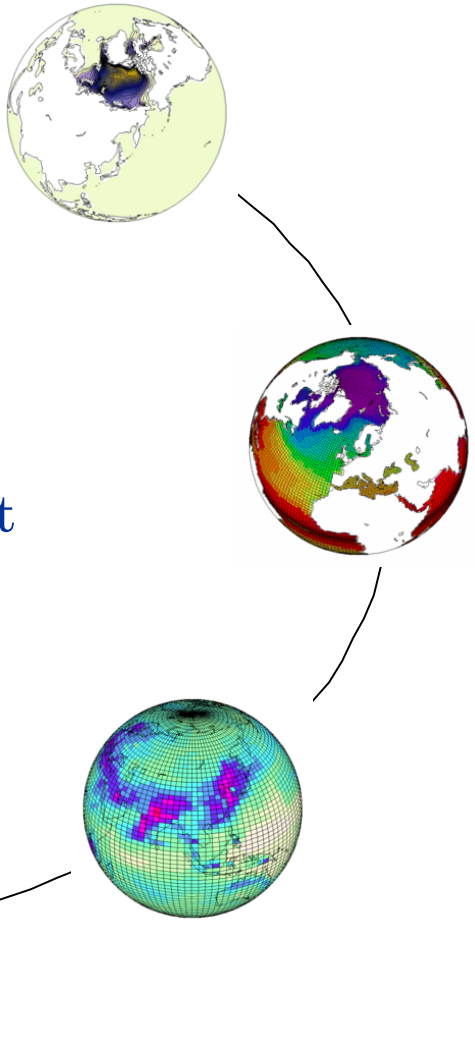
SC 2012, Salt lake City,
Session HPC applications and society
Wednesday November 14th

Modelling the Earth's climate system: data and computing challenges



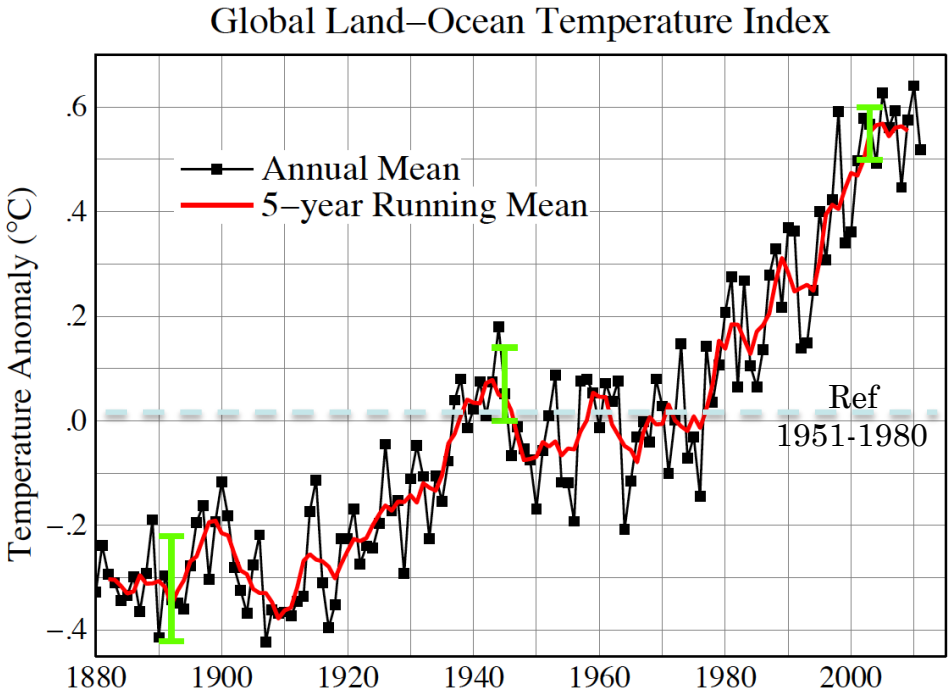
Sylvie Joussaume
CNRS, IPSL, Saclay, France
Coordinator of FP7 IS-ENES project

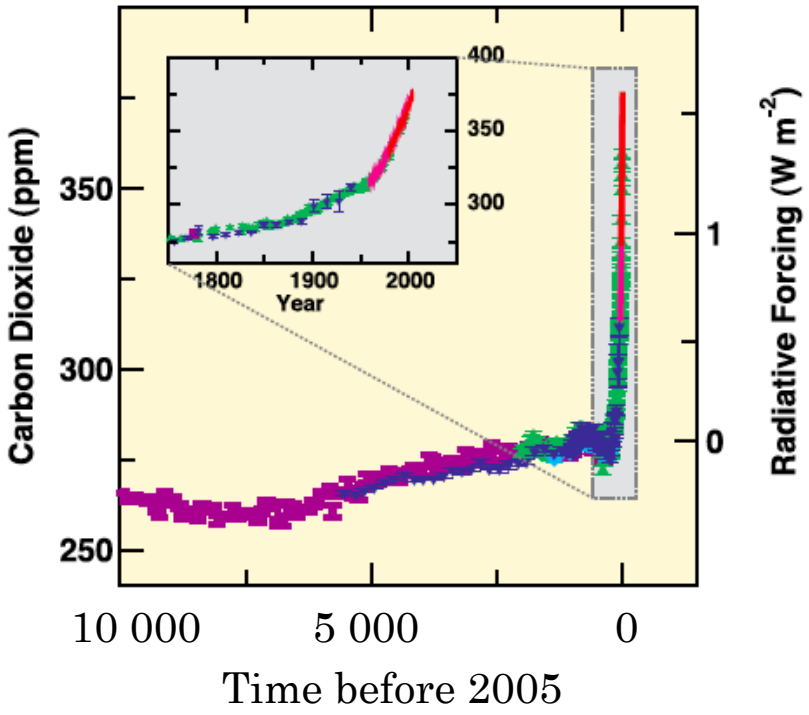
Contributions from
A. Bellucci, J. Biercamp, R. Budich, A. Dawson, M.A. Foujols, B. Lawrence, L. Linardikis, S. Masson, Y. Meurdesoif, G. Riley, K. Taylor, P.L. Vidale
& colleagues from IS-ENES project

Warming of
the climate system is unequivocal
Intergovernemental Panel on
Climate Change (2007)
+0.74°C (1996-2005)

Concentration of
greenhouse gases
due to human activities
is increasing
IPCC (2007)
2011 : CO_2 at 390 ppm

1880-2011

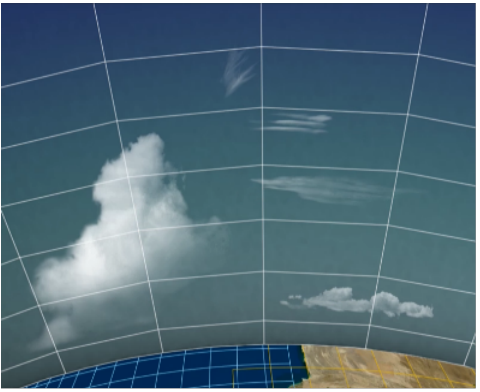
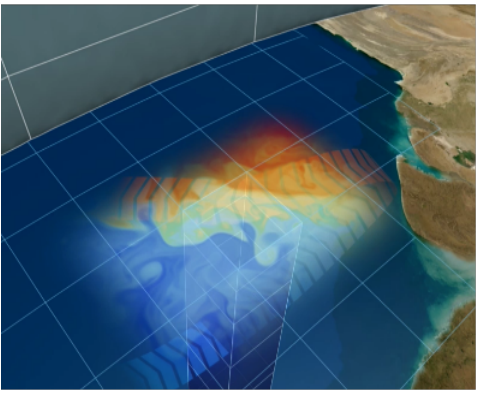
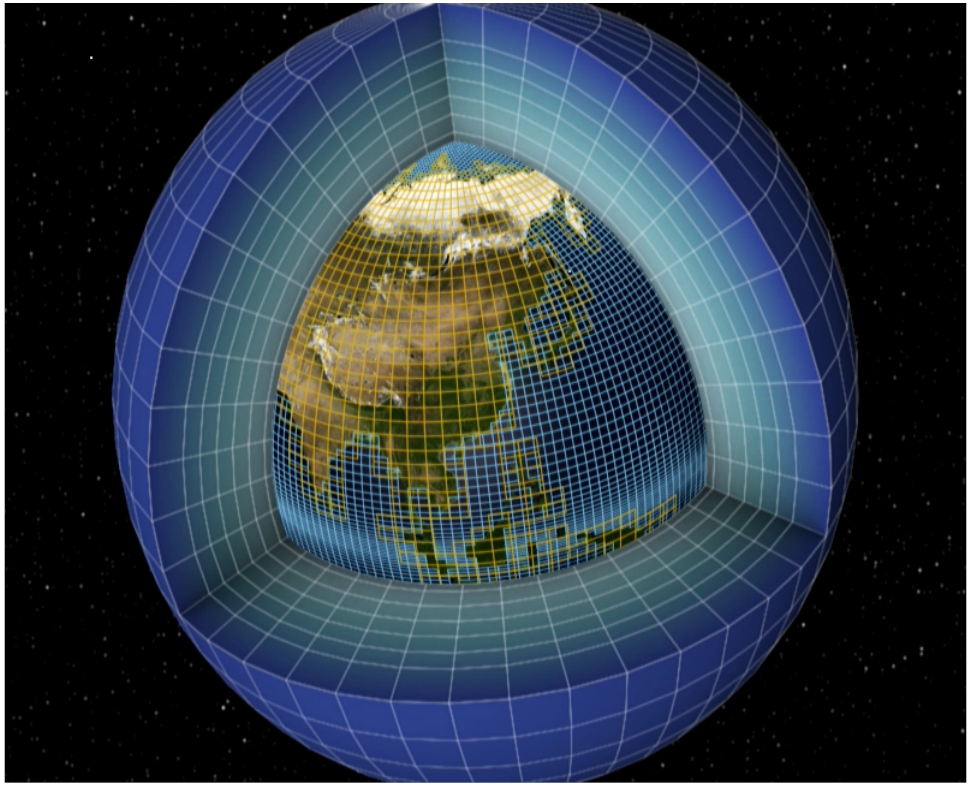
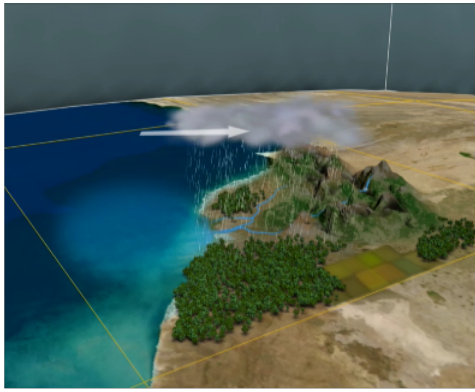
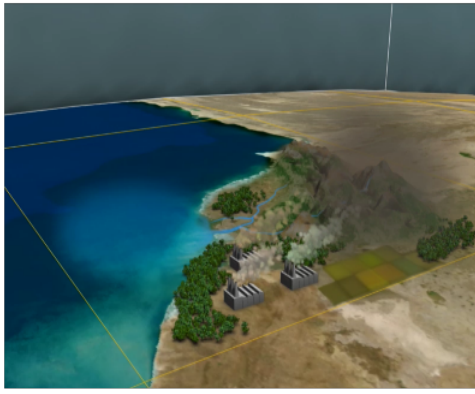




