

Data management and archiving for COP / GOP / D_PHASE

4th COPS Workshop
25./26.09. 2006 Stuttgart

Claudia Wunram
Hannes Thiemann



Data archive



Long term data archive for
COPS, GOP and D-PHASE

hosted at

World Data Centre for Climate (WDCC)

run by the group

“Model and Data” (M&D)

at

Max Planck Institute for Meteorology,

in

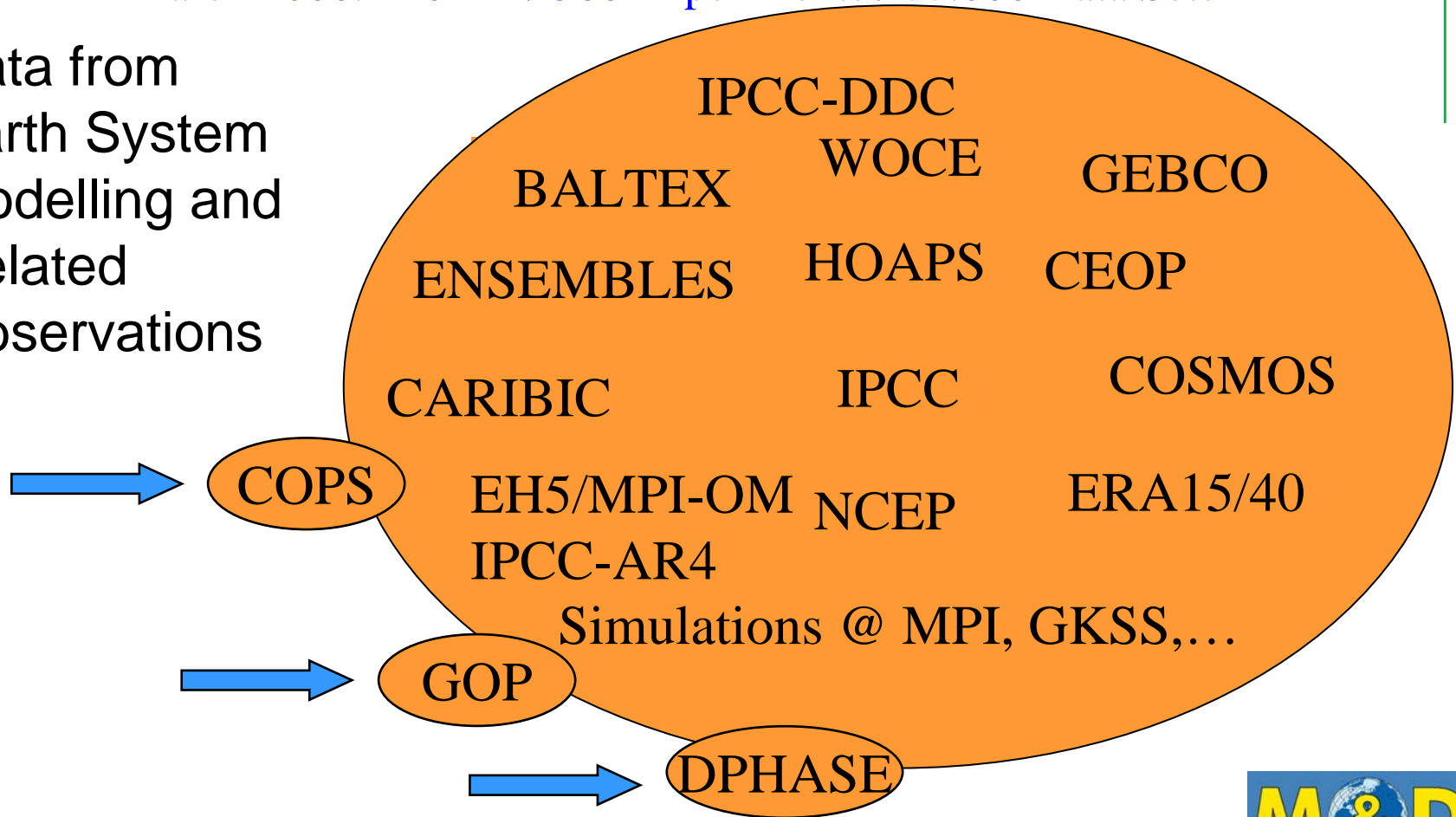
Hamburg, Germany.

