

Support of IPCC Data Distribution Centre

Data Repositories' Day
November 9, 2018, Gaborone, Botswana

B. Chen¹, M. Stockhause², M. Jukes³, A. Pirani⁴, E. Poloczanska⁵, T. Waterfield⁴
¹CIESIN/SEDAC, ²DKRZ, ³CEDA, ⁴IPCC WG1, ⁵IPCC WG2

IPCC DDC (Data Distribution Centre) – ipcc-data.org

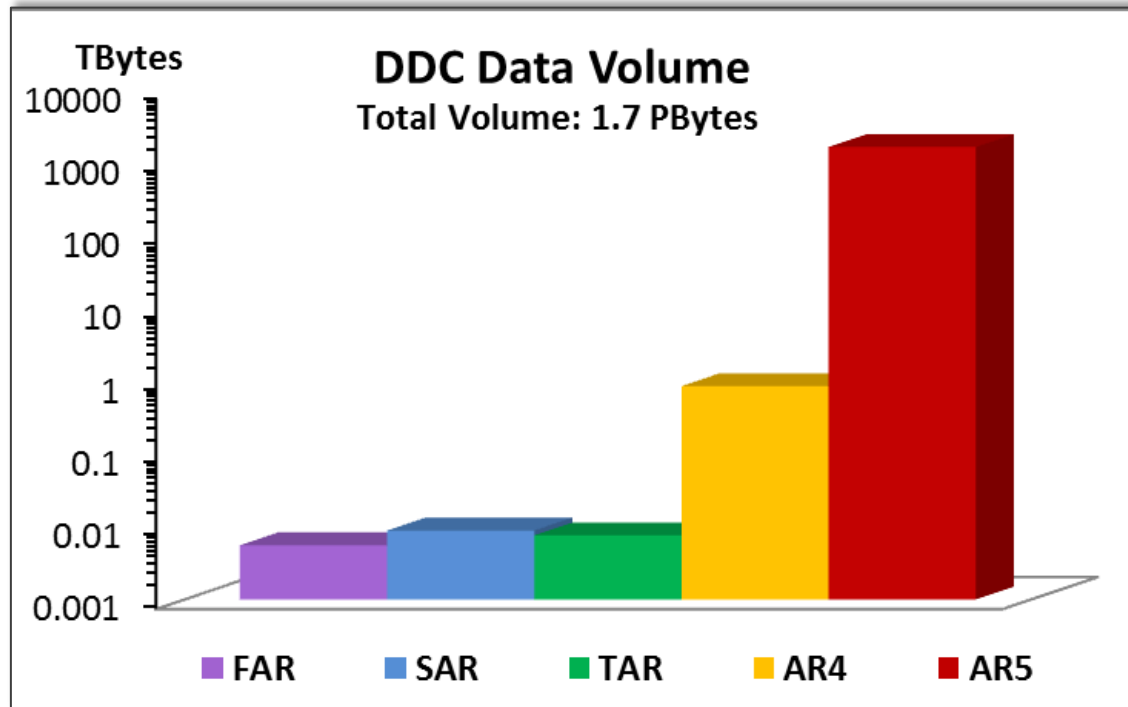
jointly managed by:

- Centre for Environmental Data Analysis (CEDA)
- World Data Center Climate (WDCC) at German Climate Computing Center (DKRZ)
- Center for International Earth Science Information Network (CIESIN) at Columbia University, NASA Socioeconomic Data and Applications Center (SEDAC)

