

# iCMIP6 Forcing Datasets Summary

20th August 2018 (v6.2.14)

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](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](#)

Here is an index to all the CMIP6 forcing datasets:

[CMIP6 Forcing Datasets Summary](#)

[Modeling center contributors to CMIP6](#)

[CMIP6 FORCING DATASETS FOR THE DECK AND HISTORICAL SIMULATIONS](#)

[Anthropogenic SLCF \(Short Lived Climate Forcing\), CO<sub>2</sub>, and CH<sub>4</sub> Emissions](#)

[Open Biomass Burning Emissions](#)

[Land Use](#)

[GHG Historical Concentrations](#)

[Stratospheric Aerosols](#)

[Ozone](#)

[Nitrogen Deposition](#)

[Solar](#)

[Aerosol Optical Properties and Relative Change in Cloud Droplet Number Concentration](#)

[AMIP Boundary Forcing](#)

## [FORCING DATASETS FOR CMIP6 ENDORSED MIPS](#)

[AerChemMIP Forcing](#)

[Future emissions \(all greenhouse gases and air pollutants\)](#)

[C4MIP Forcing](#)

[CFMIP Forcing](#)

[DAMIP Forcing](#)

[DCPP Forcing](#)

[FAFMIP Forcing](#)

[HighResMIP Forcing](#)

[LS3MIP Forcing](#)

[OMIP Forcing](#)

[PMIP Forcing](#)

[RFMIP Forcing \(of offline radiation codes\)](#)

[ScenarioMIP Forcing](#)

[Overall questions related to ScenarioMIP](#)

[Contacts:](#)

[Future emissions \(all greenhouse gases and air pollutants\)](#)

[Future Land Use](#)

[Future Ozone](#)

[Future Nitrogen Deposition](#)

[Future Solar](#)

[VolMIP Forcing](#)

## [Forcing Dataset Specifications](#)

[Proposed directory structure for input4MIPs:](#)

[input4MIPs/<mip\\_era>/<target\\_mip>/<institution\\_id>/<source\\_id>/<realm>/<frequency>/<variable\\_id>/<grid\\_label>/<version>/files.nc](#)

[Proposed filename for input4MIPs:](#)

[Documentation and miscellaneous supporting files:](#)

[Dataset formats and organization:](#)

[Dataset standards compliance:](#)

[Incoming data for review \(can be uploaded - please do not upload multi-GB files without contacting us\):](#)

[Example file:](#)

[Appendix 1: Guidance for Defining “grid” and “grid\\_label” Attributes](#)

[Appendix 2: Algorithms for Defining the “nominal\\_resolution” Attribute](#)

[Appendix 3: Relevant web links](#)

[Appendix 4: Document version information](#)

# CMIP6 FORCING DATASETS FOR THE DECK AND HISTORICAL SIMULATIONS

Status of dataset as identified below:

<status uncertain, under preparation, under review, released for CMIP6 use>

## Anthropogenic SLCF (Short Lived Climate Forcing), CO<sub>2</sub>, and CH<sub>4</sub> Emissions

Contacts:

Steven Smith [ssmith@pnnl.gov](mailto:ssmith@pnnl.gov)

Leyang Feng [fengly20@gmail.com](mailto:fengly20@gmail.com)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 2017-05-18, aircraft emission files 2017-08-30 (except SO<sub>2</sub>), and SO<sub>2</sub> aircraft emission files 2017-10-05. (the SO<sub>2</sub> aircraft data replaced 2017-08-30; the aircraft data replaced 2017-05-18; and previous releases updated the earlier versions 2017-05-19; and 2016-06-18, 2016-06-18-sectorDimV2, 2016-07-26, 2016-07-26-sectorDim)

Further information/documentation:

<https://www.geosci-model-dev.net/11/369/2018/>

Only a brief outline of the data is given here, for further details including a [README](#) file and additional supplementary information see the CEDS project web site (and the CMIP6 page there) and the GMD paper.

Common characteristics of datasets in collection:

- Used in following expts.: Historical (and satellite MIPs: AerChemMIP, ScenarioMIP and others)
- Spatial domain: Global
- Spatial resolution: 0.5° (up to 0.1° can be generated on request)
- Temporal domain: 1750 – 2014
- Temporal resolution: Monthly

Data releases:

- 1750 - 1850 released June 2016 (published on ESGF 28th June 2016)
- 1851 - 2014 released July 2016 (published on ESGF 2nd Sept 2016)
- Revised 1750 - 2014 released 1st May 2017 (published on ESGF 16th May 2017)
  - CH<sub>4</sub> from 1970 - 2014
- Corrected 1750 - 2014 released 18th May 2017 (published on ESGF 22nd May 2017)

Data format

- Sectors include: Agriculture; Energy; Industrial; Transportation; Residential, Commercial, Other; Solvents production and application; Waste; International Shipping - for further information see the README file at the project web site