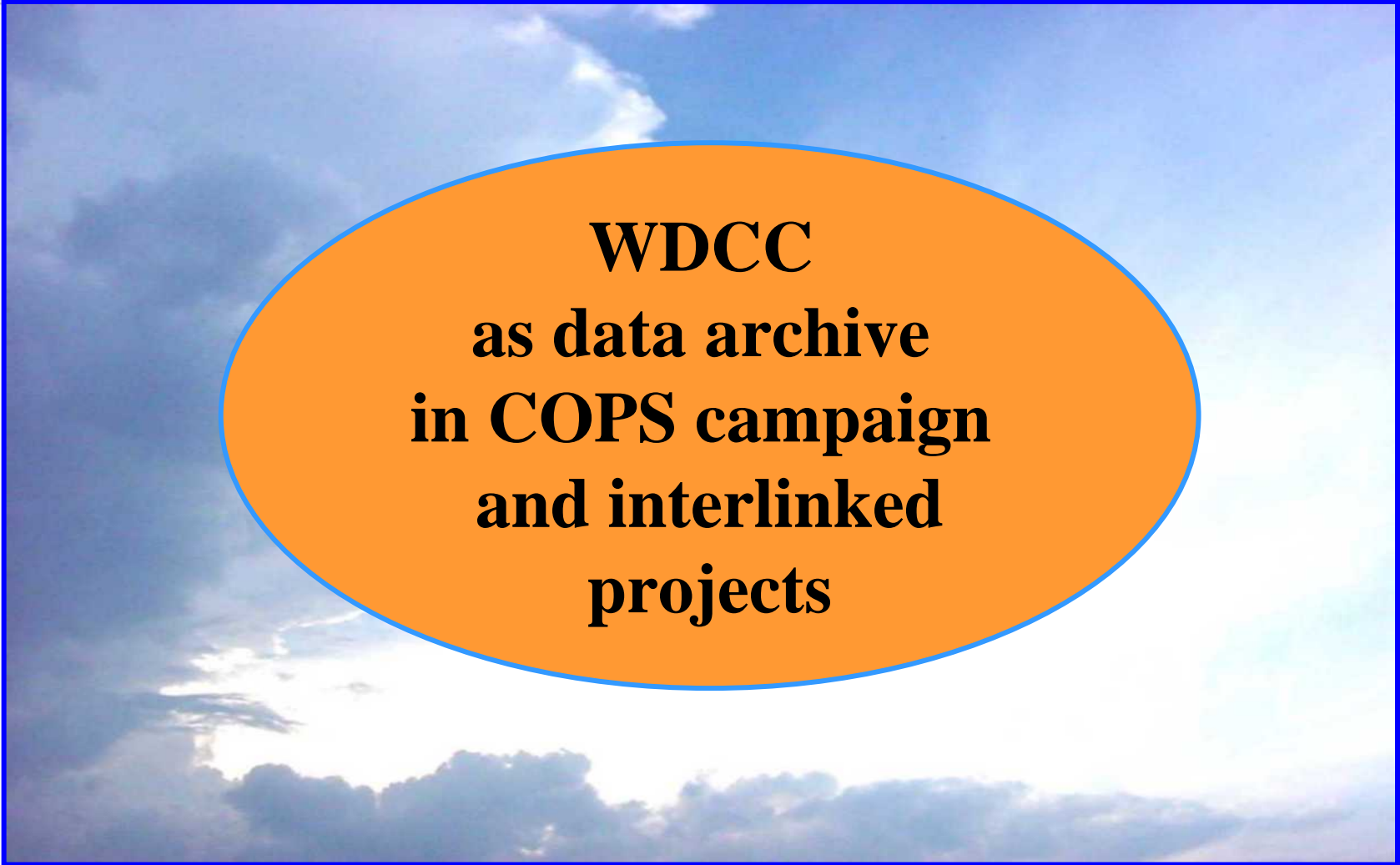


- WDCC as data archive in COPS-campaign
- Common data policy with interlinked projects
- Tasks of data archive and expected storage amounts
- Data transfer, responsibilities for quality control
- Data formats
- Meta data description
- Data structure
- Data access
- Next steps: test runs
- Outlook
- Contact info

- Mission: collect, store and disseminate data for climate research
- Approved in January 2003
- March 2006: 220 TB / 566 Experiments / 77.000 Data Sets

Data from
Earth System
Modelling and
Related
Observations





WDCC
as data archive
in COPS campaign
and interlinked
projects

- As announced in data implementation plan
- Agreed on by all PIs and M&D
- All investigators deliver promptly their data to the archive (final version 03/2008)
- M&D gives access rights according to announcements of COPS coordinator (groups and timeline)

- archive instrument data, model data, quicklooks and alerts for observation periods:
 - GOP: JAN 07 – DEC 07
 - COPS: JUN 07 – AUG 07
 - DPHASE: JUN 07 – NOV 07
- define meta data layout and handle implementation
- offer service within the frame of data storage at WDCC and help to access to data base
- no real time data handling can be done by M&D
- host data base link to external data:
 - EUMETSAT, 3D radar (DWD)
 - LMK (high resolution forecast model)

Storage amounts:

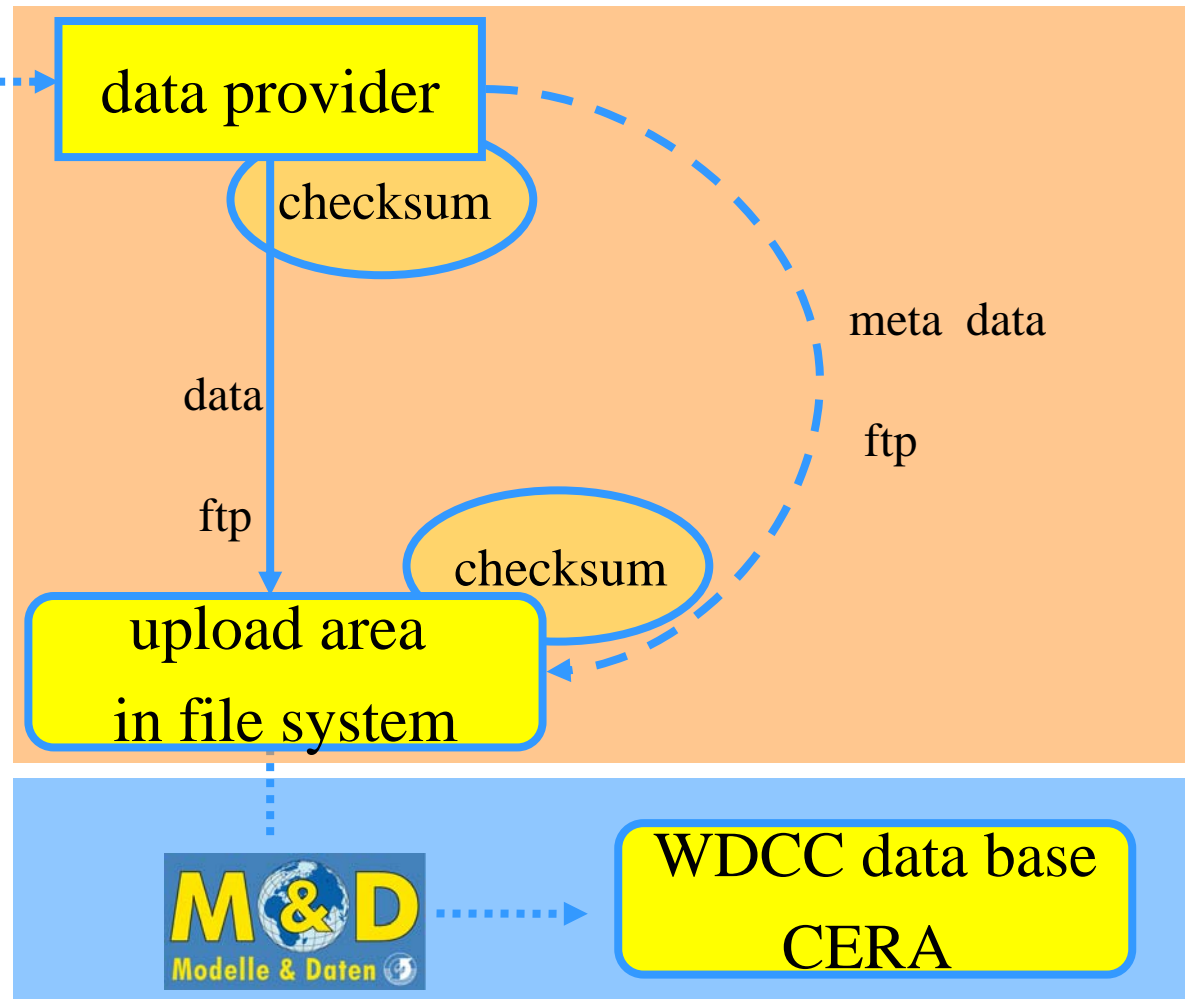
- Data storage volume for COPS, GOP and D-PHASE:
 - 20 TB
- Estimated data volume:
 - GOP: 3+ TB
 - COPS instruments: 2 TB
 - COPS models: 10 TB
 - D-PHASE: 5 TB
- Plus processing area on M&D work group server:
 - ~500 GB + CPU (visualization tasks, quick access)

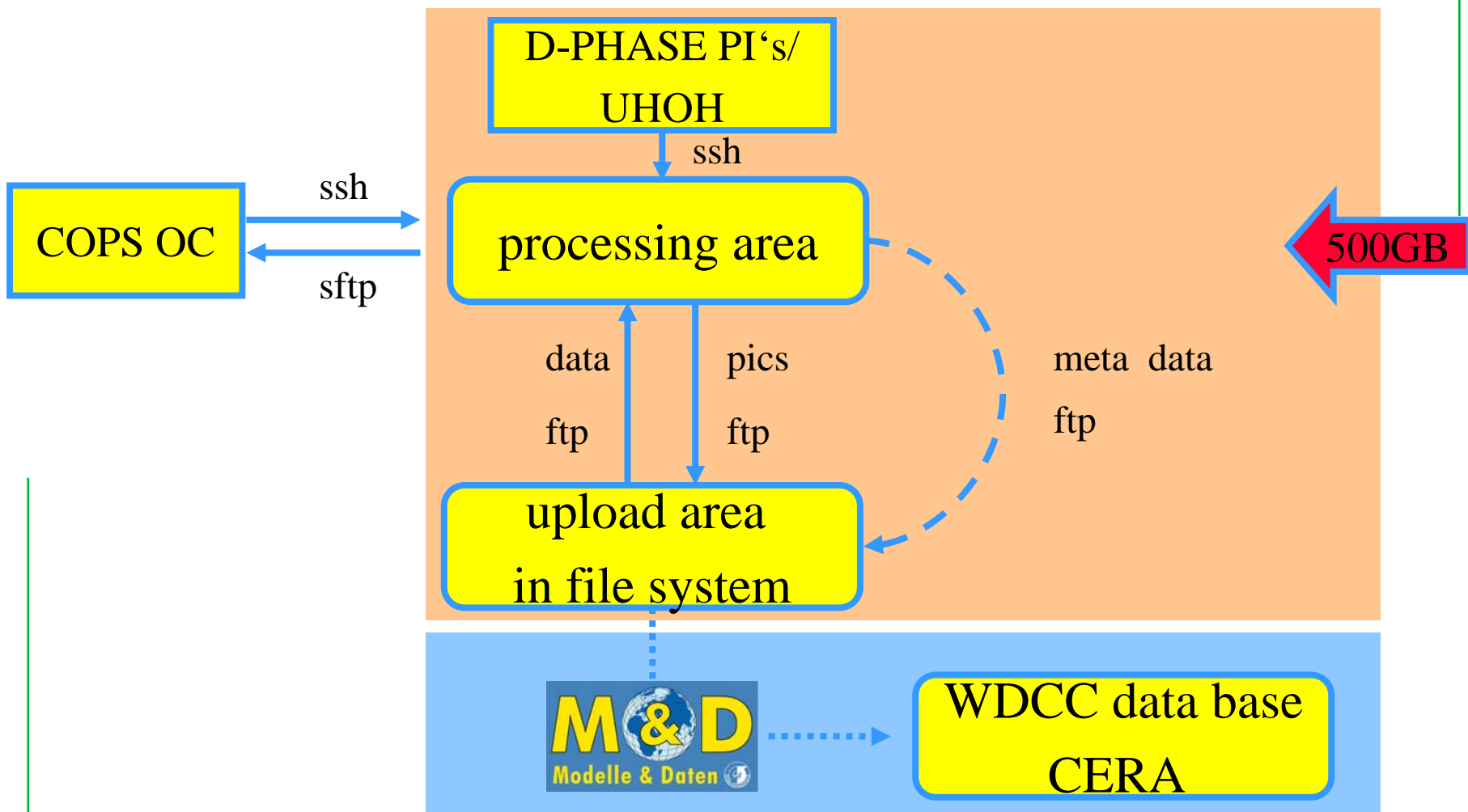
- Observation period: APR 07 to DEC 07
- Data volume: ~ 150 GB
- Data transfer: at the end of observation period
(shipped on disk, ...)

