

DataCite Data Publication

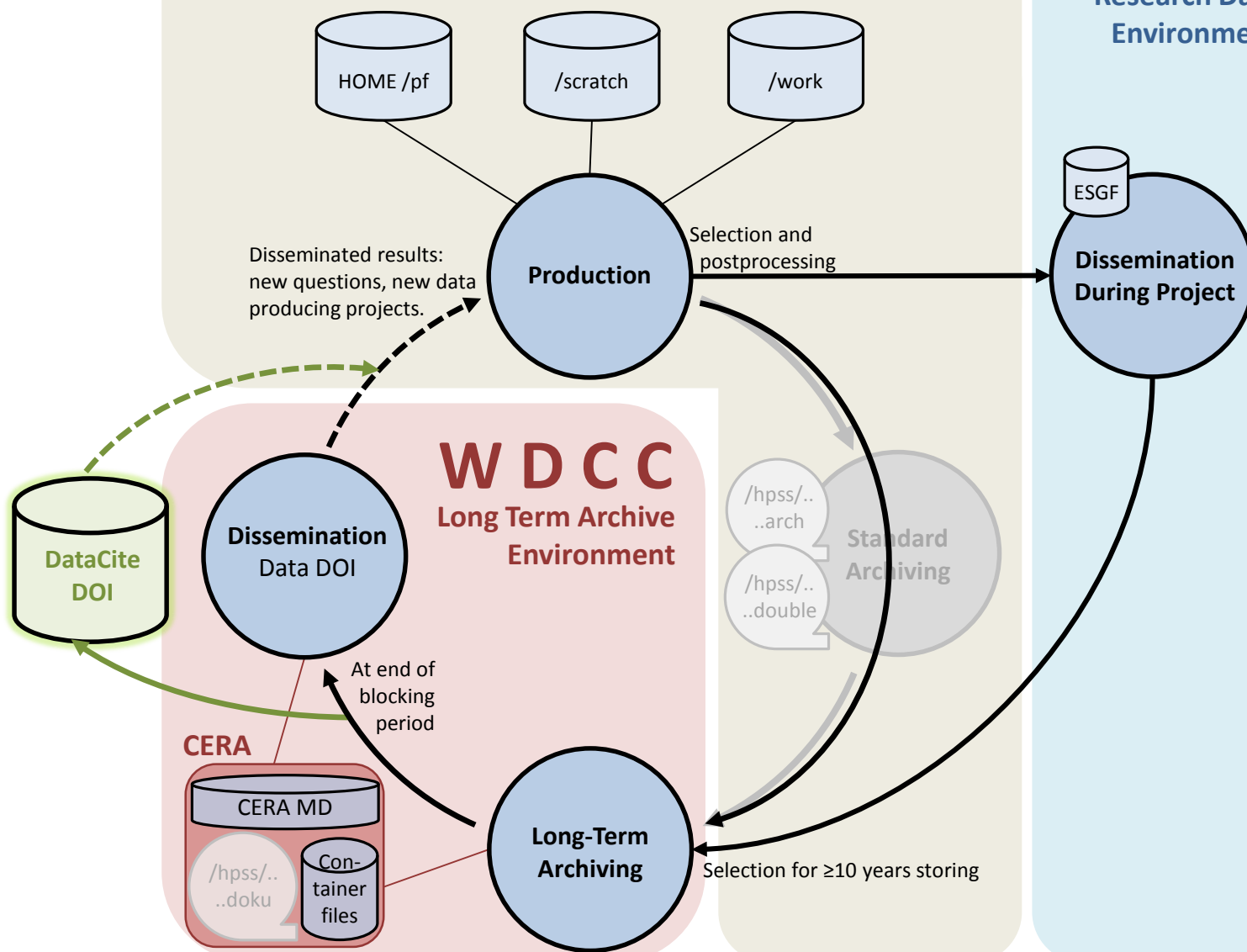
- Make your data citable -

Martina Stockhause
Deutsches Klimarechenzentrum (DKRZ)

DKRZ Projects

ESGF

Research Data Environment



- What is a data DOI?
- How to get a data DOI at WDCC/DKRZ?
- Where are data DOIs visible?