Primary Data

- Bulk emissions (BC, OC, SO<sub>2</sub>, NO<sub>x</sub>, NH<sub>3</sub>, CO, NMVOC, CO<sub>2</sub>, CH<sub>4</sub>)
- Data volume: 28.36 GB; 57 files; 9 variable; 3226.22 MB/variable; mean file size is 509.40 MB
- Aircraft emissions (BC, OC, SO<sub>2</sub>, NO<sub>x</sub>, NH<sub>3</sub>, CO, NMVOC, CO<sub>2</sub>, CH<sub>4</sub>)
  - Data volume: 23.20 GB; 57 files; 9 variable; 2639.44 MB/variable; mean file size is 416.75 MB

Supplemental Data

- Supplemental Data: Speciated VOCs

- Data volume: 59.44 GB; 161 files; 23 variable; 2646.39 MB/variable; mean file size is 378.06 MB
- Same categories as used in HTAP and CMIP5 (Lamarque et al. 2010) data
- Speciated NMVOCs add up to the bulk NMVOCs reported in the bulk emission files
- See VOC README file at the project web site for additional details
- Supplemental Data: Solid Biofuel emissions
  - Data volume: 8.72 GB; 50 files; 8 variable; 1115.50 MB/variable; mean file size is 178.48 MB
  - All species except CO<sub>2</sub>. Note that emissions from solid biofuels are already included in the bulk emission files, these supplemental files provide the sub-set of emissions produced by solid biofuels (e.g. these emissions are already included in the bulk emission files).
- Supplemental Data: Extended CH<sub>4</sub> emissions
  - Data volume:
    - i. Extended CH<sub>4</sub> bulk emissions : 177 MB; 1 files; 8 variable; 22.125 MB/variable; mean file size is 177 MB
    - ii. Extended CH<sub>4</sub> aircraft emissions : 4.14 MB; 1 files; 1 variable; 4.14 MB/variable; mean file size is 4.14 MB
  - The CEDS CH<sub>4</sub> emissions are back extended using RCP data from 1850 to 1960 at 10 years increment. There is no CH<sub>4</sub> aircraft emission, but the extended CH<sub>4</sub> aircraft emission grids are provided for file consistency.
- Emissions by country and 55 sectors
  - Available in the journal paper, zip file supplement at GMD.

## Open Biomass Burning Emissions

- Monthly estimates of open biomass burning emissions (forests, grasslands, agricultural waste burning on fields, peatlands)
- Emission species: aerosol (BC,OC) and aerosol precursor and reactive compounds (SO<sub>2</sub>, N<sub>2</sub>O, NO<sub>x</sub>, NH<sub>3</sub>, CH<sub>4</sub>, CO, NMVOC, H<sub>2</sub>)
- NMVOC consists of the sum of: C<sub>2</sub>H<sub>6</sub>, CH<sub>3</sub>OH, C<sub>2</sub>H<sub>5</sub>OH, C<sub>3</sub>H<sub>8</sub>, C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>3</sub>H<sub>6</sub>, C<sub>5</sub>H<sub>8</sub>, C<sub>10</sub>H<sub>16</sub>, C<sub>7</sub>H<sub>8</sub>, C<sub>6</sub>H<sub>6</sub>, C<sub>8</sub>H<sub>10</sub>, Toluene\_lump, Higher\_Alkenes, Higher\_Alkanes, CH<sub>2</sub>O, C<sub>2</sub>H<sub>4</sub>O, C<sub>3</sub>H<sub>6</sub>O, C<sub>2</sub>H<sub>6</sub>S, HCN, HCOOH, CH<sub>3</sub>COOH, MEK, CH<sub>3</sub>COCHO, HOCH<sub>2</sub>CHO. These NMVOCs are also provided separately

### Contacts:

Margreet van Marle [m.j.e.van.marle@vu.nl](mailto:m.j.e.van.marle@vu.nl)

Guido van der Werf [guido.vander.werf@vu.nl](mailto:guido.vander.werf@vu.nl)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 1.2 (2016-12-13; replaced v1.1)

Modellers are advised to use v1.2. While global total emissions have only slightly changed from previous versions we now include specie-specific sectoral contributions (Deforestation, Agriculture, Boreal Forests, Temperate Forests, Peat and Savanna fires). Tables with estimates of emissions of the various species and sectors can be found on <https://www.geo.vu.nl/~gwerf/GFED/GFED4/tables/> for the 1997 onwards period.

### v1.1 to v1.2 differences:

- v1.1 included only sectoral contributions of CO emissions, whereas v1.2 includes contributions for all species.

- The interannual variability between 1960 and 1997 has changed for tropical regions.
- CO<sub>2</sub> emissions are excluded to avoid double counting (see usage notes). However these are available upon request.

#### v1.0 to v1.1 differences:

- Emissions from boreal regions have changed
- Globally, fire carbon emissions are ~5% lower (~100Tg C yr<sup>-1</sup>, over 1750-2015).