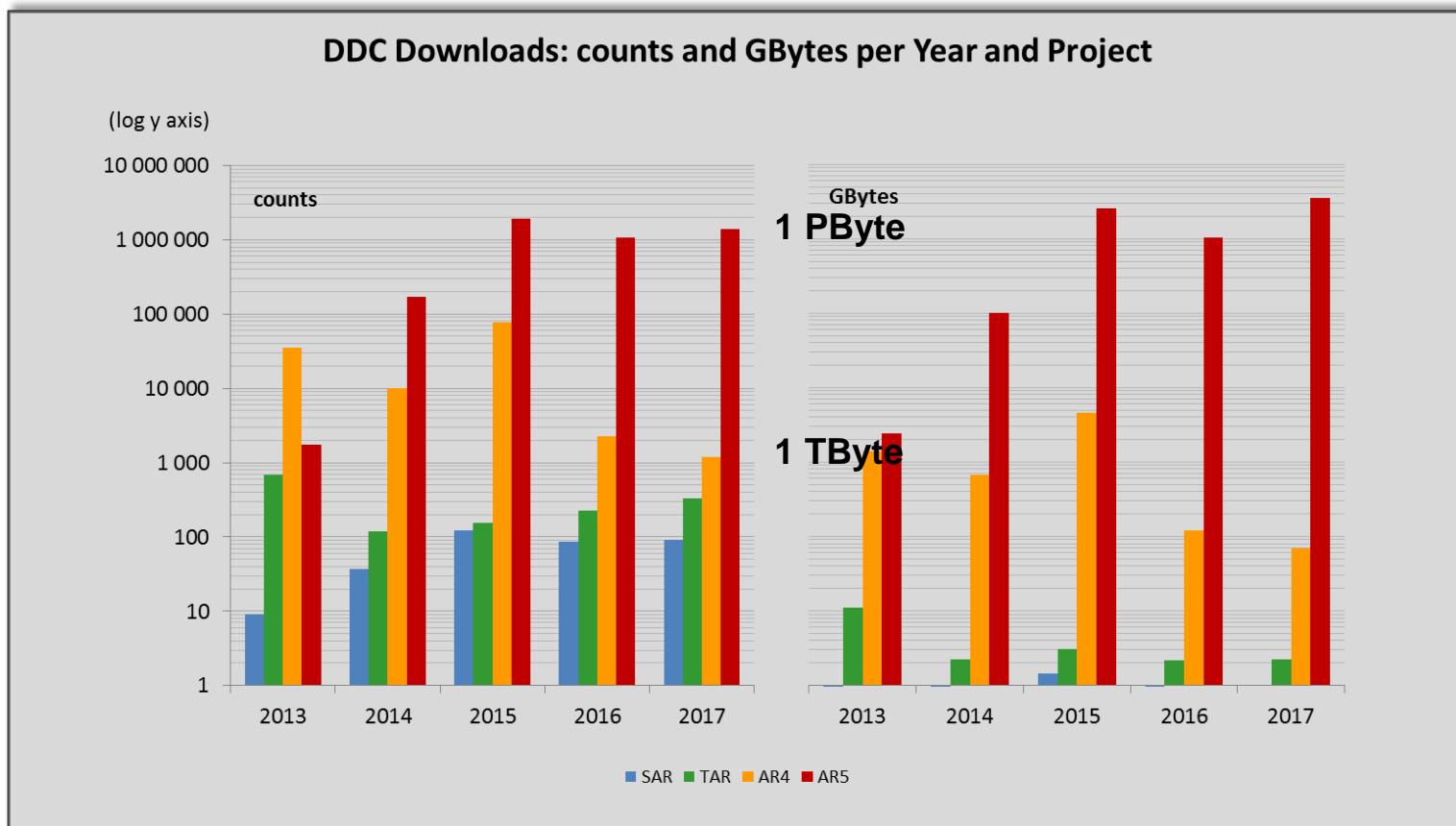


- Certified ICSU/ISC World Data System (WDS) members
- Guided by IPCC TG-Data

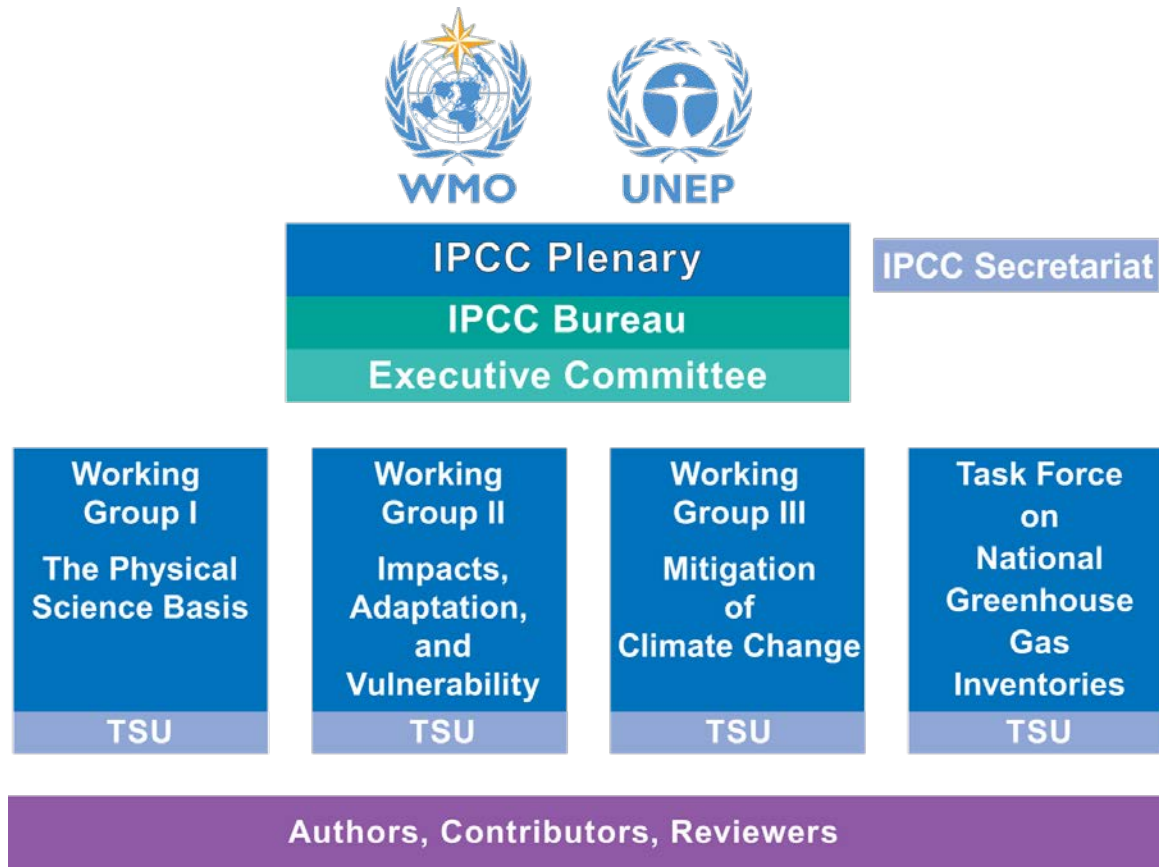
IPCC DDC at DKRZ (1): Reference Data Archive Size



IPCC DDC at DKRZ (2): Annual Data Download Rates



IPCC DDC (1): within Intergovernmental Panel on Climate Change (IPCC)



