Climate Modelling in EUDAT

EUDAT User Forum

March 7th + 8th, 2012 in Barcelona

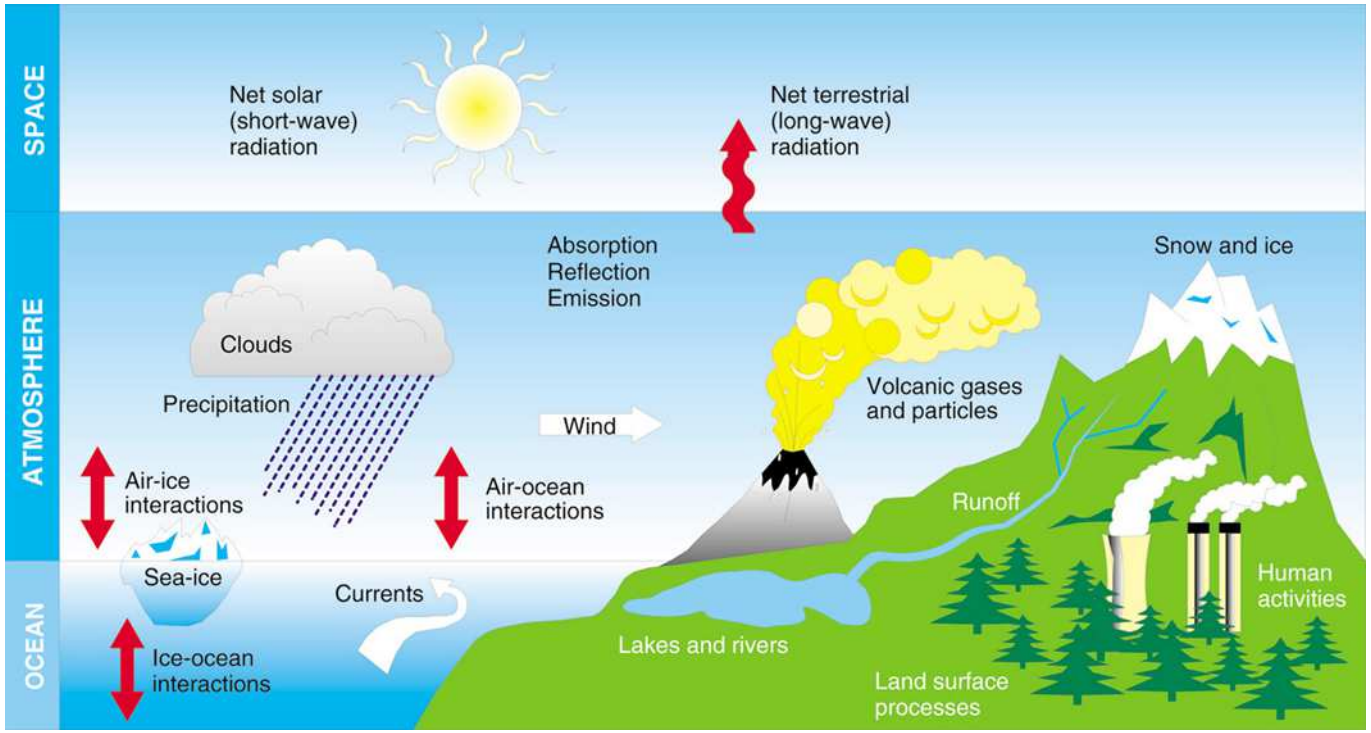
Michael Lautenschlager (DKRZ)
for the ENES Community

DKRZ
DEUTSCHES
KLIMARECHENZENTRUM

What is ENES?

