

# Climate modelling: computing and data challenges

Sylvie Joussaume

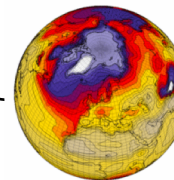
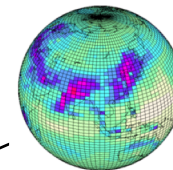
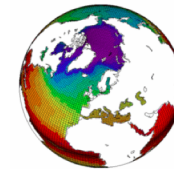
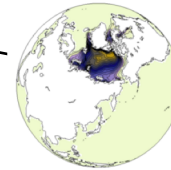
CNRS, Institut Pierre Simon Laplace

Laboratoire des Sciences du Climat et de l'Environnement

Saclay, France

Coordinator of H2020 IS-ENES3

Thanks to : Sandro Fiore (UNITN),  
Christian Pagé (CERFACS),  
Guillaume Levavasseur (IPSL)



**1- Climate change and climate modelling**

**2- Climate models and international landscape**

**3- Challenges in climate modelling**

**Conclusions**

# 1- Climate change and climate modelling

# Global warming is unequivocal

IPCC AR6 (2021)

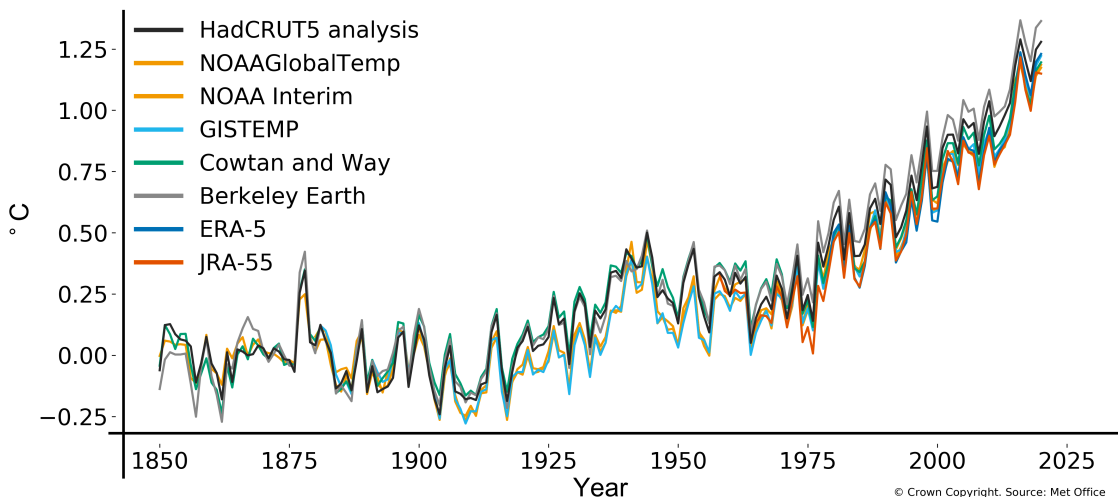
Global :

Decade 2011-2020: +1.09 warmer than 1850-1900

Each last 4 decades: warmer than any decade that preceded it since 1850

Met Office

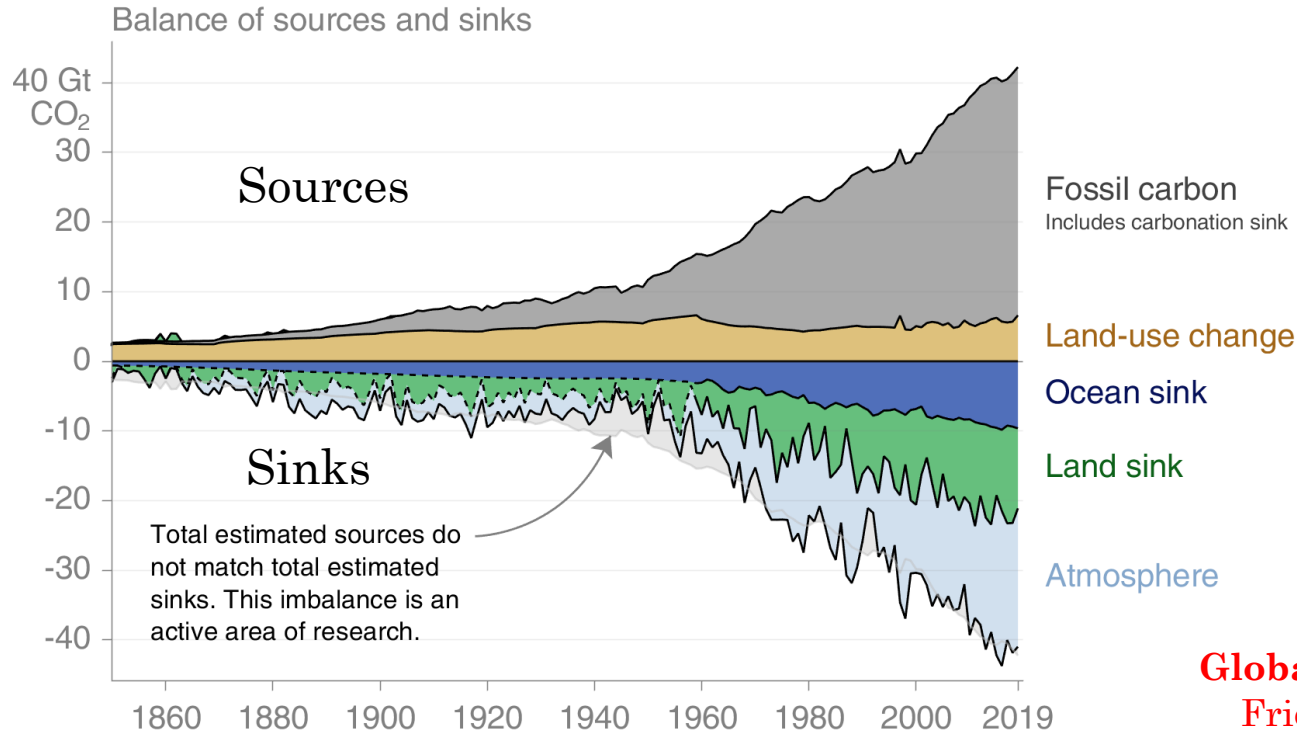
Global mean temperature difference from 1850-1900 (°C)



UK MetOffice  
2020: +1,28° C versus 1850-1900



# Increase of greenhouse gases: use of fossil fuels and demography



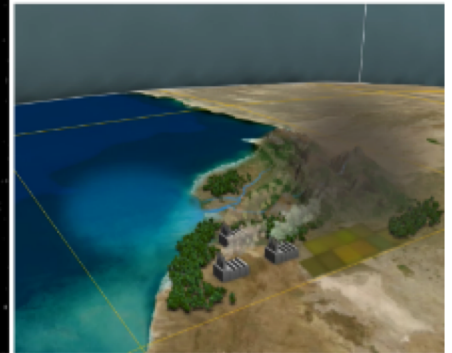
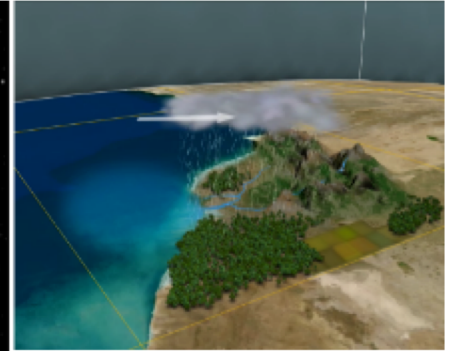
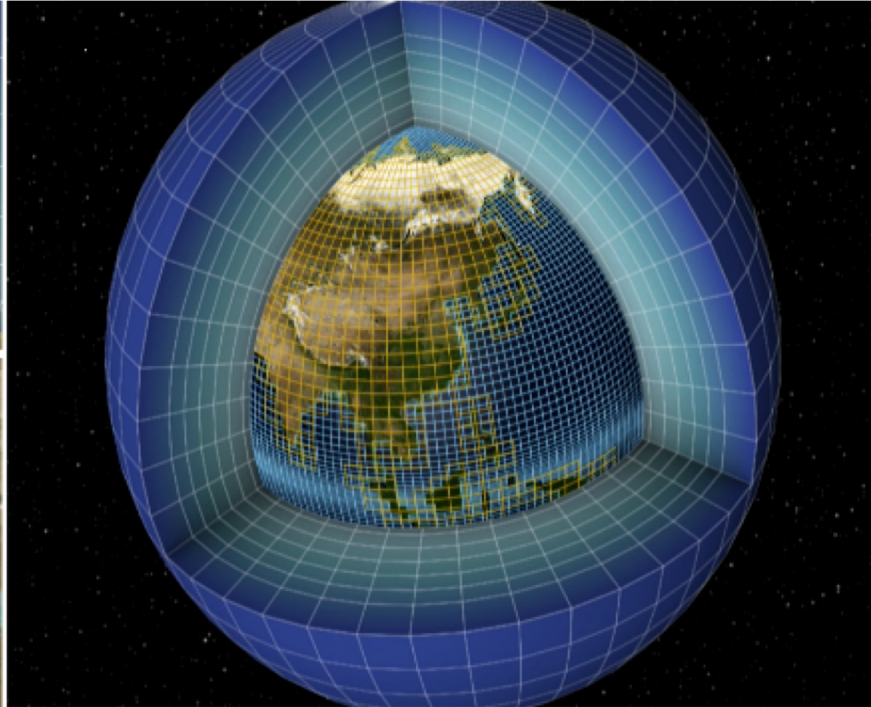
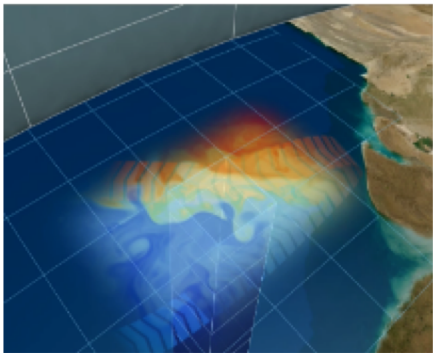
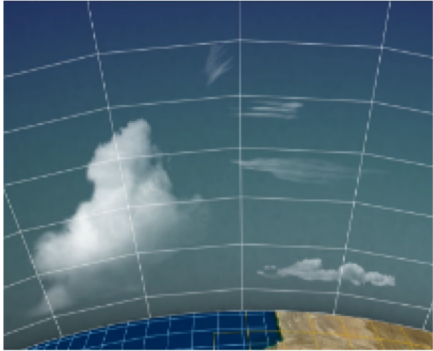
**Atmosphere**  
410 ppm in 2019  
+48% since 1750  
(277 ppm in 1750)

