



Part I: obs4MIPs (P. Gleckler)

Part II: input4MIPs (P. Durack)

ESGF F2F 2017, Sheraton Fisherman's Wharf, San Francisco

Tuesday 5th December 2017

PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON (PCMDI)
pcmdi.llnl.gov/home



U.S. DEPARTMENT OF
ENERGY

Office of
Science



An Update on obs4MIPs from an ESGF perspective: progress plans and challenges



WDAC Observations for Model Evaluation Task Team

Peter Gleckler, co-chair (PCMDI) and Duane Waliser, co-chair (JPL/NASA),
Mike Bosilovich (GSFC/NASA), Helene Chepfer (IPSL), Veronika Erying (DLR), Robert Ferraro (JPL/NASA),
Pierre-Phillipe Mathieu (ESA), Jerry Potter (GSFC), Roger Saunders (UKMO),
Jörg Schulz (EUMETSAT), Karl Taylor (PCMDI), Jean-Noël Thépaut (ECMWF)

Additional regular contributors:

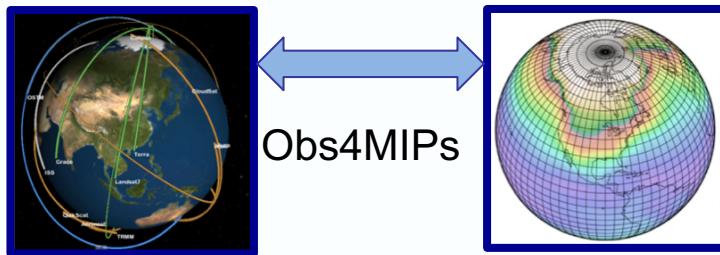
Otis Brown (WDAC), Michel Rixen (WCRP), Sophie Cloché (IPSL), Tsengdar Lee (NASA) and Renu Joseph (DOE),
Luca Cinquini (JPL) – CoG technical support
Denis Nadeau (PCMDI) – CMOR3 development, Paul Durack (PCMDI) - obs4MIPs controlled vocabulary,
Jim Biard (NCEI) and Matthias Tuma (WCRP)

... and others

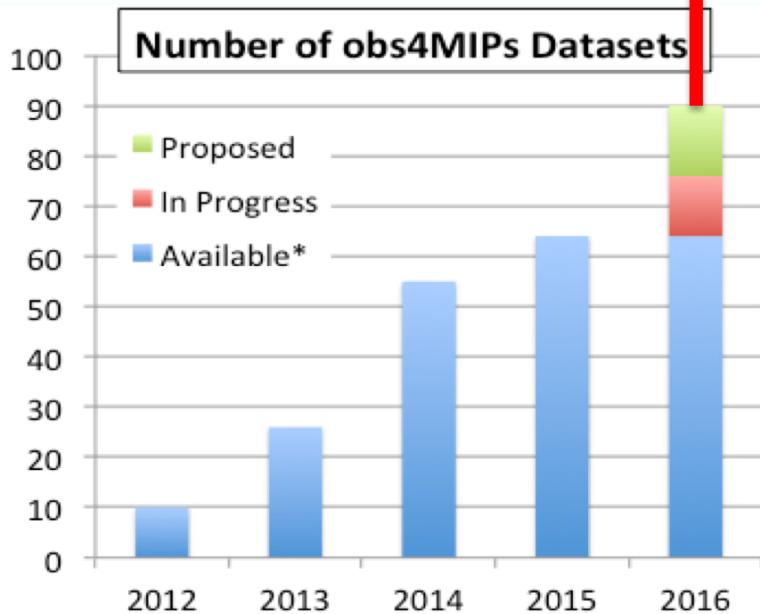
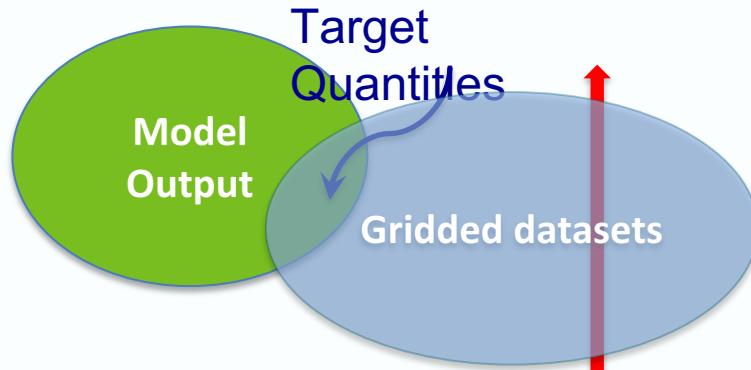
This work was supported by the U.S. Department of Energy's (DOE's) Office of Science (Biological and Environmental Research) through its Regional and Global Climate Modeling Program and was performed at Lawrence Livermore National Laboratory as a contribution to the U.S. Department of Energy, Office of Science, Climate and Environmental Sciences Division, Regional and Global Climate Modeling Program under Contract DE-AC52-07NA27344.

obs4MIPs

<https://www.earthsystemcog.org/projects/obs4mips/>



- A Project for identifying, documenting and disseminating observations for climate model evaluation.
- Data sets accessible on the ESGF alongside CMIP model output, **adhering to the same data conventions**
- Guided by the WCRP Data Advisory Council obs4MIPS Task Team



.... and growing!

