metafor Common Metadata for Climate Modelling Digital Repositories

METAFOR Revised Metadata delivered into system METAFOR Deliverable 3.2 M39

PROJECT	
Project acronym	METAFOR
Project full title	Common Metadata for Climate Modelling Digital Repositories
Grant agreement no:	211753
Funding Scheme	Combination of Collaborative Projects & Coordination and Support Actions
Call Topic	INFRA-2007-1.2.1 Scientific Digital Repositories
DOCUMENT	
Deliverable	D3.2 Month 39
Deliverable Title	Revised Metadata delivered into system
Document Identifier	METAFOR-D3.1_M33
Date	July 11, 2011
Work Package	WP3 Service Content Creation & Testing
Authors	IPSL
Document Status	Draft
Document Link	http://metaforclimate.eu/documents

Dissemination Level		
PU	Public	
PP	Restricted to other programmes participants	**X**
RE	Restricted to a group specified by the Consortium	
CO	Confidential	

Document History			
Version	Date	Comment	Author/Partner
0.1	July 11, 2011	First Draft	S. Denvil/IPSL
1.0	July 27th 2011	Final Draft with minor edits	G.Devine/UREAD

Abstract:

There is growing interest among researchers, policy-makers, businesses, and the general public in the potential impacts of climate change and to what extent undesirable consequences of climate change can be mitigated by reducing anthropogenic greenhouse gas (GHG), and specifically carbon, emissions. Having that in mind, there is a pressing need to have systems in place to both harvest and expose climate modeling metadata, i.e.

CAPACITIES e-infrastructure

information that will lend greater understanding to the outputs from climate models. As part of the METAFOR project, systems were introduced to collect comprehensive metadata for CMIP5. This document will discuss how this work is being extended to harvest a set of metadata for the ENSEMBLES project. (This deliverable complements the D3.1 deliverable but is still a self sufficient document).

Introduction

The Metafor Common Information Model, i.e. the CIM, is an ontology designed to become the ipso-facto standard for climate modeling related metadata. The CIM eco-system allows institutes to integrate the CIM into their day-to-day climate modeling processes. It achieves this by supporting various requirements: the ontology itself; validation; search; dissemination; integration (with other metadata platforms such as Earth System Grid). This deliverable will describe the ingestion of CIM-compliant metadata

One of the prerequisites for all studies of climate change and its consequences is the existence of climate simulations at the global scale from which coherent atmospheric, oceanic or surface field time-series can be extracted for further modelling or statistical studies. We chose the FP6 ENSEMBLES project as our main data source for this demonstration for five reasons:

- It provides centennial and seasonal to decadal simulations
- It provides high temporal resolution datasets covering the entire simulations; well suited for impacts studies
- It is the first project that uses the E1 mitigation scenario (Lowe et al. 2009)
- Data are publicly available through standardized interfaces
- Key WP3 partners were already involved in the ENSEMBLES project

The CMIP5 questionnaire is a web-based interactive tool for creating CIM documents about the climate models that contribute to CMIP5. The questionnaire also captures information about how those climate models were set up to run simulations for the particular CMIP5 experiments. For Ensembles, we use a simplified version of the CMIP5 questionnaire to gather information about models used to perform seasonal to decadal and centennial simulations. A TDS2CIM tool (described later) was also used to gather information regarding the files themselves.

ENSEMBLES seasonal to decadal and centennial datasets

The ENSEMBLES project was funded by the European Commission (EC), and ran from September 2004 to December 2009. ENSEMBLES was a flagship project of the EC's 6th Framework Programme (EC FP), an integrated project under the thematic sub-priority 'Global Change and Ecosystems' (contract number GOCE-CT-2003-505539).

At the core of the ENSEMBLES integrated project was the development of the first global, high-resolution, ensemble based, modelling system for the prediction of climate change and its impacts. To enable this, earth system models were combined into a multi-model ensemble system, with a common output. The purpose of RT2A was to produce sets of climate simulations with several models and to provide the multimodel results needed for the other Research Themes. The results from RT2A were used for validation (RT5), studies of feedbacks in the Earth system (RT4), as well as boundary conditions and forcing fields for regional model simulations (RT3/RT2B). The simulations covered time-scales ranging from seasons to decades and centuries. Two streams of Global Climate Model (GCM) runs were produced: the first for the ensemble prediction system and the

second using later models incorporating new features such as carbon cycle feedbacks. The development and running of the E1 stabilisation scenario was led by RT2A.