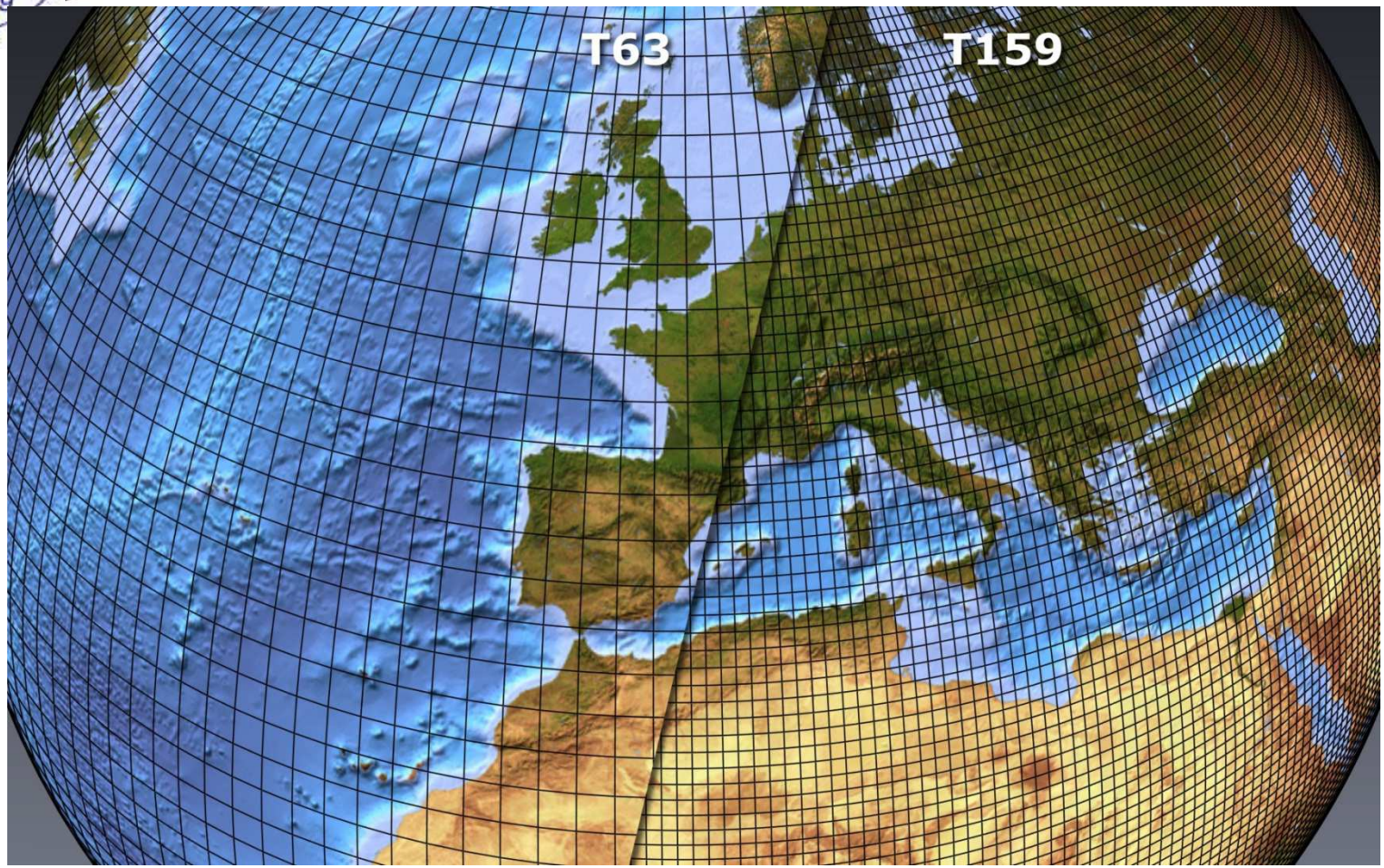
- European Network for Earth System Modelling (ENES):
  - Within Europe, ENES gathers together the European climate/Earth system modeling community. This community is strongly involved in the assessments of the Intergovernmental Panel on Climate Change (IPCC) and provides the predictions on which EU mitigation and adaptation policies are elaborated.
- Around 50 partners spread all over Europe signed the ENES MoU
- Infrastructure development within FP7 project IS-ENES
  - Runtime 2009/03/01 – 2013/02/28 with 19 partners
  - Project website: http://is.enes.org
- ENES information portal: https://verc.enes.org/
  - It provides information and services to foster intricate simulations of the climate system using high performance computers as well as the distributions and dissemination of data produced by such simulations.

