Poster
Z18

Vienna, Austria, April 2012
Monday, 23 Apr 2012, 08:00-19:30
author in attendance:17:30-19:00

The PMIP 3 Database

Jean-Yves Peterschmitt[1], Sébastien Denvil[2], Mark Morgan[2], Ashish Bhardwaj[2], and the PMIP 3 Participants Team

[1] Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Gif-sur-Yvette, France, *jean-yves.peterschmitt@lsce.ipsl.fr*
[2] IPSL, UPMC, Paris, France

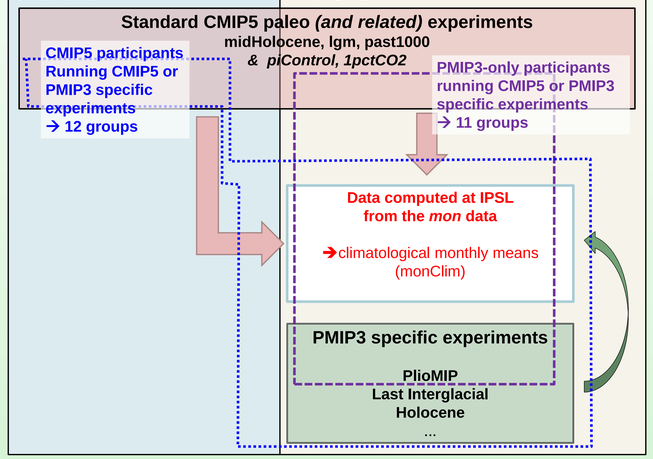
Abstract EGU2012-5982

The Paleoclimate Modelling Intercomparison Project (PMIP) is a long standing initiative that has provided an efficient mechanism for coordinating paleoclimate modelling activities that provide valuable information on the mechanisms of climate change, the identification of key feedbacks operating in the climate system and, through model evaluation, the capability of climate models to reproduce climates different from today. The third phase of PMIP (aka PMIP 3) started in 2009 and three of its key periods have been selected as part of the tier1 and tier2 of the CMIP5 set of simulations: the Last Glacial Maximum, the mid-Holocene and the Last Millennium periods. PMIP3 will also focus on mid-Pliocene (PlioMIP project) and Last Interglacial periods, and on the impact of freshwater events in the North Atlantic (such as the 8.2ka event in the early Holocene).

We will present an up-to-date overview of PMIP 3, in terms of participation to the different PMIP 3 time slices and available experiments. We will also show how we use the CMIP5 data hosting and distribution technology and the PRODIGUER server at IPSL to handle the CMIP5-like simulations of "PMIP 3 but non-CMIP5" participants, and to handle the non-CMIP5 simulations.

References:
http://pmip3.lsce.ipsl.fr/
http://forge.ipsl.jussieu.fr/prodiguer

Group and data types in the PMIP3/CMIP5 DB



Notes :
- The relative sizes of the DBs above are not correct ☺
- The CMIP 5 data and the PMIP 3 data is physically distributed over many data centers

Paleoclimate Modelling Intercomparison Project 3

PMIP3 Data: http://esgf-node.ipsl.fr/

What is the PMIP3 Database ?

- It is a **HUGE** and **standardized** collection of (sometimes huge) files
  - the DB has to be physically distributed on different servers over the world
- Some files are part of the IPCC5/CMIP5 Database so we sometimes talk about the ***PMIP3/CMIP5 Database***

- The PMIP3 experiments are **standardized**, so that everybody uses the same experimental design
  - https://pmip3.lsce.ipsl.fr/

- The data files are **standardized** with the **CMOR2 library**, so that the data can be easily exchanged and compared
  - « CMIP5 Model Output Requirements » document
    http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_output_metadata_requirements.pdf
  - list of requested variables per experiment for the **CMIP5 DB**
    http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.xlsx
    ~1100 variables for CMIP5, created with 19 **CMOR2 tables**
    CMIP5_3hr CMIP5_6hrLev CMIP5_6hrPlev CMIP5_aero CMIP5_Amon CMIP5_cf3hr CMIP5_cfDay CMIP5_cfMon CMIP5_cfOff CMIP5_cfSites CMIP5_day CMIP5_fx CMIP5_grids CMIP5_Lmon CMIP5_Lmon CMIP5_Oclim CMIP5_Olmon CMIP5_Omon CMIP5_Oyr
  - list of requested variables per experiment for the **PMIP3 DB**
    - a **subset of CMIP5**: ~470 variables in 7 CMOR2 tables
      http://www2-pcmdi.llnl.gov/cmor/tables/copy3_of_cmip5-tables/
      PMIP3_Amon PMIP3_day PMIP3_fx PMIP3_Lmon PMIP3_Lmon PMIP3_Olmon PMIP3_Omon
    - **climatological monthly means (*monClim* variables)** are automatically computed at IPSL from *Xmon* data supplied by the modeling groups: Amon → Aclim, Omon → Oclim, …
  - files are in NetCDF format and follow the CF (Climate and Forecast) convention
  - the file names and directories follow the ***DRS***
    - « CMIP5 Data Reference Syntax (DRS) » document
      http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf
  - 1 variable / file, variables can be split along the time axis
  - the variables are grouped in **datasets**