The ENSEMBLES project built ensemble prediction systems based on global climate models to generate projections of future climate. The scope of this included assembly and testing of new global climate models, development and implementation of methods to represent the effects of uncertainties in the modelling of key physical, biological and chemical processes ('modelling uncertainties'), and the use of observations to initialise and constrain the projections. Seven European climate modelling centres took part running GCMs under historic and four different scenario forcings (B1, A1B, A2, 1%CO2). All centres ran several realisations to create multi-simulation ensembles of most scenarios, which together contributed to the multi-model ensemble developed in the project.

Two streams of coordinated seasonal–decadal experiments were carried out during the project:

Stream1 covered the 1991–2001 hindcast periods for seasonal to annual range with 7-month-long hindcasts started every May and November. The November start dates were extended to 14 months in order to cover a full calendar year. Each of the groups contributing to the multi-model ensemble ran nine member ensembles sampling uncertainties in the observed initial conditions. In addition, further nine-member ensembles were run to assess the stochastic physics and perturbed parameter approaches to sampling modelling uncertainties, using the IFS/HOPE and DePreSys systems, respectively. Papers documenting the results have been written (Berner et al., 2008; Doblas-Reyes et al., 2009). The perturbed parameter hindcasts were also tested in decadal prediction mode by extending the hindcasts for all 22 start dates. Partners contributing to the multi-model ensemble also carried out test decadal projections for two start dates (November 1965 and November 1994), using the results to inform the design of the subsequent stream 2 hindcasts.

Stream2 hindcasts consisted of a comprehensive set of seasonal, annual and decadal integrations (see papers documenting (a) the results (Oldenborgh, G.J. van, F.J. Doblas-Reyes, B. Wouters and W. Hazeleger (2011) and (b) the skill in the trend and internal variability in a multi-model decadal prediction ensemble (Climate Dynamics, accepted.)) The seasonal (7-month long) and annual (14-month long) hindcasts were performed over the 46-year hindcast period 1960–2005, with start dates every 4 months (February, May, August and November). This gave a total of 184 seasonal hindcasts. Ten multi-model decadal hindcasts were carried out over the same hindcast period, starting every 5 years (1960, 1965, 1970, …, 2005) in November. The 2005 start date also provides a future prediction for 2010–2014. The seasonal–annual hindcasts again consisted of nine ensemble members per model, whereas the decadal runs were done with three members per model. The DePreSys system was used to create a large set of decadal hindcasts, initialised every November during 1960–2005. The ensemble hindcasts consisted of the HadCM3 model variant with standard parameter settings plus eight variants distinguished by multiple parameter perturbations.

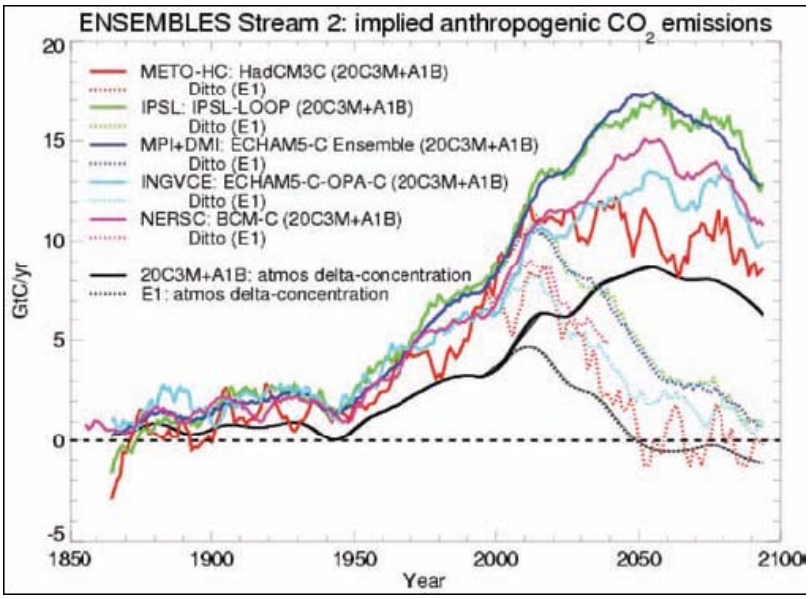
The ENSEMBLES stream 2 decadal hindcasts provided a first opportunity to assess the benefits of combining projections from different models in a coordinated experiment, following initial studies carried out with individual climate models (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009).