Hansen, http://data.giss.nasa.gov/gistemp/

Modelling the Earth's climate system

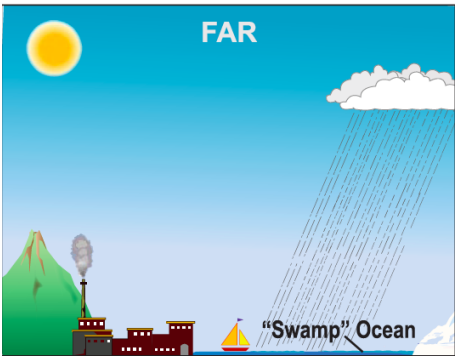
Understand & Predict
Climate Variability and Changes

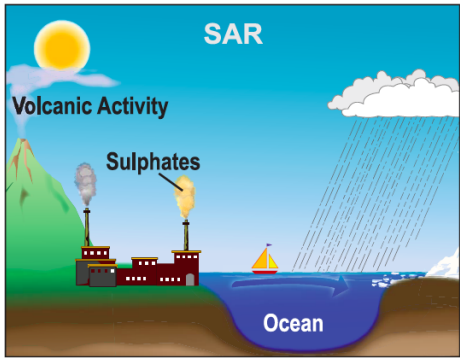
Copyright CEA

Evolution of climate modelling

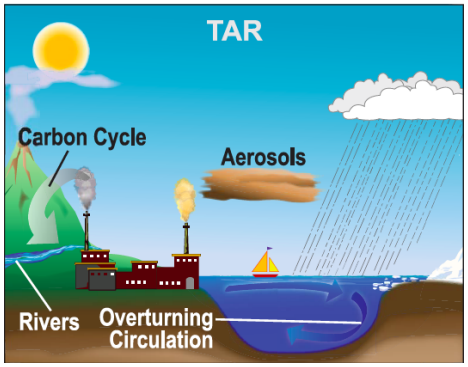
IPCC (1990)



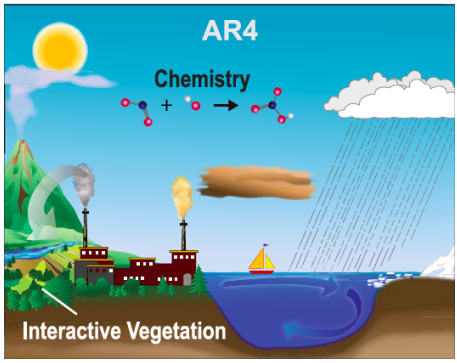
IPCC (1995)



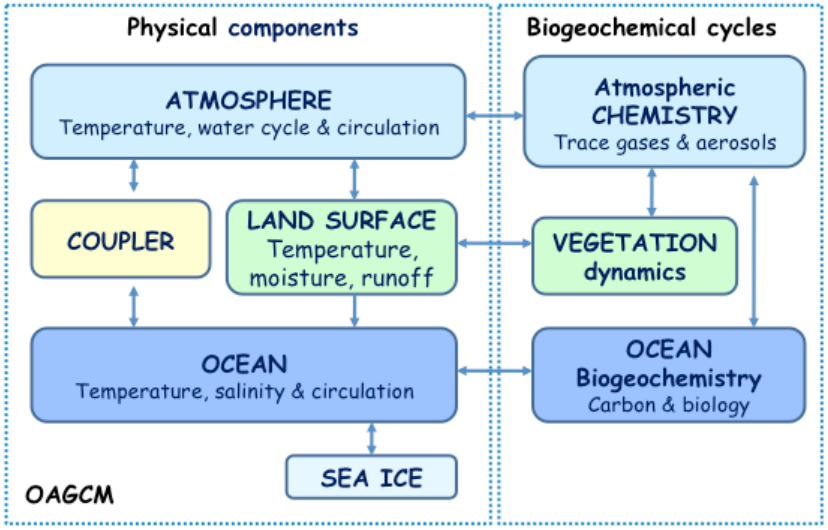
IPCC (2001)



IPCC (2007)



ESMs



Basic physical laws
Based on Navier-Stokes
Conservation of:
energy,
mass (air, water, carbon)
&
Parameterisations
clouds, radiation,
subgrid-scale processes

each ESM > 1000 man years: strong legacy

International coordination

International Coordination through WCRP

Coupled Model Intercomparison Project « CMIP »

Global Climate Models

Coordinated numerical experiments:

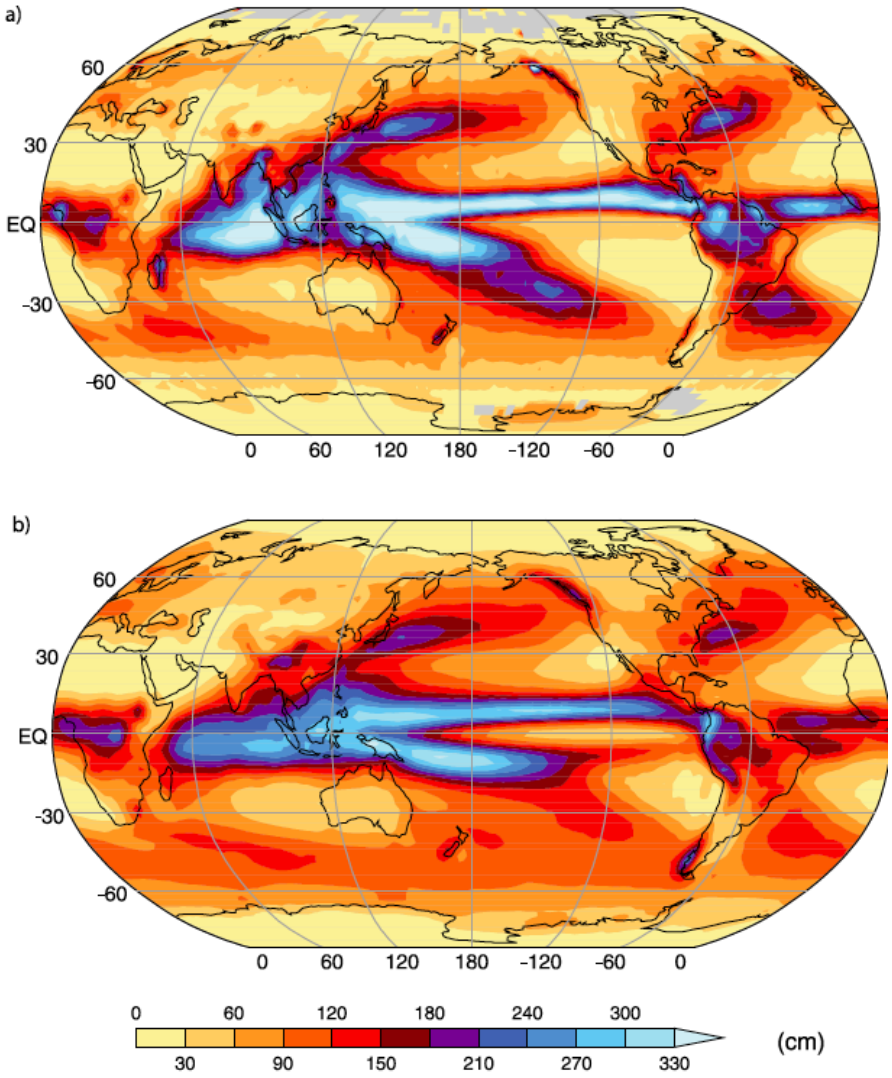
Model evaluation vs observed present & past climate
Future scenarios

Benchmarking & Multi-model analysis

Input to
Intergovernmental Panel on Climate Change (IPCC) Assessment Report (IPCC AR)

CMIP3 : IPCC AR4 (2007)
CMIP5: IPCC AR5 (to be published in 2013)

Evaluation of climate models



Annual mean precipitation (cm)

Observations 1980-1999
CMAP (Xie and Arkin)

IPCC (2007)

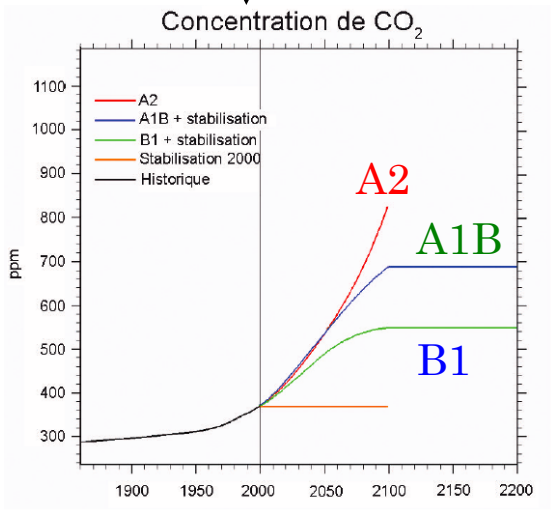
Simulated 1980-1999
Multi-model mean
CMIP3

Projections of future changes in climate

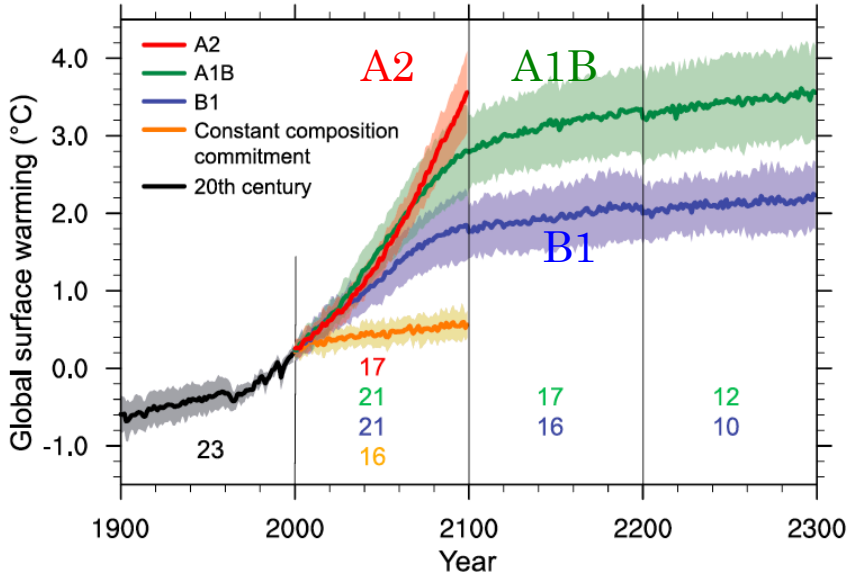
IPCC (2007)