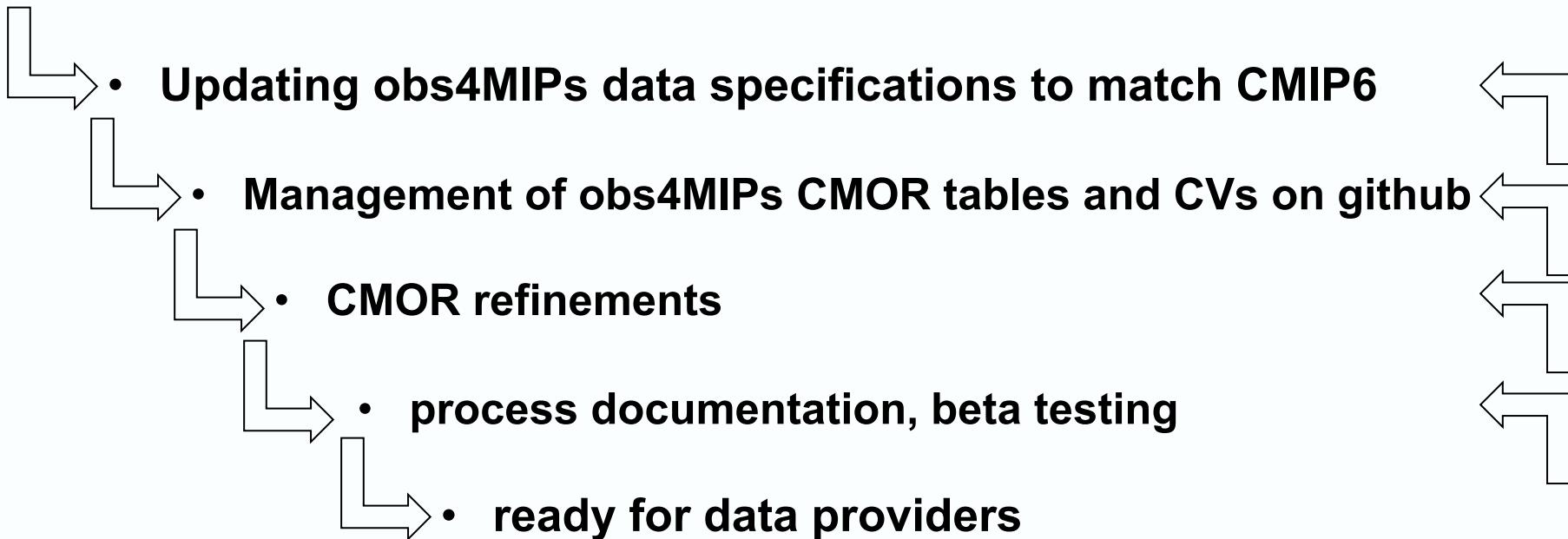
WCRP
World Climate Research Programme

Progress since ESGF F2F #6

(the not so good news)

New entries to obs4MIPs essentially stalled, awaiting...

- CMIP6 global attributes, DRS, filenames, directory structure and CVs



Progress since ESGF F2F #6

(more recently... the good news!)

- CMIP6 global attributes, DRS, filenames, directory structure and CVs ✓
 - obs4MIPs data specifications* (ODS 2.1) ✓
- Management of obs4MIPs CVs on github ✓
- Additional CMOR enhancements to deal with gridded observations ✓
(now being used for new NOAA and ESA contributions)
- Establishment of two new features ✓
 - Dataset “suitability indicators” ✓
 - Introduction of “supplemental information” ✓

* Includes placeholders for ES-DOC and DKRZ publication service

obs4MIPs Dataset Suitability & Maturity Indicators

Technical Requirements		Dataset Suitability and Maturity			Comparison Complexity
Meets obs4MIPs data technical requirements	Includes obs4MIPs technical note information	Closeness or robustness of measurement to observed reference quantity	Maturity with respect to climate model evaluation	Provision for robust uncertainty information	Complexity of Model Observation Comparison
Data suitably processed with CMOR and/or consistent with obs4MIPs standards	Complete technical note information provided	Measurement approach provides a very close relationship to observation quantity	Multiple peer-reviewed examples of application to climate model evaluation	Uncertainty information provided per retrieval/grid point	Comparison can be made directly with CMIP model output variable
Largely complete with minor metadata inconsistencies	Technical note information incomplete and/or could be improved	Measurement approach requires complex and/or non-linear retrieval methods and/or subjective inferences/definitions	One peer-reviewed example of application to climate or component model evaluation.	General uncertainty information given relative to the methodology and dataset as a whole - backed by actual field/in-situ validation exercises	Comparison requires some simple post processing of CMIP output variable(s) (e.g. vertical integral or ratio of two variables)
Non-compliant. Should be removed from database!	Technical note not provided	Measurement approach requires significant use/influence from complex or weakly constrained model and/or has significant ambiguity in definition(s)	No peer-reviewed examples of application to model evalauation	No uncertainty information provided	Comparison requires complex processing of CMIP output (e.g. "simulator", budget calculation)

Illustration Of Dataset Suitability And Maturity Indicators

Based on typical obs4MIPs dataset search

MIPS Data Search | O...  + 

earthsystemcog.org/search/obs4mips/  150%   obs4mips cog  

Hosted by  University of Colorado Boulder

Powered by 

Welcome, Peter. | You are a project administrator | Register a New Project | My Profile | Log out

Obs4MIPs

You are at the CoG-CU node

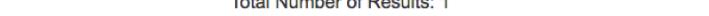
Home About Us Governance Contact Us  Last Search |  My Data Cart (1)

Institute 
Instrument 
Time Frequency 
Realm 
Variable 
Variable Long Name 
CF Standard Name 
Data Node 

Enter Text:  Search  Reset Display 10  results per page 

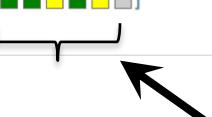
Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Search Constraints:  temperature, tes

Total Number of Results: 1
-1-
 Add all displayed results to Data Cart  Remove all displayed results from Data Cart
Expert Users: you may display the search URL and return results as XML or return results as JSON

1. obs4mips.NASA-JPL.TES.tro3.mon
Data Node: esgf-data.jpl.nasa.gov
Version: 20110608
Total Number of Files (for all variables): 3
Full Dataset Services: [Show Metadata](#) [List Files](#) [THREDDS Catalog](#) [WGET Script](#) [LAS Visualization](#) [Tech Note](#) [Supplementary Data](#) [Globus Download](#) 

 [Add to Data Cart](#)



Linked as a single entity to explanatory page of color codes

Illustration Of Dataset Suitability And Maturity Indicators

Based on typical obs4MIPs dataset search

MIPs Data Search | O... × +

www.earthsystemcog.org/search/obs4mips/ 150% obs4mips cog ↗ ☆ | ☰

Hosted by University of Colorado Boulder Powered by ESGF and CCG

Welcome, Peter. | You are a project administrator | Register a New Project | My Profile | Log out

Obs4MIPs

You are at the CoG-CU node

Home About Us Governance Contact Us Technical Support

Last Search | My Data Cart (1)

Institute +
Instrument +
Time Frequency +
Realm +
Variable +
Variable Long Name +
CF Standard Name +
Data Node +

Enter Text: temperature, tes Display 10 results per page [More Search Options]

Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Search Constraints: ~~temperature, tes~~

Total Number of Results: 1
-1-

Add all displayed results to Data Cart Remove all displayed results from Data Cart
Expert Users: you may display the search URL and return results as XML or return results as JSON

1. obs4mips.NASA-JPL.TES.tro3.mon
Data Node: esgf-data.jpl.nasa.gov
Version: 20110608
Total Number of Files (for all variables): 3
Full Dataset Services: [Show Metadata] [List Files] [THREDDS Catalog] [WGET Script] [LAS Visualization] [Tech Note] [Supplementary Data]
[Globus Download] [██████████]

 Add to Data Cart



Access to SUPPLEMENTAL DATA (in zip file; instructions available)

Example: obs4MIPs “registered content”: (Data providers enter as github issue)

```
"NOAA-NCEI-ERSST-4-0":{  
    "institution_id":"NOAA-NCEI",  
    "region":[  
        "global_ocean"  
    ],  
    "release_year":"2015",  
    "source_description":"Extended Reconstructed Sea Surface Temperatures",  
    "source_id":"NOAA-NCEI-ERSST-4-0",  
    "source_label":"NOAA-NCEI-ERSST",  
    "source_name":"NOAA NCEI ERSST",  
    "source_type":"gridded_in situ",  
    "source_variables":[  
        "tos"  
    ],  
    "source_version_number":"4.0"  
},
```

Questions..