Climate Modelling and CMIP5



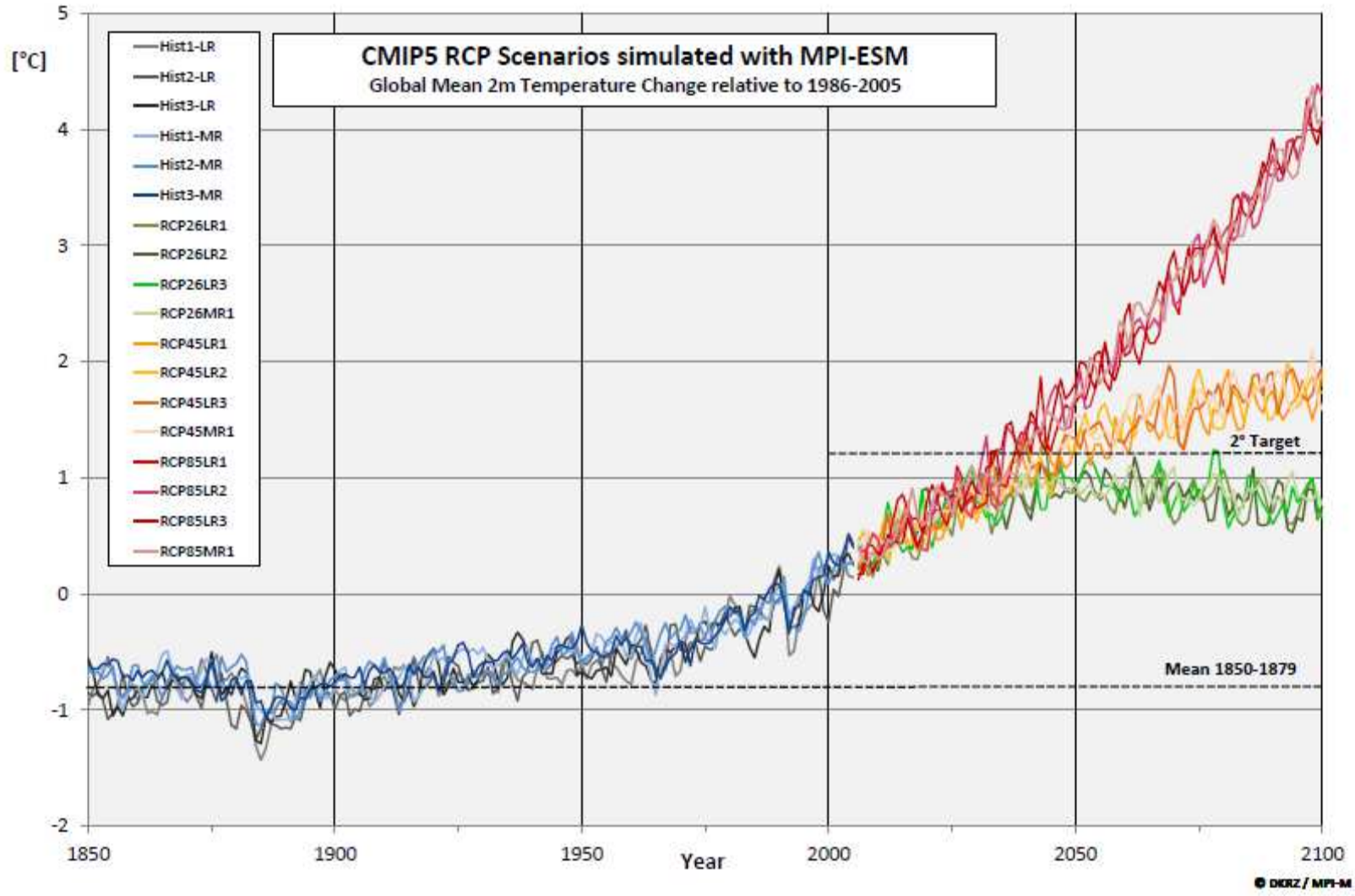
Modelling:
a)Climate phenomena → mathematical formulation
b)Mathematical formulation → algorithms