Further information/documentation: [www.globalfiredata.org](http://www.globalfiredata.org)

#### Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain: Global
- Spatial resolution: 0.25° x 0.25°
- Temporal domain: 1750-01 through 2015-12
- Temporal resolution: Monthly

#### Datasets:

- <specie\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2'
  - Per specie 2 files: 1 covering time period 175001-184912, 1 covering 185001-201512
  - Data volume: 12.8 GB total; 20 files; 1 variable per file; ~501 MB/variable (1750-1850), ~787MB/variable (1850-2016)
  - NMVOC\_bulk is the sum of all NMVOC species listed below at NMVOC\_<nmvoc\_specie\_name>
- NMVOC\_<nmvoc\_specie\_name>: biomass burning emissions per nmvoc specie
  - Per <nmvoc\_specie\_name> 2 files: 1 covering time period 175001-184912, 1 covering 185001-201512
  - <nmvoc\_specie\_name>: 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
  - Data volume: 32.2 GB; 50 files; 1 variable per file; ~501 MB/variable (1750-1850), ~787 MB/variable (1850-2016)
- <species\_name>-percentageAGRI: Percentage of CO bulk emissions related to agricultural waste burning.
  - Data volume: 27.7 GB total; 35 file; 1 variable; ~790 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
- <species\_name>-percentageBORF: Percentage of emissions related to boreal forest fires.
  - Data volume: 9.8 GB total; 35 file; 1 variable; ~279 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'

- <species\_name>-percentageDEFO: Percentage of emissions related to fires used in the deforestation.
  - Data volume: 9.9 GB total; 35 file; 1 variable; ~281 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
- <species\_name>-percentagePEAT: Percentage of emissions related to peat fires.
  - Data volume: 9.9 GB total; 35 file; 1 variable; ~281 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
- <species\_name>-percentageSAVA: Percentage of emissions related to savanna fires.
  - Data volume: 27.1 GB total; 35 file; 1 variable; ~774 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
- <species\_name>-percentageTEMF: Percentage of emissions related to temperate forest fires.
  - Data volume: 10.4 GB total; 35 file; 1 variable; ~299 MB/variable
  - Annual percentages from 1750-2016
  - <species\_name>: 'CH4', 'SO2', 'NOx', 'N2O', 'CO', 'BC', 'OC', 'NH3', 'NMVOC\_bulk', 'H2', 'C2H6', 'CH3OH', 'C2H5OH', 'C3H8', 'C2H2', 'C2H4', 'C3H6', 'C5H8', 'C10H16', 'C7H8', 'C6H6', 'C8H10', 'Toluene\_lump', 'Higher\_Alkenes', 'Higher\_Alkanes', 'CH2O', 'C2H4O', 'C3H6O', 'C2H6S', 'HCN', 'HCOOH', 'CH3COOH', 'MEK', 'CH3COCHO', 'HOCH2CHO'
- datasource: Specifies the datasource per year for the whole time period 1750-2015
  - Data volume: 116 MB; 1 file; 1 variable; 116MB/variable
  - "0: Ocean"
  - "1: Based on FIRE-MIP model output constrained by charcoal records, scaled to Global Fire Emissions Database (GFED) version 4s";
  - "2: Based on visibility records, scaled to Global Fire Emissions Database (GFED) version 4s";
  - "3: Based on FIRE-MIP model output, scaled to Global Fire Emissions Database (GFED) version 4s";
  - "4: Global Fire Emissions Database (GFED) version 4s";
- gridcellarea
  - Data volume: 39KB total bytes; 1 file; 1 variable; 39KB/variable
  - area of the 0.25° x 0.25° grid cells in m<sup>2</sup>

#### Usage notes:

- This dataset is made available as forcing dataset for the Coupled Model Intercomparison Project Phase 6 (CMIP6) analyses at the PCMDI repository (<https://esgf-node.llnl.gov/search/input4mips/>) and is obviously only needed for models that

do not simulate fire emissions. Emissions are bulk values for all biomes, ancillary datasets with contribution of emissions related to agricultural waste burning, fires used in deforestation, boreal forest fires, peat fires, savanna fires and temperate forest fires are provided per specie. The calculated sectoral emissions for the 1997-2015 period can be validated per specie with the tables found on:

<http://www.falw.vu/~gwerf/GFED/GFED4/tables/>

- Models that have their own fire model but do not simulate anthropogenic fires are advised to use only the emissions related to deforestation and agricultural waste burning. We provide the fraction of emissions associated with this
- While the large interannual variability is a key feature of global fire emissions, modelers may consider averaging out this fire signal to avoid having interannual variability in climate and in fires being out of sync
- While using: Please check if the global annual emissions for the first and last year of the file are correct. These are provided in the attributes [annual\_total\_first\_year\_Tg\_yr] and [annual\_total\_last\_year\_Tg\_yr]
- The vertical distribution for biomass burning emissions are unspecified. GFDL chose to use the distributions recommended by [Dentener et al. \(2006\)](#) (Table 4)
- Molecular weights for all species are listed in the following table:

Species (bulk)	Molecular weight (g)	Species (NMVOC)	Molecular weight (g)
CO2	44.01	C2H6 (ethane)	30.07
CO	28.01	CH3OH (methanol)	32.04
CH4	16.04	C2H5OH (ethanol)	46.07
NMHC	15	C3H8 (propane)	44.1
H2	2.02	C2H2 (acetylene)	26.04
NOx (as NO)*	30.01	C2H4 (ethylene)	28.05
N2O	44.01	C3H6 (propylene)	42.08
PM2.5	x	C5H8 (isoprene)	68.12
TPM	x	C10H16 (terpenes)	136.24
TPC (OC+BC)	12	C7H8 (toluene)	92.14
OC	12	C6H6 (benzene)	78.11
BC	12	C8H10 (xylene)	106.17
SO2	64.02	Toluene_lump	12
NH3 (ammonia)	17.03	Higher_Alkenes	12
		Higher_Alkanes	12
		CH2O (formaldehyde)	30.03
		C2H4O (acetaldehyde)	44.05
		C3H6O (acetone)	58.08
		C2H6S (dms)	62.07
		HCN (hydrogen cyanide)	27.02

\* NOx has different units in anthro vs open burning data.