Two streams of coordinated centennial were carried out during the project:

The modelling groups involved (CNRM, DMI, IPSL,METOHC, MPIMET, NERSC and FUB) performed the **Stream 1** simulations using the common set of agreed forcings for the historical simulations over the period 1860–2000, and for the three recommended IPCC SRES scenarios A2, A1B and B1 over the 21st century. Some simulations were extended beyond the year 2100 with constant atmospheric concentrations from the B1 and A1B scenarios. Additional simulations with a 1% increase of CO2 per year with stabilisation at 2×CO2 and 4×CO2 were also performed.

Stream2 simulations made use of improved coupled atmosphere ocean models, as developed in Work Package 1.1 of RT1. Differences were introduced in the coupling of an interactive carbon cycle in five Earth system models, so that the net CO2 fluxes between both atmosphere and ocean, and atmosphere and land, can be computed interactively, (depending on the prescribed atmospheric CO2 concentrations, as proposed by Hibbard et al. (2007)). The carbon cycle model components generally describe the carbon storage in different pools related to vegetation and soils, and the carbon uptake and cycling in the oceans (see figure 1).

Eight of the models can be forced with land cover changes. Land cover changes influence the physical properties of land surfaces, and imply carbon emissions in models including land carbon pools for vegetation and soils, e.g. when deforestation takes place. Aerosol and/or chemistry models were introduced in three models. Further improvements concerning details of the model formulations have also been included, so that the new set of Earth system models represents a considerable step forward towards future models with prognostic treatment of all major greenhouse gases. Several models used here represent prototypes of Earth system models that will be used for CMIP5 simulations, in support of the Fifth Assessment Report of IPCC. A paper documenting the results has been written (Tim Johns et al., 2011).