Unprecedented in  
800 000 years

**Global Carbon Budget 2020**  
Friedlingstein et al., 2020

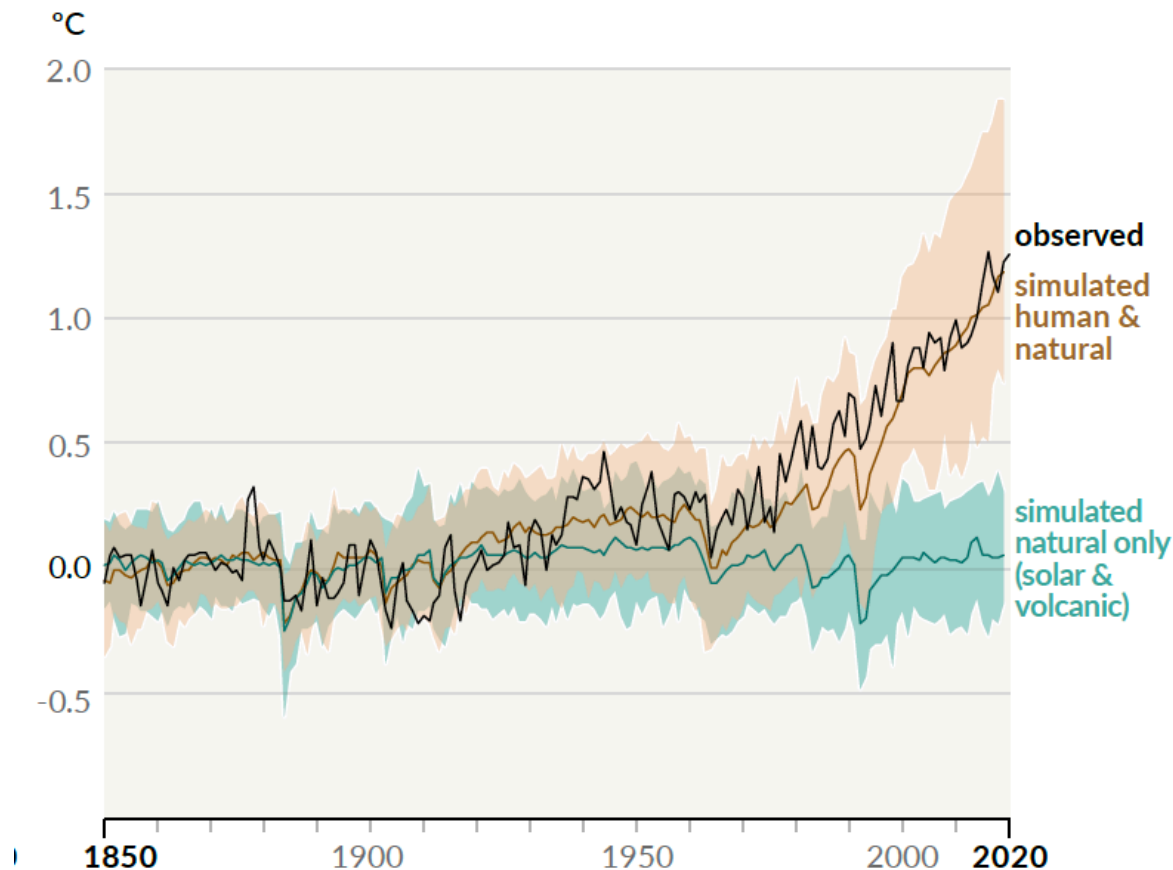
# Modelling the Earth's climate system

Understand & Predict Climate Variability and Changes



“It is unequivocal that human influence has warmed the atmosphere, ocean and land”

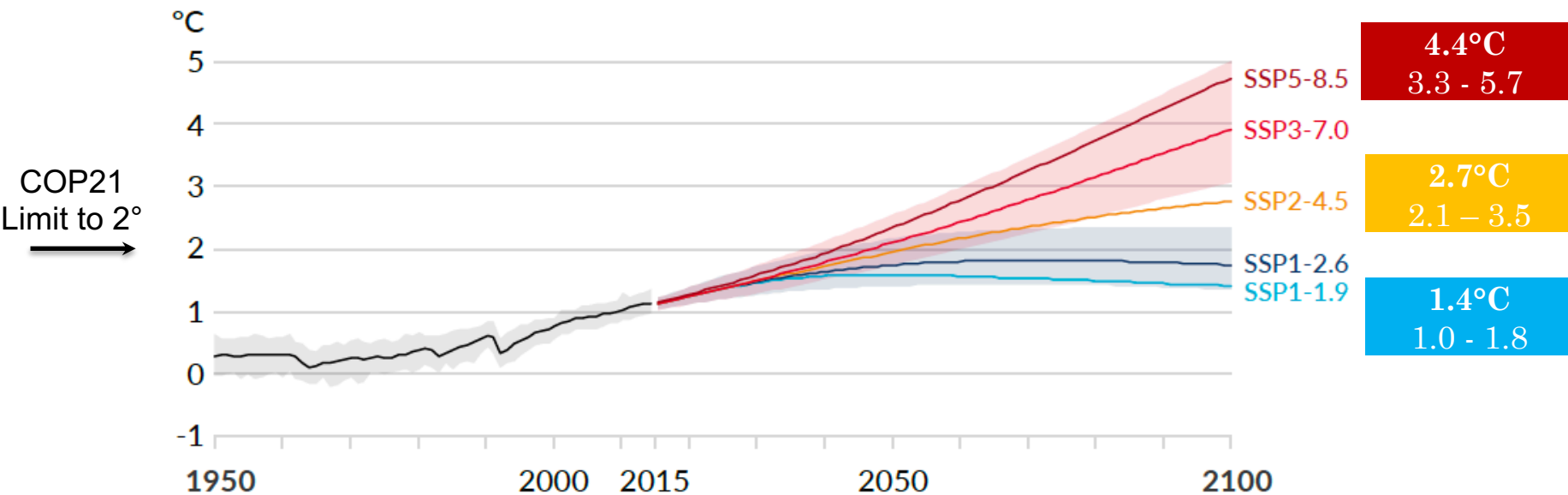
IPCC AR6 SPM (2021)



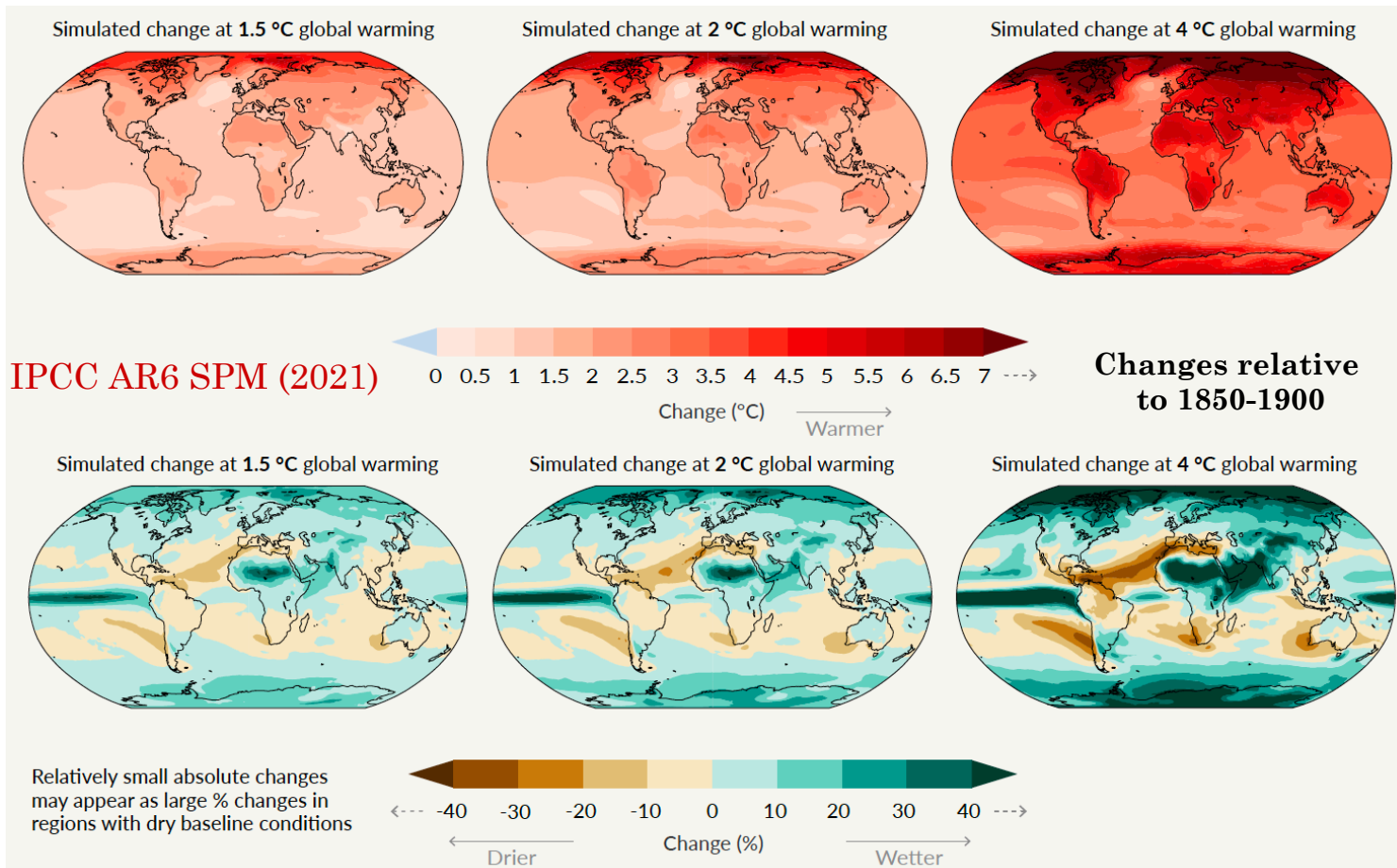
# Simulations of future climate change under different scenarios