The *DRS*

- **DRS ⇔ Data Reference Syntax**
- You **have to** understand the DRS (or at least know about it…) to find data in the DB and use it

- **CMOR directory structure**
  <activity>/<product>/<institute>/<model>/<experiment>/<frequency>/
  <modeling realm>/<variable name>/<ensemble member>/<CMOR filename>.nc

- **File names**
  <variable name>_<MIP table>_<model>_<experiment>_
  <ensemble member>[_<temporal subset>].nc
  with <temporal subset> = 'yyyy[mm[dd[hh][mm]]][-clim]'

- **Datasets**
  <activity>.<product>.<institute>.<model>.<experiment>.<frequency>.
  <modeling realm>.<MIP table>.<ensemble member>.<version>
  - a dataset logically groups all the variables of the same type (same realm, frequency and CMOR2 table) of an experiment performed by a modeling group. If you change 1 file (1 variable can be split into several files along the time axis), you have to change the version of the dataset
  - the DB access gateways help you find the datasets holding the variables you are looking for
  - the gateways show the latest version of the datasets (by default)

- **Data node directory structure**
  users won't have to deal with this structure, unless they use a tool like *synchro_data* to duplicate a subset of the DB at their site
  <activity>/<product>/<institute>/<model>/<experiment>/<frequency>/
  <modeling realm>/<MIP table>/<ensemble member>/<version number>/
  <variable name>/<CMOR filename>.nc

- The fields of the DRS can only have **predefined values** (aka *Controlled Vocabulary*)
  - **activity**: *PMIP3* or *CMIP5*
  - **product**: *output* (PMIP3) or *output1* or *output2* (CMIP5)
  - **institute**: one of the expected contributing groups (institute column in the PMIP3 participants table)
  - **model**: one of the expected models (see model id column)
  - **experiment**: standardized short name of one of the PMIP3/CMIP5 experiments
    eg : *piControl*, *midHolocene*, *lgm*, *past1000*, *1pctCO2*, …
  - **frequency**: *fx*, *yr*, *monClim* (climatological monthly mean), *mon*, *day*, *6hr*, *3hr* or *subhr*
  - **modeling realm**: *atmos*, *ocean*, *land*, landIce, *seaIce*, *aerosol*, *atmosChem*, *ocnBgchem*
  - **variable name**: the standardized short name of a variable (eg *tas*, *ua*, *va*, *landCoverFrac*, …)
  - **MIP table**: the CMOR2 table where the variable is defined (eg *fx*, *Amon*, *Aclim*, …)
  - **ensemble member**: usually *r0i0p0* for fx fields and *r1i1p1* for the rest. Other values for ensemble runs
  - **version number**: the version of the dataset where the variable was found (usually a *yyyymmdd* date string)

- Example
  PMIP3/output/IPSL/IPSL-CM5A-LR/lgm/
  monClim/atmos/Aclim/r1i1p1/v20120418/
  tas/tas_Aclim_IPSL-CM5A-LR_lgm_r1i1p1_260101-280012-clim.nc

  This variable is available in the following dataset (along with 45 other variables)
  pmip3.output.IPSL.IPSL-CM5A-LR.lgm.monClim.atmos.Aclim.r1i1p1.v20120418