HCOOH (formic acid)	47.02
CH3COOH (acetic acid)	60.05
MEK (methyl Ethyl Ketone / 2-butanone)	72.11
CH3COCHO (methylglyoxal)	72.06
HOCH2CHO (hydroxyacetaldehyde)	60.05

## Land Use

### Contacts:

George Hurtt [gchurtt@umd.edu](mailto:gchurtt@umd.edu)

Ritvik Sahajpal [ritvik@umd.edu](mailto:ritvik@umd.edu)

Louise Chini [lchini@umd.edu](mailto:lchini@umd.edu)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: v2.1h (2017-01-26; replaced v2h, 2016-10-14; replaced v1.0h)

Further information/documentation:

<https://cmip.ucar.edu/lumip>

[https://cmip.ucar.edu/sites/default/files/lumip/LUH2\\_v1.0h\\_README.pdf](https://cmip.ucar.edu/sites/default/files/lumip/LUH2_v1.0h_README.pdf)

<http://luh.umd.edu>

### Common characteristics of datasets in collection:

- Used in following expts.: historical, piControl, amip
- Domain: global (land only)
- Spatial resolution: 0.25 x 0.25 degree
- Temporal domain: 850 - 2015
- Temporal resolution: annual

### Datasets:

- multiple\_input4MIPs\_landState\_CMIP\_UofMD-landState-2-1-h\_gn\_0850-2015\_diagnostics.nc
  - 4 variables (e.g. ?)
  - Data volume: 578 MB, 1 file; 4 variables
  - Years: 850-2015
- multiple\_input4MIPs\_landState\_CMIP\_UofMD-landState-2-1-h\_gn\_0850-2015\_management.nc
  - 22 variables (e.g. fertilizer, and irrigation information for each of 5 crop types, flooded fraction of C3 annuals (for paddy rice), fraction of each crop type grown for use as bioenergy, fraction of C3 perennial and C4 perennial biomass harvested each year (e.g. for fruit trees etc), fraction of wood harvest biomass used for industrial roundwood, traditional fuelwood, or commercial biofuels )
  - Data volume: 864 MB; 1 file; 22 variables
  - Years: 850-2015
- multiple\_input4MIPs\_landState\_CMIP\_UofMD-landState-2-1-h\_gn\_0850-2015\_states.nc
  - 12 land use states recorded as fractions of grid cell area (e.g. “forested primary land”, “managed pasture”, “C4 annual crops”), and 2 land-use variables given in biomass density unit and age units
  - Data volume: 5.4 GB; 1 file; 14 variables
  - Years: 850-2015



- multiple\_input4MIPs\_landState\_CMIP\_UofMD-landState-2-1-h\_gn\_0850-2015\_staticData-quarterdeg.nc
  - grid-cell area, forest/non-forest mask, potential biomass carbon content, and country codes
  - Data volume: 681 KB; 1 file; 4 variables
- multiple\_input4MIPs\_landState\_CMIP\_UofMD-landState-2-1-h\_gn\_0850-2015\_transitions.nc
  - transitions from each of 12 land-use states to 9 target land-use states (12x9 = 108 variables) plus harvest information from 5 forest/non-forest types, recorded both as an area fraction and a carbon mass harvested (5x2 = 10 variables):
  - Data volume: 13 GB; 1 file; 118 variables
  - Years: 850-2015

Usage notes:

- See Frequently Asked Questions section on Land-Use Harmonization website: <http://luh.umd.edu/faq.shtml>

## GHG Historical Concentrations

Contacts:

Malte Meinshausen [malte.meinshausen@unimelb.edu.au](mailto:malte.meinshausen@unimelb.edu.au)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 1.2.0 (2016-07-01)

Further information/documentation: [10.5194/gmd-2016-169](https://doi.org/10.5194/gmd-2016-169) and <http://www.climate-energy-college.net/cmip6>

- Malte Meinshausen<sup>1,2,3</sup>, Elisabeth Vogel<sup>1,2</sup>, Alexander Nauels<sup>1,2</sup>, Katja Lorbacher<sup>1,2</sup>, Nicolai Meinshausen<sup>4</sup>, David Etheridge<sup>5</sup>, Paul Fraser<sup>5</sup>, Stephen A. Montzka<sup>6</sup>, Peter Rayner<sup>2</sup>, Cathy Trudinger<sup>5</sup>, Paul Krummel<sup>5</sup>, Urs Beyerle<sup>7</sup>, Josep G. Canadell<sup>8</sup>, John S. Daniel<sup>9</sup>, Ian Enting<sup>10</sup>, Rachel M. Law<sup>5</sup>, Simon O'Doherty<sup>11</sup>, Ron G. Prinn<sup>12</sup>, Stefan Reimann<sup>13</sup>, Mauro Rubino<sup>5,14</sup>, Guus J.M. Velders<sup>15</sup>, Martin K. Vollmer<sup>13</sup>, Ray Weiss<sup>16</sup> “Historical greenhouse gas surface concentrations”
- Manuscript available here: <http://www.geosci-model-dev-discuss.net/gmd-2016-169/>

