



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

A Digital Twin for Climate Change Adaptation

F.J. Doblas-Reyes representing the DE_340 group

31 May 2023



Funded by
the European Union

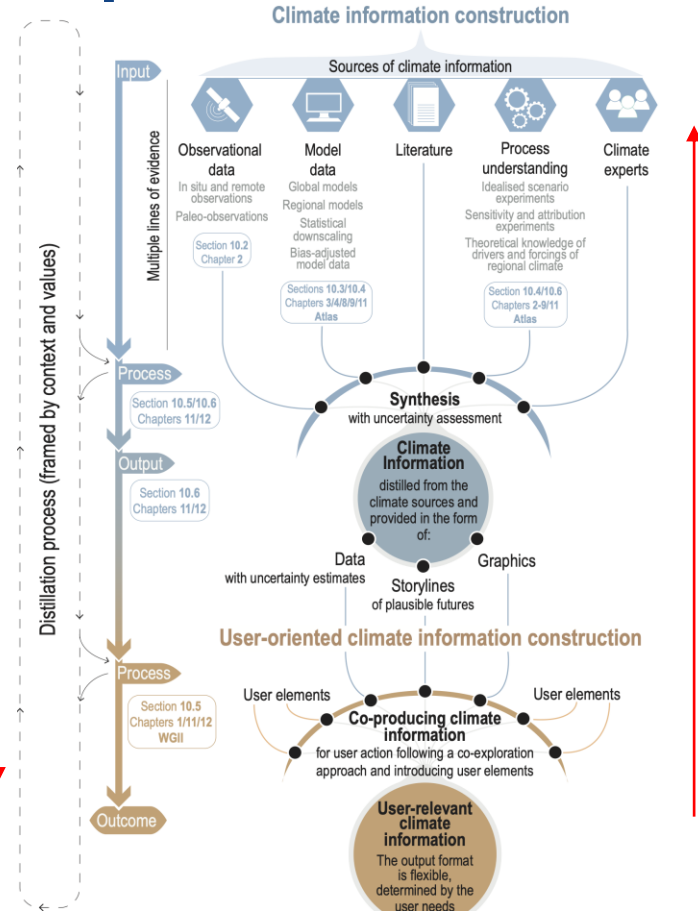
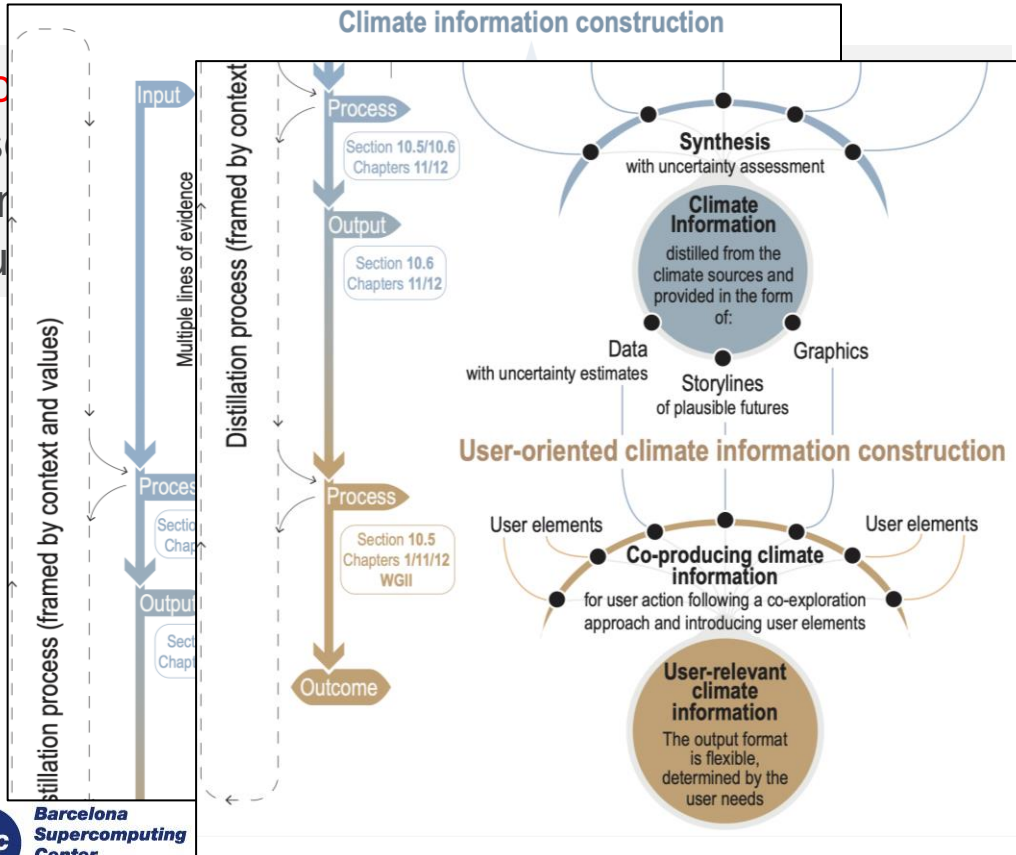
Destination Earth

implemented by



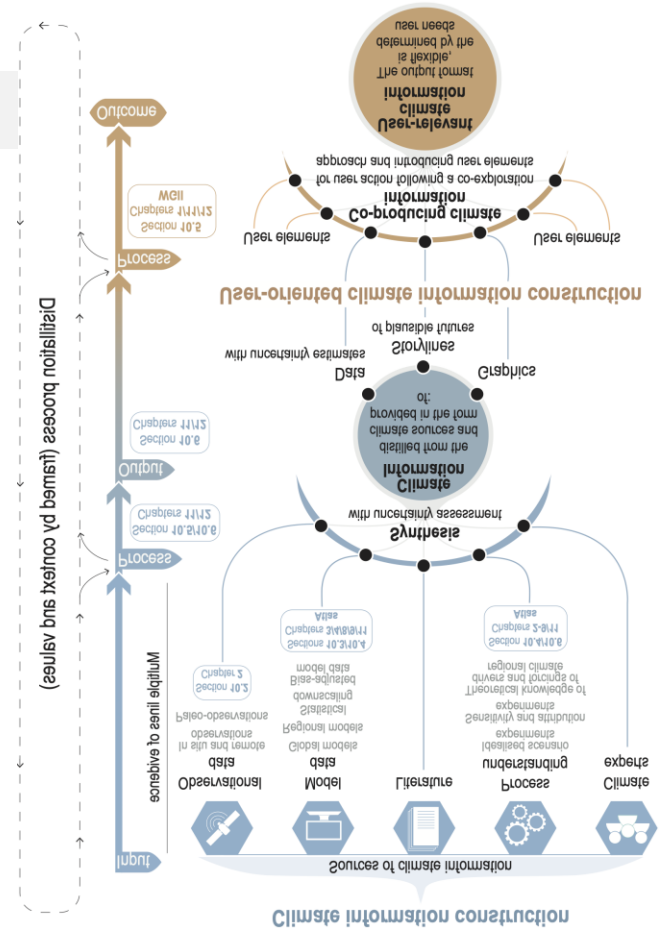
Climate information for adaptation

Adapt
cons
amor
about



Climate information for adaptation

What happens if the **user** is at the top of the process?

