation and Observable collections over a common spatio-temporal coverage area to reduce raw data measurements to analysis ready data.

The use case is modelled as a federated service employing a recognized schema compliant with OGC and/or external industry standards. A user query resolves each ECV to its source and combines the related feature and observation data into a 'decision ready dataset' for further exploration.

Example – Storyline: As a user, I want to see the effect of rising sea surface temperatures, salinity and other key ECVs on local aquaculture production for my area of interest.

In this use case, site information available through the Maine open data portal is used to define an area of interest. Related socio-economic variables for the area of interest and the topic are resolved against the state government's open data portal (GDP, employment metrics, etc.). The area of interest is used to refine the ARD datasets applicable to the area and associated ECV measurements across the time period of interest are processed and aggregated using a weighted average approach. The net result is an indicator of the relationship between the set of ECV measurements as a trend with milestones representing the harvest yields for each defined time period.

8.6.3. Challenges, Resolutions & Lessons Learned

- Spatial Resolution
- Temporal Resolution
- Pub/Sub Event Model
- Provenance [accuracy, reliability, peer-review, ...]
- Map, Binning and Global Grids
- Weighted relationships between observable properties and features of interest

8.6.4. Future Work

Catalog Services When combining EO observation datasets with in-situ observations and sampling programs, an inordinate amount of effort is required to find acceptable sources of ARD datasets. Although individual organizations tend to align with the ISO 19115 metadata standard for describing ARD datasets, there is limited support apart from manual efforts to discover aligned ARD datasets provided across multiple providers. Recently, the OGC announced an effort to establish the GeoDCAT working group. This effort, combined with efforts aligned with the OGC OMS SWG, would be beneficial if the goal is to address the requirement to harvest metadata across multiple providers in one ‘centralized’ service endpoint.

Temporal Resolution Typically when addressing spatial analysis, the temporal resolution of the datasets is assumed to be aligned. In the case of climate modelling and raw EO datasets, care must be taken to ensure the temporal resolution of the ARD aligns with the temporal dimension of in-situ observations, sampling programs, and real-world feature datasets.

Scalability Considering the volume of data to describe climate trends specific to an area of interest, the methodology of how raw data through to ARD is loaded into a client environment needs to be addressed. The integration framework in support of the above use case tends to instantiate local copies of raw data and ARD datasets into the compute environment for processing and analysis. The OGC GeoDatacube initiative is well positioned to play a role in addressing the scalability requirements although it’s unclear whether this approach addresses a loosely coupled, distributed data pipelines or requires local cacheing of datasets within the GDC processing workflow.

9

LESSONS LEARNED

LESSONS LEARNED

Participants of the various organizations and institutes that contribute to the Climate Resilience Pilot noted the following gaps or challenges do still exist and require additional work (in future) to overcome:

The Pixalytics Drought indicator utilizes data from sources such as the Copernicus Climate Data Store (CDS), Global Drought Observatory and NOAA Climate Environmental Data Retrieval (EDR)

As an example, we compared the input precipitation data obtained from the ERA5 dataset within the Registry of Open Data on AWS to the CDS API. It was found that accessing the data stored on Amazon Web Service (AWS) Simple Storage Service (S3) was faster once virtual Zarrs were set up. However, there are concerns regarding the data's provenance, as it was uploaded to AWS by an organization other than the original data provider. Additionally, the Zarr approach faced challenges when dealing with more recent years' data, as the NetCDFs stored on S3 had inconsistent chunking. To address this issue, a request has been submitted to enhance the Python kerchunk library's ability to handle variable chunking. We are pointing this out as it is not specific to this datasource; these challenges can happen to any large datasource that needs to transform into Zarrs to operate faster.

Also, through testing the CDI and NOAA APIs we were able to see that having an OGC API interface to datasets provided a more streamlined interface than directly accessing files as once code had been written it was easier to amend when an additional API was incorporated. Feedback was provided to ECMWF and NOAA on their API usage by Pixalytics, including collaborative discussions on potential improvements. In terms of the Pixalytics drought indicator output, QGIS modules have been identified to allow non-programmers to access and visualize the API outputs.

For ESRI's contribution, a few things learned in building CMRA version 1 and the last 6 months since its release:

- Survey responses and conversations with users of the CRMA app and the underlying data have confirmed that providing open and usable data is equally important to providing a well designed web application. Both are needed, for different audiences.
- There will always be groups that want their own application with their branding, even if its only slightly different than original.
- Consider meaningful units of geography to summarize by based on the intended users/end-use. The polygons needs to be suitable size for the decisions being made, but not so small as to challenge the spatial resolution of the climate projection data.
- When processing climate data into means, it's critical to capture the extremes in space, time, and members. For version 2 we will compute the min, max, and mean of each variable and scenario combination.
- It is important to communicate what variables are useful for what applications/considerations. E.g., why is 40 deg C a critical threshold? We are building this to

expand access to a potentially less scientific and less climate aware audience and more explanation will be useful.