Data transfer



- unix account
- user instruction
 - data structure
 - data upload





Data control

Data providers:

- responsible for quality of data file content and meta data content
- responsible for data transfer (checksum tests)

M&D:

- technical controls (time stamp, consistency of time series)

Accepted data formats:

model data

GRIB1, netCDF/CF

instrument data

netCDF/CF

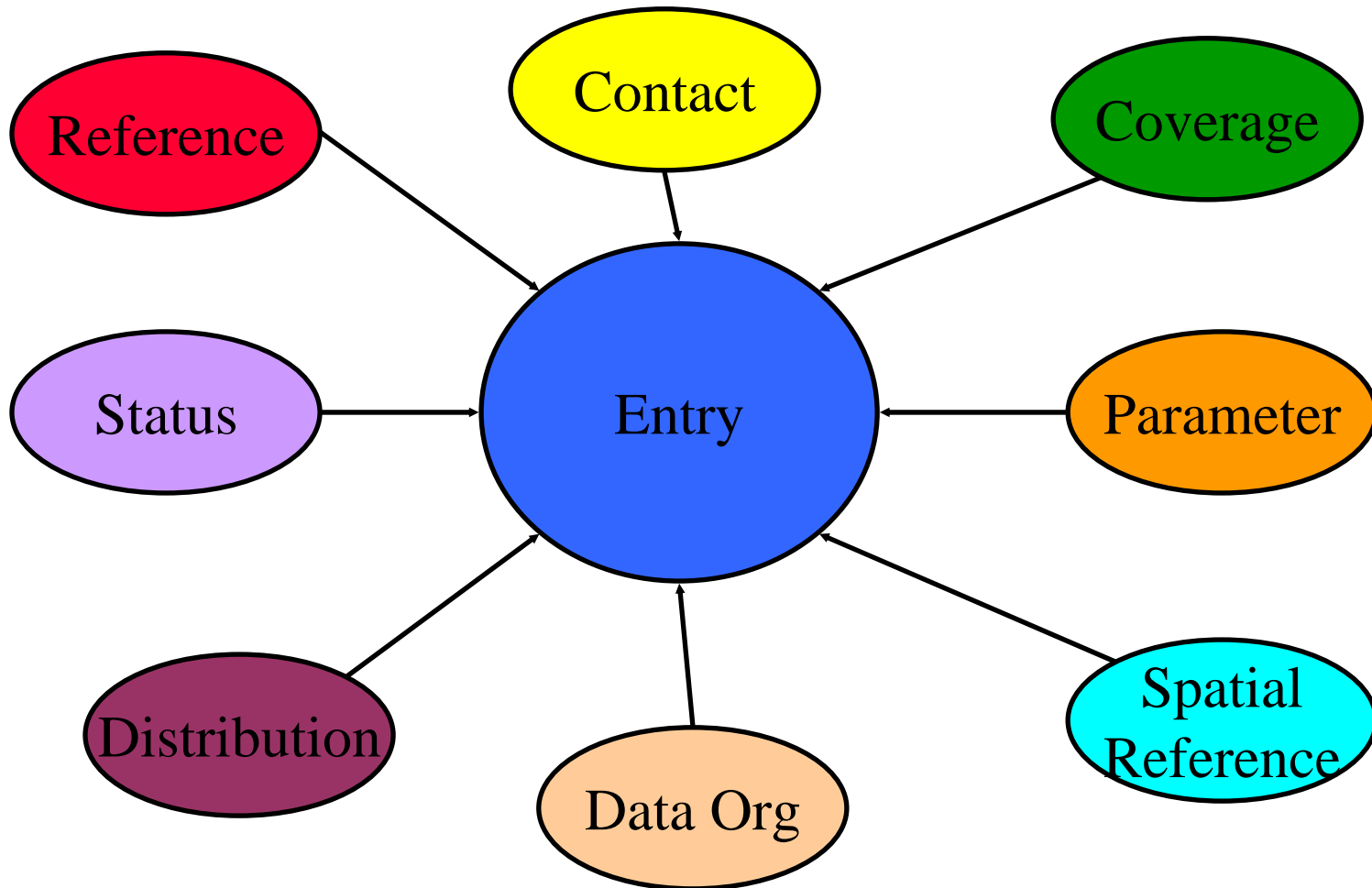
quicklooks

jpg, gif, png, eps, ...