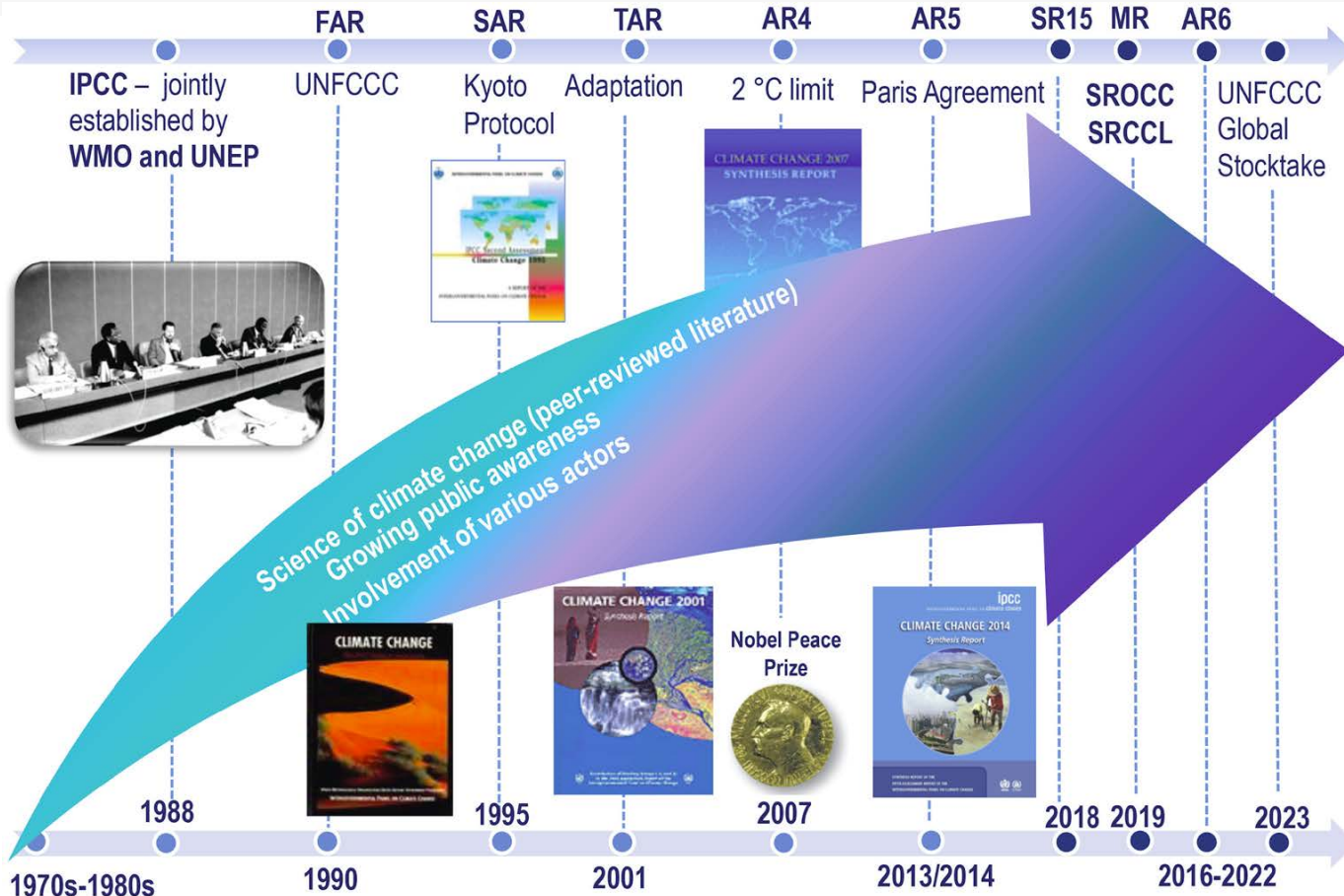
Task Groups:

established for a limited or longer duration to consider a specific topic or question