Emission scenarios

CO2

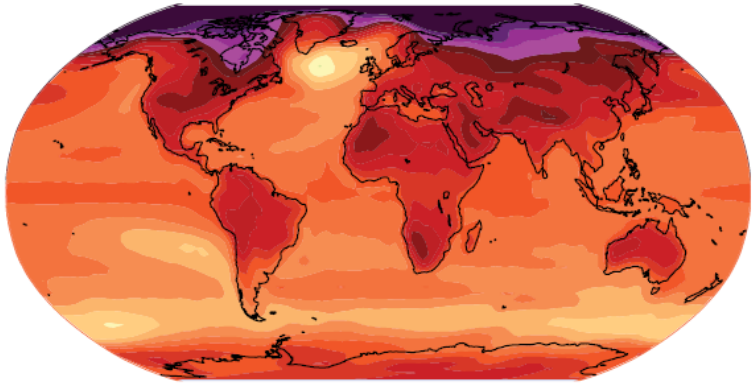


Temperature



A2

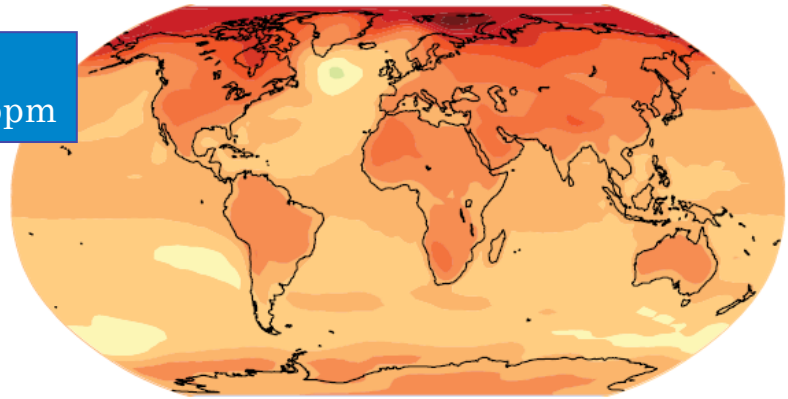
2100
CO_2 ~ 850 ppm



B1

2090 - 2099

2100
CO_2 ~ 550 ppm



CMIP5- Coupled Model Intercomparison Project
global climate models

WCRP World Climate Research Programme

Near-term experiments (10-30 years)

Decadal Prediction
(initial conditions)
Hindcasts/ Future

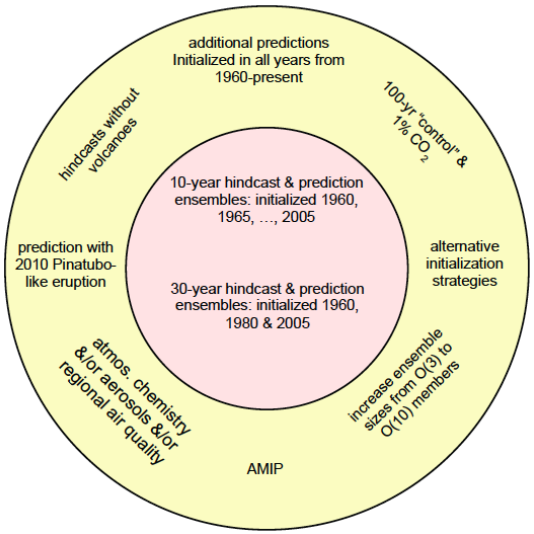
AOGCMs

Long-term experiments (century)

Future projections
Model evaluation
Understanding

AOGCMs & ESMs

Taylor et al., (BAMS, 2011)



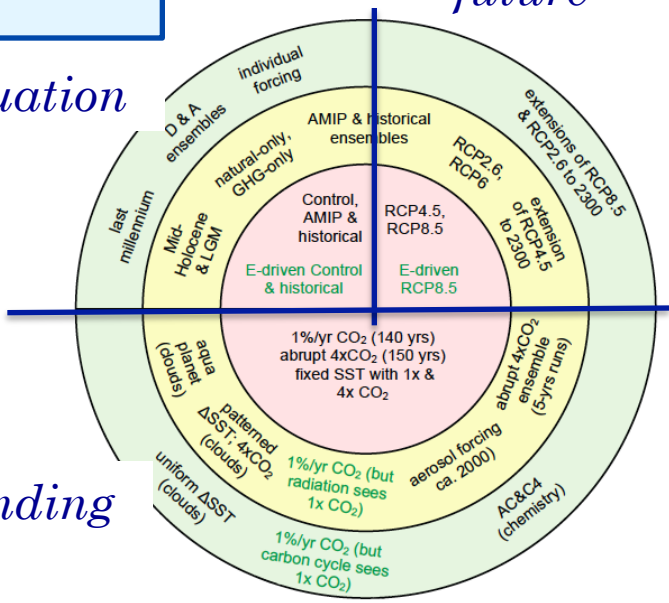
Large number of experiments

Minimum required 2300 years
Total > 10 000 years
IPSL 18 000 simulated years

future

evaluation

understanding



Status of CMIP5 experiments

27 modelling groups
58 models

1 Canada

CCCma	CanAM4 CanCM4 CanESM2
NSF-DOE-NCAR	CESM1(BGC) CESM1(CAM5) CESM1(CAM5.1, FV2) CESM1(FAST CHEM) CESM1(WACCM)
NCAR	CCSM4
NOAA GFDL	GFDL-CM2.1 GFDL-CM3 GFDL-ESM2G GFDL-ESM2M GFDL-HIRAM-C180 GFDL-HIRAM-C360
NASA GMAO	GEOS-5
NASA GISS	GISS-E2-H GISS-E2-H-CC GISS-E2-R GISS-E2-R-CC
NCEP	CFSv2-2011

6 USA

NCC	NorESM1-M NorESM1-ME
MPI-M	MPI-ESM-LR MPI-ESM-MR MPI-ESM-P
MOHC	HadCM3 HadGEM2-A HadGEM2-CC HadGEM2-ES
EC-EARTH	EC-EARTH
IPSL	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR
CNRM-CERFACS	CNRM-CM5
CMCC	CMCC-CESM CMCC-CM CMCC-CMS
INM	INM-CM4

7 in Europe

InfraStructure for the European Network for Earth System Modelling
is-enes

5 China / 1 Korea

LASG-IAP	FGOALS-gl FGOALS-s2
LASG-CESS	FGOALS-g2
GCESS	BNU-ESM
FIO	FIO-ESM
BCC	BCC-CSM1.1(m) BCC-CSM1.1
NIMR/KMA	HadGEM2-AO
NICAM	NICAM.09
MRI	MRI-AGCM3.2H MRI-AGCM3.2S MRI-CGCM3
MIROC	MIROC-ESM MIROC-ESM-CHEM
MIROC	MIROC4h MIROC5
CSIRO-QCCCE	CSIRO-Mk3.6.0
CSIRO-BOM	ACCESS1.0 ACCESS1.3

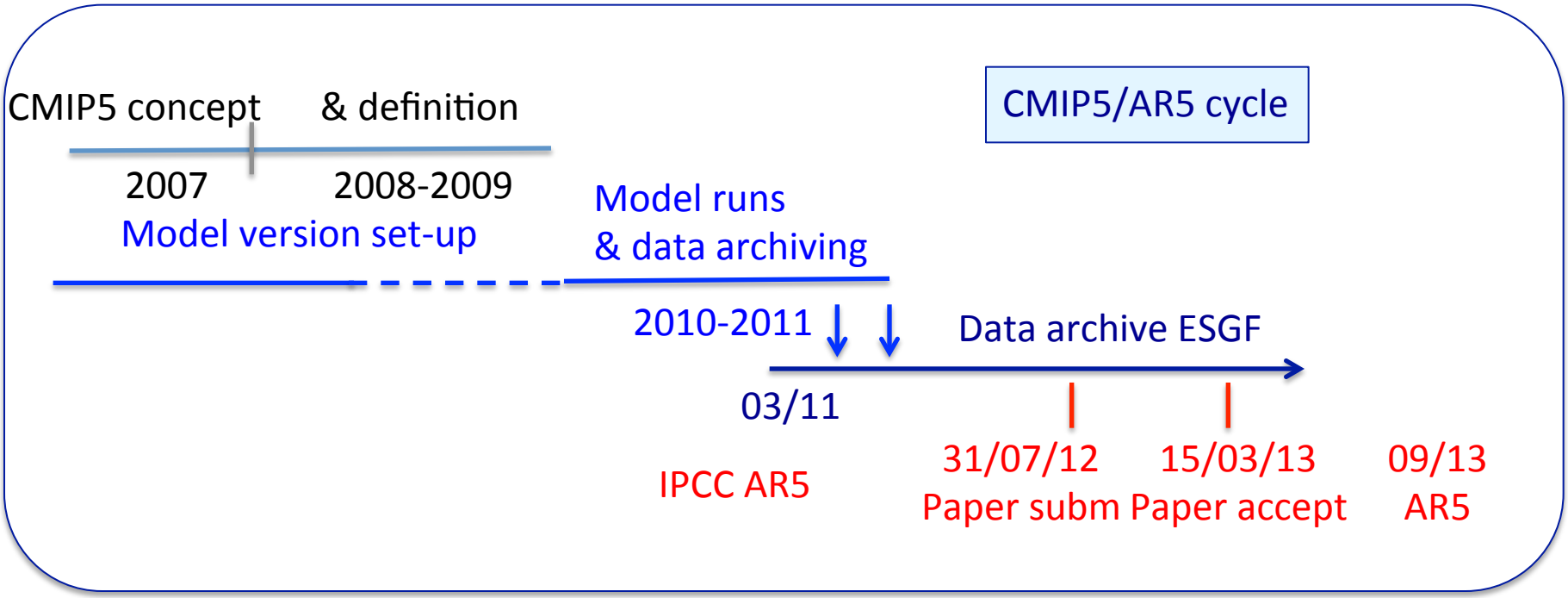
1 Russia

4 Japan

2 Australia

Coupled Model Intercomparison Project – CMIP5

International community under strong pressure



CMIP international data bases

Ease access for model analysis and impact studies

CMIP3

36 TB

Central archive at PCMDI (LLNL)