The PMIP3/CMIP5 Participants

Information as of April 19th 2012

Up to date info: https://pmip3.lsce.ipsl.fr/wiki/doku.php/pmip3:database:status

	Institute	Country	0k *piControl*	6k *midHolocene*	21k *lgm*	LM *past1000* (1000 years)	1 % CO_2 *1pctCO2* (140 years)	CMIP5	PlioMIP	Last Interglacial	Holocene	Carbon cycle	Atm	Ocn	Model id	Term of Use	Data Node	Publish to
1	**AWI**	Germany	Completed	Completed	Completed			No	Yes	Yes	No	Yes	96×48 x L19	120×101 x L40	*COSMOS-aso ?*		*DKRZ ?*	*ESG-WDCC ?*
2	**BCC**	China	**CMIP5** (500)	**CMIP5** (100)		**CMIP5**	**CMIP5**			No	No	Yes	128×64 x L26	360x232 x L40	bcc-csm1-1	Unrestricted	BCC	ESG-PCMDI
3	**BCCR**	Norway	Running Summer 2012	Running Summer 2012	Running Summer 2012	Running Summer 2012	Start April 2012 May 2012	No	Yes	Yes	No	Yes	96×48 x L26	100×116 x L32	NorESM1-L	Unrestricted	*DKRZ ?*	*ESG-WDCC ?*
4	**CAU-GEOMAR**	Germany	Completed	Completed			Running	No	Yes	Yes		No	96x48 x L19	182x149 x L31	KCM1-2-2	Non-commercial only	*DKRZ ?*	*ESG-WDCC ?*
5	**CNRM-CERFACS**	France	**CMIP5** (850)	**CMIP5** (200)	Running April 2012		**CMIP5**			No	No	No	256x128 x L31	362x292 x L42	CNRM-CM5	Non-commercial only	CNRM	ESG-PCMDI
6	**FUB**	Germany	**PMIP3** (400)		Completed			No		No	No	Yes	96×48 x L19	120×101 x L40	COSMOS-ASO	Unrestricted	IPSL *(DKRZ later?)*	ESG-BADC
7	**NOAA-GFDL**	USA	**CMIP5** (470)		Start Spring 2012 End 2012		**CMIP5**			No	No	Yes/No	144x90 x L24	360x200 x L50	GFDL-CM3	Unrestricted	GFDL	ESG-PCMDI
8	**NASA-GISS**	USA	**CMIP5** (1153)	Completed	Completed	**CMIP5**			Yes		Yes	No	144x90 x L40	288x180 x L32	GISS-E2-R	Unrestricted	NCCS	ESG-PCMDI
9	**IPSL**	France	**CMIP5** (1000)	**CMIP5** (500)	**CMIP5** (200)	Running April 2012	**CMIP5**		Yes			Yes	96x95 x L39	182x149 x L31	IPSL-CM5A-LR	Unrestricted	IPSL	ESG-BADC
10	**KNMI** *or ICHEC ?*	Netherlands	Completed	Completed			?	?	Yes	No	No	No	320×160 x L62	362x292 x L42	EC-Earth-2-2	Unrestricted	BADC *or IPSL ?*	ESG-BADC
11a	**LASG-IAP LASG-CESS**	China	**CMIP5** (900)	Completed	Completed		**CMIP5**		Yes			No	128x60 x L26		FGOALS-g2	Unrestricted	LASG	ESG-PCMDI
11b			**CMIP5** (501)	Completed ?	Started ? End ?								x L26	360x180 x L30	FGOALS-s2	Unrestricted	LASG	ESG-PCMDI
11c			Completed			Completed							72x45 x L26		FGOALS-gl	Unrestricted	LASG	ESG-PCMDI
12	**LOVECLIM**	Belgium France Netherlands	Completed	Completed	Completed	Completed		No		Yes	Yes	No	32x64 x L3	122x65 x L20	LOVECLIM1-2	Unrestricted	IPSL	ESG-BADC
13	**MIROC**	Japan	**CMIP5** (531)	**CMIP5** (100)	**CMIP5** (100)	**CMIP5**	**CMIP5**		Yes			Yes	128x64 x L80	256x192 x L44	MIROC-ESM	Non-commercial only	DIAS	ESG-PCMDI
14	**MPI-M**	Germany	**CMIP5** (1156)	**CMIP5** (100)	**CMIP5** (100)	**CMIP5**	**CMIP5**					No	196x98 x L47	256x220 x L40	MPI-ESM-P	Unrestricted	DKRZ	ESG-WDCC
15	**MRI**	Japan	**CMIP5** (500)	**CMIP5** (100)	Running April 2012	Not started July 2012?	**CMIP5**		Yes	No	No	No	320x160 x L48	364x368 x L51	MRI-CGCM3	Non-commercial only	DIAS	ESG-PCMDI
16	**NCAR**	USA	**CMIP5** (501)	**CMIP5** (301)	**CMIP5** (101)	**CMIP5**	**CMIP5**		Yes			No	288x192 x L26	320x384 x L60	CCSM4	Unrestricted	NCAR	ESG-NCAR
17	**OSUVic**	USA	Completed	Running May 2012	Running May 2012		Not started April 2012	No		No	No	No	128x64 x L10	100 x 100 x L19	OSUVic-0-3	Unrestricted	?	?
18	**CSIRO-QCCCE**	Australia	**CMIP5** (500)	**CMIP5**			**CMIP5**					No	192x96 x L18	192x192 x L31	CSIRO-Mk3-6-0	Non-commercial only	NCI	ESG-NCI
19	**UK Groups** (UBRIS/LEEDS/EDINBUR GH - Hadley)	UK	**CMIP5** (497)	**CMIP5** (-ES: 102 -CC: 35)	Not started Spring 2012	Running Summer 2012	**CMIP5**		Yes	Yes	Yes	Yes	192x145 x L38	360x216 x L40	HadGEM2-ES HadGEM2-CC	Unrestricted	BADC	ESG-BADC
20	**UNSW**	Australia	**PMIP3** (1000)	**PMIP3** (500)	Running June 2012	**PMIP3**	**PMIP3**	No	Yes	Yes	Yes	No	64x56 x L18	128x112 x L21	CSIRO-Mk3L-1-2	Non-commercial only	IPSL *(NCI later?)*	ESG-BADC
21	**UofT**	Canada	Running June 2012		Running June 2012			No		No	No	No	256x128 x L26	320x386 x L40	UofT-CCSM3	Unrestricted	?	?