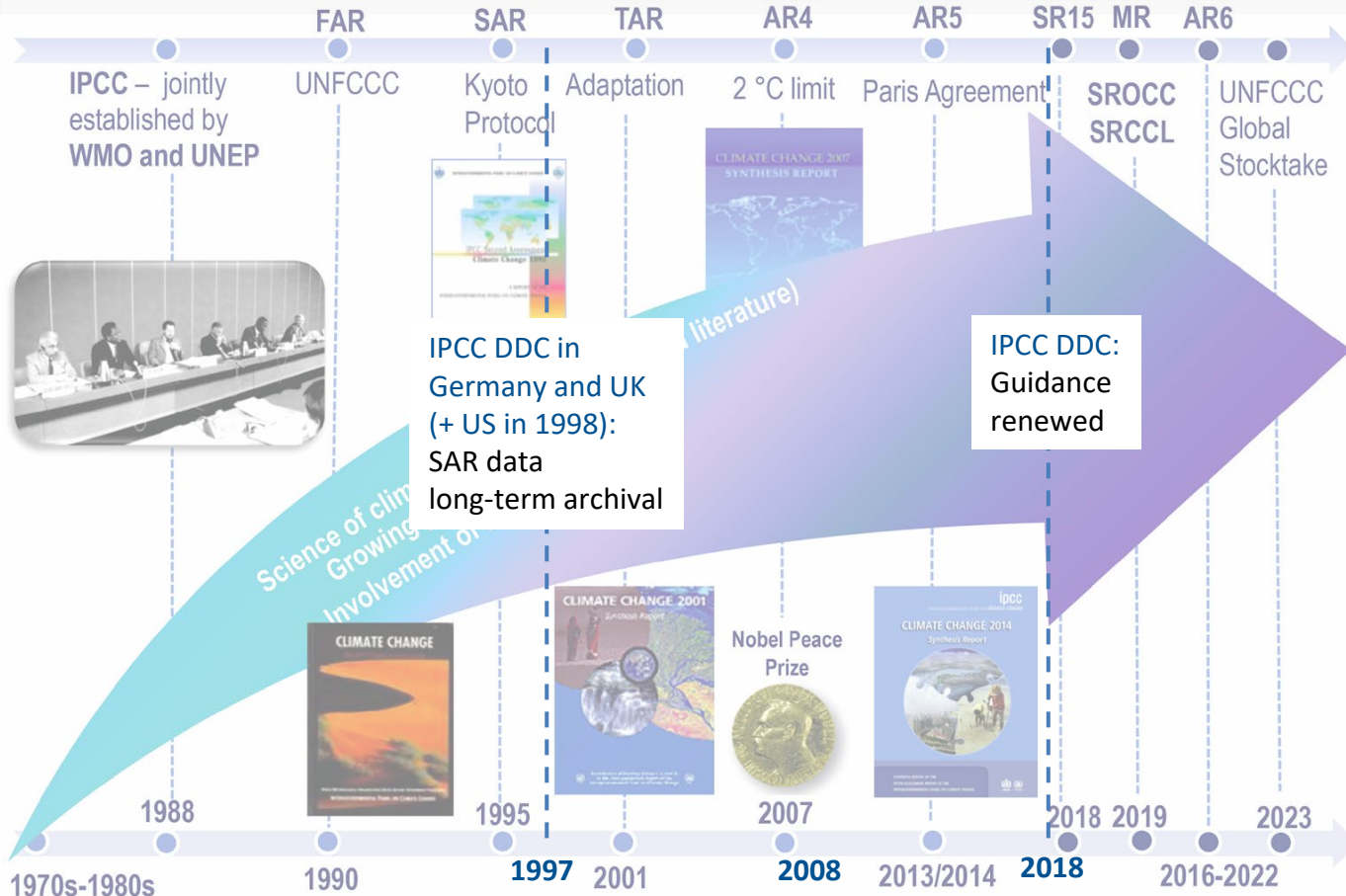


* not part of IPCC

IPCC DDC (2): History of the IPCC



IPCC DDC (2): History of the IPCC DDC



Growing awareness of the importance of data and data provenance

Guidance for IPCC's DDC ([IPCC-47/Doc. 9](#))

“The formal, institutional relationship between IPCC's Data Distribution Centre (DDC) and TG-Data and IPCC on the other hand is currently unclear.”