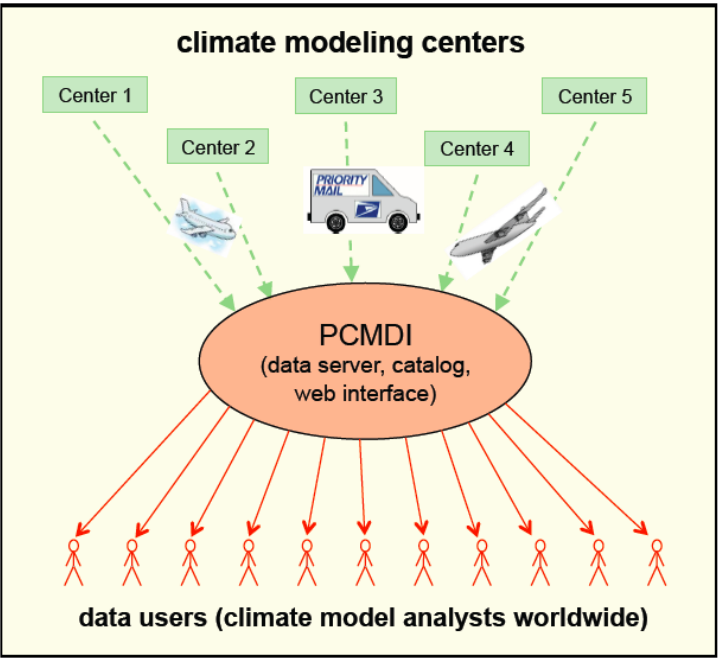
Earth System Grid

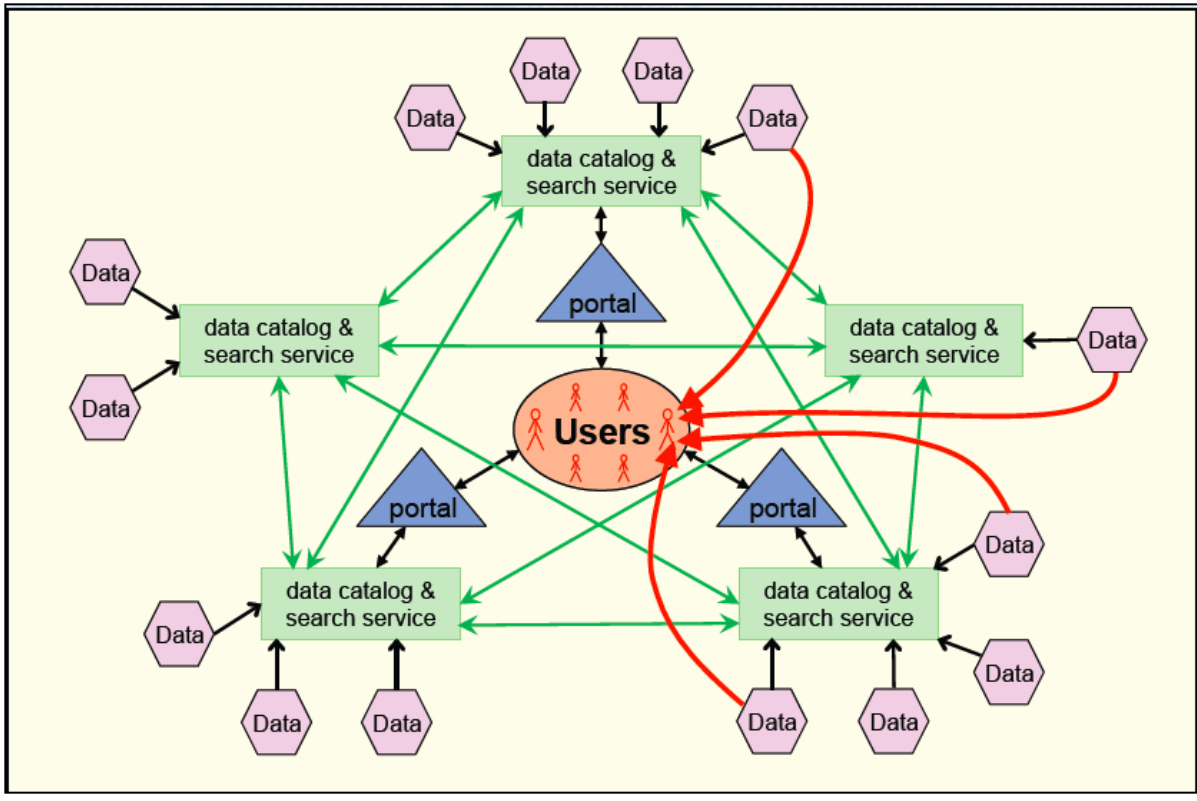
CMIP5

1.8PB (up to 3 PB)

Distributed data infrastructure

Earth System Grid Federation





Courtesy K. Taylor (PCDMI)

European coordination

ENES

European Network for Earth System modelling

http://enes.org

A network of European groups in climate/Earth system modeling

Launched in 2001 (MOU)

Ca 50 groups from academic, public and industrial world

<u>Main focus</u> :
discuss strategy to accelerate progress in climate/Earth system modelling and understanding

Several EU projects

Collaboration with PRACE

IS-ENES
Infrastructure for ENES

European projects
2009-2013; 2013-2017

Infrastructure
Models & their environment
Model data (ESGF)
Interface with HPC ecosystem

<u>Users</u> :
Climate modelling community
Impact studies

Infrastructure Strategy for the European Earth System Modelling Community 2012-2022

Drivers : Science & Society

From understanding to development of "Climate Services"

Key science questions

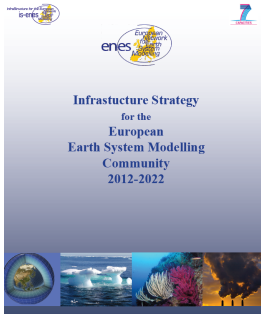
Q1. How predictable is climate on a range of timescales ?

Q2. What is the sensitivity of climate and how can we reduce uncertainties ?

Q3. What is needed to provide reliable predictions of regional climate changes ?

Q4. Can we model and understand glacial-interglacial cycles ?

Q5. Can we attribute observed signals to understand processes ?



Writing team:
J. Mitchell, R. Budich, S. Joussaume, B. Lawrence & J. Marotzke
52 contributors from BE, CZ, DE, DK, FI, FR, IT, NO, SE, SP, UK

Q1. How predictable is climate at different time scales ?

HPC
Data assimilation
Large ensemble runs
Resolution

Observations

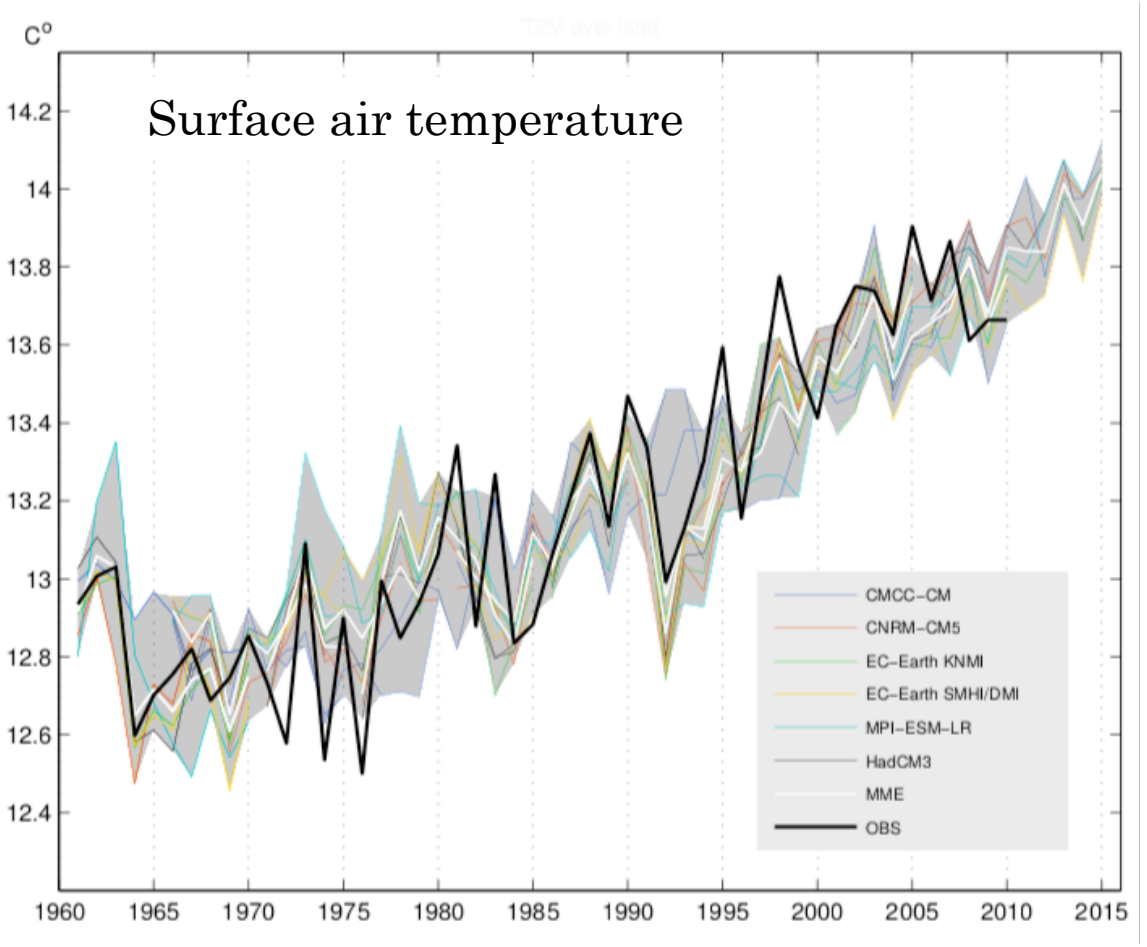
Multi model

COMBINE EU Project, courtesy of A. Bellucci (CMCC)

Multi-model decadal predictions

5 European models, 10 yr simulations at every 5 years

With ocean initial conditions



1960 2015

Q2. What is the sensitivity of climate and can we reduce uncertainties ?
Feedbacks (eg. clouds, carbon cycle), nonlinear behaviours

HPC
Ensemble experiments
(eg. process studies)

Temperature change to 2 x CO_2

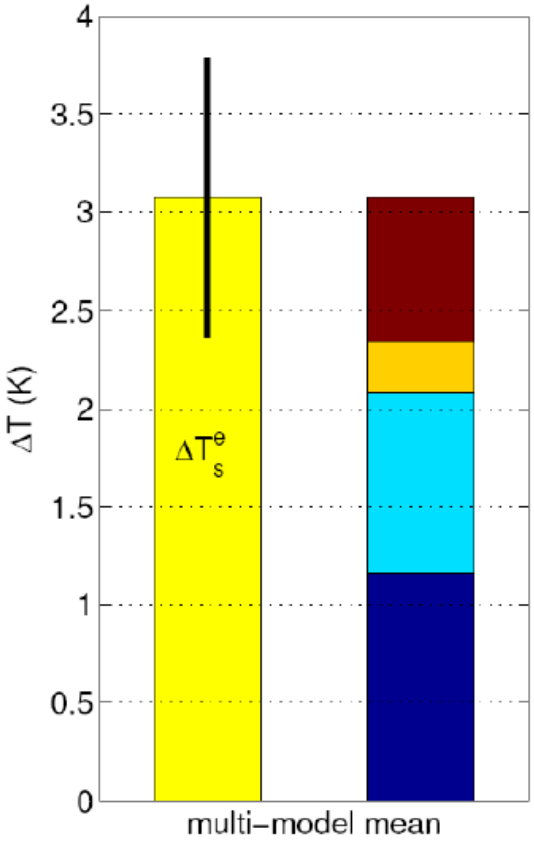
Uncertainty to cloud feedbacks

CMIP3 (AR4)

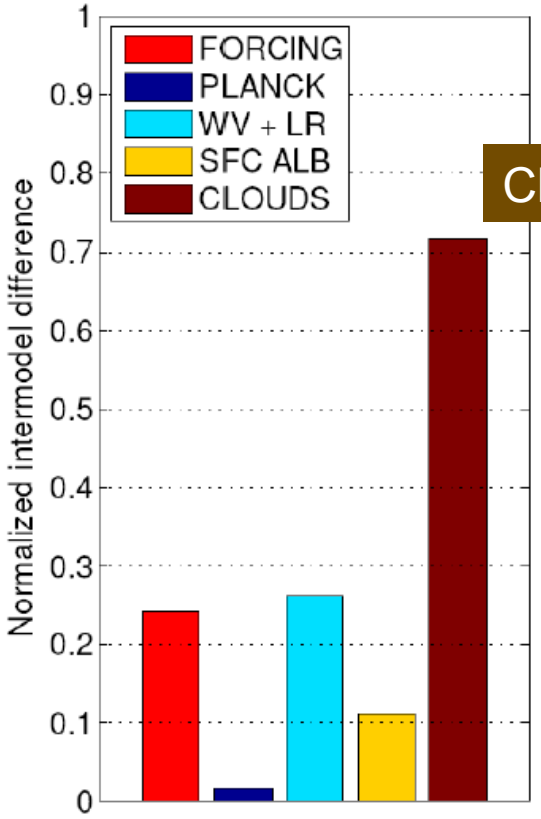
2 to 4.5 °C
mean 3°C

Dufresne & Bony
J. Climate, 2008

Multi-model mean



Inter-model

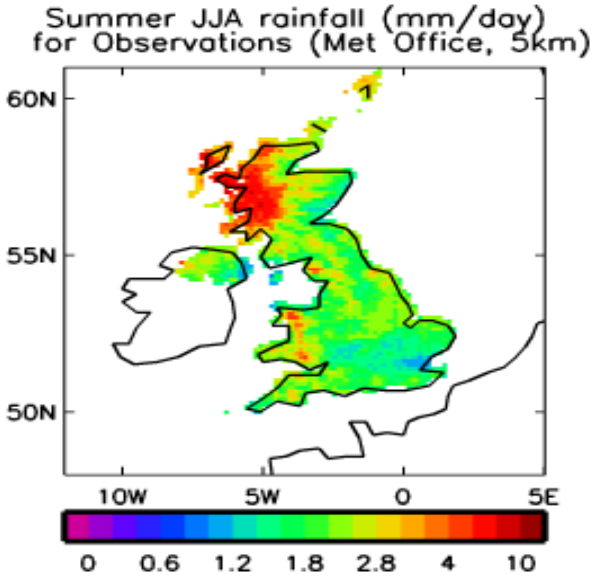


Clouds

Q3. What is needed to provide reliable predictions/projections of regional climate changes ?