Q: What are the key things that are difficult to do today and are impeding scientific progress or productivity?

A: Preparing and publishing data

Q: What are key development effort that you see are needed for the future success of your projects?

A: Streamlining the process of contributing data

Q: What is the estimated size of your distributed archive?

A: ~Terabytes

Q: What are the administrative/sponsor requirements that arise from your project?

A: Usage metrics, persistent and digital object identifiers, dataset version control

Q: What are your expected strategic roadmaps for the ESGF's short, mid and long term development efforts?

A: 1-3 years: Get caught up, including updating existing datasets, with new data specifications (ODS 2.1)

3-5 years: Get a majority of gridded data used for model evaluation research in obs4MIPs

5-10 years: It is possible, with sufficient progress, that obs4MIPs or a similar effort will revolutionize how data are coordinated across agencies, borders, and classes of data

obs4MIPs: Current ESGF strategic needs

- Smooth transition to new facets/datasets; backwards compatibility with existing data
- Remove or filter unwanted data (by whom?)
- Planning for in-situ data



input4MIPs: Boundary Conditions and Forcing Datasets for CMIP6

DOE ESGF F2F 2017, Sheraton Fisherman's Wharf, San Francisco

Paul J. Durack, Karl E. Taylor, Sasha Ames, Tony Hoang, Martina Stockhouse and many others..

Tuesday 5th December 2017

PROGRAM FOR CLIMATE MODEL DIAGNOSIS AND INTERCOMPARISON (PCMDI)
pcmdi.llnl.gov/home



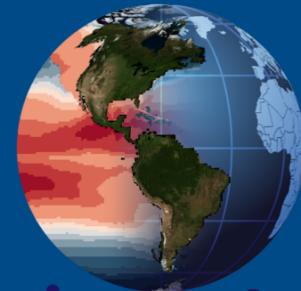
U.S. DEPARTMENT OF
ENERGY

Office of
Science



Section 0: What is input4MIPs?

Observationally-derived input datasets for
Model Intercomparison Projects (MIPs)



input4MIPs

input datasets for Model Intercomparison Projects

What is input4MIPs:

- Hosts all required CMIP (DECK) and satellite MIP experiments forcing datasets
- Currently ~1200 files, ~0.5TB, ~15 contributing institutions
- User base is small ~100 users, ~30+ modeling groups

Hosted by  Department of Energy
Lawrence Livermore National Laboratory

Powered by 
Welcome, Guest | Login | Create Account

Input4MIPs

You are at the ESGF@DOE/LLNL node

Home Contact Us Technical Support

input4MIPs: Boundary Condition and Forcing Datasets for CMIP6

input4MIPs Home

Visitors
[List All News](#)
[List ESGF Data Groups](#)

Search & Download Data ?
Simple Text Search More search options

To get input4MIPs data via ESGF:
Please enter some search text in the "Search & Download Data" box above.

input4MIPs (input datasets for Model Intercomparison Projects) is an activity to make available via ESGF the boundary condition and forcing datasets needed for CMIP6. Various datasets are needed for the pre-industrial control (piControl), AMIP, and historical simulations, and additional datasets are needed for many of the CMIP6-endorsed model intercomparison projects (MIPs) experiments. Earlier versions of many of these datasets were used in the 5th Coupled Model Intercomparison Project (CMIP5).

Summary and documentation of the available data, in addition to additional datasets being used in satellite MIPs are provided in the summary document. This document provides information about available and in-preparation datasets and points to documentation for all registered data providers:

input4MIPs summary

For more information, contact Paul J. Durack (pcmdi-cmip@llnl.gov)

Last Update: June 29, 2017, 12:23 p.m. by Paul J. Durack

Federated ESGF-CoG Nodes

- CoG-CU
- ESGF@CEDA
- ESGF@DKRZ
- ESGF@IPSL
- ESGF@NASA/JPL
- ESGF@NASA/NCCS
- ESGF@NCI
- ESGF@NOAA/ESRL
- ESGF@NSCCS.NL

Browse Projects

This All My Tags

Parent projects (1)
ESGF-LLNL

Peer projects (4)
ACME-LLNL
CMIP5
Obs4MIPs
WIP

Child projects (0)

Enter Tag

Start typing, or use the 'Delete' key to show all available tags.

input4MIPs Tags: None

ESGF sponsors and partners
DoE Office of Science | IS-ENES | NASA | NOAA | NCI | NSF

CoG version 3.9.5
ESGF P2P Version v2.5.13 master

Earth System CoG sponsors and partners
NOAA | NASA | NSF | DoE Office of Science | IS-ENES

Privacy & Legal Notice

input4MIPs | Page <https://esgf-node.llnl.gov>

PCMDI

What is input4MIPs:

- Moving target – using “live” google doc to keep track of changing datasets

The screenshot shows a Google Doc titled "CMIP6_Forceing_Datasets_Summary". The document has a sidebar on the left containing a table of contents with various sections like "CMIP6 Forcing Dat...", "Official CMIP Pa...", "Modeling center ...", etc. The main content area includes a title "CMIP6 Forcing Datasets Summary" dated "29th November 2017 (v6.2.3)", author information, a short URL (<http://goo.gl/r8up31>), and release history for v6.0, v6.1.1, and v6.2.0. It also mentions the version history and provides instructions for preparing datasets. The bottom section discusses modeling center contributors and their details.

CMIP6 Forcing Datasets Summary
29th November 2017 (v6.2.3)

Paul J. Durack, Karl E. Taylor and those preparing forcing datasets for CMIP6 (see contacts below)

A short URL to this document is: <http://goo.gl/r8up31>

Official CMIP Panel input4MIPs data collection releases:
Initial v6.0 was released 20th December 2016
v6.1.1 update was released 22nd May 2017
v6.2.0 update was released 11th September 2017

The version history of this document and datasets can be viewed [here](#)

The forcing datasets (and boundary conditions) needed for CMIP6 experiments are being prepared by a number of different experts. Initially many of these datasets may only be available from those experts, but over time as part of the “input4MIPs” activity most of them will be archived by PCMDI and served by the Earth System Grid Federation (<https://esgf-node.llnl.gov/search/input4mips/>). This “living” document will be updated as needed and serves as the official CMIP6 forcing dataset inventory and directory. The characteristics of each dataset are summarized, and links to the location of the data and documentation are provided.

Some of the datasets have not yet been released for CMIP6 use (see “status” of each below). Once a dataset has been released, we hope to avoid any further changes. If changes are required due to errors in the original data, the revision (and changes associated with this) will be documented below. All released datasets will be persistently stored at the input4MIPs project page.

Instructions for preparing forcing datasets (and an example) are provided at the [end of this document](#).