What is a DOI (1)

Definition and Benefits

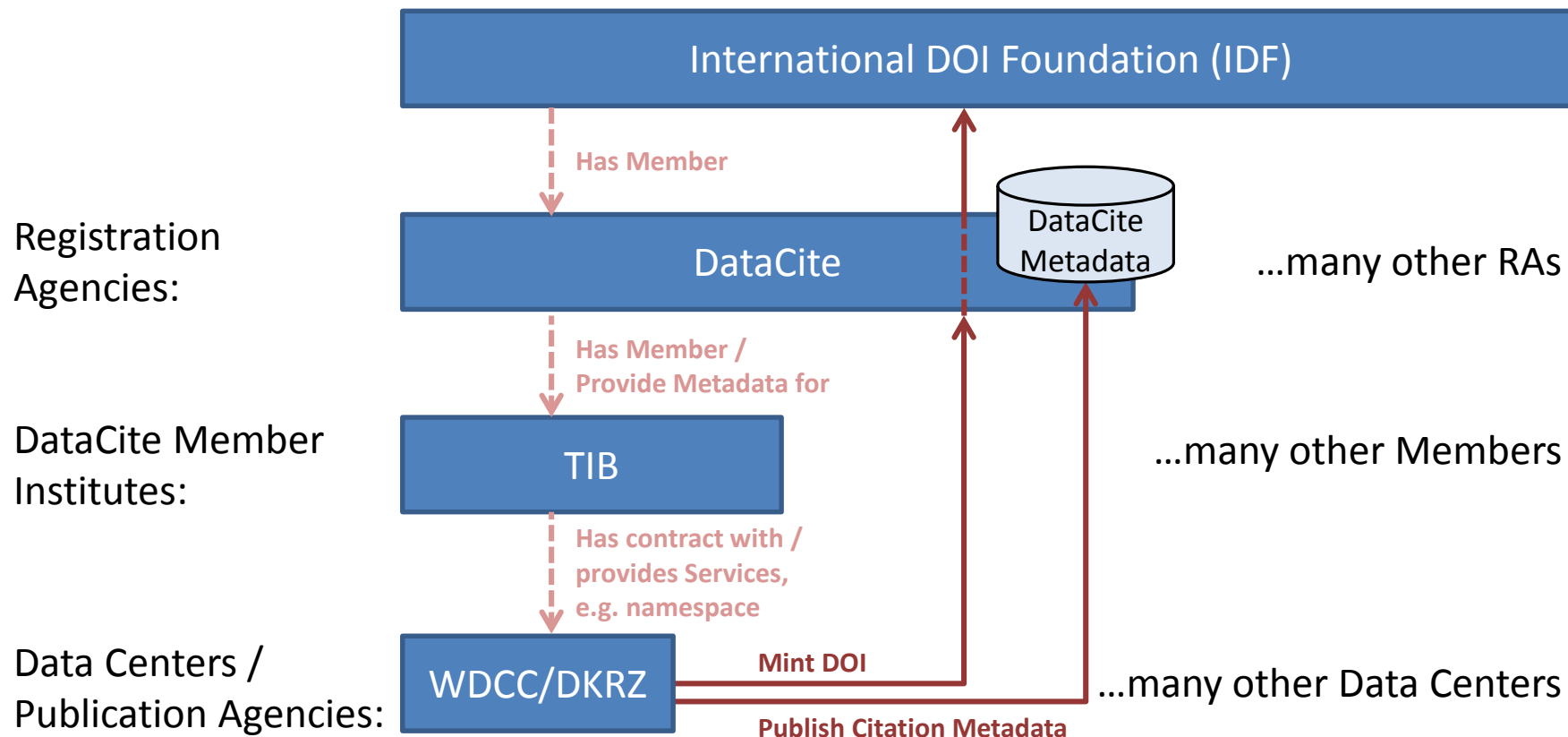
Data DOI: Digital Object Identifier
= Persistent Identifier (PID)