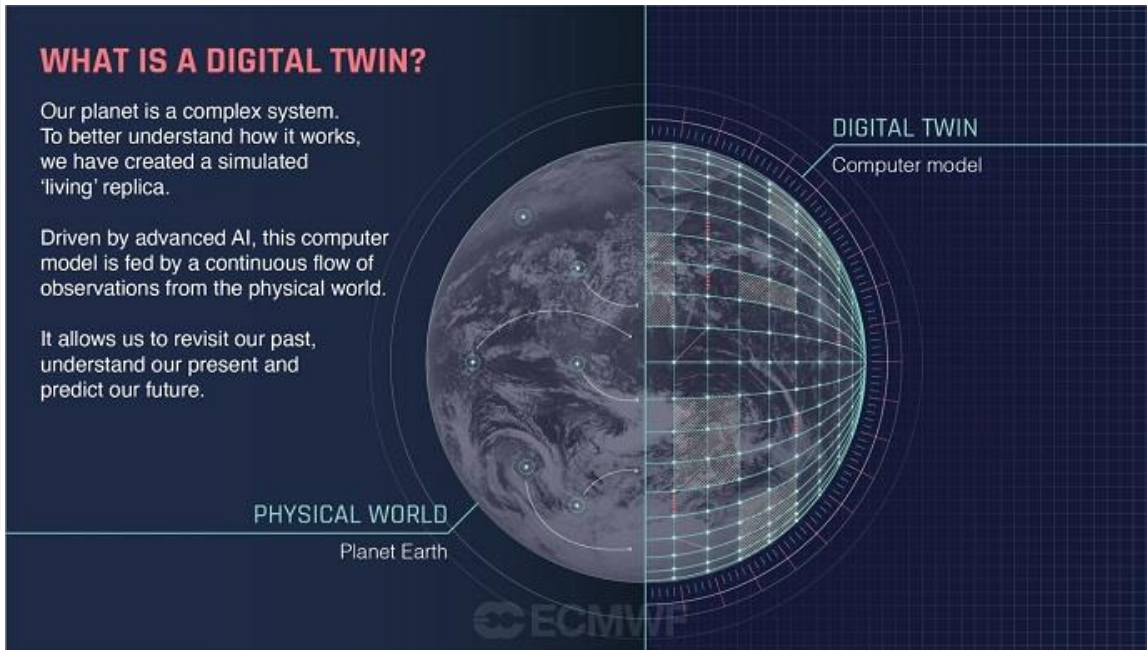


Digital twins and climate information

A **digital twin for climate adaptation** is a system that supports decision-making in adaptation to climate change using the best models available in an environment that allows **an interactive relation with the user**.

The digital twin requires (non-exhaustive list):

- a strategy to **collect user requirements**
- a well-validated set of interoperable **models**
- an operational **environment** (software and hardware)
- a **workflow** strategy
- a suitable **interface**



WHAT IS A DIGITAL TWIN?


Our planet is a complex system. To better understand how it works, we have created a simulated 'living' replica.

Driven by advanced AI, this computer model is fed by a continuous flow of observations from the physical world.

It allows us to revisit our past, understand our present and predict our future.

PHYSICAL WORLD
Planet Earth

DIGITAL TWIN
Computer model



Why a digital twin for climate adaptation?

The digital twin emerges in a busy context, with many requirements for climate information, a cacophony of sources, a growing market, increasing needs, no defined standards, and some well-positioned actors. Are the needs taken care of? What about the adequacy of the information sources? Are timing, quality and authority addressed?

Services

Climate and Environmental Risk

Lead the integration into your risk management

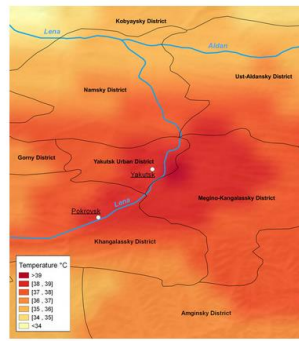
Climate risks are materially impacting organisations and play a key role in the risk management strategy. Integrating these climate risks into your risk management can be challenging. Deloitte can help your organisation with the identification, assessment and management of climate risks, including climate risk modelling, climate scenario analysis, and setting up a climate risk data infrastructure and climate risk dashboard.

Supporting UN Habitat – CRPT for climate analysis in cities

Urban areas are strongly exposed to the effects of climate change. Local authorities monitor these effects very closely, so that they can adapt in a timely manner and reduce their impact on citizens' lives: health care, transport, water management, energy supply, etc.

UN Habitat's **City Resilience Profiling Programme (CRPT)** provides national and local governments with tools for measuring and increasing resilience to multi-hazard impacts, including those associated with climate change.

Lobelia's climate engine is integrated in the CRPT, helping analyse climate trends and link to other types of data as required to allow the definition of a suitable action plan. Lobelia's toolbox can be used by cities worldwide, with current users including Asunción, Yakutsk,



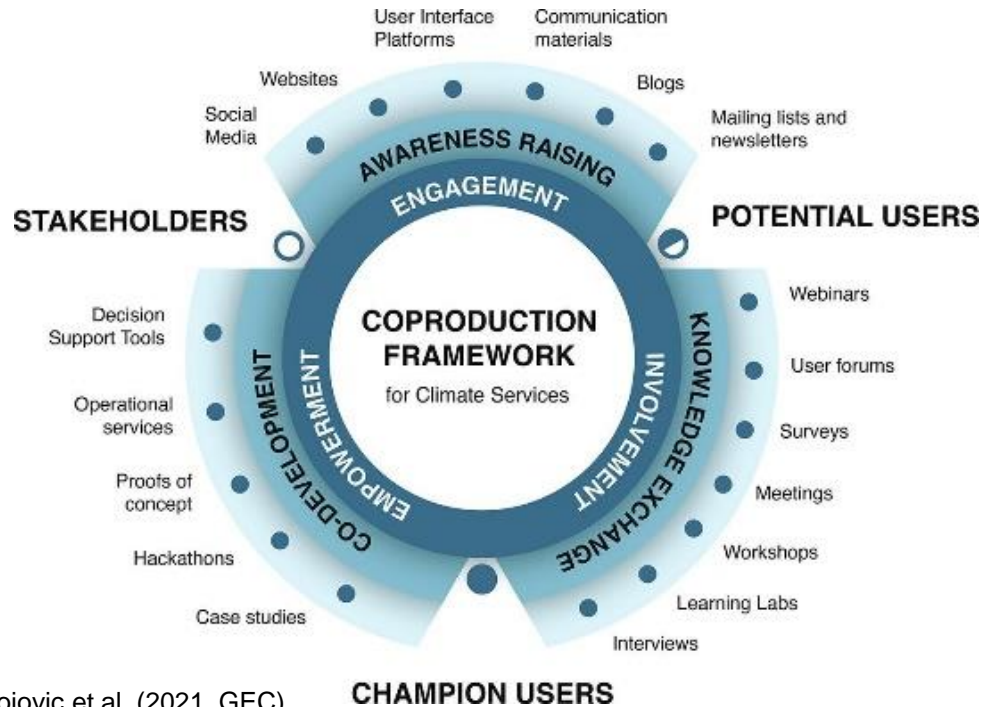
Maximum air temperature in Yakutsk, Russia (average) | 2014-2070 | RCP 4.5