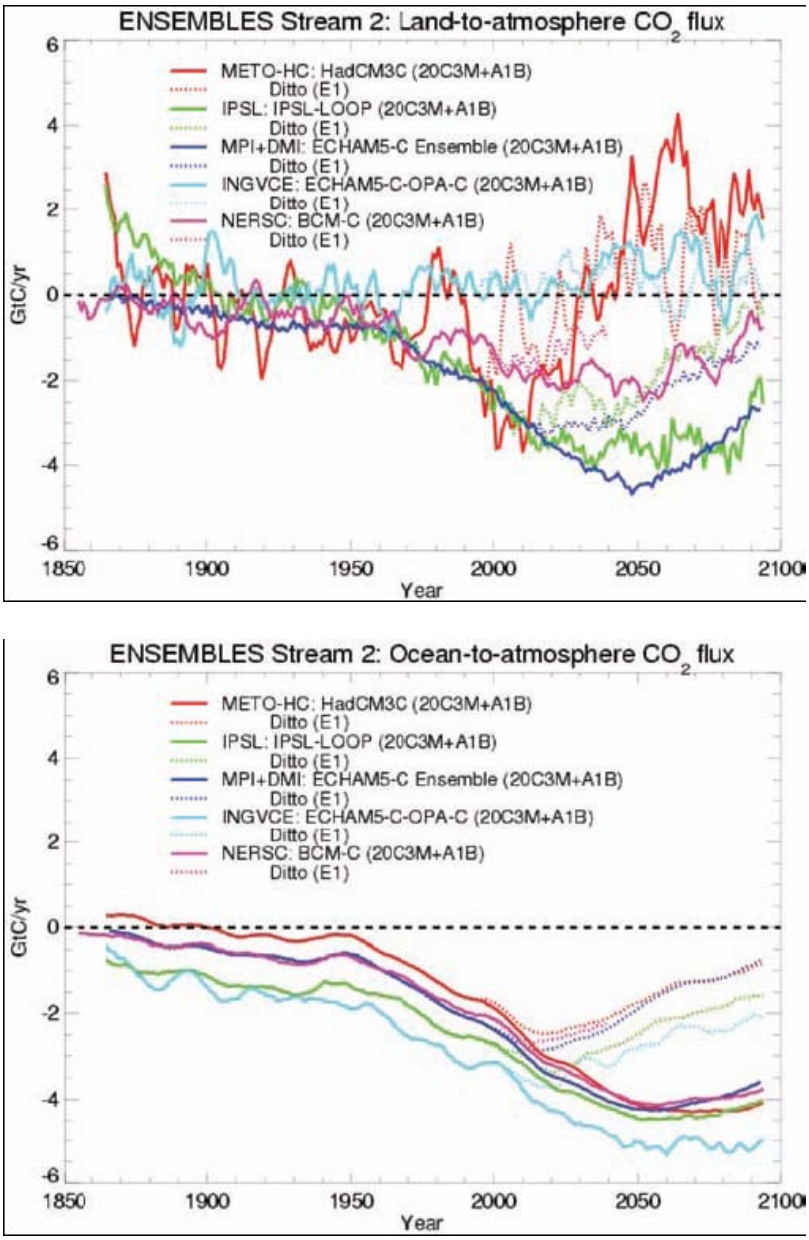


Figure 1: Implied ('permitted') anthropogenic net carbon dioxide emissions to the atmosphere (Gt C/yr) in ENSEMBLES RT2A Stream 2 runs for the 20th and 21st century (top panel) as diagnosed from the imposed change in atmospheric concentrations (black curves) and the modelled net carbon flux exchange between the atmosphere and the land surface (middle panel) and ocean (bottom panel). Note that an 11-year mean smoothing has been applied to all curves (including delta-concentrations) and that MPI+DMI: ECHAM5-C results show ensemble means of eight (20C3M+A1B) and five (E1) independent simulations, tending to smooth those results compared with other models.

METAFOR Questionnaire

The CMIP5 questionnaire, developed as part of the EU-funded METAFOR project, is a comprehensive metadata collection tool, introduced for CMIP5 to collect a wide range of climate modelling metadata, which will allow a much wider understanding of the eventual datasets produced. This includes information about the model software itself. For

ENSEMBLES it was decided to employ the same questionnaire machinery to harvest ENSEMBLES-relevant metadata. However, the rigorousness of the actual metadata collected will be simplified somewhat, and allow the user to fill out as much as they can without the more stringent checks implemented for CMIP5. In particular, a primary aim for ENSEMBLES will be to capture specific information about how the models conformed to the requirements of the ENSEMBLES experiments.

ENSEMBLES decadal:
The ENSEMBLES decadal experiment we document is the equivalent to the CMIP5 "Experiment 1.1 decadal " (found at http://q.cmip5.ceda.ac.uk/cmip5/15/experiment/203/).

-Ensemble of 10-year hindcasts and predictions. Assess model skill in forecasting climate change on time-scales when the initial climate may exert some influence.

The atmospheric composition (and other conditions) should be prescribed using some estimate of the observed greenhouse gas and anthropogenic aerosol concentrations and the A1B scenario from 2000. Total solar irradiation variations should be prescribed too. No observed volcanic aerosol load information should be used during the hindcast, and only some sort of persistence of the load at the initial date is allowed. Ocean initial conditions should be in some way representative of the observed anomalies or full fields for the start date. Land, sea-ice and atmosphere initial conditions are left to the discretion of each group. Simulations should be initialized on the 1st of November of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000 and 2005. A minimum ensemble size of 3 should be produced for each start date.