To view the datasets already archived on the ESGF input4MIPs project see:
<https://esgf-node.llnl.gov/search/input4mips/>

Modeling center contributors to CMIP6

Institutions and modeling centers that are contributing to CMIP6 are asked to update their details (institution_id, source_id/model acronym) in the CMIP6 Controlled Vocabulary (CV; https://github.com/WCRP-CMIP/CMIP6_CVs) by submitting a new issue [here](#). Existing entries can be viewed at [CMIP6 institution id.html](#) and [CMIP6 source id.html](#)

What is input4MIPs:

- Search facets – have attempted to conform as closely to CMIP6 as possible
- Learning-by-doing publication quirks will be fixed in a Jan '18 republication

Hosted by  Department of Energy
Lawrence Livermore National Laboratory

Powered by 
Welcome, Guest | Login | Create Account

Input4MIPs

You are at the ESGF@DOE/LLNL node

Home Contact Us Technical Support

Variable
Target MIP
 AerChemMIP (15)
 C4MIP (17)
 CMIP (507)
 DAMIP (1)
 DCPP (22)
 GeoMIP (15)
 HighResMIP (2)
 LS3MIP (15)
 LUMIP (15)
 OMIP (2)
 RFMIP (1)
 RFMIP CMIP (1)
 ScenarioMIP (15)

Institution
Dataset Category
Source ID
Grid Label
 gm (1)
 gn (347)
 gn-15x360deg (46)
 gr (1)
 gr-0p5x360deg (46)
 gr1-GMNHSH (92)
 gz (1)

Grid Resolution
MIP Era
 CMIP6 (536)

Time Frequency
Dataset Version Number
Data Node
Enter Text
 Display 10 results per page [More Search Options]
 Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)
The search returned 0 results.



What is input4MIPs:

- Search facets – have attempted to conform to MIP6
- Learning-by-doing publication quirks will be present

MIP6 as possible
'18 republication

The screenshot shows the input4MIPs search interface. At the top, there's a header with the Lawrence Livermore National Laboratory logo and "Department of Energy". Below the header, the title "Input4MIPs" is displayed, followed by a navigation bar with "Home" and "Contact Us". The main area contains several search facets on the left and right sides, with a central search form in the center.

Facets (Left Side):

- Variable:
 - Target MIP:
 - AerChemMIP (15)
 - C4MIP (17)
 - CMIP (507)
 - DAMIP (1)
 - DCPP (22)
 - GeoMIP (15)
 - HighResMIP (2)
 - LS3MIP (15)
 - LUMIP (15)
 - OMIP (2)
 - RFMIP (1)
 - RFMIP CMIP (1)
 - ScenarioMIP (15)
 - Institution:
 - gm (1)
 - gn (347)
 - gn-15x360deg (46)
 - gr (1)
 - gr-0p5x360deg (46)
 - gr1-GMNHSH (92)
 - gz (1)
 - Grid Resolution:
 - CMIP6 (536)
 - MIP Era:
 - CMIP6 (536)
 - Time Frequency:
 - day (3)
 - fx (6)
 - invariant (7)
 - mon (454)
 - monClim (1)
 - yr (64)
 - yrClim (1)
 - Dataset Version Number:
 - Dataset Version Number
 - Data Node:
 - Data Node

Section 1: Problem data

Expanding ESGF support for more data formats

Problematic data:

- Large single files vs temporal chunking (1 x 16GB vs 8 x 2GB time chunks)
- ~16GB (UMD) – user downloads in 3rd world countries

New! LUH2 Historical datasets now available

LUH2 v2h Release (10/14/16): The updated release of the historical land-use forcing dataset (LUH2 v2h) covers the period 850-2015 and corrects all known issues and notices identified with the previous version (LUH2 v1.0h). This dataset replaces the previously released dataset (LUH2 v1.0h). This product is the result of a series of prototypes released previously, uses the established data format, and will connect smoothly to gridded products for the future.

- Historic Data (850 - 2015 AD)
 - [states.nc](#) (5.8 GB)
 - [transitions.nc](#) (16 GB) (circled in red)
 - [management.nc](#) (1.4 GB)
- Supporting Files
 - [staticData_quarterdeg.nc](#) (1 MB)
- Data Documentation
 - [LUH2 v2h README](#)

- Will large files be an issue for users behind firewalls?
- What is the experience with download restarts with **wget** (what about rsync?)

Problematic data:

- Non-CMIP data: Multiple variables per file (UoM) – rewrite if possible

```
[durack1@oceanonly CMIP6]$ ncdump -h input4MIPs/UoM/GHGConc/CMIP/yr/atmos/UoM-CMIP-1-1-0/GHGConc/gr3-GMNHSH/v20160701/mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-1-0_gr3-GMNHSH_0000-2014.nc
netcdf mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-1-0_gr3-GMNHSH_0000-2014 {
dimensions:
lat = 1 ;
bnds = 2 ;
time = 2015 ;
variables:
double lat(lat) ;
  lat:units = "degrees_north" ;
  lat:long_name = "latitude" ;
  lat:standard_name = "latitude" ;
  lat:axis = "Y" ;
  lat:bounds = "lat_bnds" ;
double lat_bnds(lat, bnds) ;
double time(time) ;
  time:units = "days since 1850-01-01 00:00:00" ;
  time:calendar = "365_day" ;
  time:long_name = "time" ;
  time:standard_name = "time" ;
  time:axis = "T" ;
  time:bounds = "time_bnds" ;
double time_bnds(time, bnds) ;
float carbon_dioxide_GM(time) ;
  carbon_dioxide_GM:long_name = "Global Mean Mole Fraction of CO2" ;
  carbon_dioxide_GM:original_name = "CO2_GM" ;
  carbon_dioxide_GM:standard_name = "mole_fraction_of_carbon_dioxide_in_air" ;
  carbon_dioxide_GM:units = "1.e-6" ;
  carbon_dioxide_GM:cell_methods = "time: mean area: mean" ;
  carbon_dioxide_GM:lat = "0.0" ;
  carbon_dioxide_GM:lat_bnds = "-90.0, 90.0" ;
float carbon_dioxide_NH(time) ;
  carbon_dioxide_NH:long_name = "Northern Hemisphere Mean Mole Fraction of CO2" ;
  carbon_dioxide_NH:original_name = "CO2_NH" ;
  carbon_dioxide_NH:standard_name = "mole_fraction_of_carbon_dioxide_in_air" ;
  carbon_dioxide_NH:units = "1.e-6" ;
  carbon_dioxide_NH:cell_methods = "time: mean area: mean" ;
  carbon_dioxide_NH:lat = "30.0" ;
  carbon_dioxide_NH:lat_bnds = "0.0, 90.0" ;
float carbon_dioxide_SH(time) ;
  carbon_dioxide_SH:long_name = "Southern Hemisphere Mean Mole Fraction of CO2" ;
  carbon_dioxide_SH:original_name = "CO2_SH" ;
  carbon_dioxide_SH:standard_name = "mole_fraction_of_carbon_dioxide_in_air" ;
  carbon_dioxide_SH:units = "1.e-6" ;
  carbon_dioxide_SH:cell_methods = "time: mean area: mean" ;
  carbon_dioxide_SH:lat = "-30.0" ;
  carbon_dioxide_SH:lat_bnds = "-90.0, 0.0" ;

// global attributes:
:title = "UoM-CMIP-1-1-0: historical GHG concentrations: global and hemispheric means of CO2 prepared for input4MIPs" ;
:institution_id = "UoM" ;
:dataset_category = "GHGConcentrations" ;
:dataset_version_number = "1.1.0" ;
```