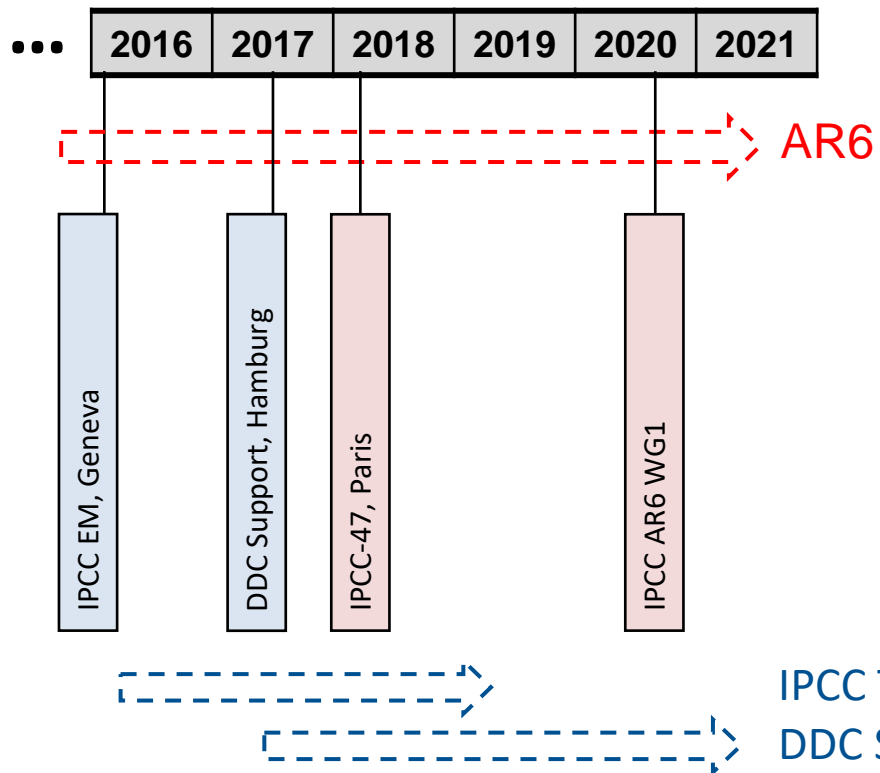
Current Objectives

1. Long-term data archival
of data underlying the reports, incl. data for key figures and tables
2. Data curation and supporting materials for data usage

(Partly) New Objectives

3. Collaboration with other data centers as IPCC DDC in a transparent manner
4. Contribute to a sustainable (infra-)structure for data on regional to local scales

IPCC DDC (4): DDC in AR6



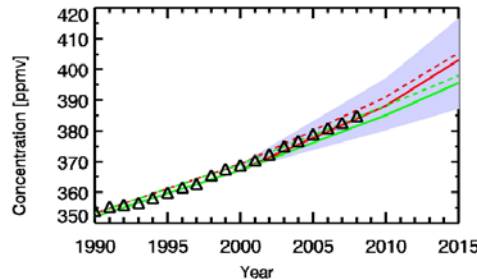
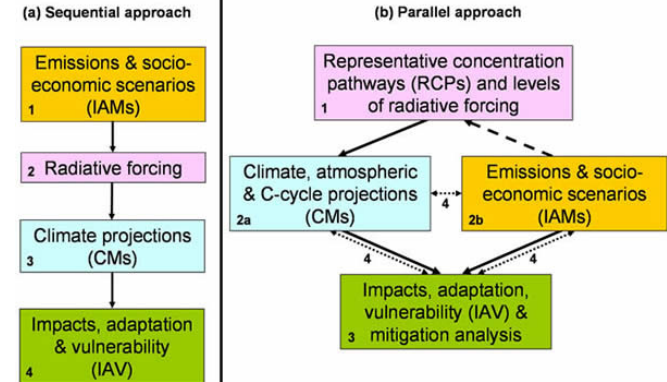
2017 - DDC Support Group:
 DDC Manager and IPCC AR6
 WG TSUs

2018 - TG-Data:
 TGICA was under review of
 IPCC for two years (2016-2018)
 resulting in a revised ToR at
 IPCC-47
 (https://www.ipcc.ch/meetings/session47/p47_decisions.pdf), but
 the TG membership decision is
 to be finalized ca. 02/2019.

IPCC TGICA transition to TG-Data
 DDC Support Group

Primary Areas of Activity

- Global Climate Model Output Data from CMIP (WDCC)
- Observational Data (CEDA)
- Socioeconomic Scenarios (SEDAC)
- Guidance Documents / Information on Data and Scenarios, Impact and Vulnerability Assessment (TGICA)



ipcc

CMIP5 data provided at the IPCC Data Distribution Centre

Executive Summary

Background

Examples of Data Stewardship Accomplishments

- Archiving Climate Model Data from CMIP

- Rescue part of FAR data (2008)
- DataCite DOIs registered since AR5
- Adding documentation (e.g. ES-DOC) and references (paper, data)