meta data

xml

CF-convention for meta data description is strongly advised:
Variable names are described by CF-standard names
-> search in data base and intercomparison



1: COPS_Campaign	2: Contact	1: Coverage	0: Location
0: Parameters	1: Keywords	0: Distribution	1: Languages
1: Lit.Reference	0: ProcessStep		

CERA-2 Metadata of the Data

DB Time of Extraction

Entry Name

Entry Acronym

Entry Id

Selection List Entry Type

webbased or local fill in

output is xml-file

#1 of 2: Contact Information Related to the Data

Selection List Type of the Contact

investigator

Description of the Contact Type

investigator (worked data out)

DOI author
administrator
contact_DOI_author
copyright
distributor
inhouse (in this institute)
investigator
leading centre (enforcing centre)
load data
metadata
metadata reviewer
n/a
not filled
originating centre
originator

OK

Cancel

Selection List Institute Info

not filled

Institute Name

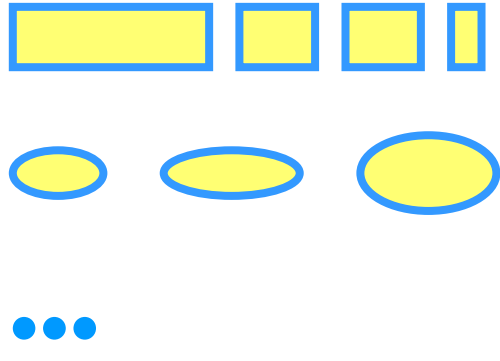
Bundesamt fuer Seeschifffahrt und Hydrografie

Data structure 1

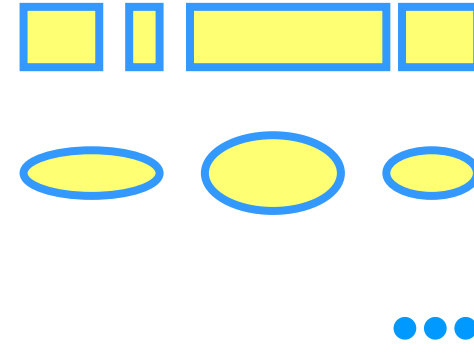
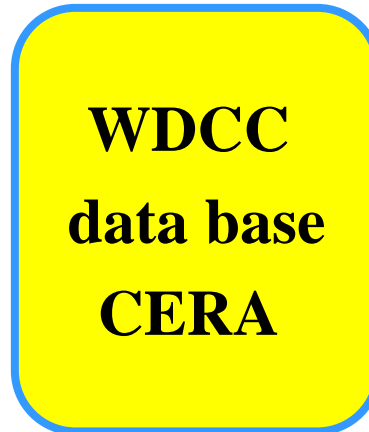
upload

download

A
B
C
...



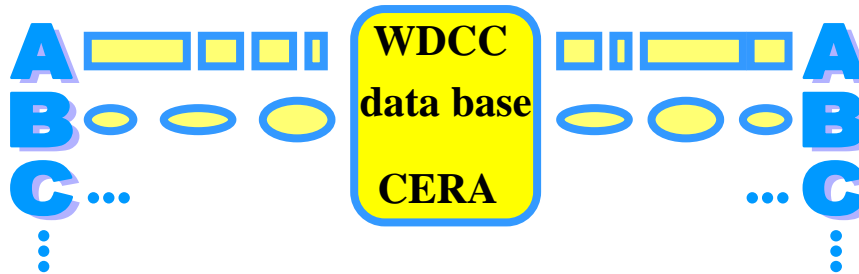
Data sets



A
B
C
...

Upload data structure
defines the access options
for downloading

Data structure 2



According to
user needs

Examples for download structure/**data set definition**:

A: focus on case studies (COPS, D-PHASE ?)

- **Specific day** -> all instruments, models, pics

B: focus on statistics (GOP ?)

- **Specific parameter** -> timeseries of observation period

C: other

- vertical model profiles / subregions



CERA user account
set access rights
according to data policy

data user

view meta data
download data via
web interface

download
data in
batch mode

CERA data base

- Define data structure model (-> investigators)
- Provide meta data formular to investigators
- **Test runs for data delivery and upload are needed**
 - Prior to campaign start of each project
 - Each data group has to deliver representative test data
 - and full meta data description
- **Test run timeline**
 - GOP: NOV 2006
 - DPHASE: FEB 2007
 - COPS: APR 2007