European Climate Risk Assessment



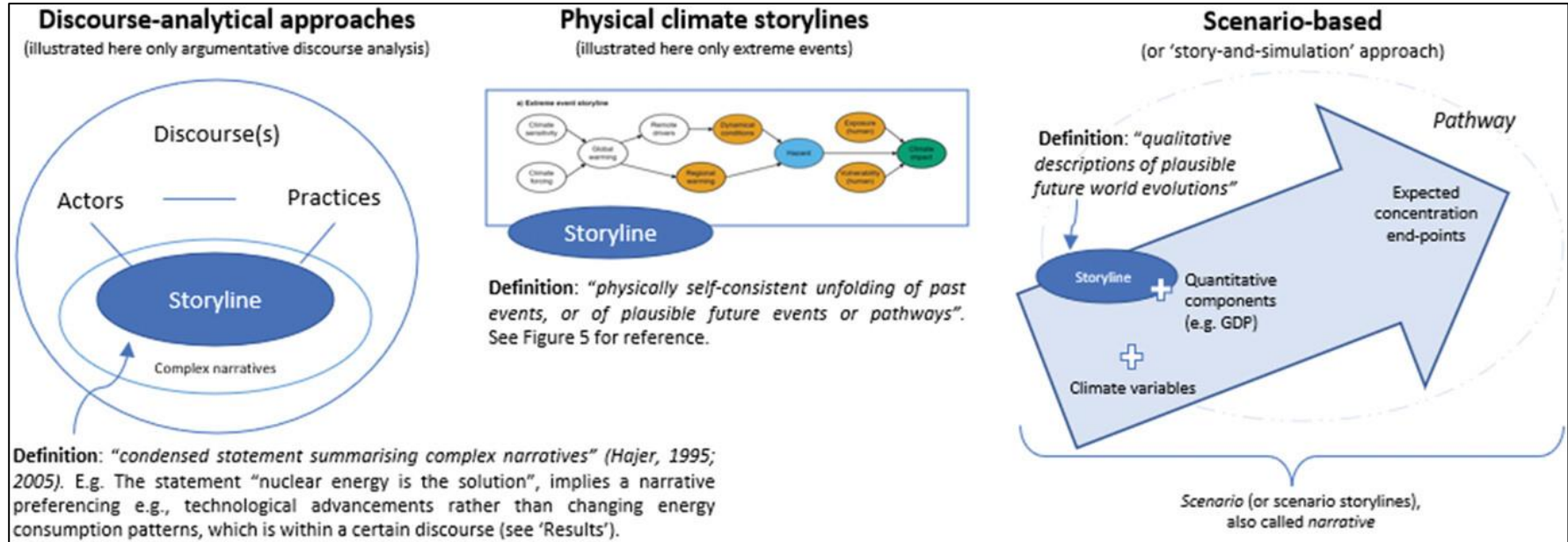
1) Start from the demand: user interaction

Social sciences and humanities play an increasingly important role in the services that provide climate information. New and varied approaches are leading to more efficient and successful links to both public administrations and the private sector.



1) User interaction: storylines

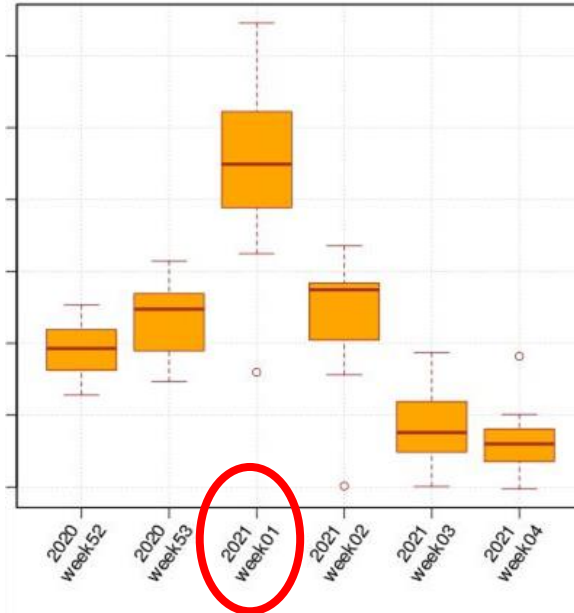
Storylines are an efficient tool to engage with the hugely varied adaptation community. It goes beyond maps and probabilities and aims at putting all actors at the same level.



1) User interaction: start from the question

In a particular case, a known retailer needs to know the impact of some specific extreme climate event in the sales of winter and mountain product. These sales are **sensitive to the combination** of snowfall occurrence, rain after snow, maximum temperature, soil conditions, etc., **plus many other non-climatic drivers**.

Turnover progression in winter and mountain sports

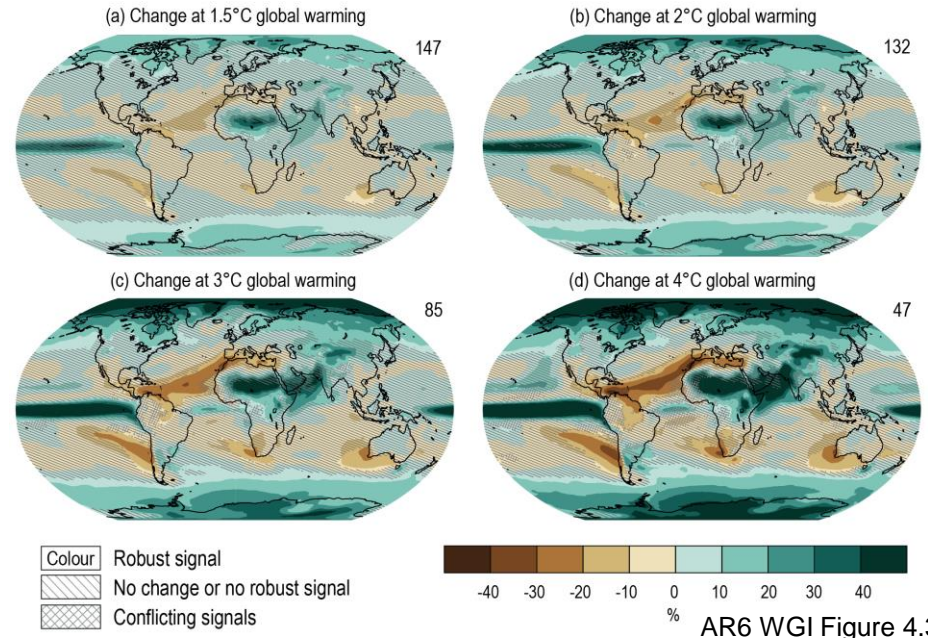
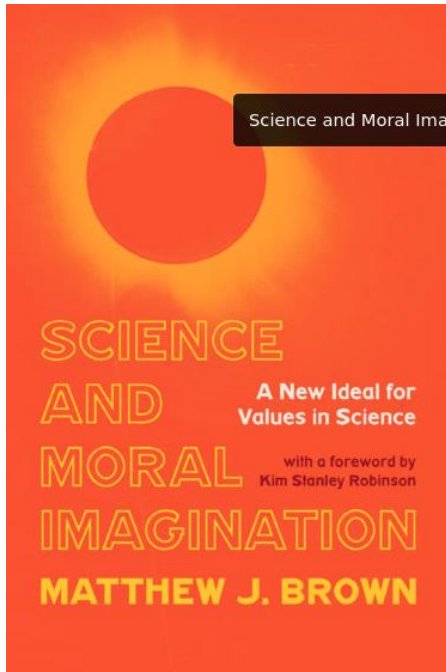