IPCC AR6 SPM (2021)

a) Global surface temperature change relative to 1850-1900



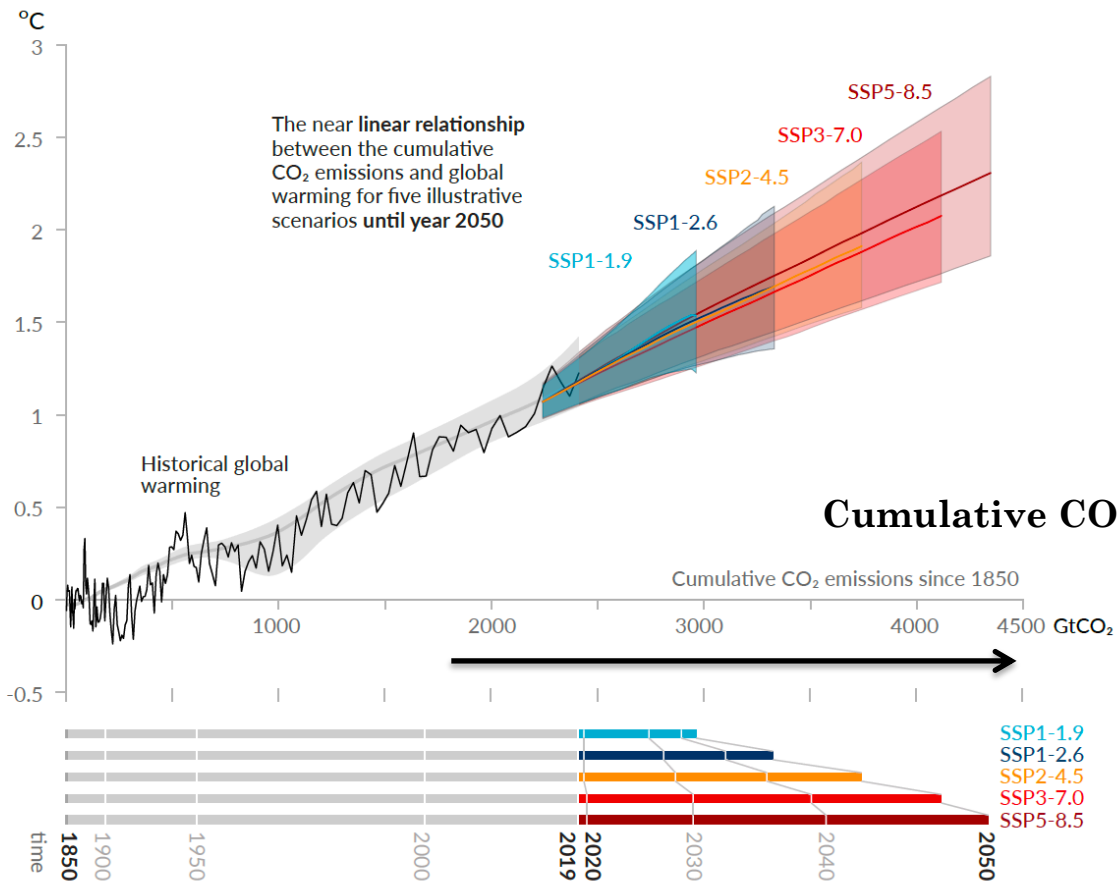
# Projections of temperature and precipitation for different global warming



# Every tonne of CO<sub>2</sub> emissions adds to global warming

Global surface temperature increase since 1850-1900 (°C) as a function of cumulative CO<sub>2</sub> emissions (GtCO<sub>2</sub>)

Temperature Change



IPCC AR6 SPM (2021)

Cumulative CO<sub>2</sub> emissions (GtCO<sub>2</sub>)

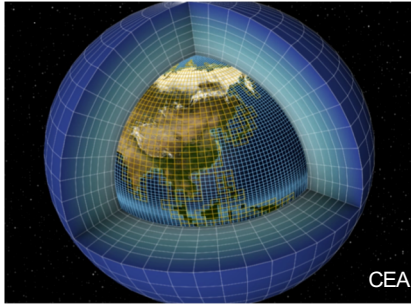
Future cumulative CO<sub>2</sub> emissions differ across scenarios, and determine how much warming we will experience

HISTORICAL  
Cumulative CO<sub>2</sub> emissions between 1850 and 2019

PROJECTIONS  
Cumulative CO<sub>2</sub> emissions between 2020 and 2050

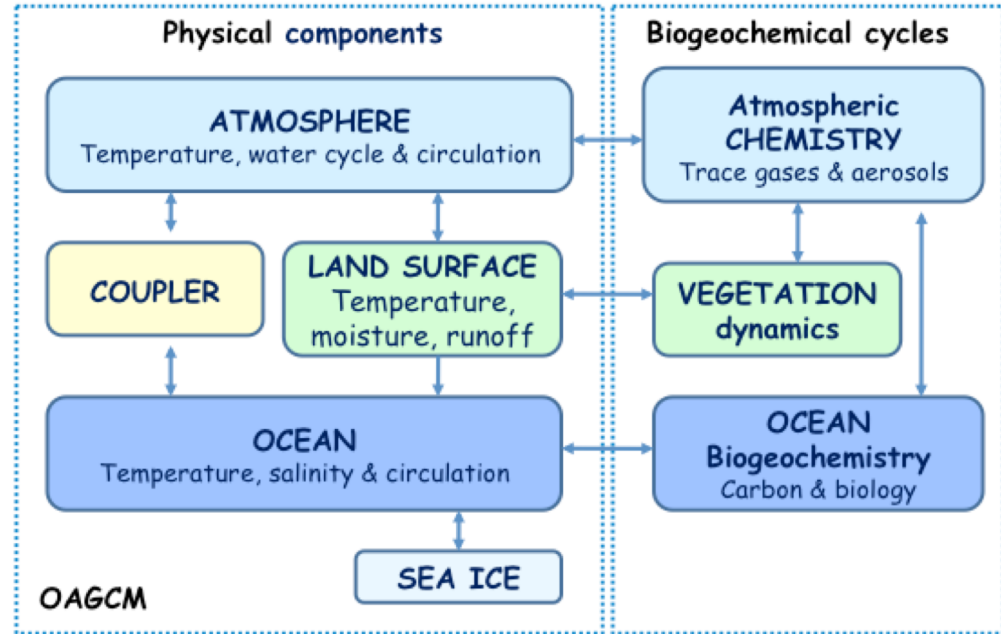
## **2- Climate models and international landscape**

# Earth's climate system models



ESM > 1000 man years  
Strong legacy

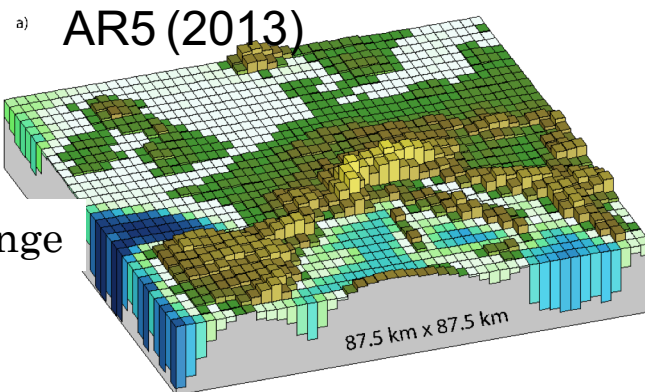
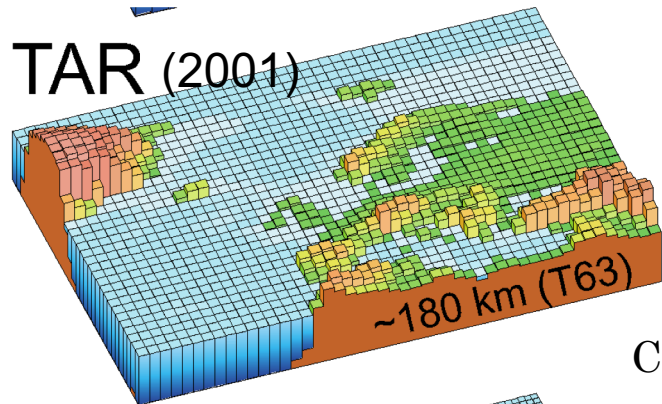
## EARTH SYSTEM MODELS



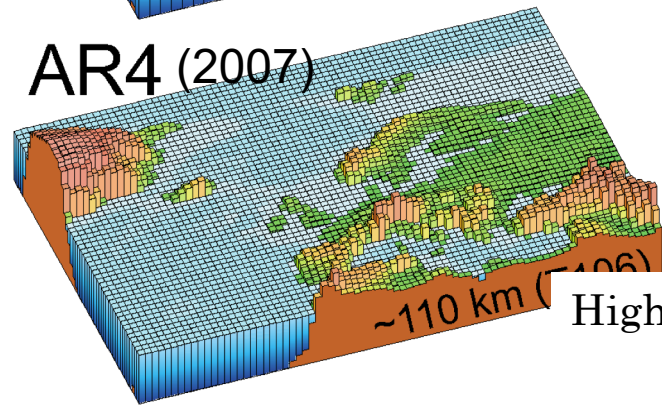
**Based on physical laws** (navier-stokes, conservation)  
& **parameterizations** (clouds, surface fluxes, radiation, sub-grid scale processes)



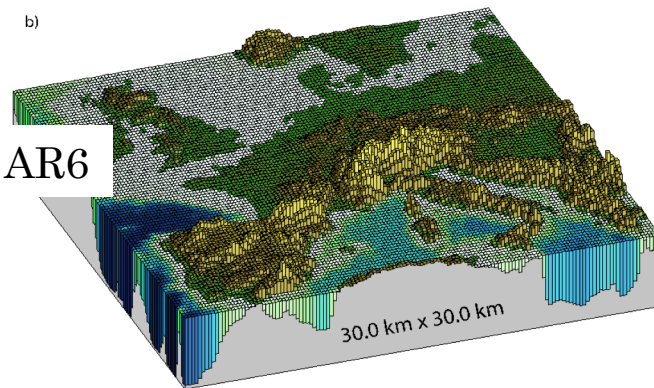
# Spatial resolution



Common range



Highest range AR6  
(2021)



# World Climate Research Program: Coupled Model Intercomparison Project, Phase 6 (CMIP6)

**CMIP6**  
(Eyring et al., GMD, 2016)

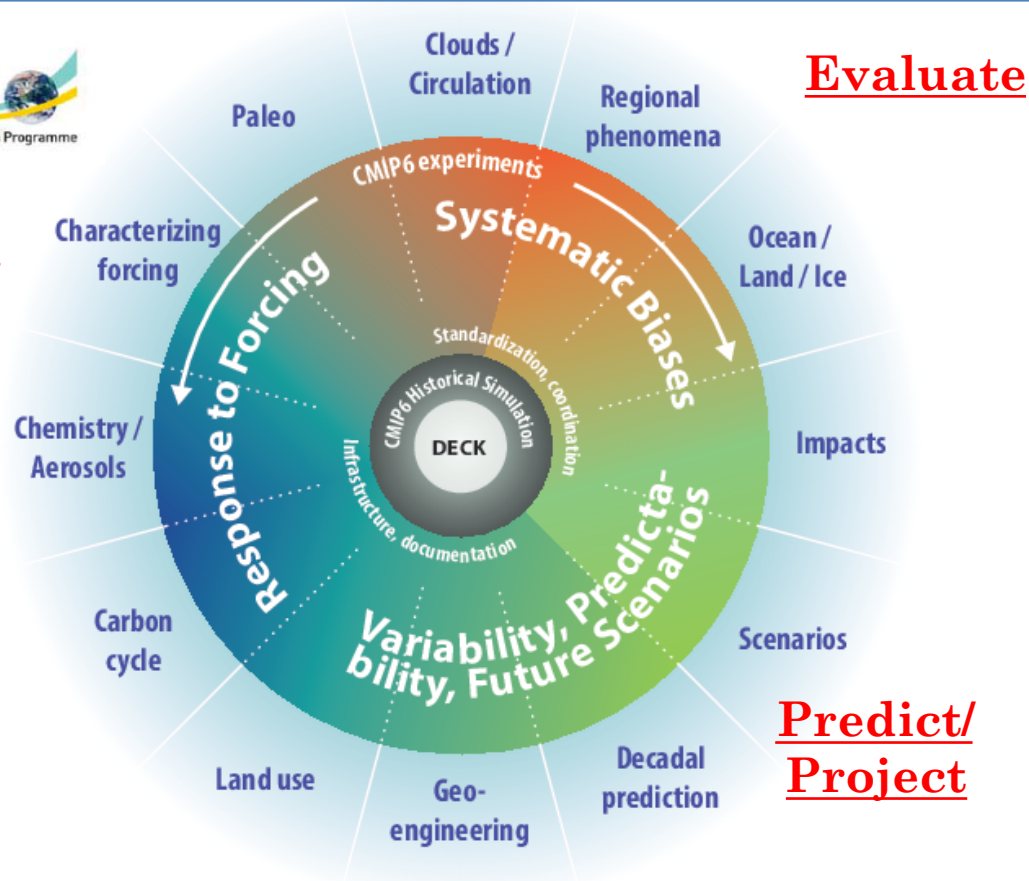
**23 endorsed MIPs**  
Model  
Intercomparison  
Projects