Requirements
1.1.ic.oc InitialCondition
Ocean Initial Conditions must represent in some measure the observed anomalies for the start date used
1.1.bc.ant_wmg BoundaryCondition Imposed changing concentrations of well mixed gases (anthropogenic)
1.1.bc.ant_aer BoundaryCondition Imposed changing concentrations or emissions of aerosols (anthropogenic)
1.1.bc.solar BoundaryCondition Imposed changing solar forcing
1.1.bc.NoVolc2005 BoundaryCondition No volcanic aerosol emissions beyond 2005
1.1.bc.seaice BoundaryCondition Sea ice observations forcing (unless modelled) must precede start date of simulation
1.1.stc.decadal_10yr SpatioTemporalConstraint Run for 10 years
Required Duration: 1960-09-01T00:00:00Z to 1970-10-31T23:59:59Z plus 10 years
Required Resolution: None
1.1.bc.a1b_wmg BoundaryCondition Imposed changing concentrations or emissions of A1B well mixed gases, including CO2
1.1.bc.a1b_sls BoundaryCondition Imposed changing concentrations or emissions of A1B short lived gas species
1.1.bc.a1b_aer BoundaryCondition Imposed changing concentrations or emissions of A1B aerosols

The ENSEMBLES centennial experiments we document are very close to the
- Experiment 3.2 historical:
http://q.cmip5.ceda.ac.uk/cmip5/15/experiment/50/

- Experiment 4.1 rcp45:
http://q.cmip5.ceda.ac.uk/cmip5/15/experiment/3/
- Experiment 4.3 rcp26:
http://q.cmip5.ceda.ac.uk/cmip5/15/experiment/44/

Experiment 20C3M:
Simulation of recent past. Evaluate model performance against present climate and observed climate change. Impose changing conditions (consistent with observations).

For the 1860 to 2000 (present day) period, the 20C3M (twentieth Century in Coupled Climate Models) model run specifies anthropogenic forcings (GHGs, aerosols as concentrations or precursor emissions, ozone, and land use change), in most cases without any variation in natural (solar and volcanic) forcings. Some additional 20C3M simulations were also conducted including solar and volcanic forcings as specified in previous AR4 simulations, to allow a better comparison with observed changes for validation purposes, and two models only ran 20C3M simulations with anthropogenic plus solar and volcanic forcing.

Requirements
2.1.ic InitialCondition
Initial conditions are from experiment 3.1.piControl
2.1.bc.ant_wmg_conc BoundaryCondition
Imposed changing concentrations of historical anthropogenic well mixed gases, including CO2
2.1.bc.volc_wmg_conc BoundaryCondition
Imposed changing concentrations of historical volcanic well mixed gases, including CO2
2.1.bc.sls BoundaryCondition
Imposed changing concentrations or emissions of historical short lived gas species
2.1.bc.nat_aer BoundaryCondition
Imposed changing concentrations or emissions of historical natural aerosols
2.1.bc.ant_aer BoundaryCondition
Imposed changing concentrations or emissions of historical anthropogenic aerosols
2.1.bc.nat_aer_pre BoundaryCondition
Imposed changing concentrations or emissions of historical natural aerosol precursors
2.1.bc.solar BoundaryCondition
Imposed historical changing solar forcing
2.1.bc.LU BoundaryCondition
Imposed historical changing land use
2.1.stc.1860_141yr SpatioTemporalConstraint
Begin in 1860 and run for 141 years
Required Duration: 1860-01-01T00:00:00Z to 2006-01-01T00:00:00Z plus 141 years
Required Resolution: None
2.1.bc.ant_aer_pre BoundaryCondition
Imposed changing concentrations or emissions of historical anthropogenic aerosol precursors

Experiment A1B
Provide estimate of future anthropogenic climate change. Future projection (2001-2100) forced by A1B.

Within the SRES set, the A1B scenario forms a medium-high emission scenario driven by high economic growth, strong globalization and rapid technology development. The scenario also assumes a material-intensive lifestyle so energy consumption grows rapidly despite population growth being relatively low (the population peaks around 9 billion in 2050 and declines to around 7 billion in 2100). The energy supply as a balance between fossil fuel and non-fossil fuel sources. The A1B scenario has been chosen as the baseline scenario for the ES2 simulations because it provides overlap with earlier climate modelling work.