- Registration of data as DOI (digital object identifier) is strongly advised
- Advantages:
 - data in final version are peer reviewed by review agency
 - citation of published data is possible like a reviewed scientific article
 - completeness of data set descriptions (metadata) is needed
 - quality of data values (precision, sequence and ranges) is needed



contact information



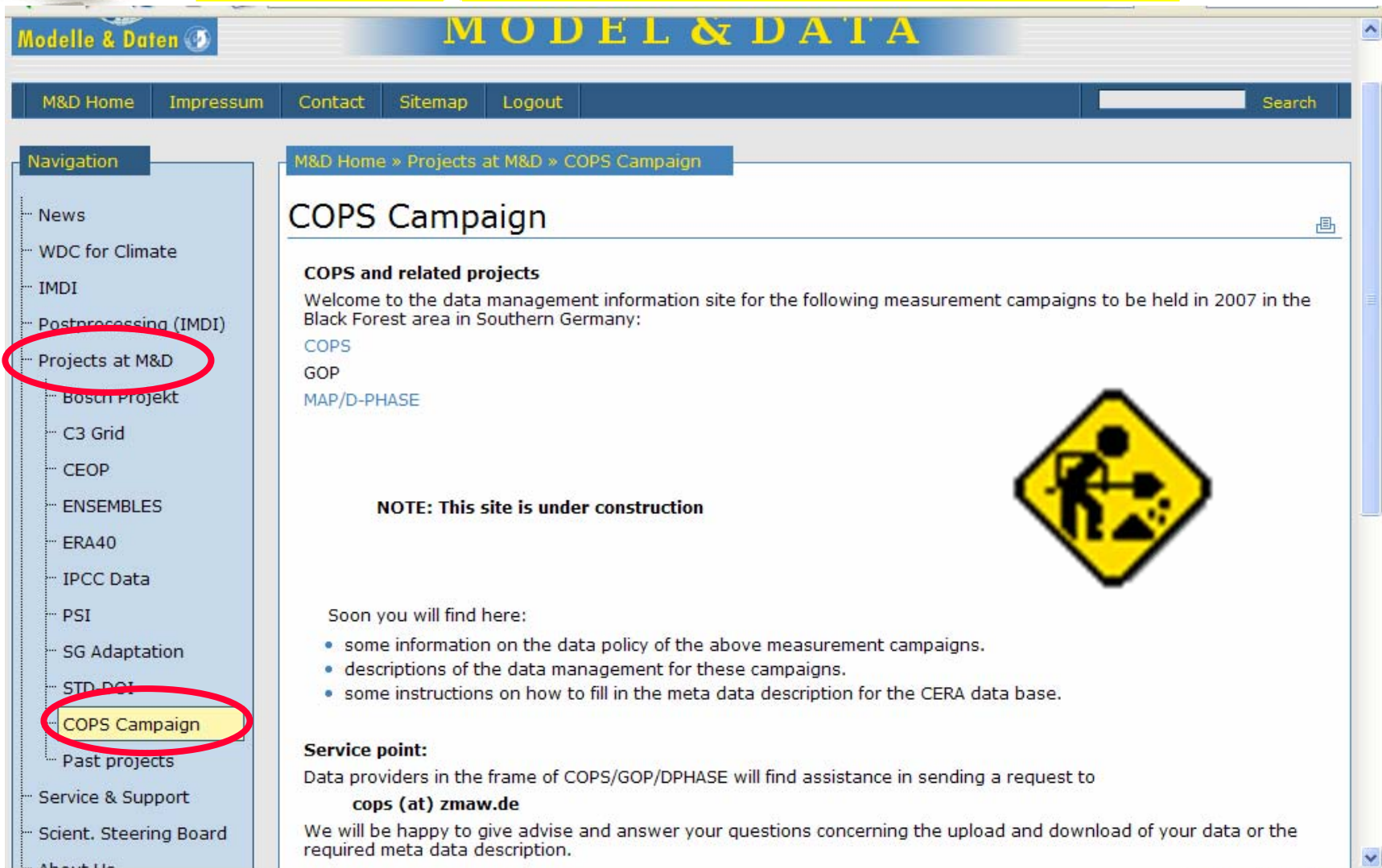
Service email adress:

cops@zmaw.de

User information on:

cops.wdc-climate.de





The screenshot shows the 'MODEL & DATA' website. The main navigation menu includes 'M&D Home', 'Impressum', 'Contact', 'Sitemap', and 'Logout'. A search bar is located on the right. The left sidebar contains a 'Navigation' menu with items like 'News', 'WDC for Climate', 'IMDI', 'Postprocessing (IMDI)', 'Projects at M&D', 'Bosch Projekt', 'C3 Grid', 'CEOP', 'ENSEMBLES', 'ERA40', 'IPCC Data', 'PSI', 'SG Adaptation', 'STD PSI', 'COPS Campaign', and 'Past projects'. The 'COPS Campaign' item is highlighted in yellow. The main content area displays the 'COPS Campaign' page, which includes a breadcrumb trail 'M&D Home » Projects at M&D » COPS Campaign', a title 'COPS Campaign', and a section 'COPS and related projects'. The text on the page reads: 'Welcome to the data management information site for the following measurement campaigns to be held in 2007 in the Black Forest area in Southern Germany: COPS, GOP, MAP/D-PHASE'. A yellow diamond-shaped warning sign with a black silhouette of a person digging is placed on the page. Below the sign, a note states: 'NOTE: This site is under construction'. Further down, it says 'Soon you will find here:' followed by a bulleted list: 'some information on the data policy of the above measurement campaigns.', 'descriptions of the data management for these campaigns.', and 'some instructions on how to fill in the meta data description for the CERA data base.'. A 'Service point:' section follows, stating 'Data providers in the frame of COPS/GOP/DPHASE will find assistance in sending a request to cops (at) zmaw.de'. The final paragraph says 'We will be happy to give advise and answer your questions concerning the upload and download of your data or the required meta data description.'