- Persistent link to data creator
- Data objects are long-term
- DOI published data is credited
= credit for data creator
 - Data citations can be included in reference lists of scientific publications (example: [doi:10.5194/cp-7-1103-2011](https://doi.org/10.5194/cp-7-1103-2011))
 - Data citations can be included in scientists' list of publications

The screenshot shows a web browser window with the URL www.clim-past.net/7/1103/2011/cp-7-1103-2011-relations.html. The page is for the journal "Climate of the Past" and features a navigation menu on the left with options like Home, Online Library CP, and Alerts & RSS Feeds. The main content area displays the article title "Upper ocean climate of the Eastern Mediterranean Sea during the Holocene Insolation Maximum – a model study" by F. Adloff et al. Below the title, there is a "Related Subject Area" section and a "Cited Articles" section. One article in the "Cited Articles" section is highlighted with an orange border: "Andersson, A., Bakan, S., Fennig, K., Grassl, H., Klepp, C., and Schulz, J.: Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data – HOAPS-3 – monthly mean, World Data Center for Climate, doi:10.1594/WDCC/HOAPS3_MONTHLY, 2007." Another article, "Arakawa, A. and Lamb, V. R.: Computational design of the basic dynamical processes of the UCLA general circulation model, Methods Comput. Phys., 17, 173–265, 1977." is also visible.

What is a DOI (2)

Organization



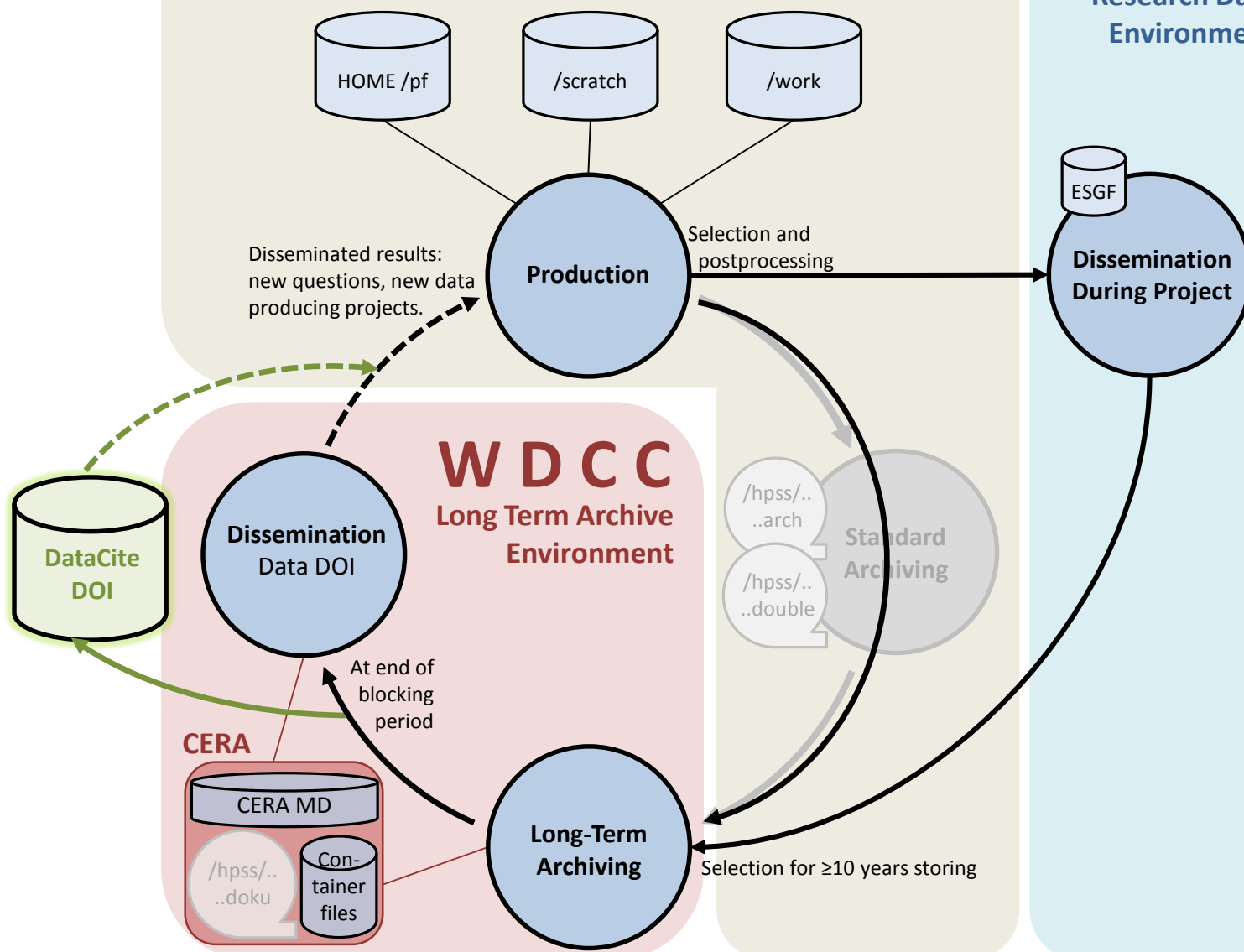
Data DOIs at WDCC (1)