**Per model:**  
20 to 50 000  
simulated years  
100s Mh  
1 – 10 PB produced

**Input to IPCC AR6  
2021**



**Understand**



**Evaluate**

**Predict/  
Project**

**30 modelling groups**

*59 models*

*All: 38/100*

## 1 Canada

Canada	CAN	CanESM5
		CanESM5-CanOE
USA	NCAR	CESM2
		CESM2-FV2
		CESM2-WACCM
		CESM2-WACCM-FV2
USA	E3SM	E3SM-1-0
		E3SM-1-1
		E3SM-1-1-ECA
USA	GFDL	GFDL-AM4
		GFDL-CM4
		GFDL-ESM4
USA	GISS	GISS-E2-1-G
		GISS-E2-1-G-CC
		GISS-E2-1-H
		GISS-E2-2-G
USA	MCM	MCM-UA-1-0

**5 USA**

DE	MPI	MPI-ESM-1-2-HAM
		MPI-ESM1-2-HR
		MPI-ESM1-2-LR
DE	AWI	AWI-CM-1-1-MR
		AWI-ESM-1-1-LR
FR	CNRM	CNRM-CM6-1
		CNRM-CM6-1-HR
		CNRM-ESM2-1
FR	IPSL	IPSL-CM6A-LR
IT	CMCC	CMCC-CM2-HR4
		CMCC-CM2-SR5
UK	HADLEY/UKESM	HadGEM3-GC31-LL
		HadGEM3-GC31-MM
		UKESM1-0-LL
NO	NorESM	NorCPM1
		NorESM1-F
		NorESM2-LM
		NorESM2-MM
EC-Earth	EC-Earth	EC-Earth3
		EC-Earth3-LR
		EC-Earth3-Veg
		EC-Earth3-Veg-LR
Russie	INM	INM-CM4-8
		INM-CM5-0

**1 Australia**



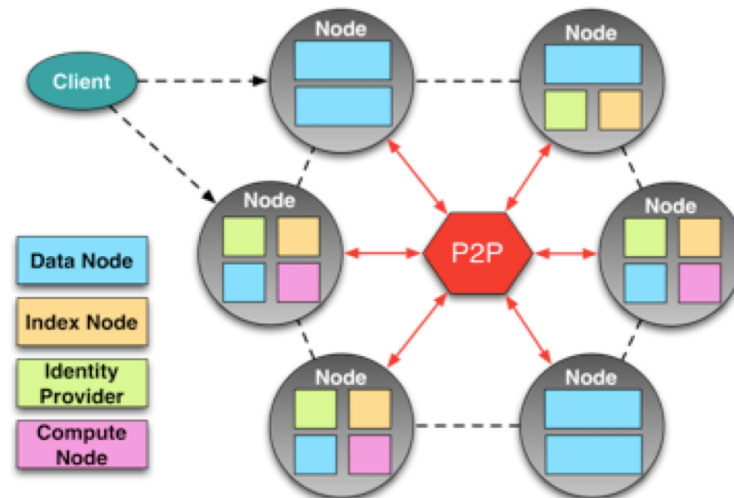
**8 in Europe**

**7 China / 1 Taiwan/ 3 Korea**

**2 Japan**

**1 Russia**

Chine	BCC	BCC-CSM2-MR
		BCC-ESM1
Chine	CAMS	CAMS-CSM1-0
Chine	CAS	CAS-ESM2-0
Chine	CIESM	CIESM
Chine	FGOALS	FGOALS-f3-L
		FGOALS-g3
Chine	FIO	FIO-ESM-2-0
Chine	NESM	NESM3
Taiwan	TaiESM	TaiESM1
Korea	KACE	KACE-1-0-G
Korea	KIOST	KIOST-ESM
Korea	SNU	SAM0-UNICON
JP	MIROC	MIROC-ES2L
		MIROC6
JP	MRI	MRI-ESM2-0
Inde	IITM	IITM-ESM
AU	ACCESS	ACCESS-CM2



## Dashboard stat

ESGF: 13 M datasets, 30.4 PB - ca 15 000 registered users

## FAIR data

Open access, common data and metadata standards  
Multi-agencies support: *DOE, NOAA, NASA, IS-ENES, NCI*