Climate Modelling and CMIP5

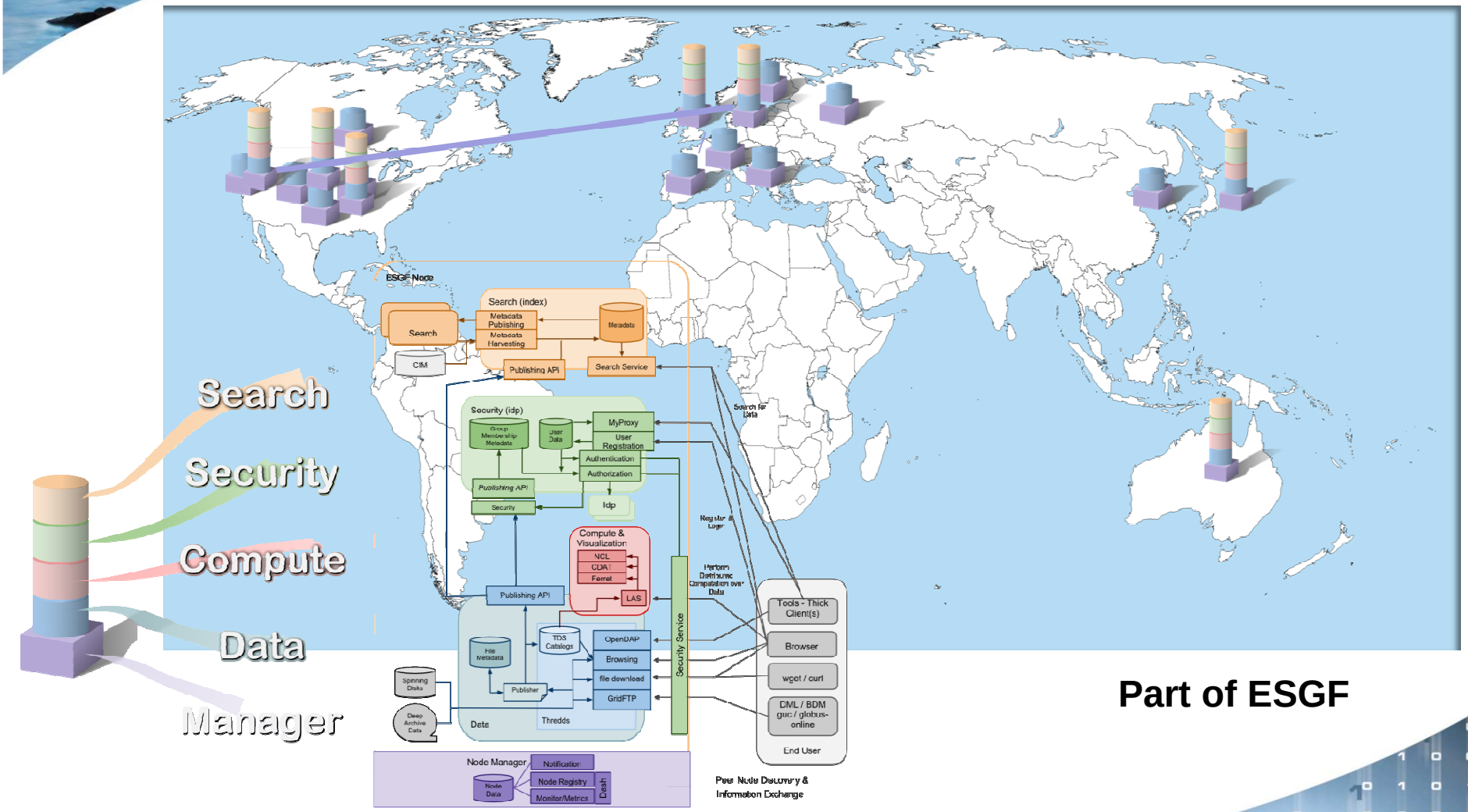


Refinement of the spatial resolution by a factor of 2 increases model data production by a factor of 8.