Common characteristics of datasets in collection:

- Used in following expts.: historical, piControl, amip, ...
- Spatial domain: global
- Spatial resolution: 0.5 degree latitudinal resolution and 15 degree latitudinal as also NH, SH, and global
- Temporal domain: Year Jan 0 to Dec 2014
- Temporal resolution: monthly (and annual for NH, SH, and global spatial means)

Datasets:

- mole-fraction-of-XXX-in-air\_input4MIPs\_GHGConcentrations\_CMIP\_UoM-CMIP-1-2-0\_gr1-G MNHSH\_0000-2014.nc: Hemispheric mean and global-mean annual-mean surface mole fraction of greenhouse gas XXX
  - Data volume: 3MB; 46 files; 46 variables; 75kB/variable
  - Per file, one greenhouse gas is provided. Annual-mean, hemispheric- and global-mean resolution (the `sector` dimension defines the spatial domain). Note that 3 equivalence species are provided, too, providing a radiative efficiency-weighted sum of greenhouses gases other than CO<sub>2</sub>, CH<sub>4</sub>, and N<sub>2</sub>O. See ‘Usage notes’ below.
- mole-fraction-of-XXX-in-air\_input4MIPs\_GHGConcentrations\_CMIP\_UoM-CMIP-1-2-0\_gr1-G MNHSH\_000001-201412.nc: Hemispheric mean and global-mean monthly-mean surface mole fraction of greenhouse gas XXX

- Data volume: 42MB; 46 files; 46 variables; 873kB/variable
- mole\_fraction\_of\_XXXX\_in\_air\_input4MIPs\_GHGConcentrations\_CMIP\_UoM-CMIP-1-2-0\_gn-15x360deg\_000001-201412.nc: Zonal mean, monthly-mean surface mole fraction of greenhouse gas XXX in 15-degree latitudinal bands
  - Data volume: 85MB; 46 files; 46 variables; 1.7MB/variable
- mole-fraction-of-XXXX-in-air\_input4MIPs\_GHGConcentrations\_CMIP\_UoM-CMIP-1-2-0\_gr-0p5x360deg\_000001-201412.nc: Zonal mean, monthly-mean surface mole fraction of greenhouse gas XXX in 0.5-degree latitudinal bands
  - Data volume: 1.5GB; 46 files; 46 variables; 35MB/variable
  - Note that this is finely-interpolated dataset is a mean-preserving interpolation of the 15-degree data. Provided for convenience to interpolate to the respective model grid. The underlying observational data assimilation procedure cannot provide information on GHG surface concentrations on such a fine latitude resolution.

#### Usage notes:

- **The greenhouse gases XXX**, for which data is provided are the 43 gases: 'CO2', 'CH4', 'N2O', 'CFC-11', 'CFC-12', 'CFC-113', 'CFC-114', 'CFC-115', 'HCFC-22', 'HCFC-141b', 'HCFC-142b', 'HFC-23', 'HFC-32', 'HFC-125', 'HFC-134a', 'HFC-143a', 'HFC-152a', 'HFC-227ea', 'HFC-236fa', 'HFC-245fa', 'HFC-365mfc', 'HFC-43-10mee', 'CH3CCl3', 'CCl4', 'CH3Cl', 'CH2Cl2', 'CHCl3', 'CH3Br', 'Halon-1211', 'Halon-1301', 'Halon-2402', 'NF3', 'SF6', 'SO2F2', 'CF4', 'C2F6', 'C3F8', 'C4F10', 'C5F12', 'C6F14', 'C7F16', 'C8F18' and 'c-C4F8'. The XXX gas names used in the files are CF-compliant for those gases that have a CF-compliant name. Namely, the XXX names are: 'carbon\_dioxide', 'methane', 'nitrous\_oxide', 'cfc11', 'cfc12', 'cfc113', 'cfc114', 'cfc115', 'hfc22', 'hfc141b', 'hfc142b', 'hfc23', 'hfc32', 'hfc125', 'hfc134a', 'hfc143a', 'hfc152a', 'hfc227ea', 'hfc236fa', 'hfc245fa', 'hfc365mfc', 'hfc4310mee', 'ch3ccl3', 'carbon\_tetrachloride', 'methyl\_chloride', 'ch2cl2', 'chcl3', 'methyl\_bromide', 'halon1211', 'halon1301', 'halon2402', 'nf3', 'sf6', 'so2f2', 'cf4', 'c2f6', 'c3f8', 'c4f10', 'c5f12', 'c6f14', 'c7f16', 'c8f18', 'c\_c4f8'
- **Equivalence concentrations:** In addition, three equivalent concentration timeseries are provided, that provide radiative efficiency weighted sums of above greenhouse gases, namely:
  - 'HFC-134a-eq'. This equivalent concentration summarizes the gases: 'HFC-134a', 'HFC-23', 'HFC-32', 'HFC-125', 'HFC-143a', 'HFC-152a', 'HFC-227ea', 'HFC-236fa', 'HFC-245fa', 'HFC-365mfc', 'HFC-43-10mee', 'NF3', 'SF6', 'SO2F2', 'CF4', 'C2F6', 'C3F8', 'C4F10', 'C5F12', 'C6F14', 'C7F16', 'C8F18', 'c-C4F8'
  - 'CFC-12-eq'. This equivalent concentration summarizes the gases: 'CFC-12', 'CFC-11', 'CFC-113', 'CFC-114', 'CFC-115', 'HCFC-22', 'HCFC-141b', 'HCFC-142b', 'CH3CCl3', 'CCl4', 'CH3Cl', 'CH2Cl2', 'CHCl3', 'CH3Br', 'Halon-1211', 'Halon-1301', 'Halon-2402'
  - 'CFC-11-eq'. This equivalent concentration summarizes the gases: 'HFC-134a', 'HFC-23', 'HFC-32', 'HFC-125', 'HFC-143a', 'HFC-152a', 'HFC-227ea', 'HFC-236fa', 'HFC-245fa', 'HFC-365mfc', 'HFC-43-10mee', 'NF3', 'SF6', 'SO2F2', 'CF4', 'C2F6', 'C3F8', 'C4F10', 'C5F12', 'C6F14', 'C7F16', 'C8F18', 'c-C4F8', 'CFC-11', 'CFC-113', 'CFC-114', 'CFC-115', 'HCFC-22', 'HCFC-141b', 'HCFC-142b', 'CH3CCl3', 'CCl4', 'CH3Cl', 'CH2Cl2', 'CHCl3', 'CH3Br', 'Halon-1211', 'Halon-1301', 'Halon-2402'
  - The equivalent species can be used as surrogate to approximate the radiative forcing effect of the summarized gases. Specifically, three options to include the radiative forcing effect of the full set of 43 gases are available:
    - Option 1:** Climate models implement a subset of 43 greenhouse gases.
    - Option 2:** Climate models implement the four most important GHGs with their actual concentrations explicitly, namely CO<sub>2</sub>, CH<sub>4</sub>, N<sub>2</sub>O and CFC-12 and