Problematic data:

- Non-CMIP data: Multiple variables per file (UoM) – rewrite if possible

```
[durack1@oceanonly CMIP6]$ ncdump -h input4MIPs/UoM/GHGConc/CMIP/yr/atmos/UoM-CMIP-1-1-0/GHGConc/gr3-GMNHSH/v20160701/mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-1-0_gr3-GMNHSH_0000-2014.nc
netcdf mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-1-0_gr3-GMNHSH_0000-2014 {
dimensions:
    lat = 1 ;
    bndz = 2 ;
    time = 2015 ;
variables:
    double lat(lat) ;
        lat:units = "degrees_north" ;
        lat:long_name = "latitude" ;
        lat:standard_name = "latitude" ;
        lat:axis = "Y" ;
        lat:bounds = "lat_bnds" ;
    double lat_bnds(lat, bndz) ;
    double time(time) ;
        time:units = "days since 1850-01-01 00:00:00" ;
        time:calendar = "365_day" ;
        time:long_name = "time" ;
        time:standard_name = "time" ;
        time:axis = "T" ;
        time:bounds = "time_bnds" ;
    double time_bnds(time, bndz) ;
    float carbon_dioxide_GM(time) ;
        carbon_dioxide_GM:long_name = "Global Mean Mole Fraction of CO2" ;
        carbon_dioxide_GM:original_name = "CO2_GM" ;
        carbon_dioxide_GM:standard_name = "mole_fraction_of_carbon_dioxide" ;
        carbon_dioxide_GM:units = "1.e-6" ;
        carbon_dioxide_GM:cell_methods = "time: mean area: mean" ;
        carbon_dioxide_GM:lat = "0.0" ;
        carbon_dioxide_GM:lat_bnds = "-90.0, 90.0" ;
    float carbon_dioxide_NH(time) ;
        carbon_dioxide_NH:long_name = "Northern Hemisphere Mean Mole Fraction of CO2" ;
        carbon_dioxide_NH:original_name = "CO2_NH" ;
        carbon_dioxide_NH:standard_name = "mole_fraction_of_carbon_dioxide" ;
        carbon_dioxide_NH:units = "1.e-6" ;
        carbon_dioxide_NH:cell_methods = "time: mean area: mean" ;
        carbon_dioxide_NH:lat = "30.0" ;
        carbon_dioxide_NH:lat_bnds = "0.0, 90.0" ;
    float carbon_dioxide_SH(time) ;
        carbon_dioxide_SH:long_name = "Southern Hemisphere Mean Mole Fraction of CO2" ;
        carbon_dioxide_SH:original_name = "CO2_SH" ;
        carbon_dioxide_SH:standard_name = "mole_fraction_of_carbon_dioxide" ;
        carbon_dioxide_SH:units = "1.e-6" ;
        carbon_dioxide_SH:cell_methods = "time: mean area: mean" ;
        carbon_dioxide_SH:lat = "-30.0" ;
        carbon_dioxide_SH:lat_bnds = "-90.0, 0.0" ;
// global attributes:
    :title = "UoM-CMIP-1-1-0: historical GHG concentrations: global and hemispheric means of CO2 prepared for input4MIPs" ;
    :institution_id = "UoM" ;
    :dataset_category = "GHGConcentrations" ;
    :dataset_version_number = "1.1.0" ;
```

Rewritten file..

```
[durack1@oceanonly CMIP6]$ ncdump -h input4MIPs/UoM/GHGConcentrations/CMIP/yr/atmos/UoM-CMIP-1-2-0/mole_fraction_of_carbon_dioxide_in_air/gr1-GMNHSH/v20160830/mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-2-0_gr1-GMNHSH_0000-2014.nc
netcdf mole_fraction_of_carbon_dioxide_in_air_input4MIPs_GHGConcentrations_CMIP_UoM-CMIP-1-2-0_gr1-GMNHSH_0000-2014 {
dimensions:
    time = UNLIMITED ; // (2015 currently)
    bound = 2 ;
    sector = 3 ;
variables:
    float time(time) ;
        time:bounds = "time_bnds" ;
        time:long_name = "time" ;
        time:standard_name = "time" ;
        time:units = "days since 0-1-1" ;
        time:calendar = "gregorian" ;
        time:axis = "T" ;
    double time_bnds(time, bound) ;
    int sector(sector) ;
        sector:bounds = "sector_bnds" ;
        sector:lat_bnds = "0: -90.0, 90.0; 1: 0.0, 90.0; 2: -90.0, 0.0" ;
        sector:long_name = "sector" ;
        sector:ids = "0: Global; 1: Northern Hemisphere; 2: Southern Hemisphere" ;
        sector:original_names = "0: CO2_GM; 1: CO2_NH; 2: CO2_SH" ;
    double sector_bnds(sector, bound) ;
    float mole_fraction_of_carbon_dioxide_in_air(time, sector) ;
        mole_fraction_of_carbon_dioxide_in_air:_FillValue = 1.e+20f ;
        mole_fraction_of_carbon_dioxide_in_air:missing_value = 1.e+20f ;
        mole_fraction_of_carbon_dioxide_in_air:long_name = "mole" ;
        mole_fraction_of_carbon_dioxide_in_air:cell_methods = "time: mean area: mean" ;
        mole_fraction_of_carbon_dioxide_in_air:units = "1.e-6" ;
// global attributes:
    :Conventions = "CF-1.6" ;
    :comment = "Data provided are global and hemispheric area-weighted means. Zonal means for 15-degree lat bands or 0.5-degree lat bands are available in gn-15x360 or gr-0p5x360 files respectively" ;
    :variable_id = "mole_fraction_of_carbon_dioxide_in_air" ;
```

Problematic data:

- Non-netcdf – tar gzipped, code snippets, text files (MPI)

Name	Date Modified	Size	Kind
▼ MACv2-SP_v1	Today, 9:37 PM	--	Folder
.DS_Store	Today, 9:37 PM	6 KB	Document
MACv2.0-SP_v1.nc	11/9/16, 2:15 PM	19 KB	Document
orography_T63.nc	11/9/16, 2:13 PM	77 KB	Document
sp_make_data_v1.ncl	11/9/16, 2:14 PM	12 KB	Document
mo_simple_plumes_v1.f90	11/11/16, 9:31 AM	19 KB	Fortran Source File
sp_driver_v1.f90	11/11/16, 10:17 AM	13 KB	Fortran Source File
README	11/11/16, 10:36 AM	2 KB	TextEdit Document
MACv2-SP_v1.tgz	11/14/16, 4:01 PM	59 KB	gzip compressed tar archive

Problematic data:

- Non-netcdf – tar gzipped, code snippets, text files (MPI-M)