History and Standards at WDCC

- WDCC assigns data DOIs since 2004
- 461 data DOIs have been published by WDCC (06.05.2014), 261 of those belong to CMIP5 experiments
- WDCC's DOI namespace: [doi:10.1594/WDCC/<WDCC-ID>](https://doi.org/10.1594/WDCC/<WDCC-ID>)
- WDCC standard for DataCite data DOIs:
 - Long-term archived data
 - Long-term availability and accessibility and usability of data for a large user group – data curation
 - [Quality assessed data](#)
 - [Informative metadata attached:](#)
sufficient for interdisciplinary usage of the data
 - [Well-organized data: experiment – dataset structure](#)

DKRZ Projects

ESGF Research Data Environment



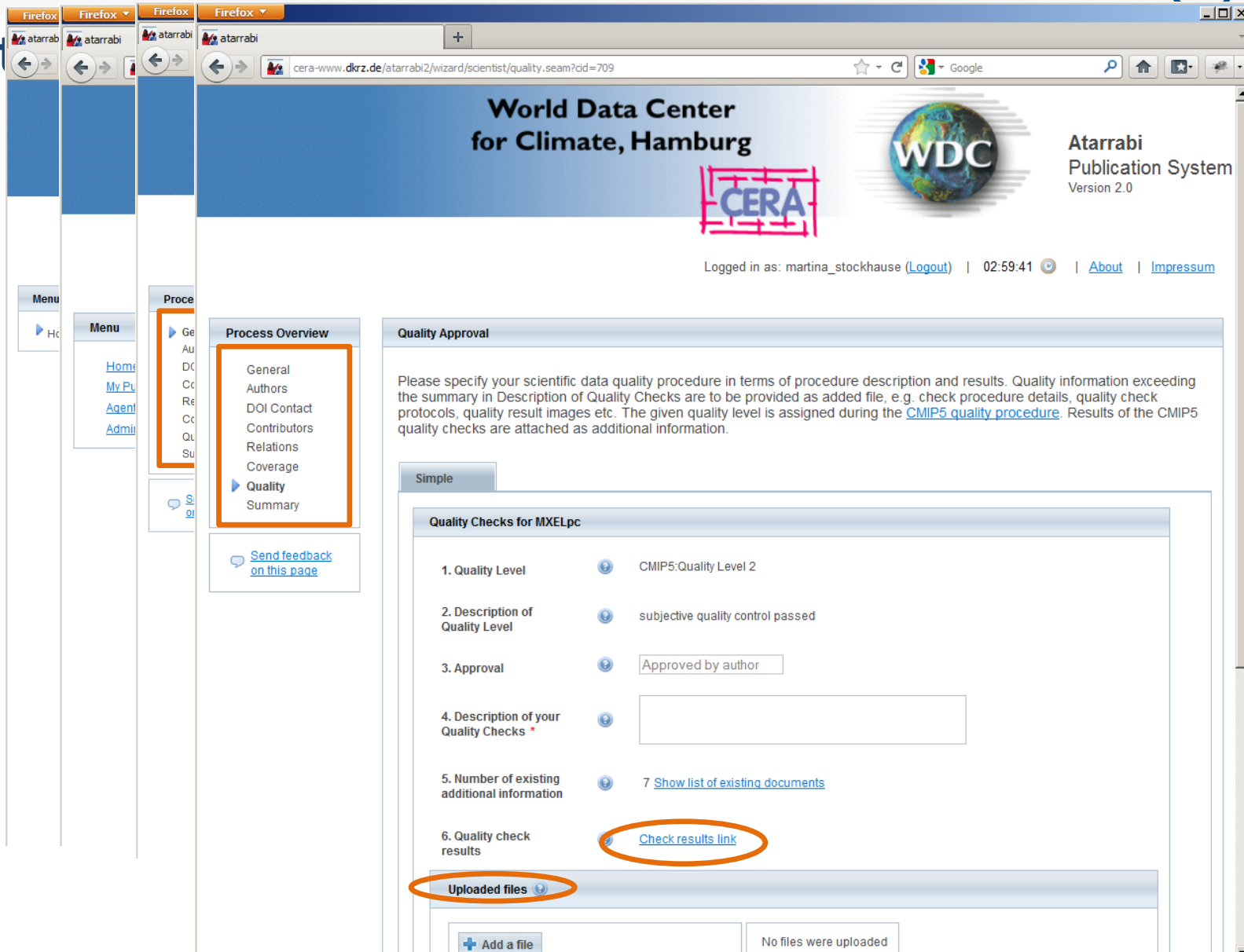
Data DOIs at WDCC (3)

DOI process at WDCC - Tasks

0. Precondition: Data is long-term archived at WDCC
1. Set-up phase with first contact
Publication Agency starts DOI Process
2. Scientific Quality Assurance (SQA) via WDCC's Data Publication System Atarrabi 2.0 (<http://cerawww.dkrz.de/atarrabi>)
3. Publication Agency performs Technical Quality Assurance (TQA)
4. Publication Agency publishes DOI via DataCite
Notification of DOI by Publication Agent

Data DOIs at WDCC - atarrabi (4)

2. Scientist



World Data Center for Climate, Hamburg

Atarrabi Publication System Version 2.0

Logged in as: martina_stockhause ([Logout](#)) | 02:59:41 | [About](#) | [Impressum](#)

Process Overview

- General
- Authors
- DOI Contact
- Contributors
- Relations
- Coverage
- Quality**
- Summary

[Send feedback on this page](#)

Quality Approval

Please specify your scientific data quality procedure in terms of procedure description and results. Quality information exceeding the summary in Description of Quality Checks are to be provided as added file, e.g. check procedure details, quality check protocols, quality result images etc. The given quality level is assigned during the [CMIP5 quality procedure](#). Results of the CMIP5 quality checks are attached as additional information.

Quality Checks for MXELpc

- Quality Level: CMIP5:Quality Level 2
- Description of Quality Level: subjective quality control passed
- Approval:
- Description of your Quality Checks *
- Number of existing additional information: 7 [Show list of existing documents](#)
- Quality check results: [Check results link](#)

Uploaded files

[+ Add a file](#) | No files were uploaded

Data DOIs at WDCC (5)

3. Technical Quality Assurance

General Checks (data vs. metadata):

- Number of datasets is consistent and > 0
- Sizes of every dataset are consistent and > 0
- Temporal coverage is consistent and the time axis is continuous.
- Variable descriptions are consistent.
- Formats are consistent and correct.
- The datasets and corresponding metadata are accessible.