2) Observations and well-validated climate models



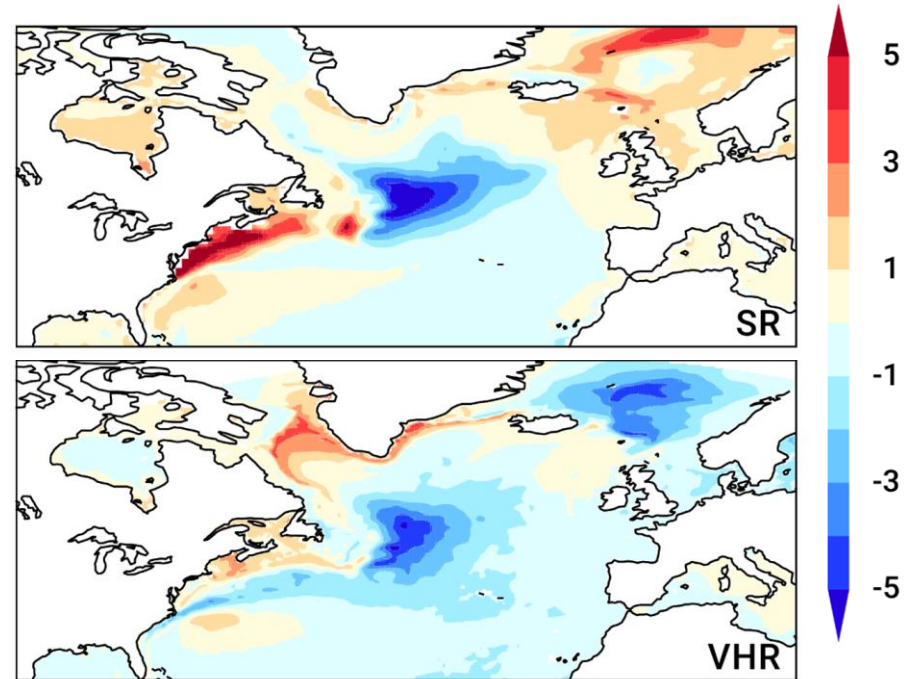
2) Current climate data sources have problems

While better schemes to interact with the users and produce **salient information** are indispensable, reliance on models that are not good enough leads to either **overconfidence** or **underconfidence**, which, in turn, leads to both **inadequate uncertainty estimation** and **insufficiently credible risk assessments**.



2) New generation of climate models: resolution

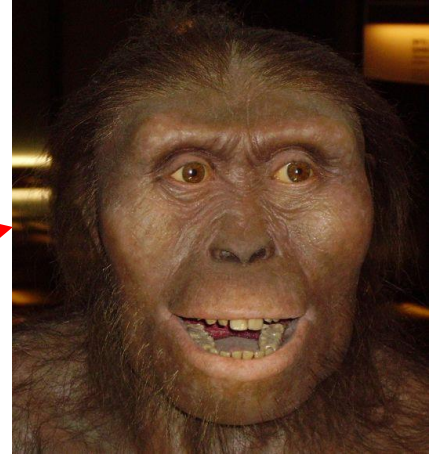
Eddy-rich, storm-resolving models (10-km and higher resolution) simulate, among other things, a decrease in SST biases over the North Atlantic with respect to traditional, standard-resolution models.



2) New generation of climate models: resolution



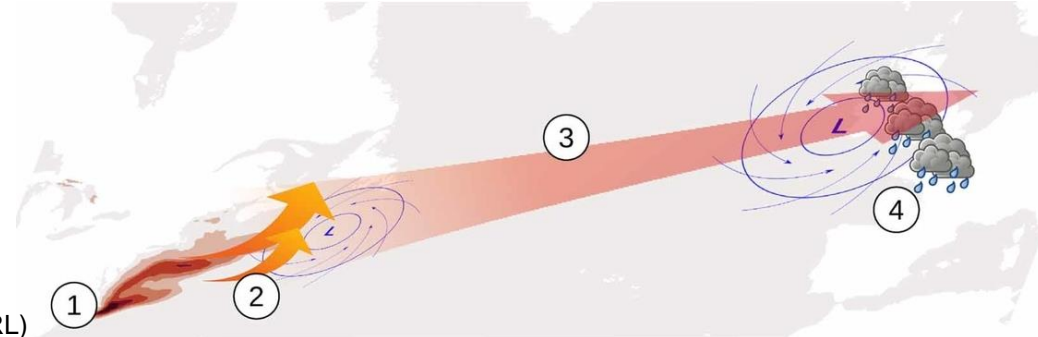
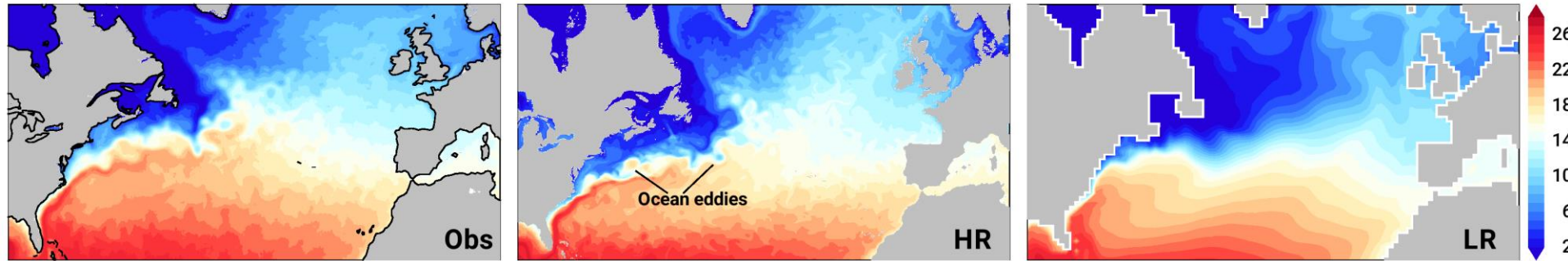
Australopithecus Lucy
([National Geographic](#))



Neanderthal ([National Geographic](#))

2) New generation of climate models: resolution

Eddy-rich, storm-resolving models (10-km and higher resolution) simulate a decrease in SST biases over the North Atlantic and a northward shift in the Gulf Stream over the XXI Century that leads to rainfall increases in Europe, not found with traditional low-resolution models.