Legend

Status expected completion date	Available in CMIP5 DB (nb years in CMIP5 DB)
No	Available in PMIP3 DB (nb years in PMIP3 DB)
Yes	

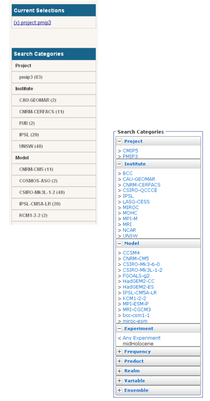
PMIP3 + PlioMIP groups: Note that the models used for PlioMIP are often not (exactly) the same as the ones used for PMIP3/CMIP5

PlioMIP-only (not LPAP, in the table):UofM

Using the PMIP3 DB

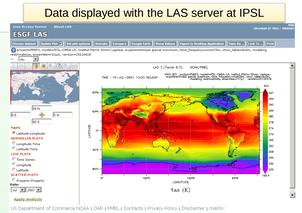
- Create an **OpenID** identifier on one of the CMIP5 access Gateways, and enroll in one of the access groups
  - http://cmip-pcmdi.llnl.gov/cmip5/data_getting_started.html
  - an OpenID is the combination of a URL-like identifier and a password
    - **https://$GATEWAY/myopenid/$USERNAME**
    - this OpenID will be requested from time to time when trying to access the data
  - the data access groups are **CMIP5 Research** (access all data) and **CMIP5 Commercial** (access only *Unrestricted* data)
    - PMIP3 uses the same OpenID and the same groups as CMIP5
- Visit one of the **Gateways** or **P2P FE** to find the data you are interested in
  - CMIP5 data is referenced on all the CMIP5 Gateways
    - **PMIP3 data is currently referenced only on the BADC Gateway**
      http://cmip-gw.badc.rl.ac.uk/
  - the Gateways and Data Nodes are being replaced by the faster **P2P system**
    - P2P FE = Peer To Peer Front end
    - **PMIP3 data is available on the IPSL P2P FE**
      - **http://esgf-node.ipsl.fr/**
- Once you have selected your data
  - download it file by file through the HTTP links
  - download several files at the same time with a wget script
  - visualize the data with the LAS server (when it is enabled)
- And now you are ready to do some science !

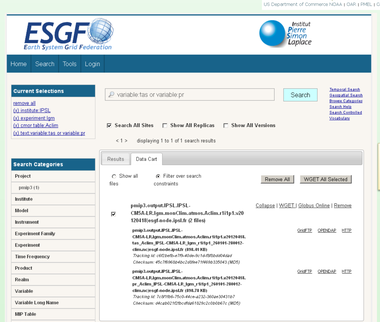
Searching **PMIP3** data on the IPSL P2P Front End



Searching PMIP3+CMIP5 **6k** data on the BADC Gateway

Data displayed with the LAS server at IPSL





PMIP3 data ready for download on the IPSL P2P Front End

- Note: due to the distributed nature of the PMIP3/CMIP5 DB, data access can be slow. You may want to use a tool to automatically mirror the data you need
  → **synchro_data**, developped at IPSL
  http://dods.ipsl.jussieu.fr/jripsl/synchro-data/README

Need more Information?

- PMIP 3 web site and wiki — https://pmip3.lsce.ipsl.fr/
- CMIP5 web site — http://cmip-pcmdi.llnl.gov/cmip5/