Project-Specific Checks, e.g. CMIP5:

http://redmine.dkrz.de/collaboration/projects/cmip5-qc/wiki/Qc_I3

Find and Resolve Data DOIs (1)

DOI data in catalogs:

- DataCite

harvested by:

- Libraries, e.g. TIB GetInfo catalog
- Scientific Publishers, e.g. Elsevier and Thomas Reuters
- Data Portals, e.g. EUDAT
- ...other

2)

DOI Landing Page

DOI for Scientific and Technical Data

10.1594/WDC/CMIP5.NCCN.Mpc

Creator (Person(s) or Institute(s))
Bentsen, Mats, Bethke, Ingo, Debernard, Jens, Drange, Helge, Heinze, Christoph, Iversen, Trond, Kirwaed, Alf, Seland, Øyvind, Tjiputra, Jerry

Publication Year
2011

Title
cmip5 output1 NCC NorESM1-M piControl, served by ESGF

Publication Agency
WDC at DKRZ

Identifier
doi:10.1594/WDC/CMIP5.NCCN.Mpc, <http://dx.doi.org/10.1594/WDC/CMIP5.NCCN.Mpc>

Contact
Trond Iversen

Contributing Institute(s)
Sjerknes Centre for Climate Research (BCCR)
Norwegian Meteorological Institute (met.no)

Hosting Institute(s)
Deutsches Klimarechenzentrum (DKRZ)
PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON UNIVERSITY OF CALIFORNIA (PCMDI)
The NCAS British Atmospheric Data Centre (BADC)

Metadata

CIM Metadata
<http://ipui.org/eng/edoo/redirects/cmip5/doi/NCC/NorESM1-M/piControl>

This metadata link is preliminary and might change in the future. In case of a broken link please check the metadata link given in the table "Published CIM metadata" at the bottom of the CIM questionnaire page (<http://q.cmip5.dea.ac.uk>)

WDC Metadata
<http://oea-www.dkrz.de/WDC/CMIP5/Entry.jsp?acronym=NCCN.Mpc>

Data Access

ESGF Data Access (openID)
<Links.jsp?acronym=NCCN.Mpc>

WDC Data Access (CERA account)
<Entry.jsp?acronym=NCCN.Mpc>

Data access links are provided to the data in the Earth System Grid Federation (ESGF: <http://esgf.org>) and to the long-term data archive of the World Data Center for Climate (WDC: <http://www.wdc-climate.de>). ESGF data access requires a valid openID (see http://cmip-pcmdi.llnl.gov/cmip5/isa_getting_started.html) and WDC data access a CERA account (see <http://oea-www.dkrz.de>).

Summary

piControl is an experiment of the CMIP5 - Coupled Model Intercomparison Project Phase 5 (<http://cmip-pcmdi.llnl.gov/cmip5/>). CMIP5 is meant to provide a framework for coordinated climate change experiments for the next few years and thus includes simulations for assessment in the AR5 as well as others that extend beyond the AR5.

3.1 piControl (3.1 Pre-industrial Control) - Version 1: Pre-industrial coupled atmosphere/ocean control run. Imposes non-evolving pre-industrial conditions.

Experiment design: http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf

List of output variables: http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.pdf

Output time series per variable in model grid spatial resolution in netCDF format
Earth System model and the simulation information: CIM repository

Entry name/title of data are specified according to the Data Reference Syntax (http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf) as activity/product/institute/model/experiment/frequency/modeling realm/MIP table/ensemble member/version number/variable name/CIMOR filename.nc.