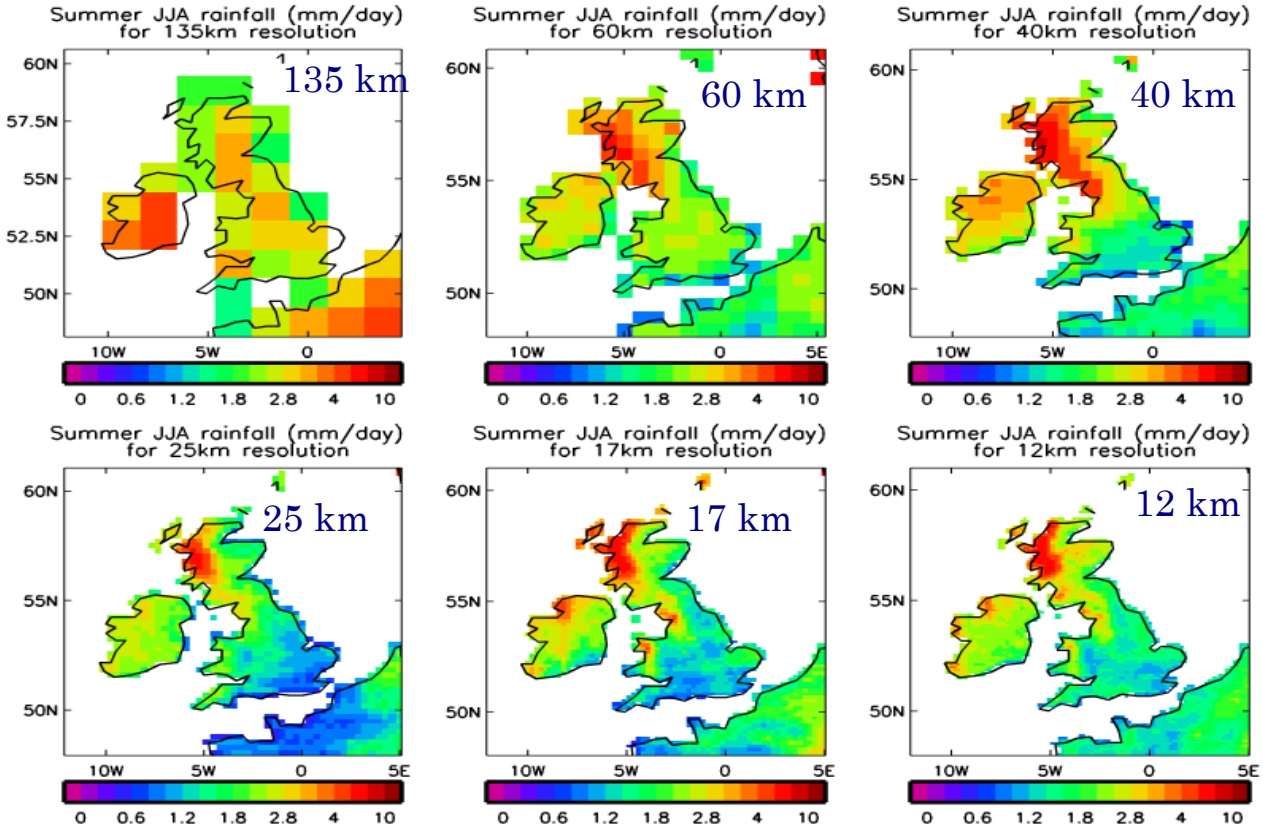
HPC:
Spatial Resolution
Ensemble runs
(internal variability, parameterisations)

Observations



Summer precipitation 2005 Simulations

global climate model HADGEM3
Resolutions 135km → 12km
PRACE UPSCALE project

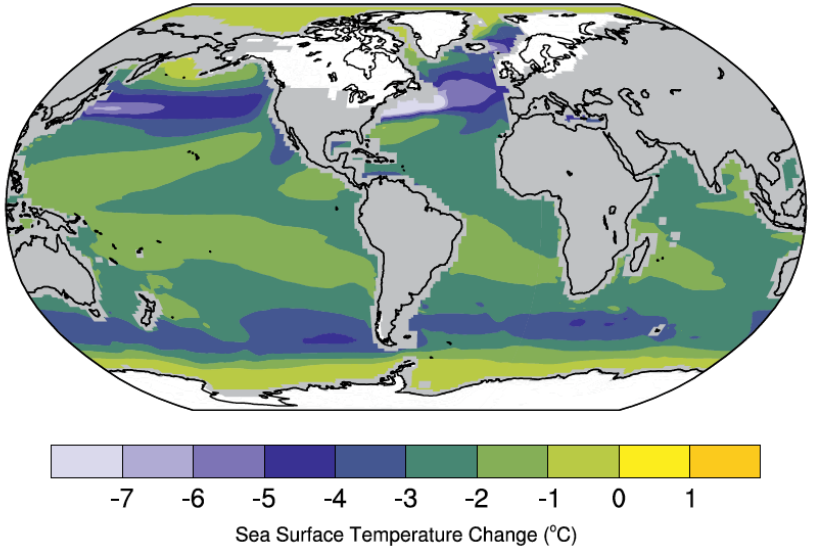


Courtesy of PL Vidale (NCAS) & M. Roberts (MO/HC)

Q4. Can we model and understand glacial-interglacial cycles and better constrain model sensitivity using the past ?

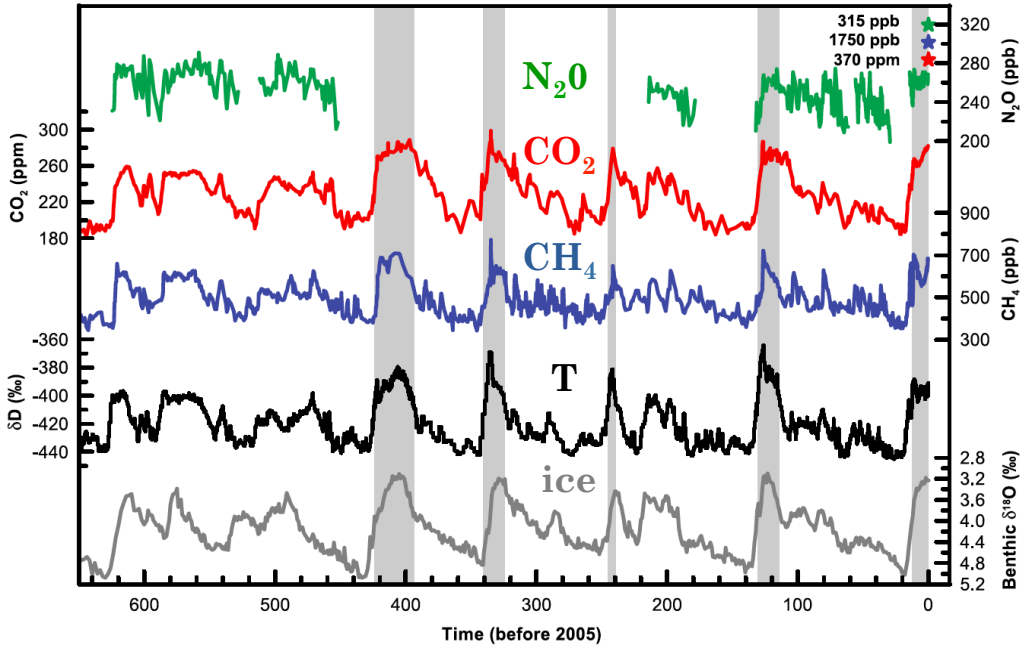
HPC
Duration
(1000 to 100 000 simulated years)
Complexity

Last Glacial Maximum (21 000 years BP)
Simulations (PMIP2)



IPCC (2007)
Braconnot et al. (Clim Past 2007)

Observed past 600 000 years



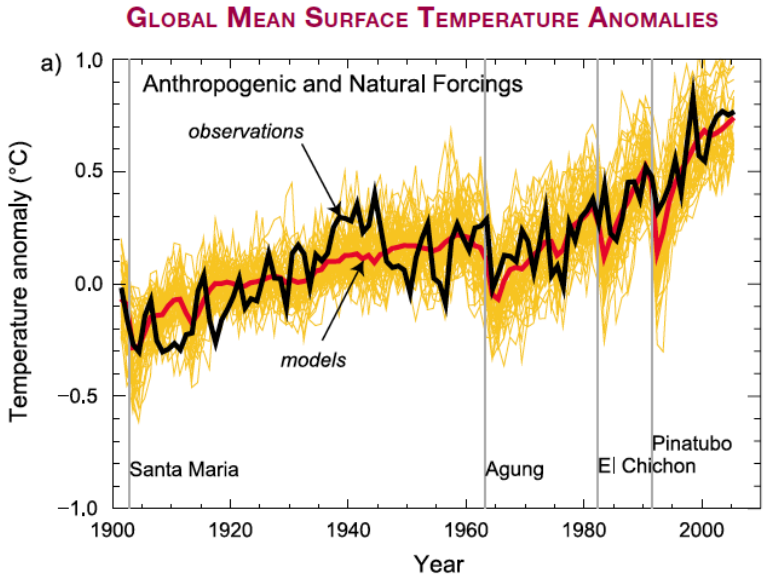
In thousand of years

Source : EPICA community members, Nature 2004

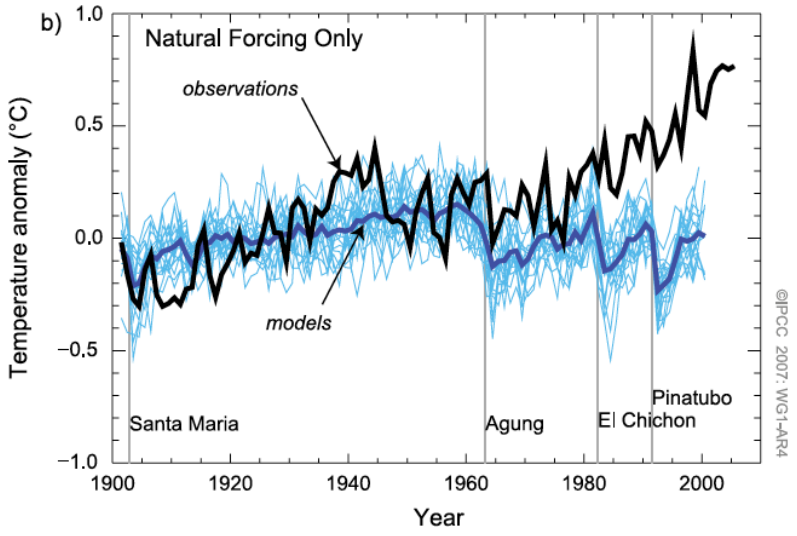
Q5. Can we attribute observed signals to understand processes ?
Globally ? Extreme events ?