3) The operational environment: pre-exascale HPC

The EuroHPC programme offers computing time (5%) to DestinE on the new pre-exascale supercomputers, MareNostrum (left) and LUMI (right). Dedicated storage and data processing platforms will be available with dedicated bandwidth.



Plus a lot of cloud-based, scalable solutions for data processing, and adequate tiered storage.

3) The operational environment: portability and scaling

New hardware, pre-exascale computing oriented



CPU

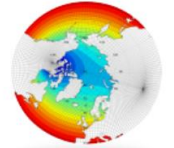


GPU

Global kilometre-scale resolution for better accuracy



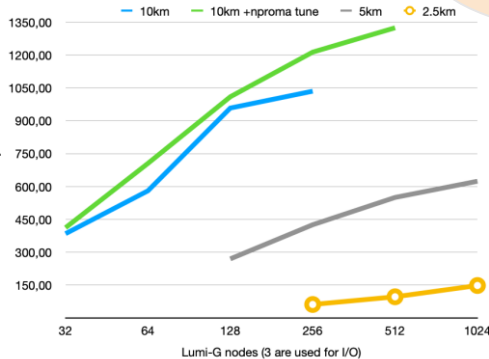
ORCA 2
550 MB of memory
8 CPU hours
10 Gigabytes of output (daily)



ORCA 1/12
414 Gigabytes of memory
90 000 CPU hours
1 Terabyte of output (daily)

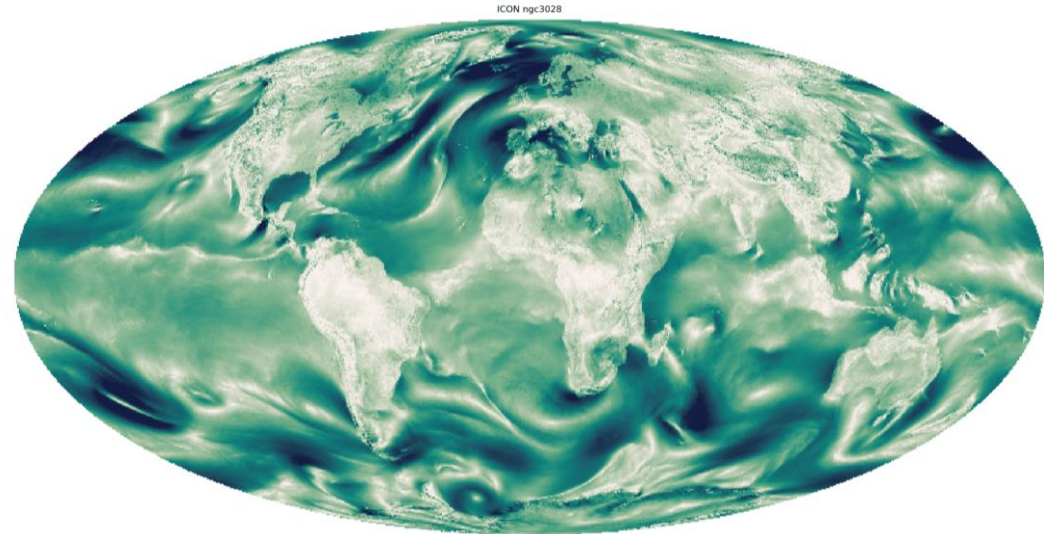
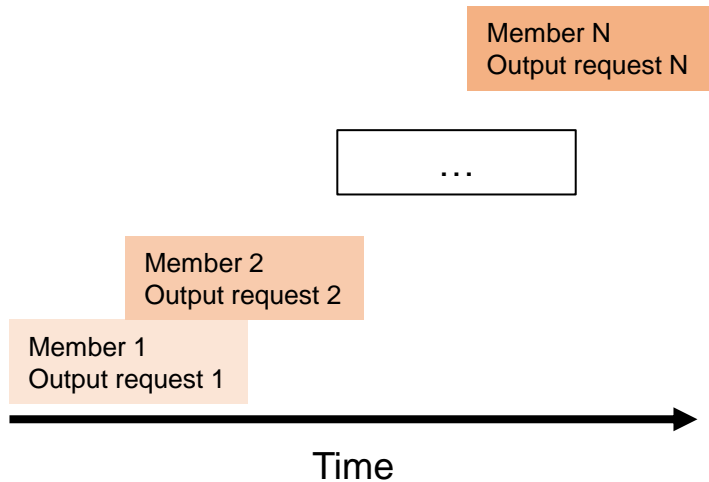
Actions for improving the computational efficiency and optimization of all tasks, including the new climate models

- Ensure that the machines are used efficiently
- Develop tests with measurable indicators of replicability, computational efficiency and code quality
- Accelerate the new models using kernels and data structures suitable for the new GPU partitions

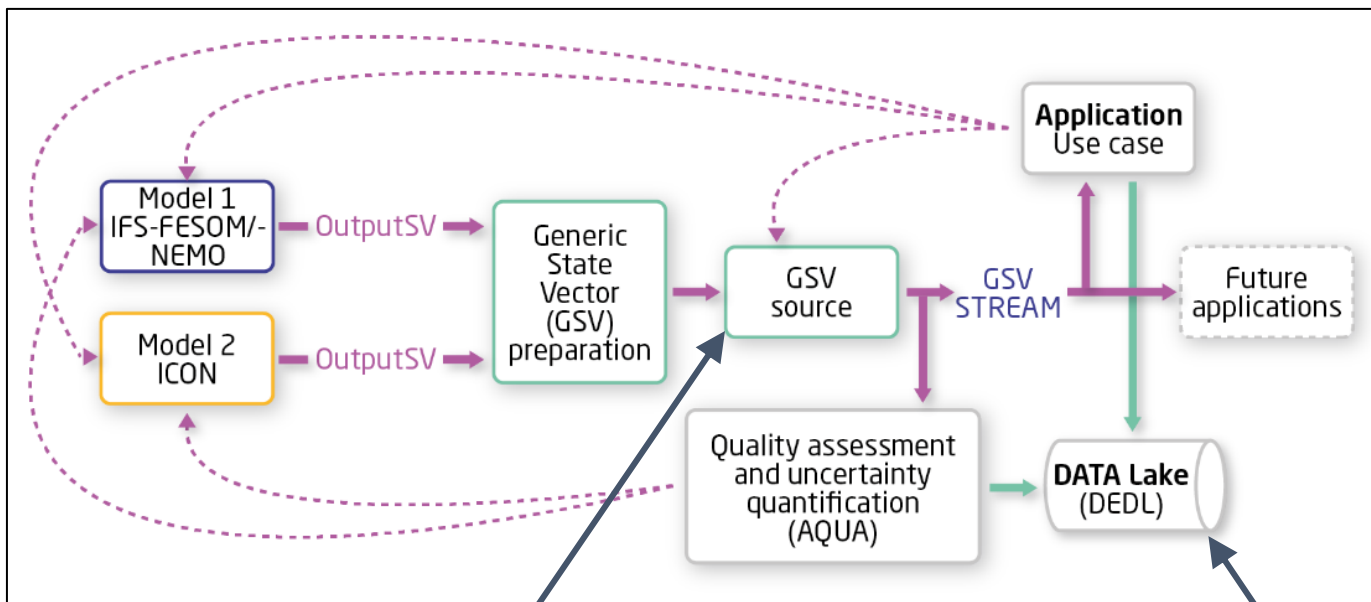


4) The workflow: climate simulations

Historical simulations from 1990 to 2020 with short spin ups, and multi-member climate projections up to 2040. Plus AMIP simulations and experiments for storylines. The members will be performed concurrently, with continuous data checks, as part of a workflow as close as possible to an operational set up.