# Coupled Model Intercomparison Project (CMIP)

## 1995 WCRP creation of the Working Group on Coupled Modelling

Foster the development and review of coupled models

**CMIP Launched in 1995 - Mainly control runs - 1 GB**

**CMIP2: Launched in 1997 – Idealised experiment 1%/year increased CO<sub>2</sub>**

**0.5 TB - Data accessible only on subproject basis - IPCC TAR (2001)**

**CMIP3: more realistic past (20th) and future simulations (scenarios) - IPCC AR4 (2007)**

**35 TB of data at PCMDI – open and free for non commercial**

**CMIP5 (2008-2013): consistent set for all experiments - IPCC AR5 (2013)**

**1.5 - 5 (with replica) PB of data**

**open data (very few closed for non commercial)**

**CMIP6 (2014-2019) common core and more independent MIPs - IPCC AR6 (2021)**

**11 (no replica) – 21 (with replica) PB**

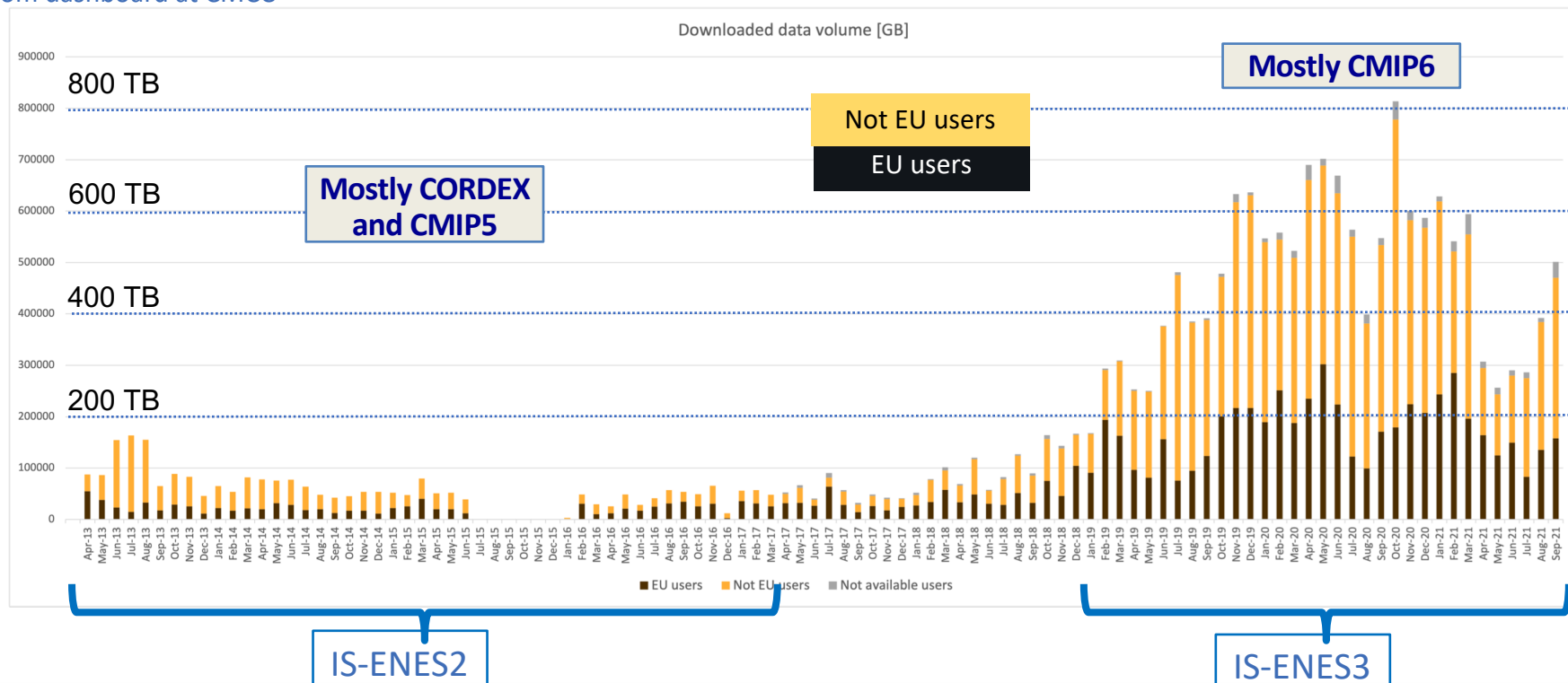
**Full open data**

Central  
storage



# Downloaded data volume over IS-ENES2 and IS-ENES3 from European ESGF data nodes (Apr 2013 – Sep 2021)

From dashboard at CMCC



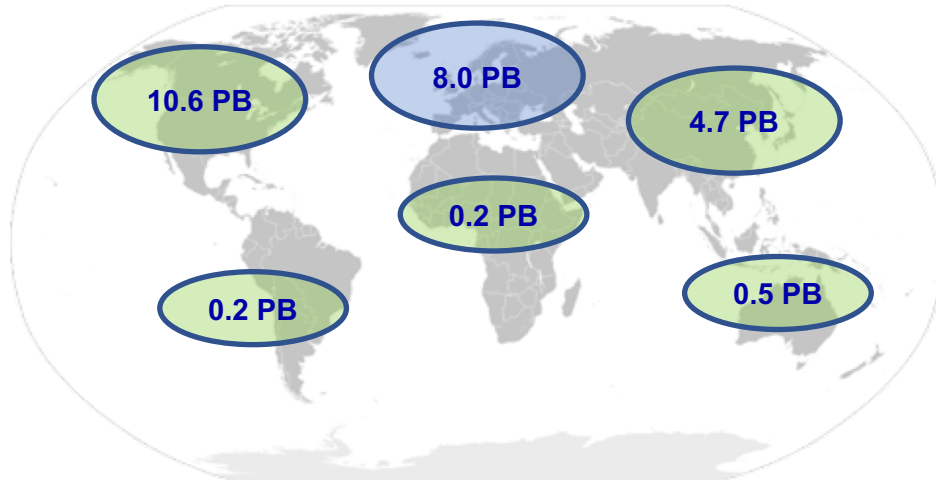
IS-ENES2

IS-ENES3

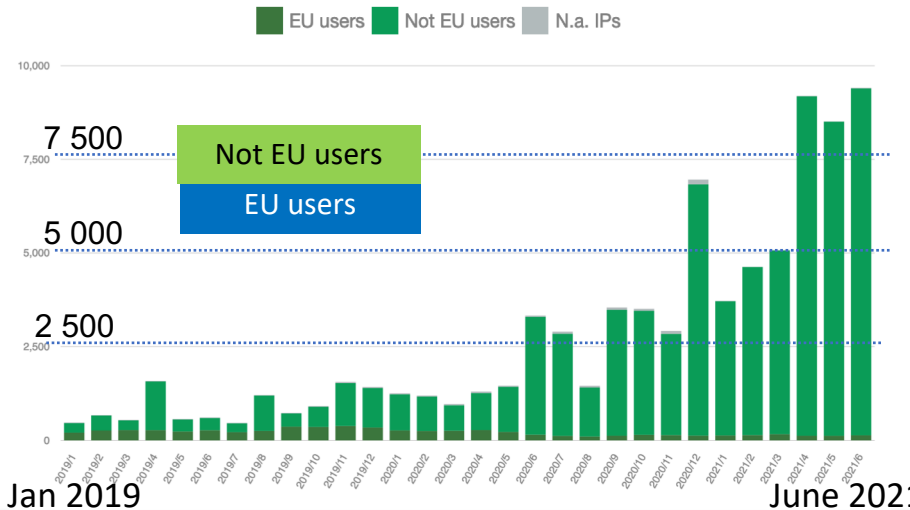




# ESGF data user distribution



**Total downloaded data volumes**  
**All ESGF nodes**  
**Geographical distribution**

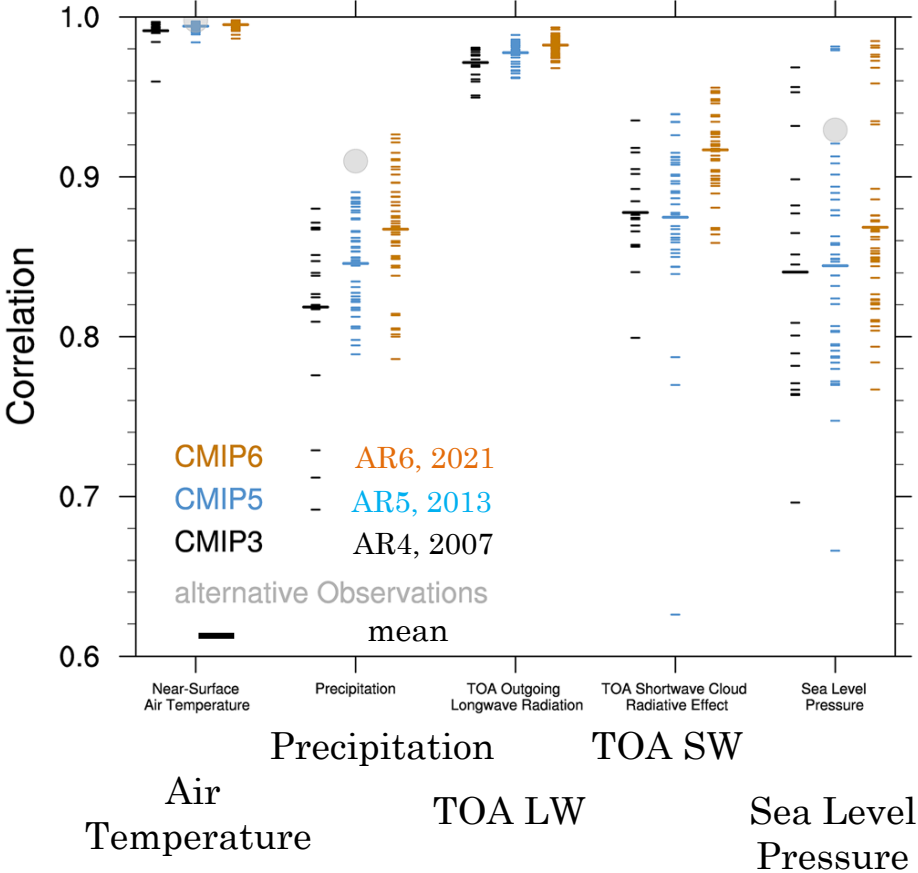


**Number of distinct clients**  
**European ESGF data nodes**

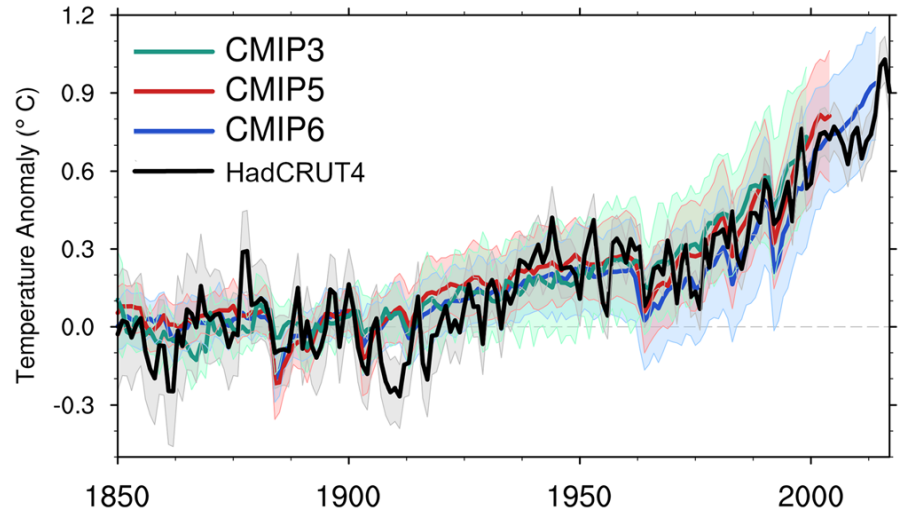


# Evaluation of CMIP6 model results

## Correlation with observations



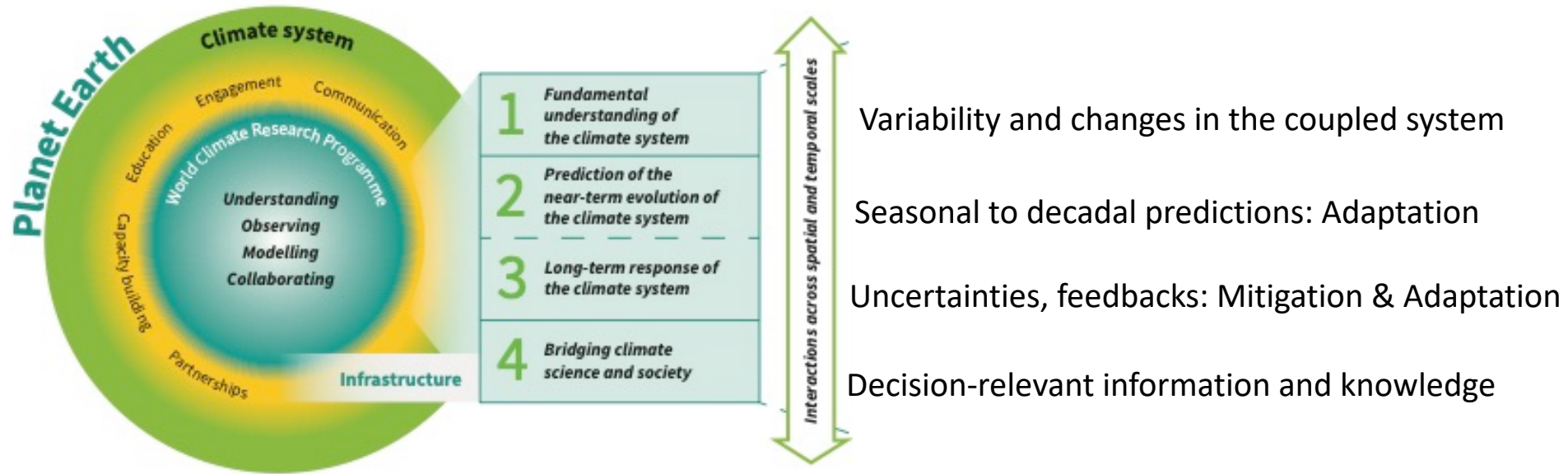
Bock et al., JGR (2020)