WDC CLIMATE Home About Documentation Utilities Register Login DKRZ

Metadata for 'ETH cmip5 rcsp45'
doi:10.2394/WDC/ETH

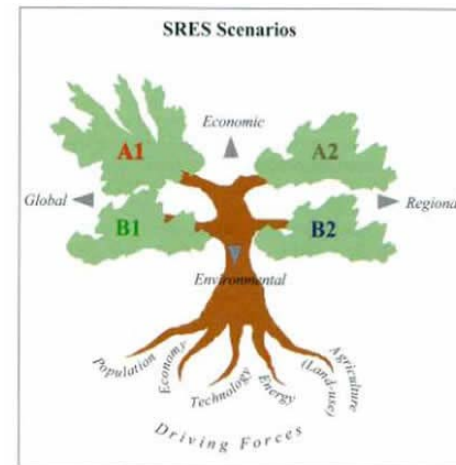
General Information Quality Contacts References Data Hierarchy

General Information

Name: ETH cmip5 rcsp45
 Abbrev: ETH-rcsp45
 Summary: rcsp45 is an experiment of the CMIP5 - Coupled Model Intercomparison Project Phase 5 (<http://pcmdi.llnl.gov/cmip5/>). CMIP5 is meant to provide a framework for coordinated climate change experiments for the next few years and that includes simulations for assessment in the IPCC AR5. The subset of the CMIP5 additional historical data were collected by ETH Zurich until 15th March 2013 in support of IPCC AR5 Working Group I (<http://www.clm-ethz.ch/CMIP5/>). These data are part of the IPCC-DOC AR5.
 rcsp45 (v.1, RC4.5) - Version 3- Future projection (2006-2100) forced by RC4.5. RC4.5 is a representative concentration pathway which represents the results of a relative forcing of 4.5 W m⁻² at year 2100, relative to pre-industrial conditions. RC4.5 are time-dependent, dependent on the scenario of emissions and concentrations of radiatively active gases and aerosols.
 Experiment design: http://pcmdi.llnl.gov/cmip5/metadata/experiment_design.html
 List of output variables: http://pcmdi.llnl.gov/cmip5/metadata/output_variables.html
 Output: time series per variable at model grid spatial resolution in netCDF format
 Each dataset model and the simulation information: CMIP-metadata
 Data are disclosed and entry names are set according to the ETH Zurich archive layout: 'experiment/ID/table/variable/model/ensemble/member/CMIP_filename.nc'
 Contact: Prof. Felix Zwiers (felix.zwiers@ethz.ch)
 Keyword(s): climate simulation, CMIP5, IPCC, IPCC AR5, IPCC-DOC, rcsp45
 Location(s): World (global), Longitude 0 to 360 latitude -90 to 90
 Spatial Coverage: Longitude 0 to 360 latitude -90 to 90 Resolution: -0.250 to 0.015°
 Temporal Coverage: 2006-01-01 to 2100-12-31 (calendar time coverage of data)
 11.07.16 (CMIP5/AR5/ETHZ) Felix
 Role: [netCDF](#)
 Program: cmip5rcsp45
 Creation Date: 2013-09-20

- Socioeconomic Scenarios

- Rescue of early IS92 scenarios (2nd Assessment)
- Preservation of SRES scenarios
- Support for transition to RCPs, SSPs



Past and Ongoing Challenges

- Data management not integral to the IPCC report drafting and review process
- IPCC authors swamped and unable to address data issues within assessment cycle
 - Identification of data (subset) underlying assessment reports, and tracking of data producing figures and tables hasn't been implemented
 - additional effort was needed for archival in DDC
- Varied levels of attention given to data stewardship by the IPCC, especially across multiple assessment cycles
- Prioritizing what data from IPCC reports needs to be managed and supported in the long-term