- It is important to have a holistic integrated approach to data, tools, workflow documentation, and communities of practice.
- As we expand to a global system, issues of local downscaling and local weighting of models comes into question. Downscaled models do not exist in most of the world at a fine enough resolution to responsibly answer the type of questions users want to ask.

For IFC, the following lessons learned are highlighted:

- Definition of common documentation guidelines for terms and licences of endpoints would facilitate approval in cybersecurity-heavy environments. Individual security reviews for each participants and their endpoints is a tedious process. Having common terms as defined in OGC's master agreement, alongside applicable open source licences, would facilitate endpoint whitelisting.
- In insurance companies, catastrophe (CAT) modeling is often looking at past events to establish probabilistic indices and maps [11]. Increasingly, climate projections are making their way into the CAT models. There is an opportunity for the OGC to advance standards in how these models are conceived, packaged, distributed and executed.

The pilot project aims to achieve multiple objectives, one of which is to reduce the obstacles that users face when accessing CDS/ADS (Atmospheric Data Store) data and services. By identifying these barriers or gaps from the users' perspective, the pilot can adapt and evolve accordingly. This approach ensures that the project engages a broader user community and facilitates their interaction with CDS/ADS.

To provide a clear direction for developers and users, the pilot intends to establish a universal and well-defined climate service workflow. This workflow will serve as a roadmap, guiding individuals through the entire process from raw data to actionable information. By offering a structured framework, the project aims to enhance efficiency and streamline the utilization of climate services.

Several enhancements were planned for the project, including improvements to the performance of the Sentinel-2 data cube. Climate data can be incorporated in this, and vegetation fuel type classification as well. This to support a wildfire risk assessment workflow. These enhancements contribute to expanding the capabilities and functionalities of the pilot project.

In regards to Analysis ready data, ARD principles can be applied to climate time series, not just earth observation (EO). Good ARD should be useful for a range of scenarios and useful to answer a range of analytic questions. ARD usually involves some degree of filtering, simplification and data aggregation without losing the essential information necessary to support decision making.

During the DP21 phase, a solid foundation was established for exploring data cube extraction and conversion to ARD using the FME data integration platform. In this pilot, a number of new approaches were explored for tasks such as data extraction, simplification, and transformation. Additionally, different methods were investigated for selecting, splitting, aggregating, and

summarizing time series. The primary objective was to generate ARD capable of answering questions related to climate trends and readily consumable by GIS and other geospatial applications.

The initial ARD approach to derive temperature and precipitation contours or polygons inherited from the work in DP21 on flood contours involved too much data simplification to be useful. Classification into temperature or participation bands resulted in an effective loss of detail, oversimplifying the data to the point where it no longer held enough variation over local areas to be useful. In discussion with other participants, it was determined that converting multidimensional data cubes to vector time series point data served the purpose of simplifying the data structure for ease of access, but retained the environmental variable precision needed to support a wider range of data interpretations for indicator derivation. It also meant that as a data provider we did not need to anticipate or encode interpretation of indicator business rules into our data simplification process. The end user is free to run queries to find locations and time steps for specific temperatures or precipitation ranges of interest.

Initially it was thought that classification rules need to more closely model impacts of interest. For example, the business rules for a heat wave might use a temperature range and stat type as part of the classification process before conversion to vector. However, this imposes the burden of domain knowledge on the data provider rather than on the climate service end user who is much more likely to understand the domain they wish to apply the data to and how best to interpret it.

In the absence of more sophisticated models, looking at the delta between future forecast and historical averages served as an interesting experiment for highlighting potential climate change impact hotspots. Past and future were differenced both spatially and temporally for equivalent time steps (monthly). These deltas may serve as a useful starting point for climate change risk indicator development. They also can serve as an approach for normalizing climate impacts when the absolute units are not the main focus. This may give local planners and managers more options to explore and analyze local areas and times of concern related to climate model scenario outputs.

More analysis needs to be done with higher resolution time steps – weekly and daily. At the outset monthly time steps were used to make it easier to prototype workflows. Daily time step computations will take significantly more processing time. Future pilots should explore ways of better supporting scalability of processing through automation and cloud computing approaches such as the use of cloud native formats (STAC, COG, ZARR etc).

Environmental climate variables (ECVs) have traditionally been discussed in the context of earth observation (EO) data. For the purposes of this and other OGC pilots, it also seems that ECVs could just as easily relate to the environmental variables stored in climate model outputs such as data cubes. Both store ECVs, its just that traditional EO ECVs relate to past or present observations while climate model ECVs relate to future potential or forecast ECV values. Either way, having a standardized understanding of what is meant by ECVs may go some way towards developing a better understanding of ARD in relation to climate change impact management.

Further experimentation is required to enhance the project's capabilities. This experimentation encompasses various aspects, including analytic techniques, statistical methods, simplification processes, and publication methodologies. Additionally, the project aims to explore cloud-native approaches such as NetCDF to COG conversion and the utilization of APIs. These ongoing

experiments contribute to refining the project's methodologies and expanding its range of applications.