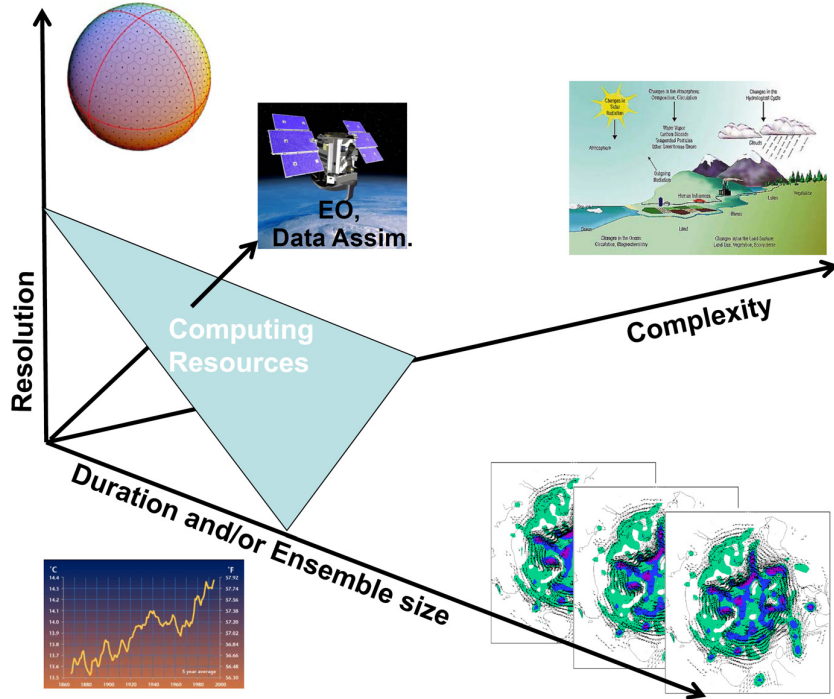




## **3- Challenges in climate modelling**

# WCRP Strategic Plan 2019-2028





## Resolution:

Resolution x 2 > Computing x 8

## Complexity :

carbon cycle, aerosols, chemistry, biosphere

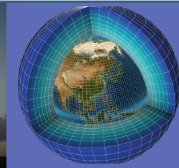
## Duration: need for long simulations

Multi-decadal to multi-centennial

At least 5-10 SYPD

## Ensemble size: document internal

variability, quantify uncertainties



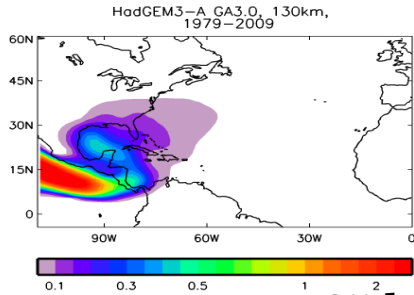
# Spatial resolution

Key for regional patterns

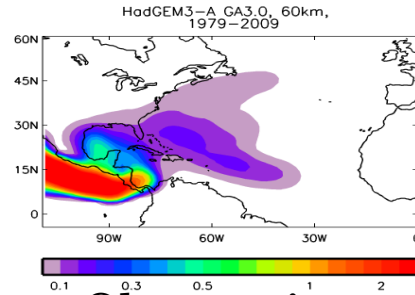


HighResMIP

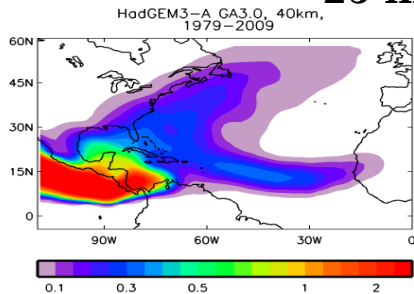
130 km



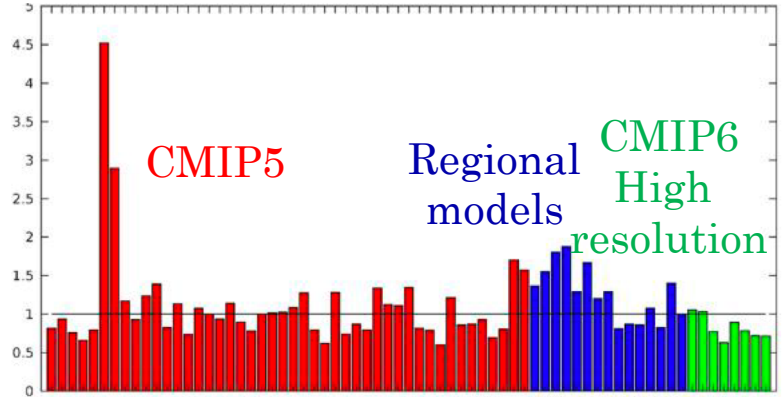
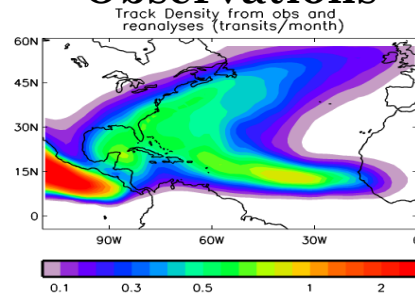
60 km



25 km



Observations



Bias in precipitation variability  
Danubius river

Simulated tropical cyclones  
Roberts et al. (2015)

# Spatial resolution and computation

Roberts et al., GMD, 2019

Model name	CMIP6 resolution (atmosphere–ocean) (km)	Initial condition	Total years (spinup years)	Nodes (atmosphere–ocean)	Max turnaround (years per day)	Output per year (TB)
LL	250–100	LL-spinup (30 years)	1130 (30)	12–2	4	0.13
MM	100–25	MM-spinup (30 years)	680 (30)	50–24	1.3	0.73
HM	50–25	MM-spinup (30 years)	117 (0)	90–24	0.5	2.8
MH	100–8	MH-spinup (30 years)	205 (30)	34–171	0.45	2.0
HH	50–8	MH-spinup (30 years)	100 (0)	90–171	0.4	4.5

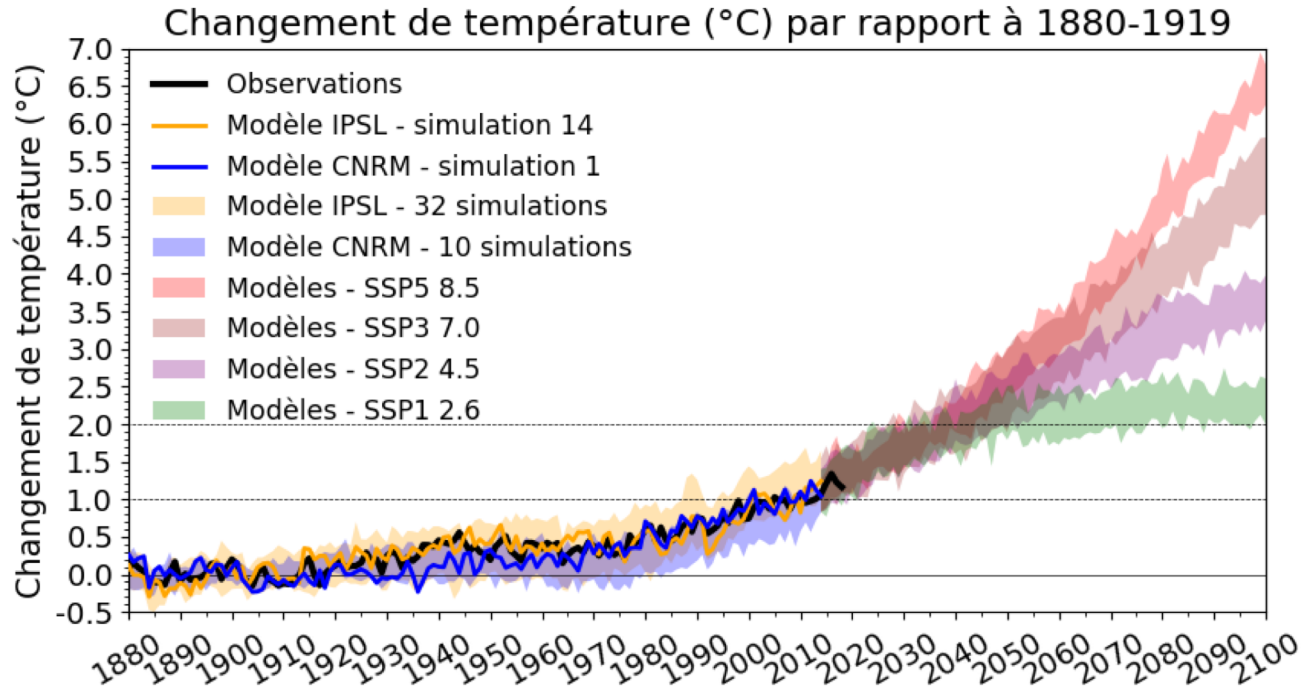
Resolution  
x 5

Nodes  
x 10

Speed  
/ 10

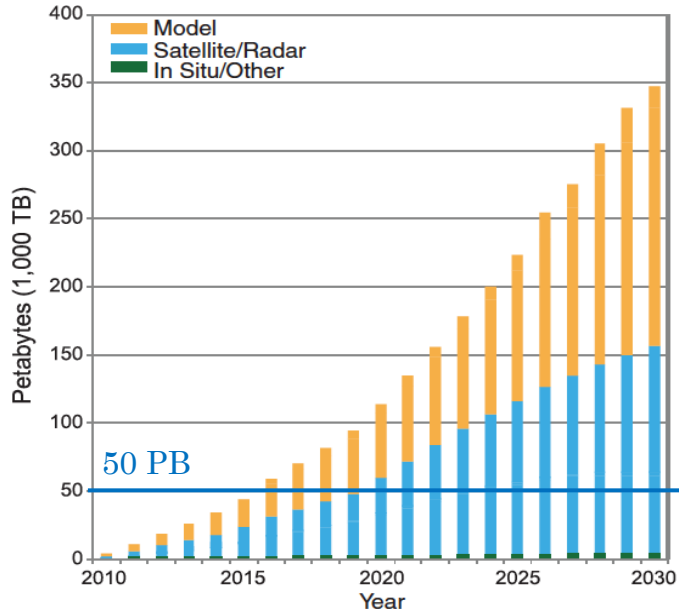
Data  
x 35

# Need for ensemble simulations

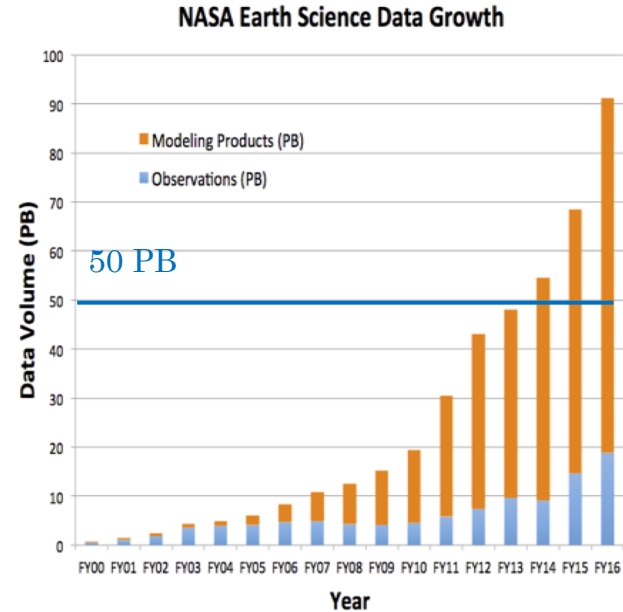


# Data deluge

Overpeck et al. (Science 2011)