Hosted by Department of Energy Lawrence Livermore National Laboratory

Powered by ESGF and CoG Welcome, Guest | Login | Create Account

Input4MIPs

Home Contact Us

You are at the ESGF@DOE/LNL node Technical Support

Name

- MACv2-SP_v1
 - .DS_Store
 - MACv2.0-
 - orography
 - sp_make_c
 - mo_simple
 - sp_driver_
 - README
- MACv2-SP_v1

Variable Target MIP Institution

MPI-M (1)

Dataset Category Source ID Grid Label Grid Resolution MIP Era Time Frequency Dataset Version Number Data Node

Enter Text: Search Reset Display 10 results per page [More Search Options]

Search Constraints: MPI-M Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Total Number of Results: 1 -1- Please login to add search results to your Data Cart Expert Users: you may display the search URL and return results as XML or return results as JSON

1. [input4MIPs.MPI-M.aerosolProperties.RFMIPmon.MPI-M-MACv2-SP.none.none](#)
Description: MACv2-SP-1.0 parameterization of anthropogenic aerosol optical properties and an associated Twomey effect
Data Node: aims3.llnl.gov
Version: 20170201
Total Number of Files (for all variables): 8
Full Dataset Services: [Show Metadata] [Hide Files] [THREDDS Catalog] [WGET Script]

Total Number of Files: 8		
MACv2.0-SP_v1.nc	Single File Access: HTTP Download	
1 Checksum: 85a805424de1fb0b27b5a15623c73e11cddb8984be4fec28d4f45e2f320fab0f Size: 18960 [More File Metadata]		
MPI-M-MACv2-SP_ESGF PublishingInfo.json	Single File Access: HTTP Download	
2 Checksum: 744fc234ba73d0d0e8da8edf7e9f6bb18cf70533764b595180ed50b0ff717b13 Size: 1382 [More File Metadata]		
README	Single File Access: HTTP Download	
3 Checksum: 89ded6bd4610bbece4adca78214e5797bd67ea69bbf3c8a01a192cae5f891a Size: 2289 [More File Metadata]		
gmd-10-433-2017-supplement-title-page.pdf	Single File Access: HTTP Download	
4 Checksum: 1610ee0da3f3c835ab121081125e8f974c6a44d7917962d4bd15de1cbae6bcf4 Size: 53747 [More File Metadata]		
mo_simple_plumes_v1.f90	Single File Access: HTTP Download	
5 Checksum: 1fc26c89ab47a6740727014cd37069d1b6432a97c1be7dbe641a9e9f04a1e5e6 Size: 19443 [More File Metadata]		
orography_T63.nc	Single File Access: HTTP Download	
6 Checksum: 57dd8c17098fc8f171771ca168b6918b9379f050a15a126feba2fc4cea2458e6 Size: 76996 [More File Metadata]		
sp_driver_v1.f90	Single File Access: HTTP Download	
7 Checksum: e348f0af924de9c41a471ae33f073385d4fd19e6341a39026d56be2f9520570 Size: 13066 [More File Metadata]		
sp_make_data_v1.nc	Single File Access: HTTP Download	
8 Checksum: a1b922f2cbee260555bbe61075e6eff8a77475c2af43bf518b6ecd9e962595ce Size: 11530 [More File Metadata]		

ESGF sponsors and partners DoE Office of Science | IS-ENES | NASA | NOAA | NCI | NSF

CoG version 3.9.5 ESGF P2P Version v2.5.13 master

Earth System CoG sponsors and partners NOAA | NASA | NSF | DoE Office of Science | IS-ENES

Privacy & Legal Notice

File File ent ed tar archive

input4MIPs | Paul J. Durack

<https://esgf-node.llnl.gov>



Section 2: Deprecating datasets

How to gracefully “hide” obsolete datasets,
but leave them available for trace-ability

Deprecating datasets (gracefully):

- Useful for datasets that are deprecated (rather than unpublishing)
- PCMDI AMIP dataset is updated every 6 months – 1.1.0 vs 1.1.1 vs 1.1.2 vs 1.1.3..

Hosted by  Department of Energy
Lawrence Livermore National Laboratory

Powered by 

Welcome, Guest | Login | Create Account

Input4MIPs

You are at the ESGF@DOE/LLNL node

Home Contact Us Technical Support

Variable Target MIP Institution

PCMDI (22)

Dataset Category Source ID Grid Label Grid Resolution MIP Era Time Frequency

Dataset Version Number

1.1.0 (5)
 1.1.1 (5)
 1.1.2 (6)
 1.1.3 (6)

Enter Text: ? Search Reset Display 10 results per page [More Search Options]

Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Search Constraints: ~~PCMDI~~

Total Number of Results: 22
-1- 2 3 Next >>

Please login to add search results to your Data Cart

Expert Users: you may display the search URL and return results as XML or return results as JSON

1. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-0.mon.siconcbcgs1x1](#)
Description: PCMDI-AMIP 1.1.0 dataset prepared for input4MIPs
Data Node: aims3.llnl.gov
Version: 20160609
Total Number of Files (for all variables): 1
Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)]
2. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-0.mon.areacellog1x1](#)
Description: PCMDI-AMIP 1.1.0 dataset prepared for input4MIPs
Data Node: aims3.llnl.gov
Version: 20160609
Total Number of Files (for all variables): 1
Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)]
3. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-0.mon.tosbcsg1x1](#)
Description: PCMDI-AMIP 1.1.0 dataset prepared for input4MIPs

Deprecating datasets (gracefully):

- Useful for datasets that are deprecated (rather than unpublishing)
- PCMDI AMIP dataset is updated every 6 months – 1.1.0 vs 1.1.1 vs 1.1.2 vs 1.1.3..

Hosted by  Department of Energy
Lawrence Livermore National Laboratory

Powered by 

Welcome, Guest | Login | Create Account

Input4MIPs

You are at the ESGF@DOE/LLNL node

Home Contact Us Technical Support

Variable Target MIP Institution

PCMDI (12)

Dataset Category Source ID Grid Label Grid Resolution MIP Era Time Frequency Dataset Version Number

1.1.2 (6) 1.1.3 (6)

Data Node

Enter Text: Display 10 results per page [More Search Options]

Show All Replicas Show All Versions Search Local Node Only (Including All Replicas)

Search Constraints: ~~PCMDI~~

Total Number of Results: 12
-1- 2 Next >>

Please login to add search results to your Data Cart

Expert Users: you may display the search URL and return results as XML or return results as JSON

1. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-2.mon.siconcbc.gn](#)
Description: PCMDI-AMIP 1.1.2 dataset prepared for input4MIPs
Data Node: aims3.llnl.gov
Version: 20170419
Total Number of Files (for all variables): 1
Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[Globus Download](#)]
2. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-2.mon.siconcbc.gn](#)
Description: PCMDI-AMIP 1.1.2 dataset prepared for input4MIPs
Data Node: aims3.llnl.gov
Version: 20170419
Total Number of Files (for all variables): 1
Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[Globus Download](#)]
3. [input4MIPs.PCMDI.SSTsAndSealce.CMIP.PCMDI-AMIP-1-1-2.fx.areacello.gn](#)
Description: PCMDI-AMIP 1.1.2 dataset prepared for input4MIPs

Section 3: DOI minting and DKRZ citation service

**Making contributed data findable and
citable**

DKRZ Citation Service active:

- Useful for dataset discovery and citation - brings recognition to data providers
- PCMDI AMIP dataset is updated every 6 months – 1.1.0 vs 1.1.1 vs 1.1.2 vs 1.1.3..