4) The workflow: climate models and data consumers



Volatile (lasts days), common grid, native resolution, high frequency, all variables

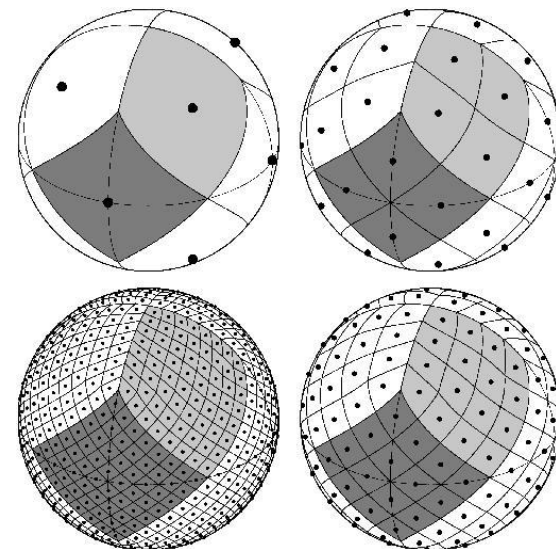
Permanent, lossy compression, interpolated

4) The workflow: GSV

The **generalised state vector** (GSV) is data repository for data in GRIB2 format interpolated in a common grid (HealPIX, the closest to the native resolution) located in the FDB. The GSV is accessible through the GSV API (with MARS-like syntax).

The challenges for the GSV are

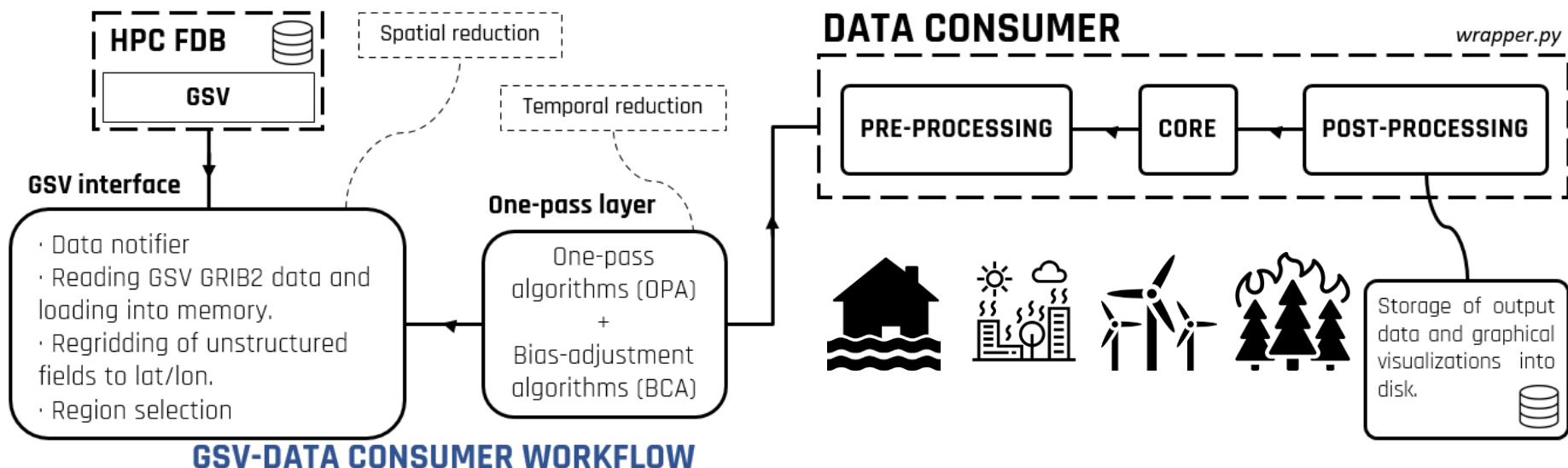
- The need to get official WMO recognition for its description.
- Technical flexibility to accommodate new variables.
- A flexible metadata scheme to add new variables and experiments that allows existing tools to interact with the data.
- Efficiency: regridding, data volume, concurrent access, etc.



Source: <https://healpix.sourceforge.io/>

4) The workflow: set up and data flow

The **streaming** concept is key in the digital twin because it gives users access to the full model state vector (standardised) for a limited time to generate unprecedented diagnostics.

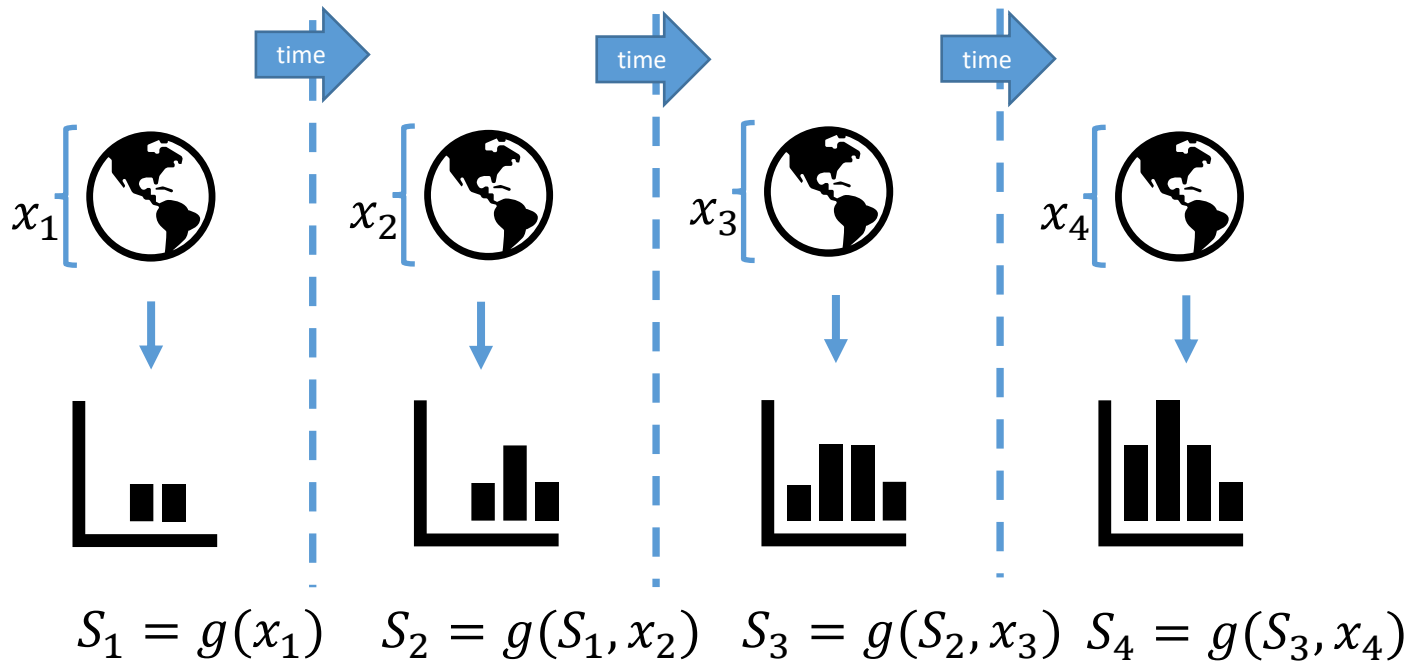


4) The workflow: streaming and one-pass algorithms

The **streaming** concept is key in the digital twin because it gives users access to the full model state vector at high frequency and the native resolution (km).