Tsengdar Lee, Icas17

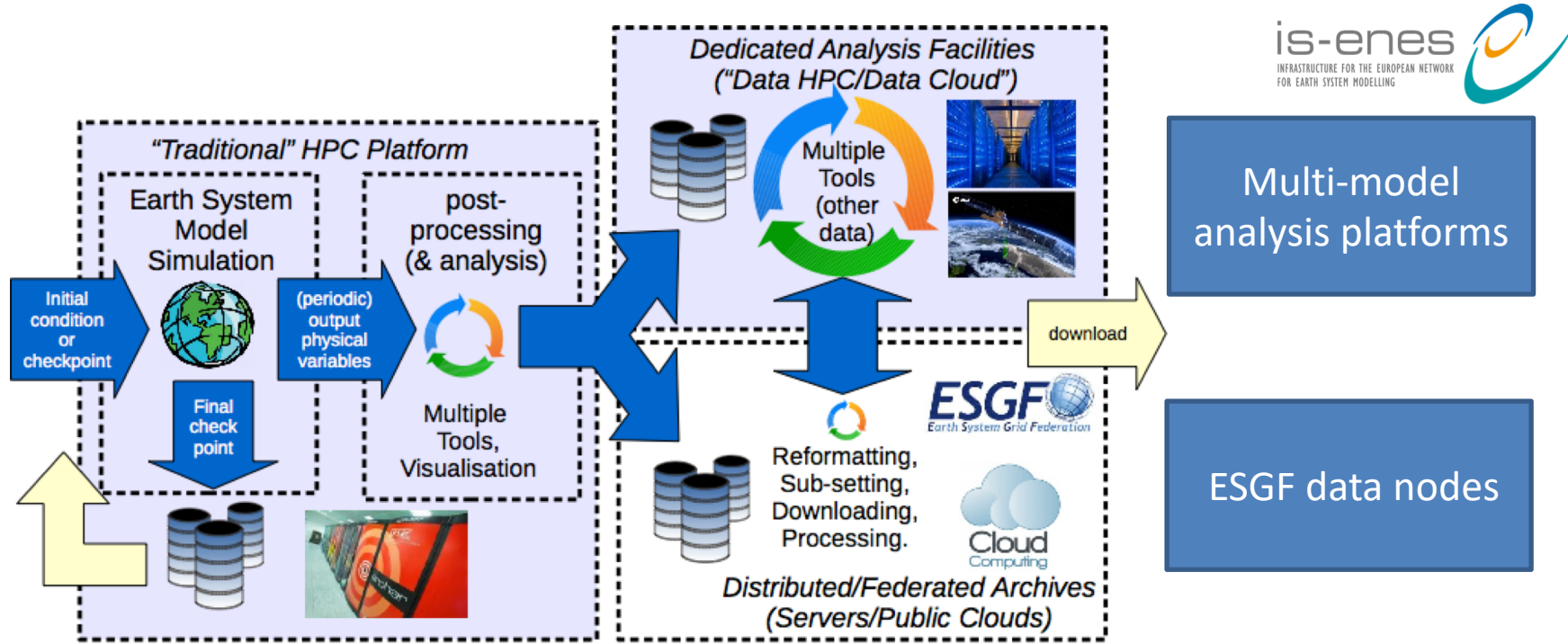


**Data workflow: increasing complexity**

Parallel IO / « on the fly » processing / Compression / Reduction of data

**Computing near data**

# Multiple Types of Storage & Data Interaction



Multiple Roles, at least:

Model Developer, Model Tinkerer, Expert Data Analyst, Service Provider, Data User







@GIS climat environnement societe

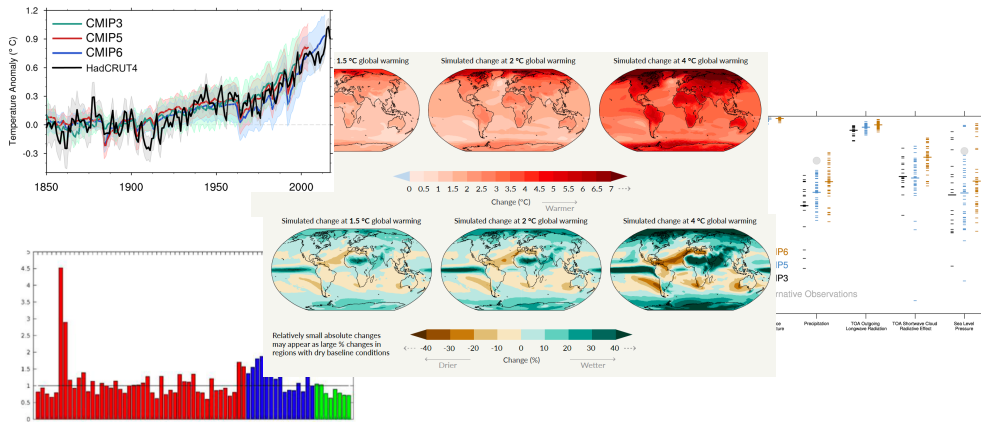
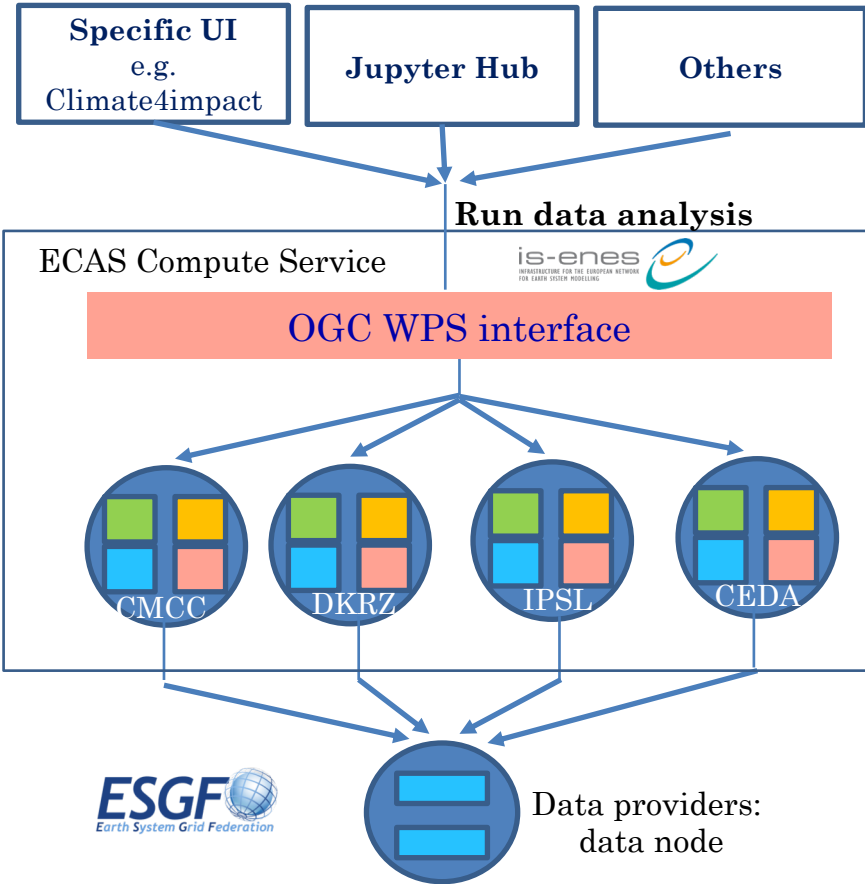
**Climate science:**  
understand and predict  
past, present and future climates

**Climate impacts:**  
sectors (agriculture, hydrology, energy,  
health ...)

**Climate services:**  
for mitigation & adaptation decision

## Multi-model analyses:

- Data at multiple sites
- Multiple computation needs:  
Averaging, Sub-setting,  
Climate diagnostics and indices,  
Downscaling at regional scale,  
Bias correcting to force impact models  
Processing & Visualisation tools

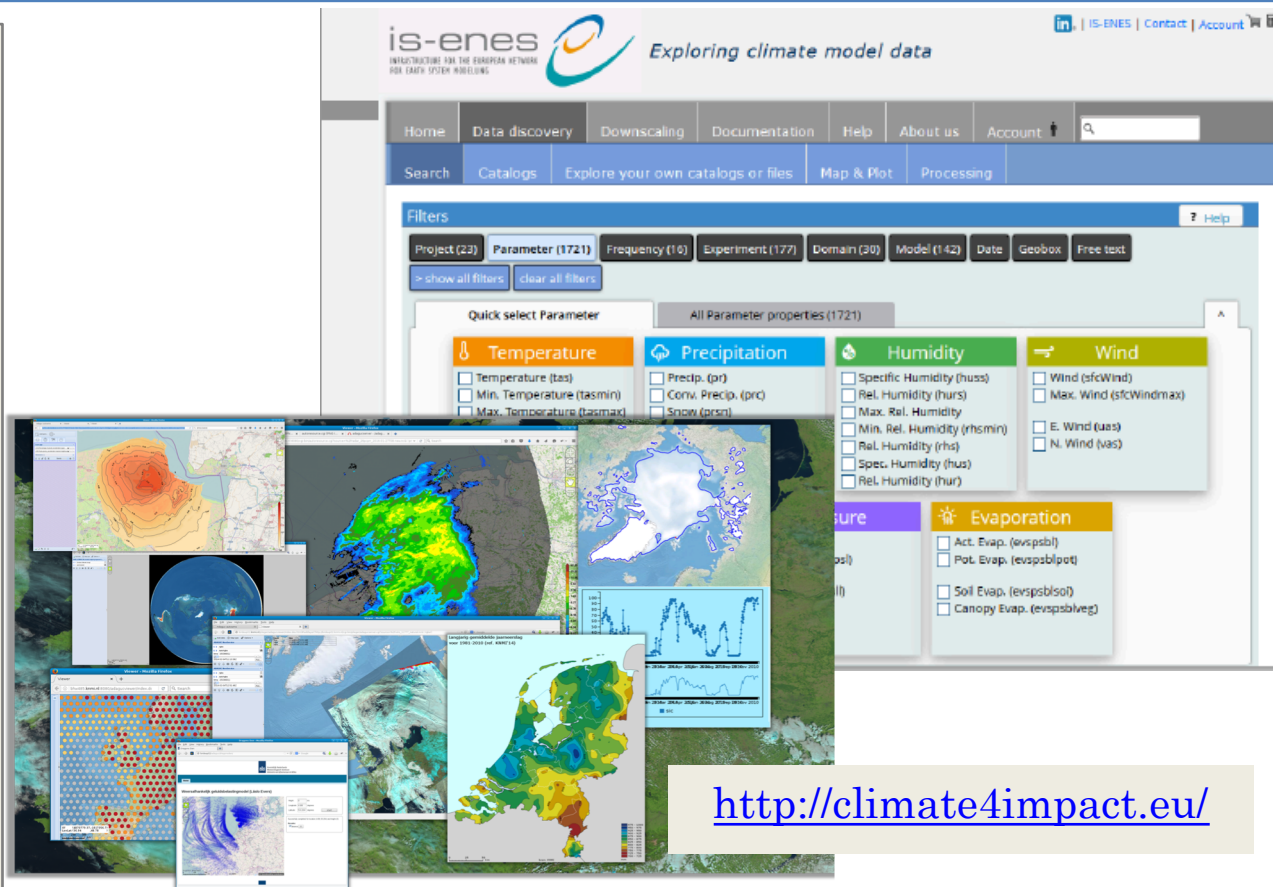


**Platform for  
Impact researchers**  
Explore climate data  
& perform analysis

## Climate4impact portal

Including:

- In-depth documentation and guidance
- Use cases from impact researchers
- **Perform calculations / Data processing: WPS**
- **Downscaling portal**