Model & Data: W...

Adresse <http://www.mad.zmaw.de/>

M&D Modelle & Daten

MODEL & DATA

M&D Home Impressum Contact Sitemap Login Search

Navigation

- News
- WDC for Climate
- Postprocessing (IMDI)
- Projects at M&D
- Service & Support
- Scient. Steering Board
- About Us
- Tool Bar

Quick Links

- CERA - Portal
- CERA - Browse Experiments

External Links

- DKRZ
- MPI-M

M&D Home

Welcome to the Model & Data Homepage

The Model and Data group (M&D) is hosted at the [Max Planck Institute for Meteorology](#) in Hamburg. Group's mission is to provide central support for the German and European climate research community. Emphasis is on development and implementation of best practice methods for Earth System modelling and related data management.

Latest News

2006-03-27
Introduction course 2006
Services and Facilities of DKRZ and Model & Data
[\[read more >\]](#)
Cat: Meetings & Workshops

2006-01-30
7. Deutsche Klimatagung
Klimatrends: Vergangenheit und Zukunft
[\[read more >\]](#)

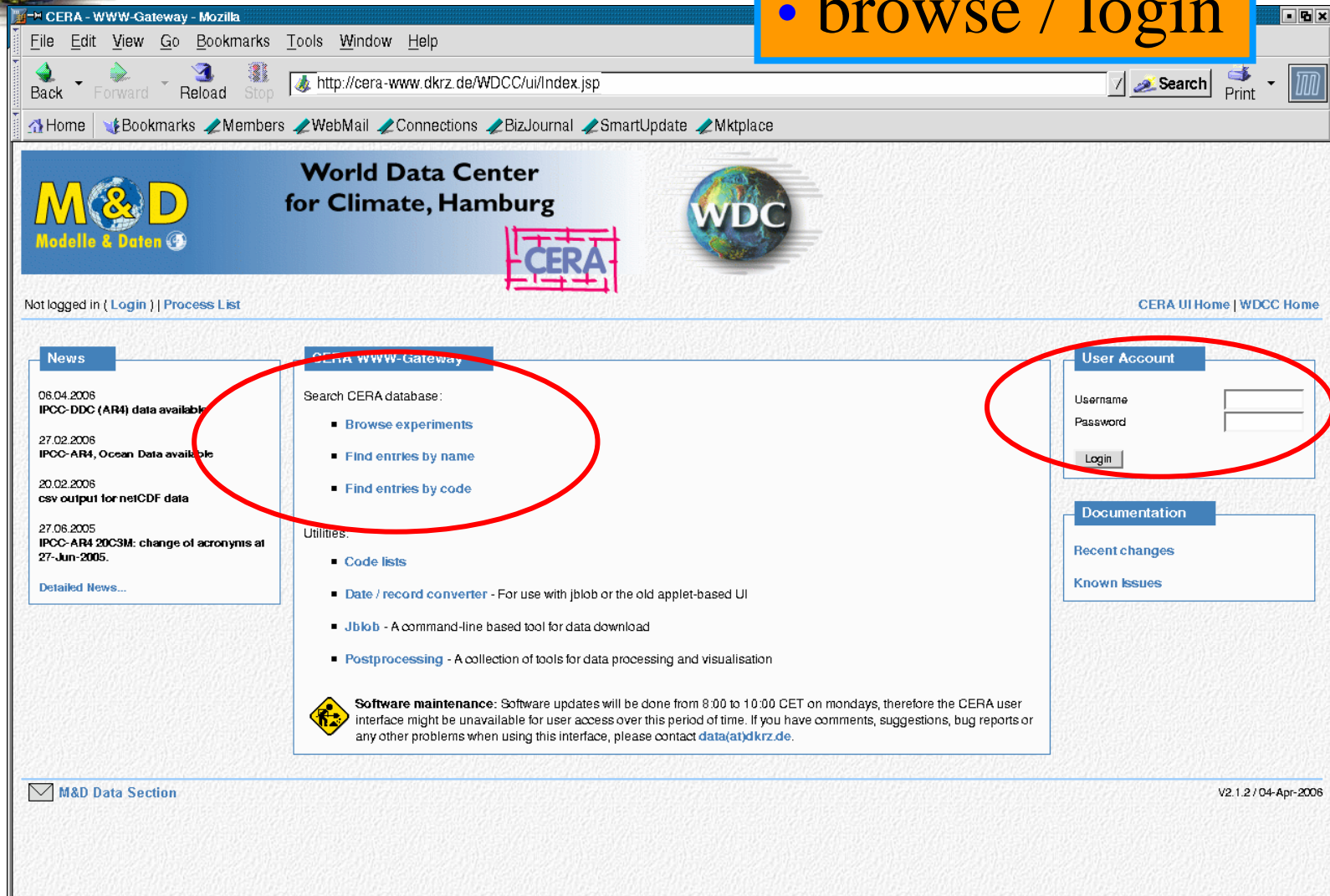
2006-01-18
PSI Hamburg Meeting
PSI Hamburg Meeting on Data on 13. and 14. February 2006
[\[read more >\]](#)
Cat: Data, Meetings & Workshops

News Feed
[XML](#) [RSS 2.0](#)