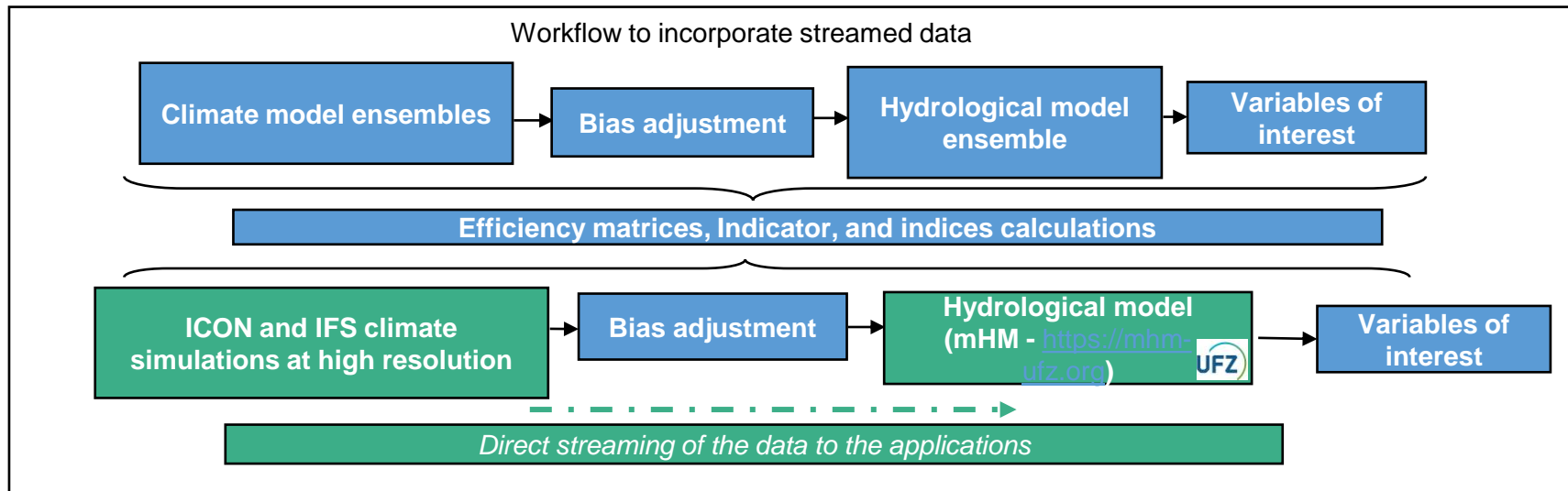
The **one-pass algorithms** are data-reduction functions that can compute (continuously updated) summary statistics or post-processed variables from the streamed data.

To compute e.g., statistics from high-frequency (hourly) data (like ocean features under a moving tropical cyclone) or online bias-adjusted data



4) The workflow: streaming and applications

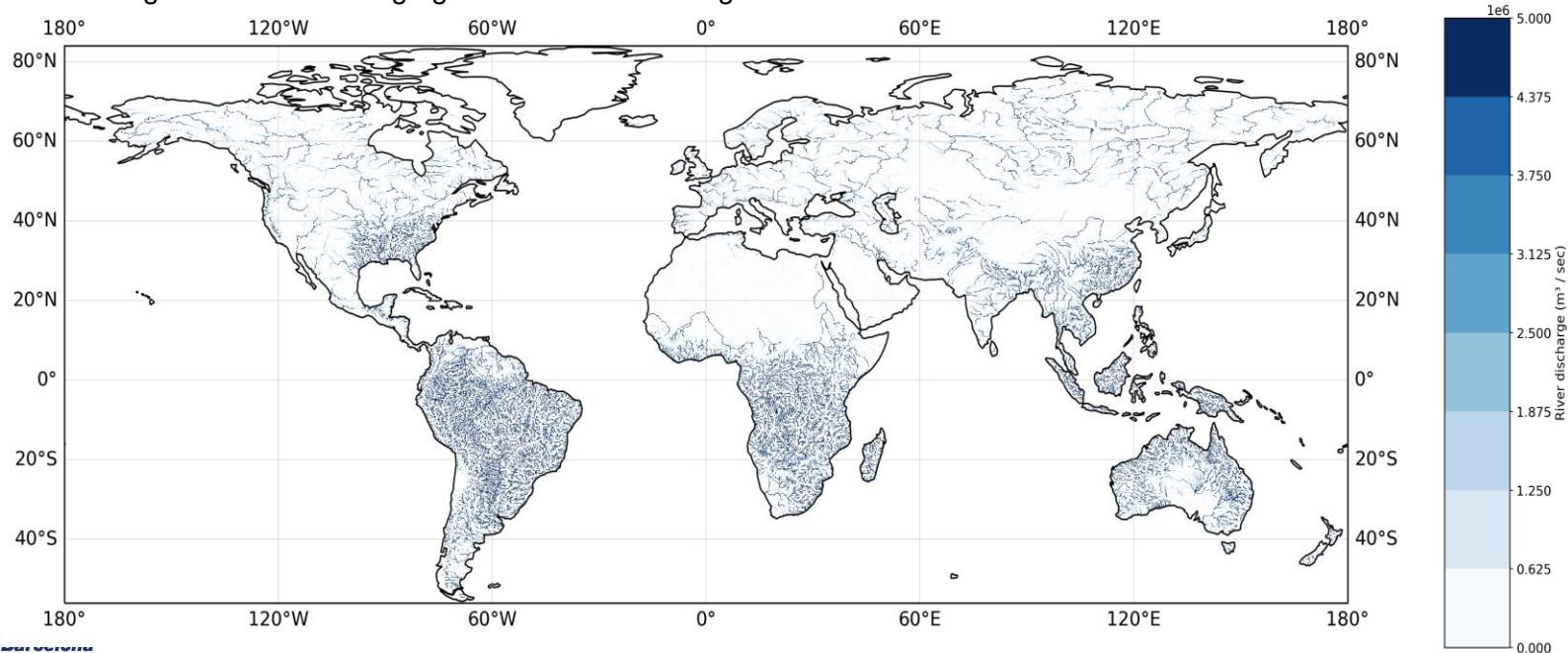
Several prototype applications are working with the NextGEMS data to illustrate the concepts until the DE_340 models start delivering data.