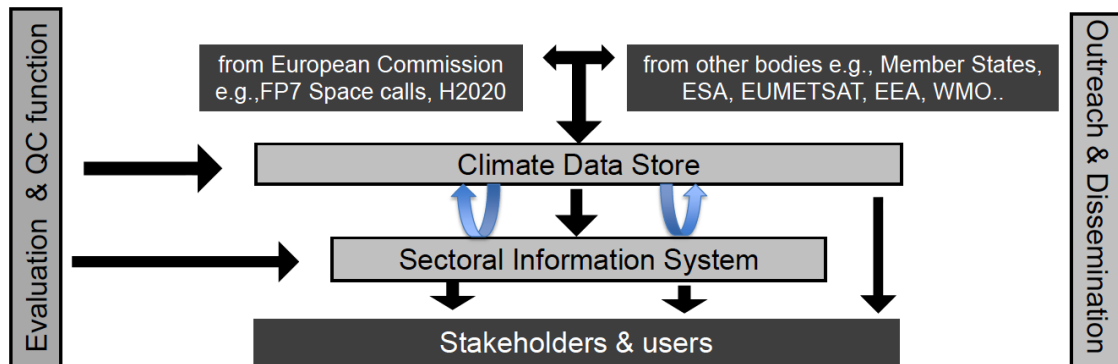


The screenshot displays the Climate4impact portal interface. At the top, the logo for 'is-enes' (Infrastructure for the European Network for Earth System Modelling) is visible, along with the tagline 'Exploring climate model data'. The navigation bar includes tabs for 'Home', 'Data discovery', 'Downscaling', 'Documentation', 'Help', 'About us', and 'Account'. Below the navigation bar, there are search and filter options, including 'Search', 'Catalogs', 'Explore your own catalogs or files', 'Map & Plot', and 'Processing'. A 'Filters' section shows active filters for 'Project (23)', 'Parameter (1721)', 'Frequency (10)', 'Experiment (177)', 'Domain (20)', 'Model (142)', 'Date', 'Geobox', and 'Free text'. A 'Quick select Parameter' dropdown is open, showing categories: 'Temperature' (with sub-options: Temperature (tas), Min. Temperature (tasmin), Max. Temperature (tasmax)), 'Precipitation' (with sub-options: Precip. (pr), Conv. Precip. (prc), Snow (psrn)), 'Humidity' (with sub-options: Specific Humidity (huss), Rel. Humidity (hurs), Max. Rel. Humidity, Min. Rel. Humidity (rhsm), Rel. Humidity (rhs), Spec. Humidity (hus), Rel. Humidity (hur)), 'Wind' (with sub-options: Wind (sfcWind), Max. Wind (sfcWindmax), E. Wind (uas), N. Wind (vas)), and 'Evaporation' (with sub-options: Act. Evap. (evspsbl), Pot. Evap. (evspsblpot), Soil Evap. (evspsblso), Canopy Evap. (evspsblveg)).

<http://climate4impact.eu/>



# Copernicus Climate Change Service (C3S) at ECMWF



<https://cds.climate.copernicus.eu/>

## Climate Data Store

Reanalyses

Satellite data (ESA CCI)

Seasonal Forecasts

**Projections (Global/Regional)**

## Projections: broker on **ESGF**

relying on the expertise from **IS-ENES partners**

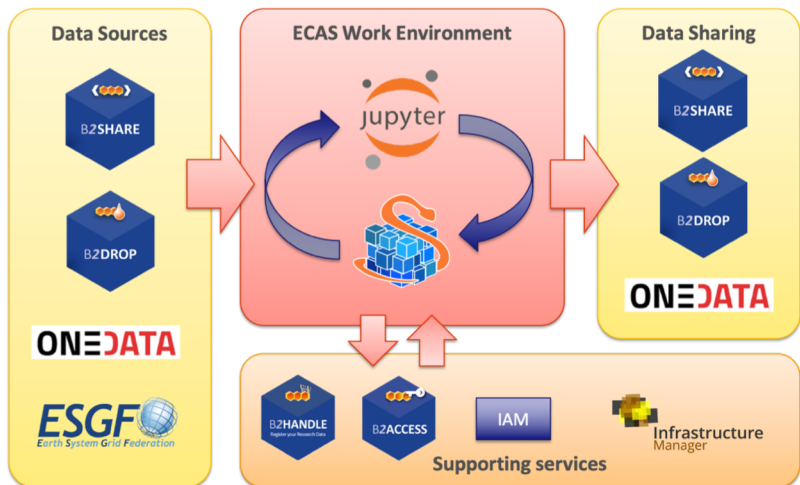
Subset of simulations (historical and projections), models & variables  
for

GLOBAL: CMIP5, CMIP6

&

REGIONAL: CORDEX including additional simulations for Europe

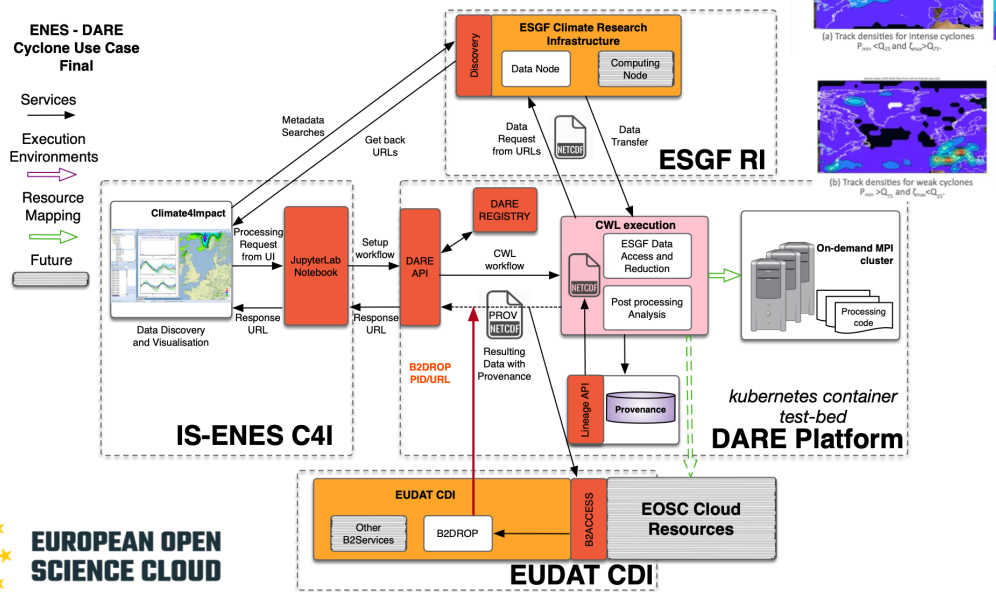
## ENES Climate Analytics Service (ECAS) an EOSC-Hub Thematic Service



EUROPEAN OPEN  
SCIENCE CLOUD

CERFACS, KNMI

## DARE Platform: Use of cloud resources



GA 777536

01/2018-03/2021

CMCC, DKRZ

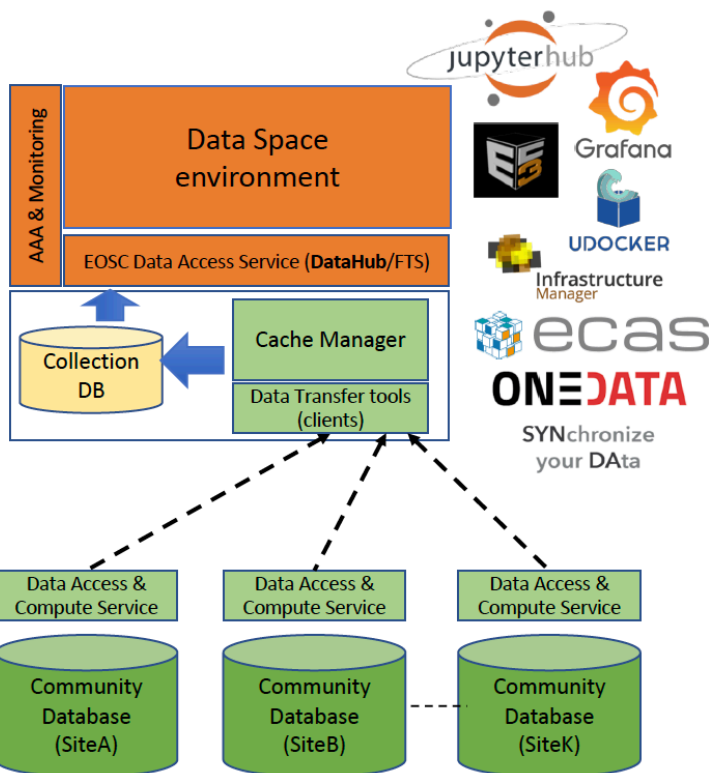
DARE

01/2018-12/2020



GA No 777413

<https://enesdataspace.vm.fedcloud.eu/>



01/2021- 06/2023



<http://is.enes.org/>



The screenshot shows the ENES Data Space web interface. The header includes the URL <https://enesdataspace.vm.fedcloud.eu/> and navigation links for Home, Notebooks, User guide, and Login. The main content area features a title 'ENES Data Space' and a description: 'The ENES Data Space delivers an open, scalable and cloud-enabled data science environment for climate data analysis on top of the EOSC Compute Platform. It provides both storage and computational capabilities. It consists of a JupyterHub instance jointly with a large set of pre-installed Python libraries for running data manipulation, analysis, and visualization, and a data publication service enabling file browsing and data access for scientific datasets. The ENES Data Space hosts (open) data from the ESGF federated data archive on compute cloud to support meteorological and industrial researchers in realistic climate model analysis experiments.' Below the text is a grid of various data visualizations, including maps, line graphs, and heatmaps. At the bottom, there is a search bar with the text 'Find resource...', a dropdown menu for 'All resour...', and a search icon. The footer includes the 'EUROPEAN OPEN SCIENCE CLOUD' logo, navigation links for Resources, Processing & Analysis, Data Analysis, Image/Data Analysis, and ENES Data Space, and a button to 'Access the resource'.





Environmental RIs

*ENRISK proposal*



ENES RI



Clouds



Data Space  
(& HPC)



Digital  
Twins  
*Intertwin  
Proposal*



**EUROPEAN OPEN  
SCIENCE CLOUD**

Wider research community

**Copernicus**



Decision  
& policy  
oriented

**Destination  
Earth**

**Landscape**  
Tentative schema

## Conclusions

**Climate models: at the core of climate information for mitigation and adaptation  
But also needed for understanding**

**To address climate research challenges, need for a sustained European climate modelling infrastructure** : *support data infrastructure, development of codes, sharing of expertise*  
**Key for IPCC assessment reports and Copernicus C3S**

### Data and computing challenges:

#### HPC:

- Range of needs, such as spatial resolution and large ensemble
- Prepare for future architectures although using legacy codes & complex workflows

#### Data:

- Manage large amounts of data near HPC
- Ease the access and democratize the analysis of large climate datasets (i.e. multi-model ensembles) – ESGF 2.0, IS-ENES and collaborations
- Enable complex analyses, e.g. AI-based analysis of very large ensemble of datasets

**Collaborations are important (ESGF, C3S, EGI ...)**  
**European infrastructures: EuroHPC & EOSC should help**





01/2019– 12/2022

<https://is.enes.org/>



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 824084*



Follow us on Twitter !  
**@ISENES\_RI**



Join the community  
on ZENODO !  
**is-enes**