Results


CMIP5 RCP Scenarios simulated with MPI-ESM
Global Mean 2m Temperature Change relative to 1986-2005
[°C]
Hist1-LR
Hist2-LR
Hist3-LR
Hist1-MR
Hist2-MR
Hist3-MR
RCP26LR1
RCP26LR2
RCP26LR3
RCP26MR1
RCP45LR1
RCP45LR2
RCP45LR3
RCP45MR1
RCP85LR1
RCP85LR2
RCP85LR3
RCP85MR1
2° Target
Mean 1850-1879
Year
5
4
3
2
1
0
-1
-2
1850
1900
1950
2000
2050
2100
© DKRZ / MPI-M

ENES Data Federation



Part of ESGF

ENES Data Federation

Key characteristics

- **Data:** Earth System Model Data and Observation Data (mostly netcdf4 / hdf5 data)
- **Data Volumes:** 2011: ~ 0.5 PByte 2012: > ~2 PByte .. (mainly CMIP5)
- **Repositories:** HPC center attached and long term funded data centers (e.g. BADC, DKRZ)
- **Users**: Climate system scientists, climate impact scientists, „added value portals“: provisioning of derived data products for e.g. for policy makers
- **Infrastructure:** based on ESGF SW stack, SSO support, replication and versioning support, web and rich client access support, AAI is OpenId, SAML and PKI based
- **Status:** In production as part of international CMIP5 federation since Mid 2011
- **Perspective:** rapidly growing, will be basis for new climate modelling projects
- **Key future requirement**: Archives must support server side and data sub-setting (and pre-processing) in order to reduce network transfer
- **Data Processing:** modular software stack CDO (Climate Data Operators)

ENES Data Federation

- Distributed Archive ←→ User Requirements / User Access
- Two distinct „User Roles“:
- End User: want to access data
- Data Producer: want to make data visible / accessible through federation

End User:

•Climate scientists: o(1000)
  - model output evaluation,comparison,..
  - High data download volumes per user (Tbytes..)

- Climate impact researchers: o(10.000)
  - Specific lower volume data requests (specific variable, time, space constraints)
  - Server side data subsetting required
  - Derived data product management/storage is important

Data Producer:

•Modelers: o(100)
  - Want to provide results to large community via standardized metadata and data interfaces
  - Versioning and Data Replication and QC support needed

ENES Data Federation

- ENES data federation is integrated in international CMIP5 data federation and emerging national data federations
  - CMIP5 funded by the participating national partners
  - ENES funded by IS-ENES FP7 program – will go beyond CMIP5

- CMIP5 status:
  - 9 Gateways (Portals) – growing, 18 data nodes (Archives)
  - > 2 Pbytes of data expected, by now ingested ~ 1PByte
  - Used in production since Mid 2011, parallel Test-infrastructure
  - Based on ESGF open source SW stack
  - Status of deployments: see http://esgf.org/wiki/Cmip5Status

- ENES status:
  - 3 Gateways (Portals): DKRZ, BADC, 6 European data nodes
  - Replication between DKRZ and BADC
  - Downloads: ~150 TByte since June
  - Prepare to X Pbytes Regional Model Data to come
  - Access: https://verc.enes.org/data

Relation ENES and EUDAT

- What ENES provides to EUDAT
  - Best practises for management of large data volumes
  - Best practises for international data federations e.g.
    - ESGF software stack
    - Data quality control procedures
    - DataCite scientific data publication
  - Provide climate (model) data for interdisciplinary scientific use

- What ENES expects from EUDAT
  - Access to climate relevant data from other disciplines
  - Adaptation of ENES data federation to best practices from other disciplines
  - Shared development and maintenance of common, cross-community data services e.g.
    - Data replication
    - Security infrastructure (AAA)
    - Transparent data identification (metadata seacrh) and access across changing storage locations (PID and Handle System)