Requirements

3.1.ic InitialCondition
Initial conditions are from the end of experiment 20C3M
3.1.bc.wmg BoundaryCondition
Imposed changing concentrations or emissions of A1B well mixed gases, including CO2
3.1.bc.sls BoundaryCondition
Imposed changing concentrations or emissions of A1B short lived gas species
3.1.bc.aer BoundaryCondition
Imposed changing concentrations or emissions of A1B aerosols
3.1.bc.aer_pre BoundaryCondition
Imposed changing concentrations or emissions of A1B aerosol precursors
3.1.bc.LU BoundaryCondition
Imposed changing A1B land use
3.1.stc.2001_100yr SpatioTemporalConstraint
Begin in 2001 and run for 100 years
Required Duration: 2001-01-01T00:00:00Z to 2100-01-01T00:00:00Z plus 100 years
Required Resolution: None

Experiment E1

Provide estimate of future anthropogenic climate change. Future projection (2001-2100) forced by E1.

The experiment contrasts the A1B baseline with a corresponding aggressive mitigation scenario E1 (Lowe et al. 2009) developed with the IMAGE 2.4 IAM. Meinshausen et al. (2006) indicate that stabilization of GHG concentrations at 450 ppmv (CO2-equivalent) would provide a 20–75% probability of stabilizing temperatures below a 2 K warming target. Starting from an A1B baseline, a ‘‘peaking’’ scenario was developed which initially peaks at around 530 ppmv CO2-e and then decreases gradually to approach 450 ppmv from above during the twenty-second century. Den Elzen and van Vuuren (2007) show that peaking scenarios may be preferable to stabilization scenarios, on the basis of cost-effectiveness considerations, for reaching long-term temperature targets. The GHG concentration data for this experiment has been calculated using the IMAGE model.

Requirements

4.1.ic InitialCondition
Initial conditions are from the end of experiment 20C3M
4.1.bc.wmg BoundaryCondition
Imposed changing concentrations or emissions of E1 well mixed gases, including CO2
4.1.bc.sls BoundaryCondition
Imposed changing concentrations or emissions of E1 short lived gas species
4.1.bc.aer BoundaryCondition
Imposed changing concentrations or emissions of E1 aerosols
4.1.bc.aer_pre BoundaryCondition

Imposed changing concentrations or emissions of E1 aerosol precursors
4.1.bc.LU BoundaryCondition
Imposed changing E1 land use
4.1.stc.2001_100yr SpatioTemporalConstraint
Begin in 2001 and run for 100 years
Required Duration: 2001-01-01T00:00:00Z to 2100-01-01T00:00:00Z plus 100 years
Required Resolution: None

TDS2CIM tools

How the METAFOR tools (especially tds2cim) gathered information from ENSEMBLES seasonal to decadal and centennial datasets.

Using models inside experiments produce a large amount of data files. These files are stored in traditional data archives. Several file formats like GRIB, netCDF, HDF or others are used for climate data files. The most common file formats are the netCDF and the GRIB format.

NetCDF especially in the CF convention contains (use) metadata describing those data which are stored in the file. These metadata are available for automatic capture and can be automatically transferred into the CIM metadata schema.

Regarding ENSEMBLES datasets, the aim was to develop a database system in a common format, allowing easy access by all the partners to selected results of the global ensemble simulations. Typically, an atmosphere–ocean coupled simulation can generate about half a terabyte of data for a 100-year simulation if daily fields are stored. A similar amount is found in ensemble seasonal simulations. Model results are stored in the MARS storage system of ECMWF, and of the Climate and Environmental Retrieval and Archive (CERA) database system at the World Data Centre for Climate hosted by the Data Management Group in Hamburg (DKRZ). Common lists of variables and the need for a common format were outlined in the early stages of the project, depending on the requirements of the scientific community taking part in the other Research Themes. Project data will still be updated and available online after the project has ended.

We have then two different data sources for the purpose of our demonstration.