- summarize the effect of all other 39 gases in an equivalence concentration of CFC-11. For this purpose, we provide CFC-11-eq concentrations ('full equivalence').
- iii. **Option 3:** Like option 2, but with a different split up of gases other than CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O. Climate models implement the three most important GHGs with their actual concentrations explicitly, namely CO<sub>2</sub>, CH<sub>4</sub>, and N<sub>2</sub>O and summarize the radiative effect of the ozone depleting substances in a CFC-12-eq concentration and the radiative effect of all other fluorinated gases in a HFC-134a-eq concentration. For this purpose, we provide CFC-12-eq and HFC-134a-eq concentrations ('full equivalence')

Additional notes:

- Approximations for vertical extensions of the surface concentration are provided for models without internal schemes. Please see the manuscript: Meinshausen et al. (2017) "Historical greenhouse gas concentrations for climate modelling (CMIP6)", available here: <https://www.geosci-model-dev.net/10/2057/2017/>

Known issues:

- No major issues known.
- Minor issues that could be addressed in future versions. See limitations section in manuscript: [10.5194/gmd-2016-169](https://www.geosci-model-dev.net/10/2057/2017/)

## Stratospheric Aerosols

Surface area density (sad) of aerosols; 1850 - 2014 (temporal extent)

The forcing data (extinction coefficients, single scattering factors, asymmetric factor g for all bandwidths for the radiation module are also provided. The extinction coefficients, single scattering albedo and asymmetrical factors of the stratospheric aerosol, including not only volcanic but also **natural and anthropogenic** contributions

Contacts:

Beiping Luo [beiping.luo@env.ethz.ch](mailto:beiping.luo@env.ethz.ch)  
Thomas Peter [thomas.peter@env.ethz.ch](mailto:thomas.peter@env.ethz.ch)  
Larry Thomason [l.w.thomason@nasa.gov](mailto:l.w.thomason@nasa.gov)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Specific model-configured downloads are available from:

[ftp://iacftp.ethz.ch/pub\\_read/luo/CMIP6/<YOUR\\_MODEL>](ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/<YOUR_MODEL>) (modeling centers should contact Beiping to get data generated for the correct model spectral bands)

Status: released for CMIP6 use

Latest version: v3.0.0 (2017-09-15; updated v2.1.0, 2017-04-12; updated v2, 2016-06-02)

Further information/documentation:

[ftp://iacftp.ethz.ch/pub\\_read/luo/CMIP6/Readme\\_Data\\_Description.pdf](ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/Readme_Data_Description.pdf)

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain: global YZ (latitude x height)
- Spatial resolution: global YZ (latitude, height; 90S and 90N and from 5 to 39.5km in 0.5km resolution)
- Temporal domain: 1850 - 2014
- Temporal resolution: monthly, and monthly and annual climatology (1850-2014)