Currently, the participants have implemented the first Drought Index (SPI) using precipitation data from the Copernicus Climate Data Store (CDS). However, they are open to incorporating additional data sources as per the project's requirements. This flexibility ensures that the pilot project remains adaptable to evolving needs and can utilize diverse datasets to enhance its outputs.

In summary, the pilot project seeks to overcome barriers and engage a wider user community by facilitating access to CDS/ADS data and services. A well-defined climate service workflow will guide developers and users through the entire process, ensuring efficiency and effectiveness. Enhancements to the Sentinel-2 data cube, the inclusion of climate data and vegetation fuel type classification, and the development of a wildfire risk assessment workflow will expand the project's capabilities. By applying ARD principles and refining classification rules, the project aims to generate valuable insights into climate trends. Ongoing experimentation and the exploration of different methods contribute to the project's continuous improvement.

10

FUTURE WORK

FUTURE WORK

Being the first OGC Climate Pilot, there has been significant underpinning work on the component elements that has supported an improved understanding of what is currently possible and what needs to be developed. Future pilots will focus on supporting the filling-in of identified gaps and definition of best practices guidelines to support and enable broader international partnerships.

During the pilot, participants agreed to the following items were specific actions where future work would be needed:

- Further integration of the contributor components so that full workflows, from raw data to visualization and communication, can be tested.
- Exploring additional scenario tests including comparisons with historical norms, e.g. calculating the difference between historical maximum temperatures and projected maximum temperatures.
- More analysis with higher resolution time steps — weekly and daily. At the outset monthly time steps have been used to make it easier to prototype workflows. Daily time step computations will take significantly more processing time.

For ESRI's contribution, the first version of CRMA was well received, it is widely used by the intended users and there is high interest by many others. Before the first version was released we had requests for other countries, and customizations of the project.

Due to many customization requests version 2 is being developed from inception with the intent for all code, from data processing python to web application Javascript to be available in Github repositories with documentation of typical customization workflows.

- Use other climate projection data
- Compute other indices
- Summarize to other geographies
- Customize the web application

The project is not only a solution, but a pattern for others to adapt to their data, geography, goals.

Version 2 data development is underway and will include more indices, both imperial and metric units, and min/max/mean for statistics instead of only areal mean. We will update all modeling to CMIP6, and expand from US to global. Anticipated release in Q4 2023.



A

ANNEX A (INFORMATIVE) REVISION HISTORY

A

ANNEX A (INFORMATIVE) REVISION HISTORY

DATE	RELEASE	AUTHOR	PRIMARY CLAUSES MODIFIED	DESCRIPTION
2023-03-28	0.1	G. Schumann; A.J. Kettner	all	First draft of ER
2023-03-29	0.2	Nils Hempelmann	adapt to new ER schema	revision draft of ER



BIBLIOGRAPHY



BIBLIOGRAPHY

- [1] Ben Domenico: OGC 10-092r3, *NetCDF Binary Encoding Extension Standard: NetCDF Classic and 64-bit Offset Format*. Open Geospatial Consortium (2011). https://portal.ogc.org/files/?artifact_id=43734.
- [2] Akinori Asahara, Ryosuke Shibasaki, Nobuhiro Ishimaru, David Burggraf: OGC 14-084r2, *OGC® Moving Features Encoding Extension: Simple Comma Separated Values (CSV)*. Open Geospatial Consortium (2015). <https://docs.ogc.org/is/14-084r2/14-084r2.html>.
- [3] Akinori Asahara, Ryosuke Shibasaki, Nobuhiro Ishimaru, David Burggraf: OGC 14-083r2, *OGC® Moving Features Encoding Part I: XML Core*. Open Geospatial Consortium (2015). <https://docs.ogc.org/is/14-083r2/14-083r2.html>.
- [4] OGC: OGC 11-165r2: CF-netCDF3 Data Model Extension standard, 2012
- [5] Standardized Big Data Processing in Hybrid Clouds. In: Proceedings of the 4th International Conference on Geographical Information Systems Theory, Applications and Management – Volume 1: GISTAM, pp. 205–210. SciTePress (2018).
- [6] Sepulcre-Canto, G., Horizon, S., Singleton, A., Carrao, H. and Vogt, J. Development of a Combined Drought Indicator to detect agricultural drought in Europe. Nat. Hazards Earth Syst. Sci., 12, pp. 3519–3531. (2012). doi:10.5194/nhess-12-3519-2012
- [7] Cammalleri C, Micale F, Vogt J. A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency. Hydrological Processes, 30(2), pp. 289–301. JRC96439. (2016). <https://hess.copernicus.org/articles/21/6329/2017/>
- [8] Lawrence Livermore National Laboratory: NetCDF CF Metadata Conventions – <http://cfconventions.org/>
- [9] ESIP: Attribute Convention for Data Discovery (ACDD) – <http://wiki.esipfed.org/index.php/>