DOI for Scientific and Technical Data 'input4MIPs.PCMDI.SSTsAndSeaIce.CMIP.PCMDI-AMIP-1-1-2'
doi:10.22033/ESGF/input4MIPs.1161

General information Creators

General Information

Name input4MIPs.PCMDI.SSTsAndSeaIce.CMIP.PCMDI-AMIP-1-1-2
Abstract CMIP6 Forcing Datasets (input4MIPs).
The forcing datasets (and boundary conditions) needed for CMIP6 experiments are being prepared by a number of different experts. Initially many of these datasets may only be available from those experts, but over time as part of the 'input4MIPs' activity most of them will be archived by PCMDI and served by the Earth System Grid Federation (<https://esgf-node.llnl.gov/search/input4mips/>). More information is available in the living document: <http://goo.gl/r8up31>

Subjects • input4MIPs.PCMDI.SSTsAndSeaIce.CMIP.PCMDI-AMIP-1-1-2 (DRS)
• CMIP6
• forcing data
• climate

Rights • Creative Commons Attribution 4.0 International License (CC BY-SA 4.0)

License • input4MIPs forcing data for CMIP6 is evolving in the sense that altered datasets might be added as new versions. The author list and the title are not final, either. Cite this data collection including the latest dataset version according to the Data Citation Guidelines (<http://bit.ly/2g8CuQM>). Individuals using the data must abide to the terms of use for CMIP6 data (<https://pcmdi.llnl.gov/CMIP6/TermsOfUse>). Details on any license restrictions are recorded as global attributes in the files.

Cite this data

Citation Durack, Paul J.; Taylor, Karl E. (2017). PCMDI AMIP SST and sea-ice boundary conditions version 1.1.2. Version YYYYMMDD^[1]. Earth System Grid Federation. <http://doi.org/10.22033/ESGF/input4MIPs.1161>

[1] Please use the latest dataset version or if not available the latest data download date as version in your data citation.

Data Access

https://esgf.dkrz.de/search/esgf-dkrz/?project=input4MIPs&institution=PCMDI&dataset_category=SSTsAndSeaIce&target_mip=CMIP&source=PCMDI-AMIP-1-1-2
https://esgf-node.llnl.gov/search/input4mips/?project=input4MIPs&institution=PCMDI&dataset_category=SSTsAndSeaIce&target_mip=CMIP&source=PCMDI-AMIP-1-1-2

Metadata Export
[XML](#) [JSON](#)

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

Impressum / Legal notice

[http://doi.org/
10.22033/ESGF/
input4MIPs.1161](http://doi.org/10.22033/ESGF/input4MIPs.1161)

Section 4:

Integrated with the ES-DOC model documentation system

Making input4MIPs versioned data correctly citable by modeling groups

Keeping track of data versions:

- Contributed datasets have been growing
- Multiple versions since December 2016, with 3 CMIP panel “official” releases and a multitude of updates in between official releases

The screenshot shows a Microsoft Word document with a table of contents on the left and an 'Appendix 4: Document version information' section on the right. A red arrow points from the text in this appendix to the list of versions below.

Appendix 4: Document version information

The document version number consists of 3 integers separated by ". ". The first integer is "document applies to CMIP6". The second integer will be incremented if errors are found in DECK/historical forcing data that had been previously released and approved by the CMIP digit will be incremented whenever an amendment or update is made to this document, download details for an existing contributing datasets (to the current 6.x release) are ame digit will also be incremented if the version number of a contributed satellite MIP dataset these changes most commonly related to preparing data to conform to the input4MIPs; required for publishing on ESGF. Note that red font is used to highlight those version corrections to the CMIP DECK and historical forcing data were made.

6.0.0 (20th December 2016) - Initial DECK and historical forcings dataset release (This was also noted as v1.0 in the email and attachment sent out through email by Veronika Eyring)

6.0.1 (12th January 2017) - [Solar data \(Katja Matthes/Bernd Funke\)](#) made available through input4MIPs

6.0.2 (23rd January 2017) - [RFMIP data \(Robert Pincus\)](#) made available through input4MIPs

6.0.3 (25th January 2017) - [DCPP data \(Christophe Cassou\)](#) made available through input4MIPs

6.0.4 (27th January 2017) - [Land use data \(Ritvik Sahajpal\)](#) made available through input4MIPs

6.0.5 (21st February 2017) - Updated forcing dataset specifications to include documentation (and miscellaneous supporting files) that will be made available through input4MIPs

6.0.6 (6th April 2017) - Updated URLs from <https://pcmdi.llnl.gov> to <https://esgf-node.llnl.gov> after webserver updates at LLNL

6.0.7 (6th April 2017) - [ScenarioMIP](#) details updated; Minor type clarifications made

6.0.8 (12th April 2017) - [Stratospheric aerosol data \(Beiping Luo\)](#) made available through input4MIPs

6.0.9 (20th April 2017) - Updated example file ncdump reflecting latest CMIP6 controlled vocabulary

6.0.10 (21st April 2017) - [AMIP Boundary forcing v1.1.2 \(PCMDI\)](#) made available through input4MIPs

6.0.11 (1st May 2017) - [C4MIP/OMIP historical carbon isotopes \(Heather Graven\)](#) made available through input4MIPs

6.0.12 (15th May 2017) - [HighResMIP SSTs and sea-ice data \(Malcolm Roberts\)](#) made available through input4MIPs

6.0.13 (16th May 2017) - [Anthropogenic SLCF \(Short Lived Climate Forcings\) Emissions data v2017-05-01 \(Steven Smith\)](#) made available through input4MIPs

6.1.0 (17th May 2017) - This version is the updated official CMIP6 forcings dataset collection. It supersedes

Keeping track of data versions:

- Contributed datasets have been growing
- Multiple versions since December 2016, with 3 CMIP panel “official” releases and a multitude of updates (~27 versions) between official releases
- Attempt to manage version info using json dictionaries for use by ES-DOCS

The screenshot shows a GitHub repository page for 'input4MIPs-cmor-tables'. The top navigation bar includes links for Features, Business, Explore, Marketplace, Pricing, and a sign-in/sign-up option. The repository name 'PCMDI / input4MIPs-cmor-tables' is displayed, along with metrics for Watch (4), Star (0), and Fork (0). Below the header, there are tabs for Code, Issues (1), Pull requests (0), Projects (0), and Insights. A dropdown menu shows the current branch is 'master'. The 'Versions' tab is selected, showing a list of JSON files: '6.2.0.json', '6.2.1.json', '6.2.2.json', and '6.2.3.json'. Each file entry includes a commit message, the latest commit hash 'f2a78af', and the time since the commit ('5 days ago' or 'a month ago').