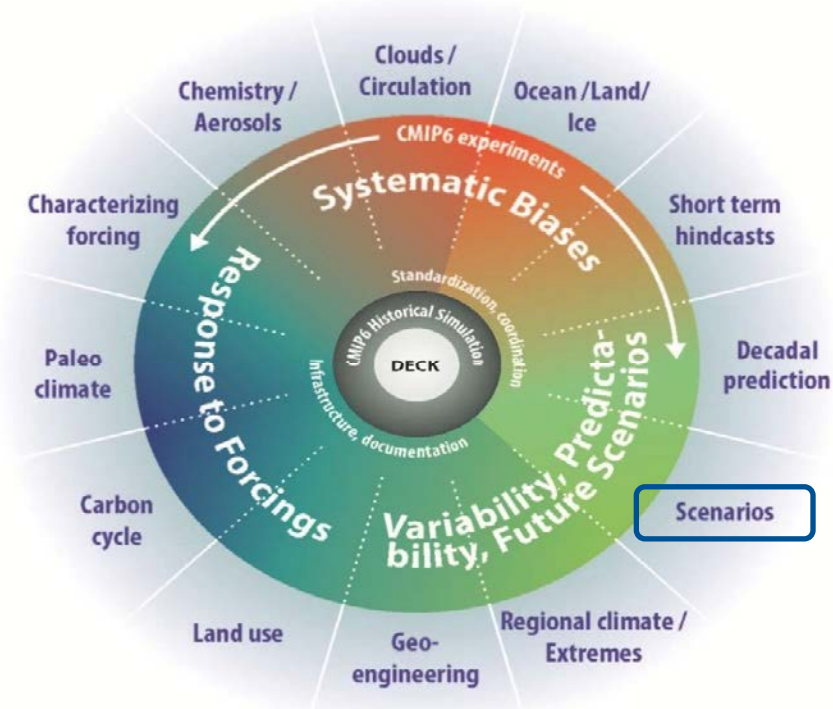
Past and Ongoing Challenges

- Ensuring clear distinction between data that may or may not be formally peer-reviewed and data outputs associated with IPCC-reviewed products
- Use of socioeconomic scenarios assessed by IPCC (e.g. SR1.5 Scenario Database) vs. scenario development by the scientific community (outside IPCC)
- Need for downscaling and translating DDC data holdings into forms more useful for impact/vulnerability communities
- Increasing diversity in scientific objectives in the report and in the reuse of the data together with the growing data volume requires the development of more sophisticated services
- Overall lack of resources for coordinated and long-term data management

DDC in IPCC AR6 (1): Coupled Model Intercomparison Project 6

CMIP6 Estimations (2018-10-11):

- 24 MIPs+DECK with ca. 292 experiments are part of CMIP6 (CMIP5: 1 MIP; 101 exp.)
- 43 institutions with 109 models are currently registered for CMIP6 (CMIP5: 27 institutions; 60 models)
- Estimated data volume: ca. 20 PB (CMIP5: 1.6 PB)
- ca. 25 ESGF data nodes world-wide will disseminate data



(Source: Eyring, V. et al. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, [doi:10.5194/gmd-9-1937-2016](https://doi.org/10.5194/gmd-9-1937-2016).)

Potential Role of the WDS

- WDS provides long-term data stewardship framework and mission to balance short-term needs of ongoing international assessments like the IPCC
- Provides standards and best practices for selection, appraisal, and stewardship of data arising from international assessment, e.g. Core Trust Seal
- Potential for bringing broader expertise and resources to bear on IPCC needs, while meeting IPCC requirements for rigor, objectivity, transparency, etc.
- WDS-ITO might bring technical expertise / solutions for federated data archives such as IPCC DDC + collaboration partners?
- Improved integration/interoperability between DDC and key input data, models, and scenarios
- Co-operation in developing data services e.g. server-side processing?

Summary / Outlook

- IPCC DDC has provided long-term access to data underlying the IPCC ARs since 1997.
- Additional aspects became relevant (renewed DDC Guidance):
 - Collaboration with external repositories, and
 - the importance of data on regional to local scales.
- DDC Support (collaboration between DDC and WG TSUs) aims to improve
 - the support for the IPCC WGs and its authors,
 - the data archival process for the DDCs, and
 - data documentation (add provenance information)
- The future of the IPCC DDC and the DDC Support group depends on the future TG-Data and its co-chairs.
- We are reaching out to the WDS community to gauge potential interest, expertise, and resources that could help address these needs