A large set of atmosphere and ocean variables from the multimodel, stochastic physics and perturbed parameter experiments s2d (seasonal to decadal) integrations are centrally stored at ECMWF for quality control, basic forecast quality assessment, and dissemination. The atmospheric variables are archived on ECMWF's Meteorological Archival and Retrieval System (MARS) in GRIB (gridded binary) format. The fields are stored following a set of atmospheric conventions, based on the experience gained in DEMETER (Palmer et al., 2004) and the operational European multi-model seasonal forecasts. The encoding of the ocean variables is carried out using rules based on newly developed conventions, with storage of CF-compliant NetCDF files into the ECFS server.

The ENSEMBLES s2d data have been made available over the internet without charge for use in research, education and commercial work. Two dissemination systems, one based on MARS and another one based on the Open-source Project for a Network Data Access

Protocol (OpenDAP), have been developed to help users to access the ENSEMBLES data in the most efficient way for their specific requirements.

DKRZ Data Management Group has established a website to enable easy access to the ENSEMBLES-related multi-decadal simulations. After providing the necessary information to the data-providing centres, metadata for most experiments have been completed. The web site is continually kept up-to-date to include in the CERA database the datasets provided by the modelling groups. ENSEMBLES data are archived in the Climate and Environmental Retrieval and Archive (CERA) of the World Data Centre System for Climate (WDCC) run by the DKRZ Data Management Group. Access is given by http://ensembles.wdc-climate.de. The experiments for Stream 1 and Stream 2 are briefly described and access is given to each scenario variable of each contributing institute.

There are two ways to produce metadata which shall run into the Metafor repository database (eXist, database as data source for the Metafor portal, see figure 2).

1. **Access the ECMWF thredds data server** and map the information into the CIM dataObject document
This functionality is an extension of the tds2cim tool. This function parses the ECMWF thredds data server, accesses the file OpenDAP server one by one, retrieves the available information and maps it into the appropriate fields of a CIM dataObjects. The result file is uploaded into the eXist database. This tds2cim tool is implemented in the Metafor portal as function to use as a data ingest source. After registration the registered thredds data server is scanned in the provided time steps and all available data are mapped and uploaded into the repository. The mapping to an existing simulation is done inside the repository by another service.

2. **Access the CERA Metadata** by a provided XSQL output
This mapping functionality uses the Oracle database XSQL capability to export from CERA metadata RDBMS tables by XSQL into a XML. The result XML is styled by a XSLT style sheet into a CIM dataObject document. The link for doing this is:
http://anticyclone.dkrz.de:8080/xsql/cera_map_cim.xsql?request=dataObject&ac=HADGEM_SRA1B_1_DM_evspsbl
- 'request' means the type of output xml format, here CIM dataObject.
- 'ac' means a CERA dataset acronym which is available from the CERA experiment browser GUI.

Upload into the Metafor repository will be done by the upload functionality used for the tds2cim tool. The connection between a simulation and data is made inside the repository (eXist database) by a service (like a cron job). This service is looking for simulations without data and data without simulations (inside the database). If it's possible to connect a simulation to data (e.g. by the DRS: Data Reference Syntax) the connection is made by inserting the simulation uuid into the sourceSimulation field of one or more CIM dataObject.