[\[show all news >\]](#)

Internet

• browse / login



File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://cera-www.dkrz.de/WDCC/ui/Index.jsp> Search Print

Home Bookmarks Members WebMail Connections BizJournal SmartUpdate Mktplace

M & D
Modelle & Daten

World Data Center
for Climate, Hamburg

WDC

Not logged in ([Login](#)) | [Process List](#) [CERA UI Home](#) | [WDCC Home](#)

News

08.04.2006
IPCC-DDC (AR4) data available

27.02.2006
IPCC-AR4, Ocean Data available

20.02.2006
csv output for netCDF data

27.06.2005
IPCC-AR4 20C3M: change of acronyms at 27-Jun-2005.

[Detailed News...](#)


CERA WWW-Gateway

Search CERA database:

- [Browse experiments](#)
- [Find entries by name](#)
- [Find entries by code](#)

Utilities:

- [Code lists](#)
- [Date / record converter](#) - For use with jblob or the old applet-based UI
- [Jblob](#) - A command-line based tool for data download
- [Postprocessing](#) - A collection of tools for data processing and visualisation

 **Software maintenance:** Software updates will be done from 8:00 to 10:00 CET on Mondays, therefore the CERA user interface might be unavailable for user access over this period of time. If you have comments, suggestions, bug reports or any other problems when using this interface, please contact data@dkrz.de.

User Account

Username

Password

Documentation

[Recent changes](#)

[Known Issues](#)

[M&D Data Section](#) V2.1.2 / 04-Apr-2006

- select experiment

Browse CERA Experiments

Select keyword(s) and/or project

GISS-AOM
GISS-EH
GISS-ER
glacial inception
glacial maximum
GLDAS
greenhouse gases
HADCM2
HADCM3
HADGEM

Select keyword(s)

Clear selection

FORCE_RM
GEBCO
GKSS
GLOWA-IMPETUS
HOAPS
IPCC-AR4_ECHO-G
IPCC-DDC_AR4
IPCC-DDC_SAR
IPCC-DDC_TAR
IPCC_HH

Select project

Clear selection

Project information

COPS

Experiments

AVHRR
Baltex Meteorological Data Centre (BMDC) from 31.03.2003
CC01CI01 - CTL: THE CONTROL INTEGRATION
CC01GG01 - GHG: THE GREENHOUSE GAS INTEGRATION
CC01GS01 - GHS: THE SULPHATE AEROSOL AND GREENHOUSE GAS INTEGRATION
CC01GS02 - GHS: THE SULPHATE AEROSOL AND GREENHOUSE GAS INTEGRATION
CC01GS03 - GHS: THE SULPHATE AEROSOL AND GREENHOUSE GAS INTEGRATION
CEOP Inter Transferability Studies: daily mean values of ECPC Regional Spectral Model
CEOP_BMRC_MOLTS_RAW
CEOP_CPTEC_GRID_RAW
CEOP_CPTEC_MOLTS_RAW
CEOP_ECMWF_GRID_RAW
CEOP_ECPCRII_GRID_RAW

- select data set

CERA - Datasets for experiment CNRM_CM3_SRESA1B_1 - Mozilla Firefox

http://cera-www.dkrz.de/WDCC/ui/DatasetList.jsp?acronym=CNRM_CM3_SRESA1B_1

Datasets for experiment CNRM_CM3_SRESA1B_1

Sort by Name Go

	Name	Progress
<input type="checkbox"/>	CNCM3_SRA1B_1_G_huss: surface specific humidity	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_oro: surface altitude	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_pr: total precipitation	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_psl: sea level pressure	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_rsds: surface downwelling shortwave radiation	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_sftgif: fraction of grid cell covered with glacier	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_sftlf: land area fraction	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_tas: 2m surface air temperature	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_uas: zonal surface wind speed	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_G_vas: meridional surface wind speed	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_N_oro: surface altitude	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_N_sftgif: fraction of grid cell covered with glacier	complete, review pending
<input type="checkbox"/>	CNCM3_SRA1B_1_N_sftlf: land area fraction	complete, review pending

Fertig

- view meta data

Not logged in (Login) | Process List CERA DT Home | WDC Home

Go to: [Topic](#) | [Code](#) | [Unit](#) | [Aggregation](#) | [Temporal Structure](#) | [Spatial Data Org.](#) | [Spatial Structure](#) Action: [Show all parameters](#)

CERA parameter information for dataset CLM_C20_1_D2_3h_T_2M

Topic

Parent (level 1)

Parent (level 2)

Topic

Topic Description

Code

Number

Type

Acronym

Description

Unit

Acronym	K
Name	Kelvin
Description	Kelvin

[Back to top...](#)

Aggregation

Description	3 hours interval
-------------	------------------

[Back to top...](#)

Temporal structure (period 1 of 1)

	Start	Increment	Additional time information	
Year	1960	0	Dimension type	real time
Month	1	0	Acronym	CLM_C20_D2_3H
Day	1	0	Name	CLM_C20_D2_3H
Hour	0	3	Description	CLM_C20_D2_3H
Minute	0	0		
Second	0	0	Start date	01/01/1960 00:00 (DD/MM/YYYY HH:MM)
			End date	31/12/2000 21:00 (DD/MM/YYYY HH:MM)
Length	119808 of 119808	record(s)		
UTC Diff	0.0	hour(s)		



End