HPC
Large ensemble runs

For extreme events:
Spatial resolution
Fast availability of computing power



IPCC (2007)



observations

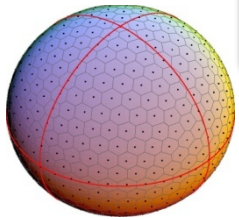
Simulations
Natural & anthropogenic forcings

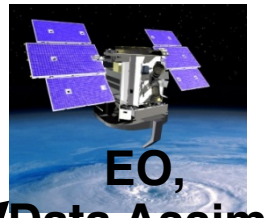
Simulations
Natural forcings

Infrastructure Strategy Roadmap

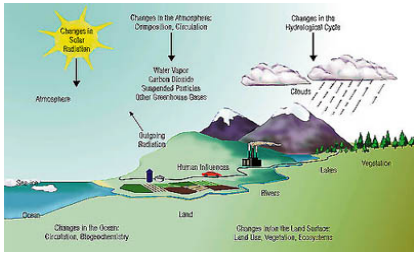
Needs for HPC

x 3 ≈ x 27
x 100 ≈ x 10^6





EO,
Data Assim.



x 5-10

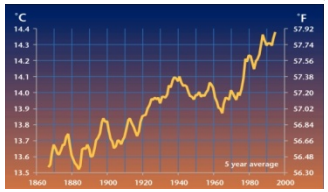
Resolution

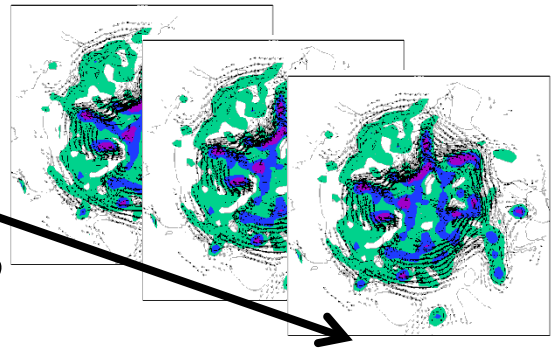
Computing Resources

Complexity

Duration and/or Ensemble size

From Jim Kinter,
the World Modelling Summit, 2008



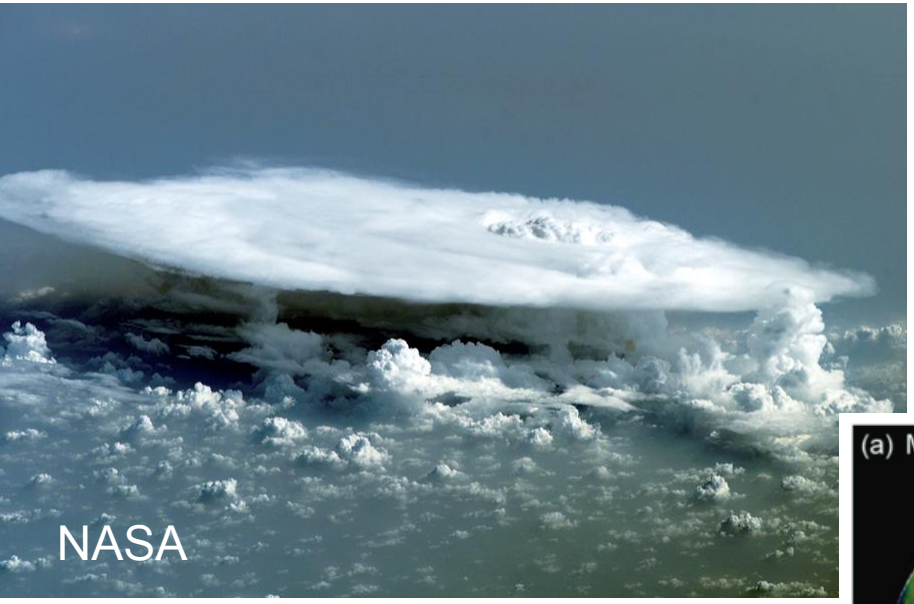


ensemble: x 10

duration: x 10-100

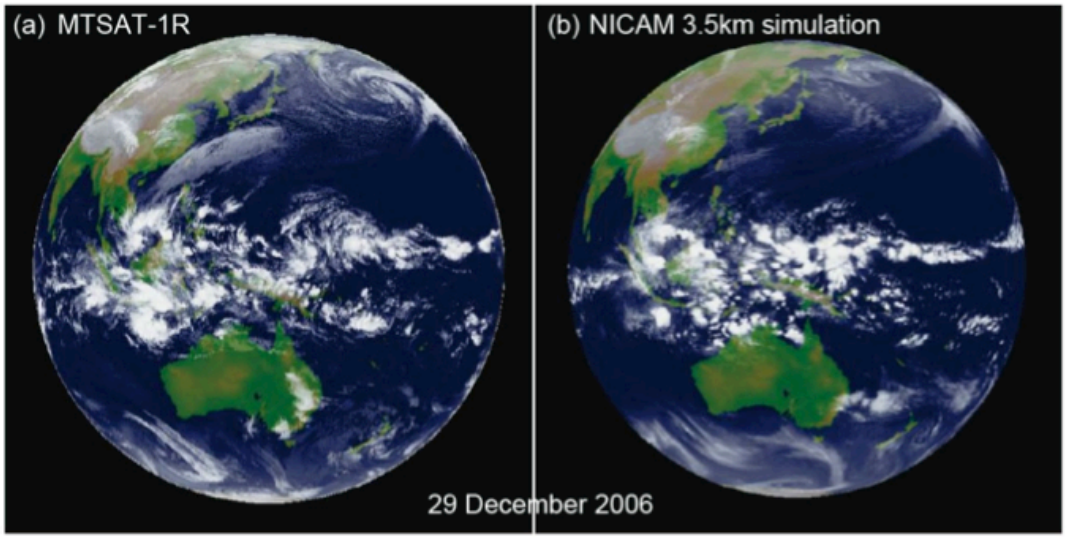
Infrastructure Strategy Roadmap

A grand challenge :
Towards ≈ 1 km scale for atmosphere
resolving deep convective clouds
in global climate models



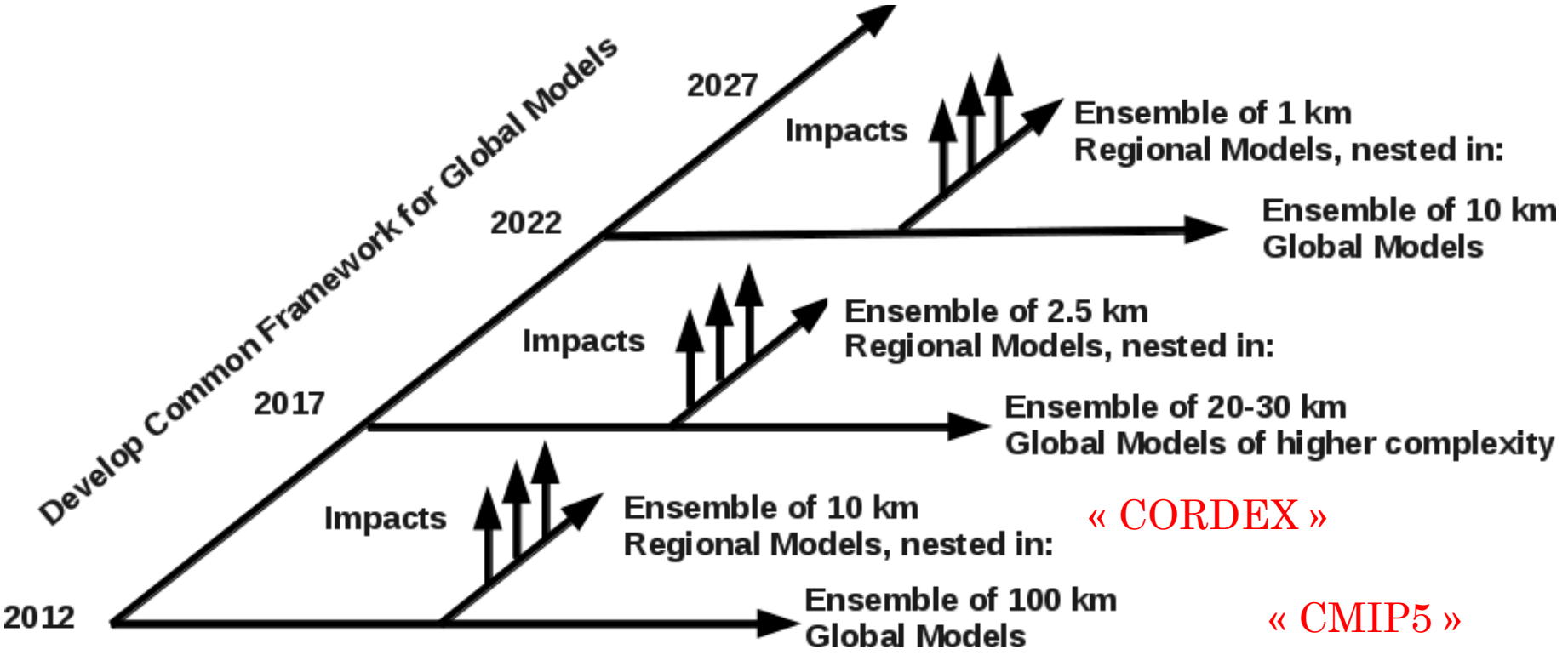
Miura et al, Science 2007
3.5 km, few months,
Earth Simulator