#### Datasets:

- multiple\_input4MIPs\_aerosolProperties\_CMIP\_IACETH-SAGE3lambda-3-0-0\_gn\_185001-201412.nc (monthly)
  - Data volume: 134MB;
- multiple\_input4MIPs\_aerosolProperties\_CMIP\_IACETH-SAGE3lambda-3-0-0\_gn\_185001-201412-monthly-clim.nc (monthly climatology 1850-2014)
  - Data volume: 713KB
- multiple\_input4MIPs\_aerosolProperties\_CMIP\_IACETH-SAGE3lambda-3-0-0\_gn\_1850-2014-clim.nc (Annual mean value for the **control run**, averaged over 1850-2014)
  - Data volume: 63KB

#### Usage notes:

- The SAD (version V2, 2016-06-02) is corrected using the long term OPC data over Laramie by Terry Deshler below 20 km. The present data set is strictly reliable at and above the tropopause. The values below the tropopause are less reliable. Below the instantaneous local model tropopause, the tropospheric data set should be used. For more detail and the procedure for dealing with the strat-trop interface, see the [Readme\\_Data\\_Description.pdf](#)
- **Version 3 (2017-10-06) information forthcoming**

## Ozone

Ozone volume mixing ratios [mol mol<sup>-1</sup>]  
1850 (control); 1850-2014; 2015-2100

#### Contacts:

Michaela Hegglin [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Further information can be found at:

<http://blogs.reading.ac.uk/ccmi/forcing-databases-in-support-of-cmip6/>

#### Status:

released for CMIP6 use:

- 1850 (control) latest version: 1.0 (2016-07-11)
- 1850-2014 (historical) latest version: 1.0 (2016-07-11)

#### Common characteristics of datasets in collection:

- Used in following expts.: CMIP6 DECK, endorsed MIPs, ...
- Spatial domain: global (3D)
- Spatial resolution: 96x144 (latxlon); 66 (pressure levels between 1000 and 0.0001 hPa)
- Temporal domain: Jan-Dec 1850; Jan 1850-Dec 2014; Jan 2015-Dec 2100
- Temporal resolution: monthly

#### Datasets:

- 1850 (control; Please use first year of 1850-1899 file)
  - Filename: vmro3\_input4MIPs\_ozone\_CMIP\_UReading-CCMI-1-0\_gr\_185001-189912.nc
  - Data volume: 1.7 GB; 1 file; 1 variable; 1.7 GB/variable
  - Comment: The monthly mean ozone fields represent multi-annual averages of multi-year pre-industrial control simulations
- 1850-2014 (historical)
  - Filename: vmro3\_input4MIPs\_ozone\_CMIP\_UReading-CCMI-1-0\_gr\_[185001-201412].nc
  - Data volume: 7.23 GB; 4 files; 1 variable; 7.23 GB/variable
- 2015-2100 (future)

- o TBD

Additional notes:

- Please subscribe to the ozone-database user list by emailing [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk) if you wish to use the data and receive updates on the status and other information of the different parts of the database.
- A description of the database will be published in GMD: *Hegglin, M. I., D. Kinnison, D. Plummer, et al., CCMI ozone database (1850-2100) in support of CMIP6, GMD, in preparation.*

## Nitrogen Deposition

Contacts:

Michaela Hegglin [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk)  
Douglas Kinnison [dkin@ucar.edu](mailto:dkin@ucar.edu)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Further information can be found at:

<http://blogs.reading.ac.uk/ccmi/forcing-databases-in-support-of-cmip6/>

Status:

released for CMIP6 use

- 1850 (control) latest version: 2.0 (2016-12-07; replaced v1.0)
- 1850-2014 latest version: 2.0 (2016-12-07; replaced v1.0)

Common characteristics of datasets in collection:

- Used in following expts.: CMIP6 DECK, endorsed MIPs, ...
- Spatial domain: global (2D)
- Spatial resolution: 96x144 (latxlon)
- Temporal domain: Jan-Dec 1850; Jan 1850-Dec 2014; Jan 2015-Dec 2100
- Temporal resolution: monthly

Variables included:

- DRYNHX, WETNHX, DRYNOY, WETNOY
- N1850 (control); 1850-2014; 2015-2100

Datasets:

- 1850 (control; Please use first year of 1850-1899 file)
  - o Filenames:  
drynhx\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-185012-clim.nc  
drynoy\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-185012-clim.nc  
wetnhx\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-185012-clim.nc  
wetnoy\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-185012-clim.nc
  - o Data volume: 2.7 MB; 4 files; 4 variables; 668 KB/variable
  - o Comment: The monthly mean nitrogen deposition fields represent multi-annual averages of multi-year pre-industrial control simulations
- 1850-2014 (historical)
  - o Filenames:  
drynhx\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-201412.nc  
drynoy\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-201412.nc  
wetnhx\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-201412.nc

wetnoy\_input4MIPs\_surfaceFluxes\_CMIP\_NCAR-CCMI-2-0\_gn\_185001-201412.nc

- o Data volume: 438 MB; 4 files; 4 variables; 109.5 GB/variable
- 2015-2100 (future)
  - o TBD

Usage notes:

- <note 1>

Additional notes:

- Please subscribe to the ozone-database user list by emailing [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk) if you wish to use the N-deposition data and receive updates on the status and other information of the different parts of the database.