4) The workflow: streaming and applications

Several prototype applications are working with the NextGEMS data to illustrate the concepts until the DE_340 models start delivering data.

High-resolution average global streamflow using ICON climate simulations from the NextGEMS dataset



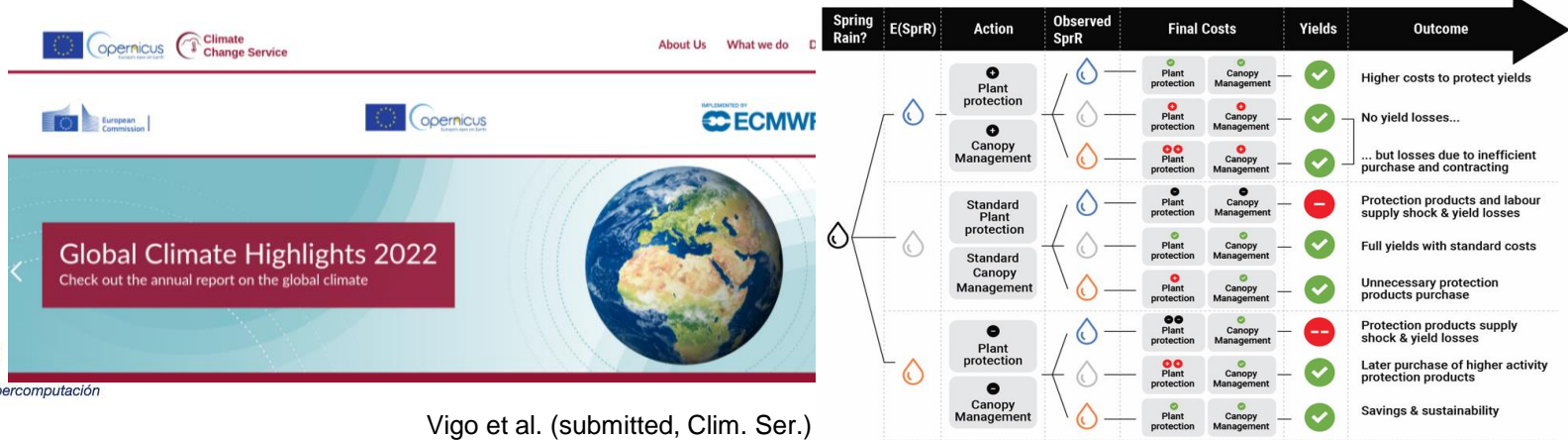
5) The interface

The interface should address the user queries and provide answers that are contextualised, qualified, on time, and traceable. It can be a

- sophisticated but traditional portal (C3S)
- trained model that engages in a conversation (ChatGPT-like)

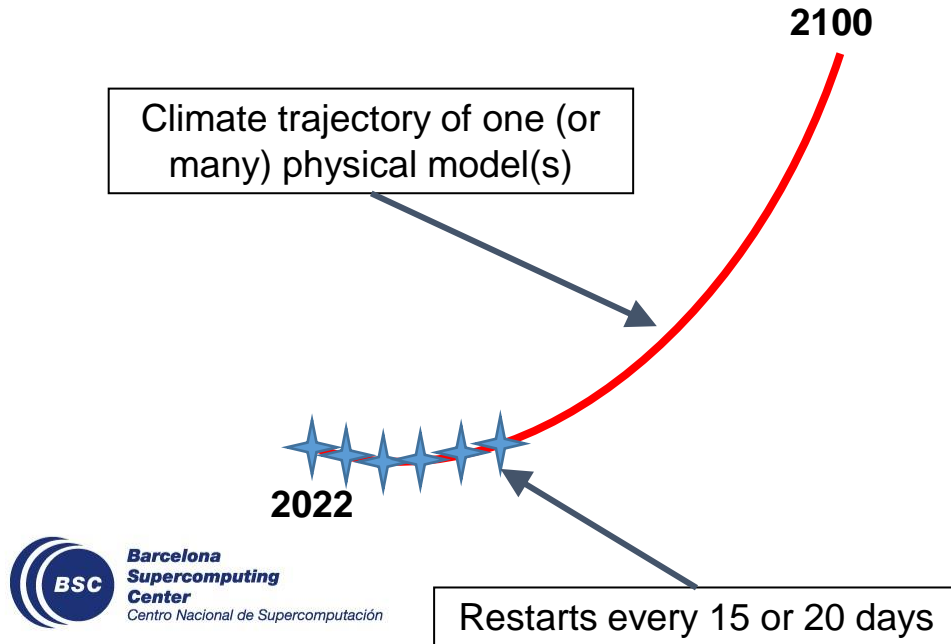
But, at a minimum, the interface should include

- interactive, targeted and traceable products according to service standards
- training (webinars, workshops, MOOCs) and guidance
- quality assurance and user-oriented impact assessment



A necessary complementary role for AI: emulation

An ML-based weather forecast model could be trained on the trajectory of a (preferably several) process-based climate model (with high-frequency output). The model restarts can be used to recreate almost instantly (in inference mode) the model output requested by any user leading to real-time **interactivity**.



The ML-based emulator can provide **downscaled** (superresolution), **targeted** (user indicators), and **bias adjusted** data with **large ensembles** (if uncertainty handling is trustworthy).

It is debatable if it will be able to simulate the impact of emission scenarios not yet explored by the process-based climate model.

A necessary complementary role for AI: interface

The temptation is to consider a solution ChatGPT-like. But ChatGPT is as good as the information it uses to train its model. Domain-oriented solutions are being developed.

F what is the probability of coastal flooding in Barcelona and other cities in the Mediterranean region?

As an AI language model, I do not have access to real-time information on specific events. However, I can provide general information on coastal flooding in Barcelona and other cities in the Mediterranean region.

ChatClimate – grounded on the latest IPCC Report

We make Climate Risk understandable and climate information more accessible to the broader communities.

GPT-4's capabilities and limitations create significant challenges. A careful study of these challenges is an important part of the report. This report includes an extensive system card (a card that we foresee around bias, disinformation, over-reliance on the model). It also describes interventions we made to mitigate these risks, including adversarial testing with domain experts.

Enter your question

Submit

Model ⓘ

Mode ⓘ

GPT-3.5 Turbo

Hybrid

If you're concerned about the risk of coastal flooding in Barcelona, I recommend consulting with local authorities or experts in the field who can provide you with more detailed and accurate information based on the latest research and data available.

Summary

- A digital twin for the adaptation to climate change is under construction in the framework of the Destination Earth initiative.
- The digital twin aims at completely revisiting the way climate information has been provided so far in the context of the existence of CMIP and CORDEX, focusing on putting the user at the helm in an operational context.
- This can only work out if there is an honest intention of the different relevant domains to work together in a true multidisciplinary environment.
- Note that the digital twin is **NOT** just about climate model resolution.



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Thanks

francisco.doblas-reyes@bsc.es