Satellite obs NICAM model



Infrastructure Strategy Roadmap

Steps towards 1 km scale global climate models

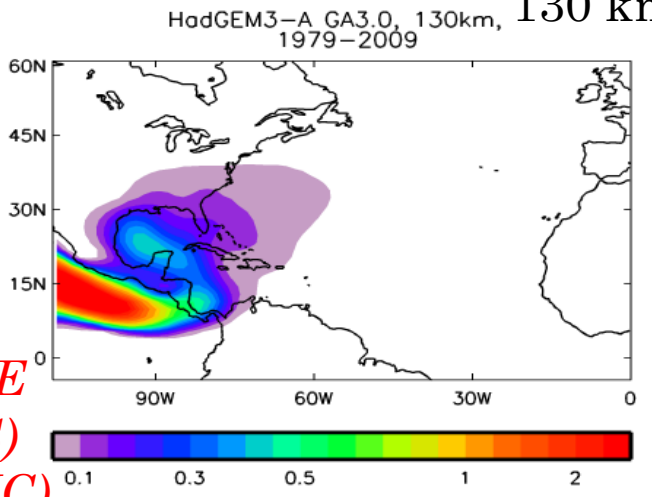


Increasing spatial resolution:
an important step to represent extreme events

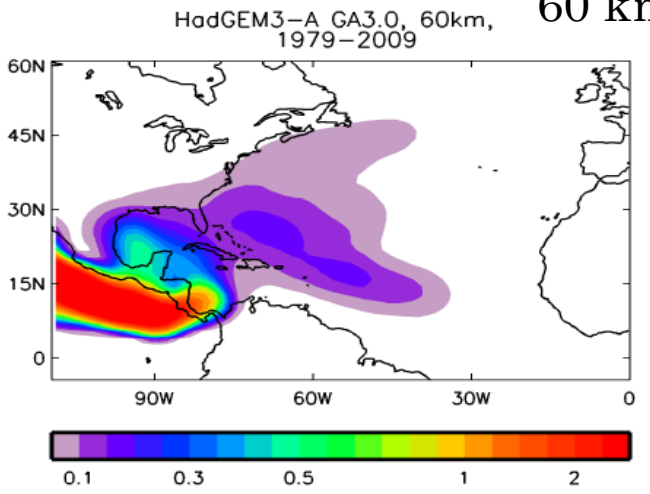
Impact on simulated tropical cyclones

HADGEM3
PRACE UPSCALE
PL Vidale (NCAS)
M. Roberts (MO/HC)

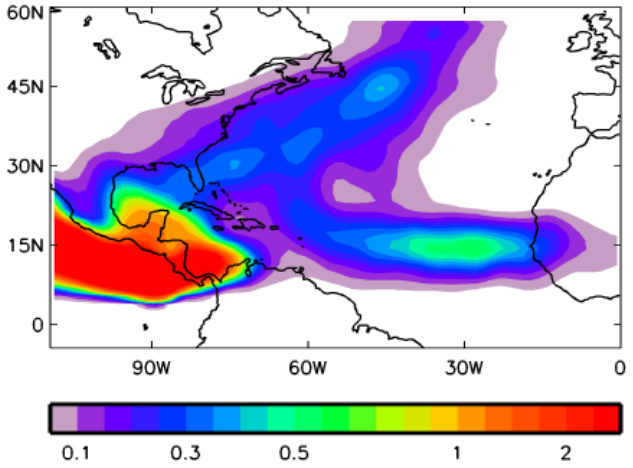
130 km



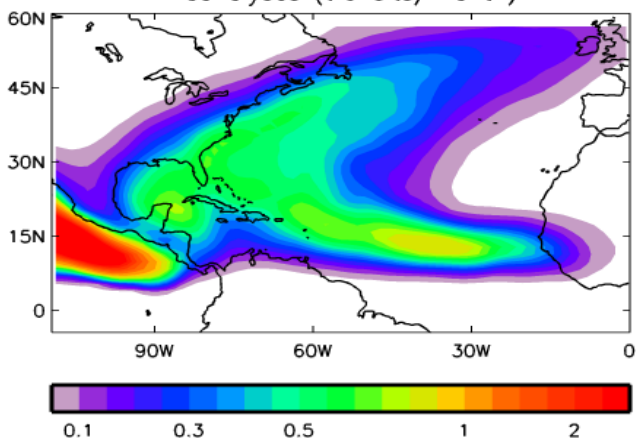
60 km



25 km



Obs



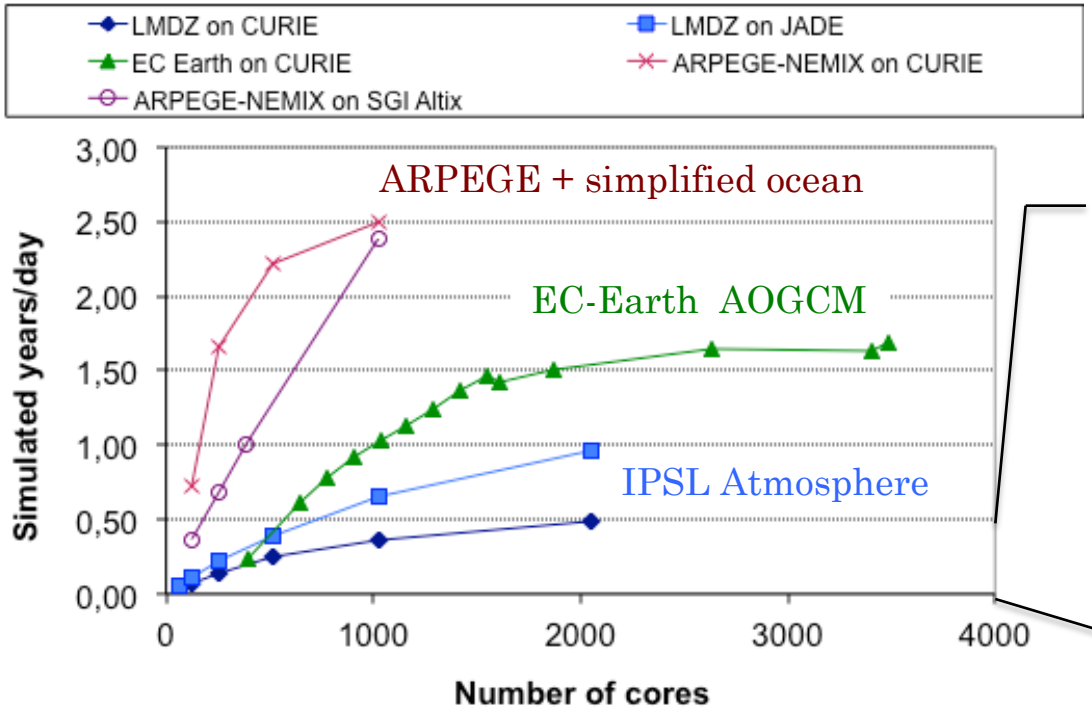
HPC and data issues

Scalability issue

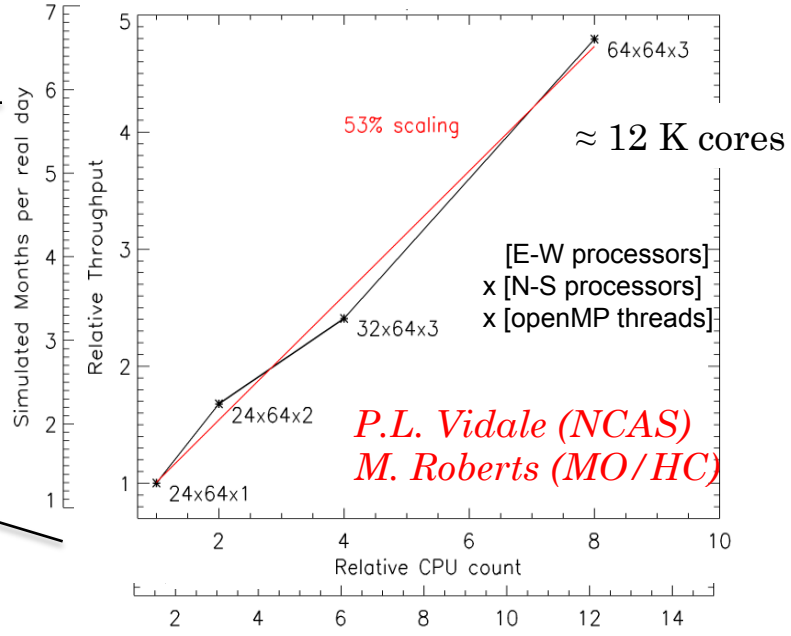
Joint Weather and Climate Research Programme

A partnership in climate research

Scalability tests at resolution 25-30 km for the atmosphere



HadGEM3 – CRAY XE6 Hector



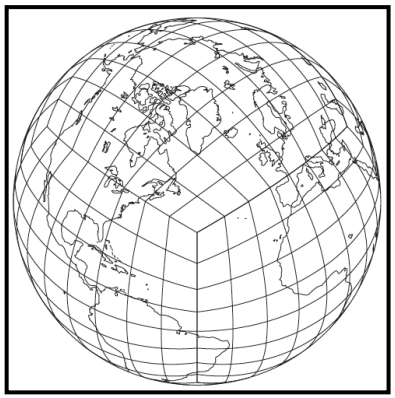
CPUs/1000

G. Riley et al., IS-ENES

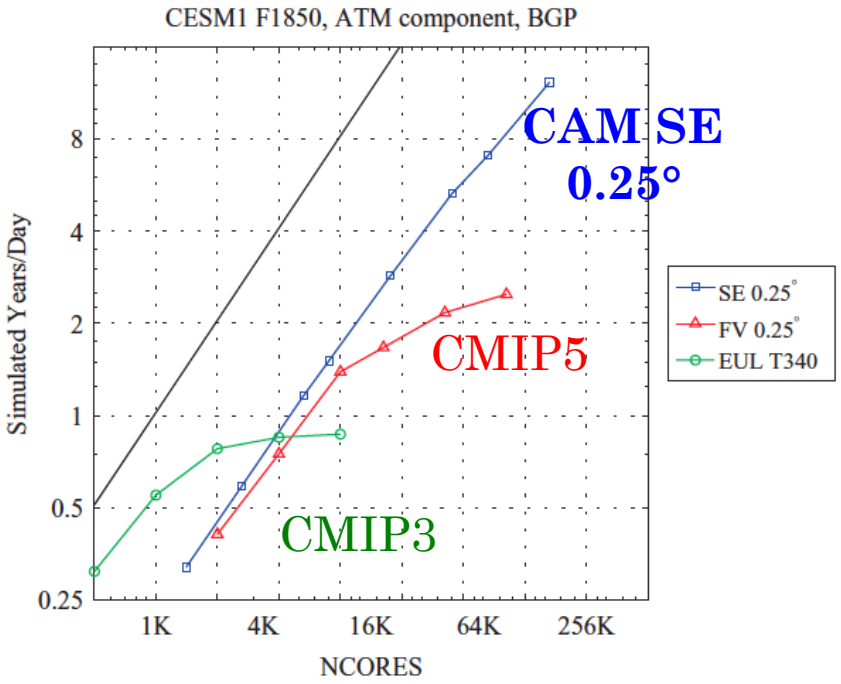
Need to revisit dynamical cores

On-going international projects:
G8 exascale project ICOMEX
Dynamical Core MIP

Cubed-sphere
(CAM-SE)



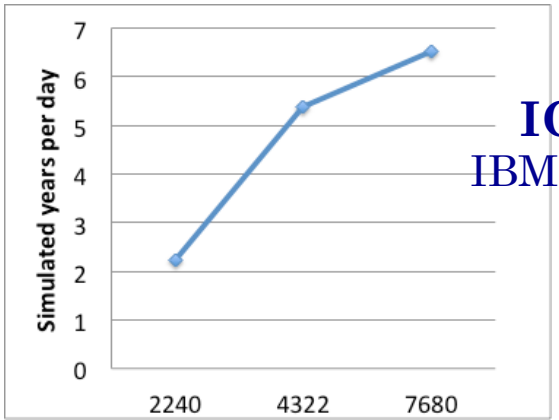
CESM1, 0.25°, BGP



Collaboration NCAR-Sandia,
Dennis et al. (IJ HPC appl, 2012)