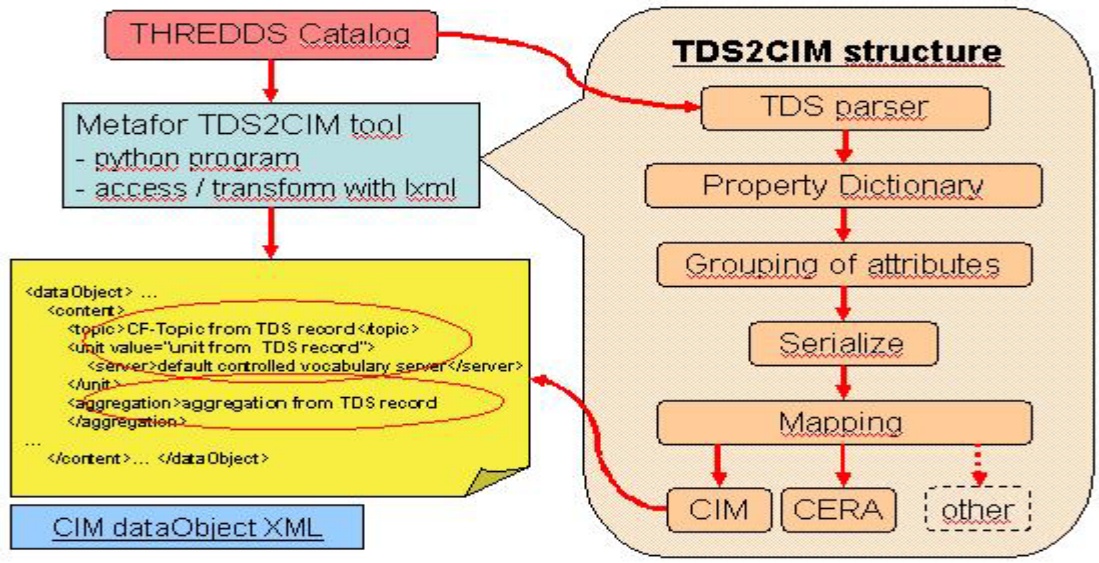


Figure 2:TDS2CIM structure.

Conclusions

Due to the complexity of the climate system and the variety of experimental designs a climate modeling software can be configured to conform to, it is crucial to offer to end users a comprehensive representation of this information. With respect to this point, description of the experimental design and description of the available data are critical point in the demonstration context.

The simplified CIM metadata describing the EU ENSEMBLES project that has been defined, produced and ingested into the system will be accessible through the METAFOR portal. http://www.purl.org/org/esmetadata/cim/portal

METAFOR provides essential pieces and framework to support parties building climate information system. From the end user perspective, the generic features of both the CMIP5 questionnaire and the tds2cim tool are essential as they are the major sources of CIM information up to now.

Bibliography

[1] ESG publisher scripts: http://www2-pcmdi.llnl.gov/Members/bdrach/.personal/esg-publication-scripts/

[2] Metafor CMIP5 questionnaire: http://q.cmip5.ceda.ac.uk/cmip5/

[3] CERA data model: http://www.mad.zmaw.de/wdc-for-climate/cera-data-model/

[4] World Data Center for Climate (WDCC): http://www.mad.zmaw.de/wdc-for-climate/

[5] CERA database: http://www.mad.zmaw.de/wdc-for-climate/cera-database/

[6] CERA metadata with internal experiment

Example for a dataObject describing a complete ENSEMBLES experiment
http://anticyclone.dkrz.de:8080/xsql/cera_map_cim.xsql?request=datObject&ac=ENSEMBLES_CNCM3_SRA1B_1_D

Or one dataset of that experiment:
http://anticyclone.dkrz.de:8080/xsql/cera_map_cim.xsql?request=datObject&ac=CNCM3_SRA1B_1_DM_clwvi

[7] CIM controlled vocabulary, CV: http://metaforclimate.eu/trac/wiki/ticket/192

[8] MindMap CV capturing: http://metaforclimate.eu/trac/wiki/tickets/245

[9] Linux MindMap tool (freemind):
http://freemind.sourceforge.net/wiki/index.php/Main_Page

[10] CF (Climate and Forecast) description: http://cf-pcmdi.llnl.gov/

[11] Metafor CIM tools and services:
http://metaforclimate.eu/trac/browser/cim_software/trunk/docs/design/METAFOR_D5_5.odt

[12] THREDDS: http://www.unidata.ucar.edu/projects/THREDDS/

[13] DRS: http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf

[14] CMIP5 questionnaire roadmap:
http://metaforclimate.eu/trac/attachment/ticket/721/CMIP5QuestionnaireRoadmapBeta5.pdf

[15] TDS2CIM tool / service: http://metaforclimate.eu/trac/wiki/tickets/707

[16] ENSEMBLES final report: http://ensembles-
-u.metoffice.com/docs/Ensembles_final_report_Nov09.pdf