Quality

Accuracy

In addition to CMORs automated quality checks we performed manually checks on a data sample for each output variable. These comprise:

- visual inspection using netview
- check whether minimum values are in physical range and consistent with variable units
- direction of flux variables and positive attribute
- consistency of data with vertical axis definition (i.e. whether data is flipped)
- land/ocean/sea ice masking
- rotation of vector data - SQA - Scientific Quality Assurance 11/10/2011 10:28:03; Reported errors by scientist;
- u/v/w velocity vectors are rotated (corrected in data version v20120124)
- grFrazil, grCongel, snoToIce, snoem1, tme1, and bme1 are too large by a factor of 1000 (corrected in data version v20120227);
- Reported Error by Ingo Bethke (08/09/2012): npp contains 'gross_primary_productivity_of_carbon' above ground instead of 'net_primary_productivity_of_carbon'

Author's remark (12/21/2012) regarding 'Amon co2mass': This variable is subject to a scaling error. Please multiply with 44/28 to obtain the total atmospheric CO2 mass in kg, or divide by 5.1125e+18 to obtain the global mean CO2 concentration in ppmv.

Consistency

Quality Control Levels in CMIP5:

- * Level 0: Spot checks on selected data
- * Level 1: CMOR2 and ESG publisher conformance checks
- * Level 2: Data consistency checks
- * Level 3: Double- and cross-checks of data and metadata and data publication as DataCite DOI

QC Result Access: <http://oea-www.dkrz.de/WDC/CMIP5/QCResult.jsp?experiment=cmip5/output1/NCC/NorESM1-MpiControl>

Completeness

[CMIP5 AuthorComment-We are still in the process of filling out the METAFOR CMIP5 questionnaire. Therefore its contents might be subject to change. Publications based on NorESM1-M are under preparation.] [CMIP5 PubAgentComment-CIM simulation document published on 23.02.2012]

Specification

[CMIP5 QualityLevel=3] [CMIP5 QualityFlag=approved by author] according to CMIP5 quality control rules <http://cmip5oc.wdc-climate.de> [CMIP5 DateOfQualityControl=2011-10-10] [CMIP5 QualityControl2Comment=QC Level2 assigned at 2011-10-11 11:51 UTC; assignment confirmed by NCC (Ingo Bethke)] [CMIP5 CimMetadata=CMIP5 questionnaire data not found or not accessible] [CMIP5 TQA=done according to criteria for QCL3 defined at https://redmine.dkrz.de/attachments/projects/cmip5-qc/wiki/Qc_criteria_for_QC_L3DOI-publication] [CMIP5 SQA=checked by author] [CMIP5 DateOfQualityControl=2011-12-01]

Please note

WDC Climate as DOI publishing agency grants the validity of metadata and data accessed at WDC Climate (WDC). WDC recommends to check all downloaded data files by their tracking_id with the CMIP5 Data Validation Service at http://oea-www.dkrz.de/WDC/CMIP5/Tracking.jsp?tracking_id=tracking_id. The tracking_id can be found in the file headers (e.g. using ndump -h).

Terms of use

See <http://cmip-pcmdi.llnl.gov/cmip5/terms.html> and <http://cmip-pcmdi.llnl.gov/cmip5/citation.html>.

This page is hosted at WDC, please send technical inquiries to data@dkrz.de.

Citation

Contact

Link to metadata

Link to data

Summary

Quality

Terms of Use

Firefox

DataCite | Helping you

search_datacite

DataCite

Search

Filter

- allocator
- datacentre
- prefix
- resourceType
- contributor
- creator
- publicationYear
- publisher
- language
- refQuality
- has_metadata

URN

Contr

Conta

Rese

Rese

Hosti

Hosti

Hosti

Othe

x-dat

x-dat

rdf>

turtl

x-bib

x-res

html

Perspective of Data DOIs at WDCC

Work in progress at WDCC:

- Development of a general quality level system for all WDCC data using the self-assessment approach of a Maturity Matrix
- Integration of Person PIDs in CERA, e.g. ORCID
- Revision of the Data Publication System - Atarrabi 3.0

Engagement in national and international initiatives:

- RDA/WDS WG Data Publication Workflows:
coordination of data and paper publication workflows
- National project 'Komfor': improve the quality and availability of research data in the earth sciences

Documentation and Contact

Citable Data – DataCite data DOI service at WDCC

Documentation: <http://www.dkrz.de/daten-en/Datapublication>

Contacts:

- general: data@dkrz.de
- personal:
 - Heinke Höck
 - Martina Stockhause