Icosahedric grids
ICON (DE)
DYNAMICO (FR)

Court. L. Linardikis (MPI-M)



ICON 35 km
IBM Power P6 nodes

Need to revisit I/O

Storage of history files:
Variables (x, y, z, t) & diagnostics
at every simulated 3 hr to 24 hr

Examples:

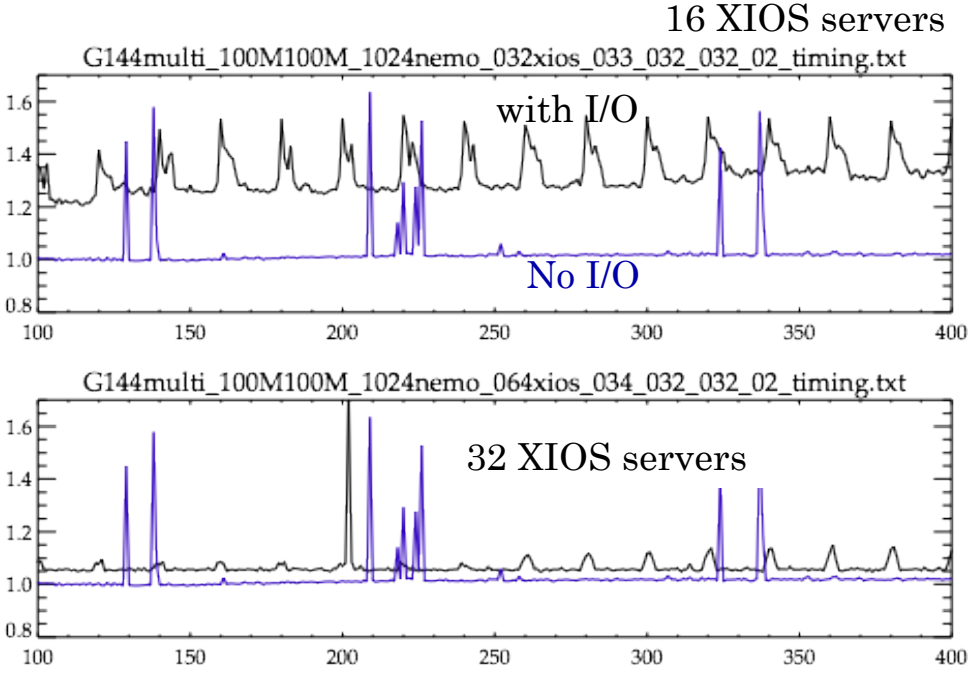
Resolution ≈ 150 km (CMIP5) :
13 TB for 150 simulated years
(IPSL)

Resolution ≈ 25 km
(UPSCALE project)
2 TB for 1 simulated year (NCAS)

transfer of 1-1.5 TB /day
total storage ≈ > 600 TB

Extreme case:
NEMO ocean model (4322x2882)
36 simulated hours
NEMO : 1024 proc. MPI

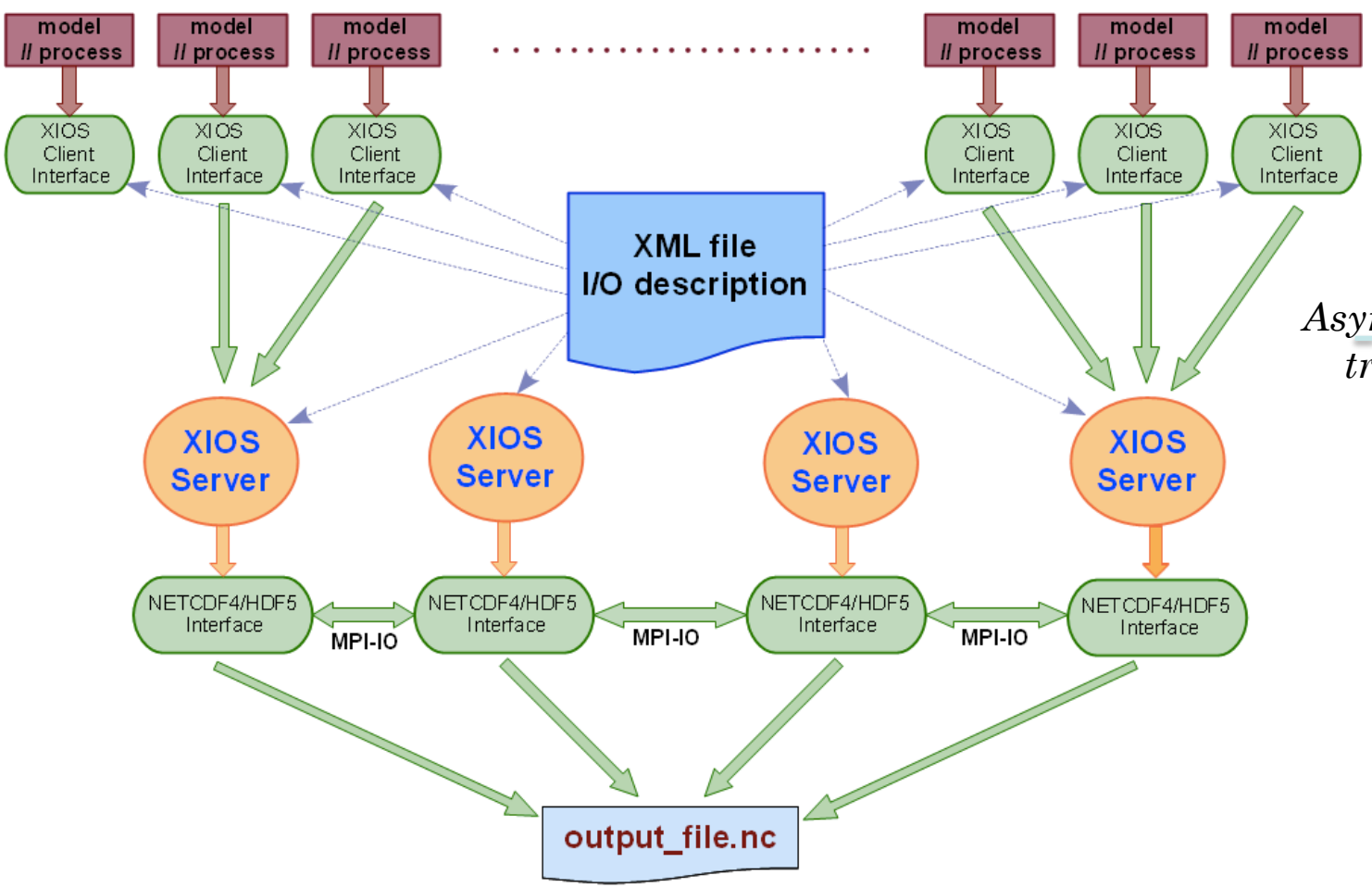
I/O every hour (246 GB)
with no I/O



Y. Meurdesoif & S. Masson (IPSL)

Parallel IO server

XIOS *Y. Meurdesoif (IPSL)*



Model // process

XIOS Client interface

« client »

Asyncronous transfer

« server »

XIOS Server

NETCDF4/ HDF5 interface

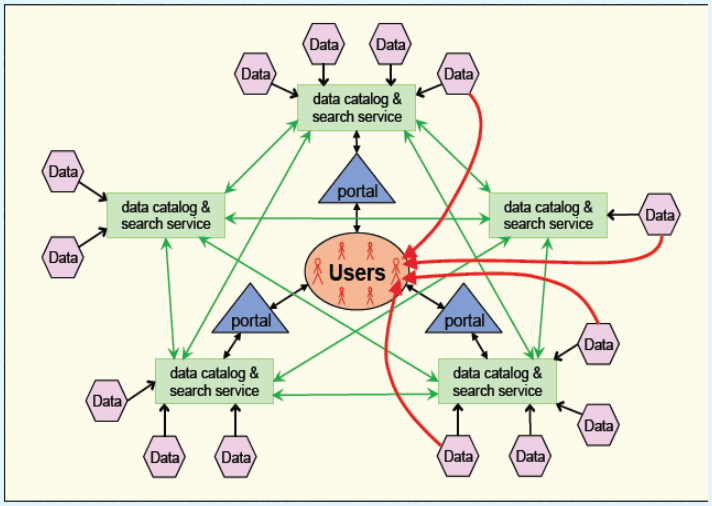
output_file.nc

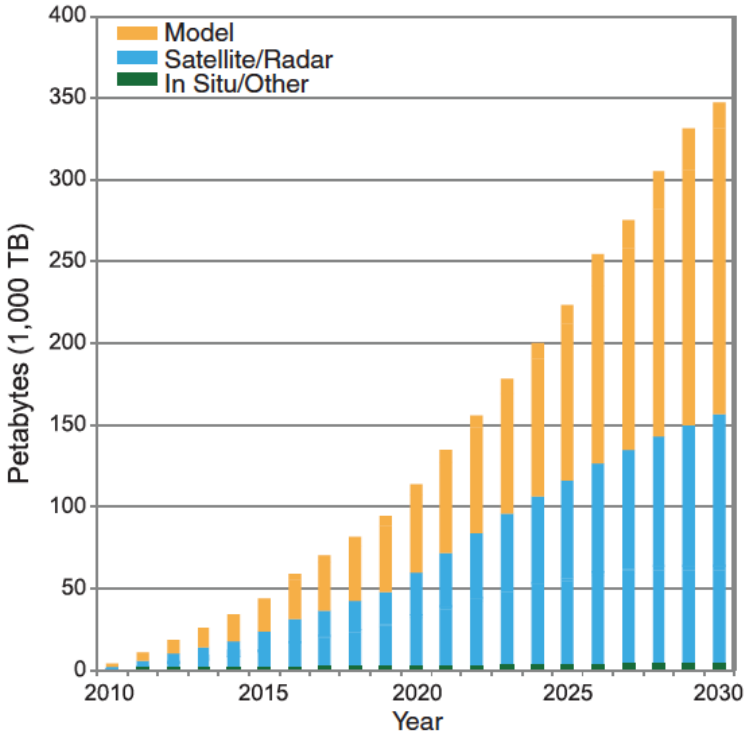
Data issue

On-going international project:
G8 ExArch

CMIP5 ≈ 2 to 3 PB
CMIP6 x 30 ?

Earth System Grid Federation





Overpeck et al. (Science 2011)

Model data expected to grow exponentially (resolution, number of simulations)

Strong demand from society : « Climate Services »

Need to have analysis and computation where data are

Conclusions

- **Increasing demand** from society:
  more reliable predictions of climate change for adaptation

- **Needs for HPC:**
  Scalability: Increased resolution, Number of experiments, complexity
  Duration of experiments - HOW ?

- **HPC facilities adapted to our needs** :
  IO, Data storage, physical network
  Stability of computing environment: development/evaluation/ production runs

- **Issue of preparing future generation of climate models for exascale**
  Computing (exascale) & society (regional, decadal)

- **International dimension** is important (WCRP, ESGF)

- **Increasing role of coordinated infrastructure** in climate modelling:
  models, data and computing (ENES, US Report NAS)

Thank you
SeaWIFS Project (NASA/GSFC et Orbimage)