File	Commit Message	Latest Commit
6.2.0.json	Update from upstream; Update version format	a month ago
6.2.1.json	Add version 6.2.1 to Versions	a month ago
6.2.2.json	Update versions for 6.2.2	5 days ago
6.2.3.json	Update versions for 6.2.3	5 days ago

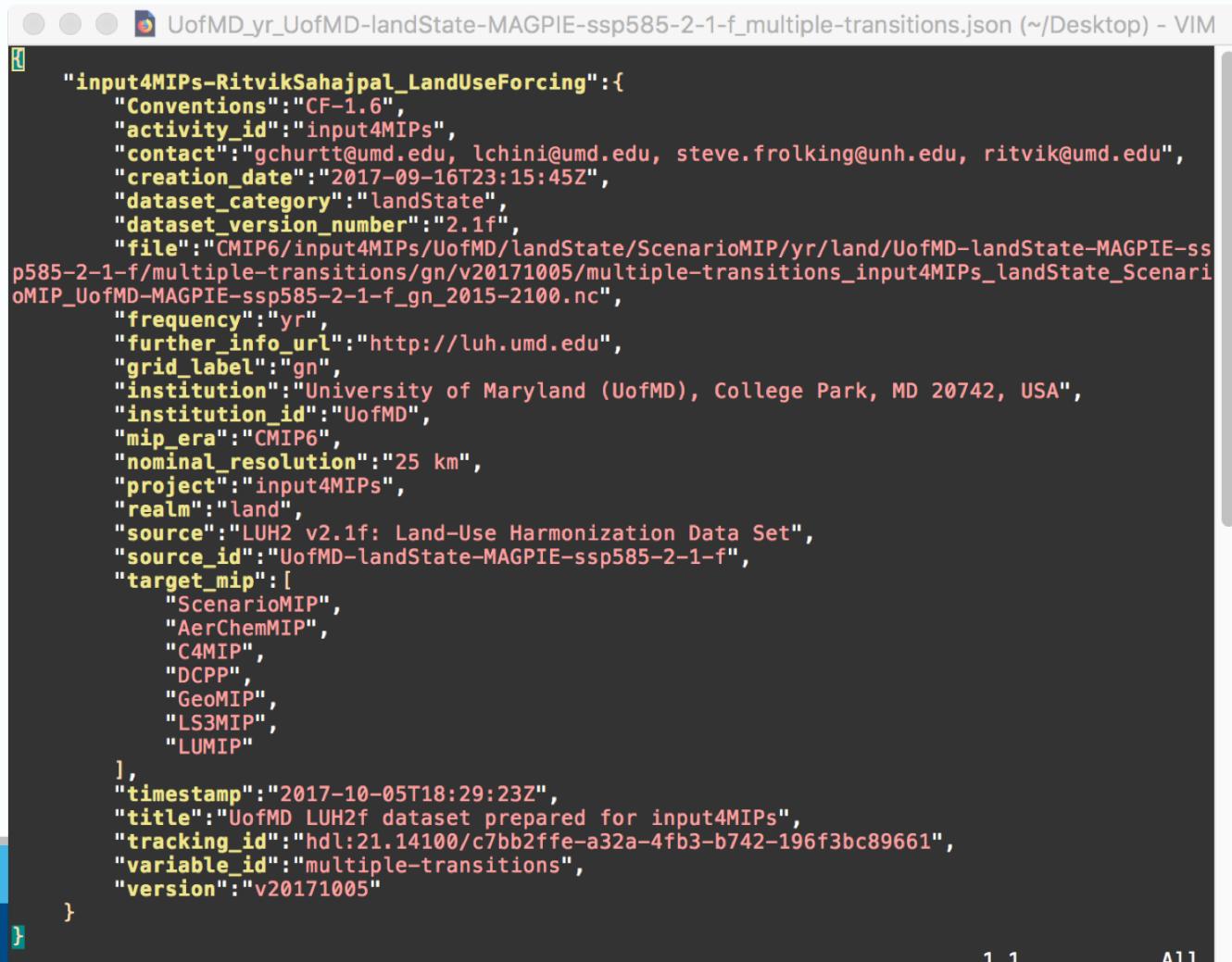


Section 5:

Ensuring consistent publication metadata across the project

If data doesn't follow identical metadata conventions:

- Maintaining consistent publication entries is a challenge



```
"input4MIPs-RitvikSahajpal_LandUseForcing":{  
    "Conventions":"CF-1.6",  
    "activity_id":"input4MIPs",  
    "contact":"gchurtt@umd.edu, lchini@umd.edu, steve.frolking@unh.edu, ritvik@umd.edu",  
    "creation_date":"2017-09-16T23:15:45Z",  
    "dataset_category":"landState",  
    "dataset_version_number":"2.1f",  
    "file":"CMIP6/input4MIPs/UofMD/landState/ScenarioMIP/yr/land/UofMD-landState-MAGPIE-ss  
p585-2-1-f/multiple-transitions/gn/v20171005/multiple-transitions_input4MIPs_landState_Scenari  
oMIP_UofMD-MAGPIE-ssp585-2-1-f_gn_2015-2100.nc",  
    "frequency":"yr",  
    "further_info_url":"http://luh.umd.edu",  
    "grid_label":"gn",  
    "institution":"University of Maryland (UofMD), College Park, MD 20742, USA",  
    "institution_id":"UofMD",  
    "mip_era":"CMIP6",  
    "nominal_resolution":"25 km",  
    "project":"input4MIPs",  
    "realm":"land",  
    "source":"LUH2 v2.1f: Land-Use Harmonization Data Set",  
    "source_id":"UofMD-landState-MAGPIE-ssp585-2-1-f",  
    "target_mip": [  
        "ScenarioMIP",  
        "AerChemMIP",  
        "C4MIP",  
        "DCPP",  
        "GeoMIP",  
        "LS3MIP",  
        "LUMIP"  
    ],  
    "timestamp":"2017-10-05T18:29:23Z",  
    "title":"UofMD LUH2f dataset prepared for input4MIPs",  
    "tracking_id":"hdl:21.14100/c7bb2ffe-a32a-4fb3-b742-196f3bc89661",  
    "variable_id":"multiple-transitions",  
    "version":"v20171005"  
}
```

Section 6: Replication with SYNDA

Multiple variable files do not conform to
CMIP-standards and..

SYNDA replication:

- **SYNDA has hit problems** downloading input4MIPs data with more than a single variable per file
- **Contributed data passed the CF-check**, but isn't CMIP "format" single variable per file
- **input4MIPs data isn't derived from models**, the DRS structure differs from the CMIP data
- Many of the “obs” products contributed have interacting variables – to use the data you need more than a single variable in memory
- ***esg.ini configuration changes seem to have solved the issue**

Thank you

Peter J. Gleckler

t +1 925 422 7631

e gleckler1@llnl.gov

w pcmdi.llnl.gov

Paul J. Durack

t +1 925 422 5208

e pauldurack@llnl.gov

w pcmdi.llnl.gov