## Solar

Contacts:

Katja Matthes [kmatthes@geomar.de](mailto:kmatthes@geomar.de)

Bernd Funke [bernd@iaa.es](mailto:bernd@iaa.es)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 3.2 (2017-01-03; [2016-10-24, replaced v3.1])

Further information/documentation: <http://solarisheppa.geomar.de/cmip6>

Common characteristics of datasets in collection:

- Used in following expts.: CMIP6 DECK and endorsed MIPs
- Spatial domain: (only variables describing particle-induced ion-pair production rate) global, zonal mean (in geomagnetic coordinates)
- Spatial resolution: variable
- Temporal domain: 1850-01-01 – 2299-12-31

Datasets:

- solarforcing-ref-day\_input4MIPs\_solar\_CMIP\_SOLARIS-HEPPA-3-2\_gn\_18500101-22991231.nc
  - o Data volume: 6.1 GB; 1 files; 9 variables
  - o Temporal resolution: day
- solarforcing-ref-mon\_input4MIPs\_solar\_CMIP\_SOLARIS-HEPPA-3-2\_gn\_18500101-22991231.nc
  - o Data volume: 84 MB; 1 file; 6 variables
  - o Temporal resolution: mon
- solarforcing-picontrol-fx\_input4MIPs\_solar\_CMIP\_SOLARIS-HEPPA-3-2\_gn\_18500101-18730128.nc
  - o Data volume: 213 kB; 1 file; 9 variables
  - o Temporal resolution: fx

Usage notes:

- The solar forcing is provided for radiative properties, i.e., total solar irradiance (TSI) and solar spectral irradiance (SSI), and F10.7cm radio flux, as well as particle forcing, i.e., geomagnetic indices  $A_p$  and  $K_p$ , and ionisation rates to account for effects of solar protons, electrons and galactic cosmic rays.
- piControl solar forcing is constructed of time-averaged historical data corresponding to 1850-1873 (solar cycle 9+10) mean conditions.

Additional notes:

- Particle forcing is provided only in the daily-resolved dataset.

## Aerosol Optical Properties and Relative Change in Cloud Droplet Number Concentration

### Contacts:

Bjorn Stevens [bjorn.stevens@mpimet.mpg.de](mailto:bjorn.stevens@mpimet.mpg.de)  
Stephanie Fiedler [stephanie.fiedler@mpimet.mpg.de](mailto:stephanie.fiedler@mpimet.mpg.de)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Previous version: <http://www.geosci-model-dev-discuss.net/gmd-2016-189/> (externally hosted)

Status: released for CMIP6 use

Latest version: MACv2-SP (2017-02-01)

Further information/documentation:

<http://www.geosci-model-dev.net/10/433/2017/gmd-10-433-2017.html>

### Common characteristics of datasets in collection:

- Used in following expts.: RFMIP, see: <http://www.geosci-model-dev.net/9/3447/2016/gmd-9-3447-2016.pdf>
- Spatial domain: global
- Spatial resolution: flexible
- Temporal domain: 1850 - 2100 (currently no data for years after 2015)
- Temporal resolution: monthly

### Datasets:

- Simple Plume Version 1.0 Release for anthropogenic aerosol optical properties and an associated Twomey effect
  - Data volume: 103 KB; 2 files; 4 variables
  - The Simple Plume Aerosol Climatology consists of two files:
    - i. `mo_simple_plumes_v1.f90`: An F90 module which a driving program needs to be ported to for execution.
    - ii. `MACv2.0-SP_v1.nc`: The input data for the `mo_simple-plumes-v1` fortran module.
  - Supplementary files for understanding and using the climatology:
    - i. `sp_driver_v1.f90`: A driver program for creating climatological output offline
    - ii. `sp_make-data_v1.ncl`: The ncl script used to create the input data (`MACv2.0-SP_v1.nc`) for the simple plume model
    - iii. `orography_T63.nc`: A file containing orographic heights at a T63 resolution.
    - iv. Sample yearly files created using the driver program may also be included, depending on the distribution.

### Usage notes:

- The Simple Plume Aerosol Climatology describes anthropogenic SW aerosol optical properties and the proportional change in cloud-droplet effective radius resulting from anthropogenic aerosol perturbations as a function of latitude, longitude, height above the ground, wave-length, and time (between 1850 and 2100).
- It can be directly implemented for usage in the radiation calculation of a climate model.

### Additional notes:

- Main Idea of MACv2-SP:

- To have a simple description of the anthropogenic aerosol optical properties (including Twomey effect) that enables the identification of robust responses to an aerosol perturbation across models.
- To describe the climatology in a way that is transparent, traceable, and facilitates experimentation.
- The current implementation uses fixed single scattering albedo (SSA) and asymmetry parameters, which leads to some small apparent discontinuities; for instance, in SSA near the South American Coast. Because the radiation code uses the optical depth weighted SSA this issue is much less relevant than it would otherwise seem and not evident in simulations using the climatology, even when performing analysis based on double radiation calls.

## AMIP Boundary Forcing

Sea surface temperature and sea-ice concentration datasets for driving atmospheric-only simulations

Contacts:

Paul J. Durack [pcmdi-cmip@llnl.gov](mailto:pcmdi-cmip@llnl.gov)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 1.1.4 (2018-04-27; replaced 2017-10-31, 1.1.3; replaced 2017-10-19, 1.1.2; replaced 2016-10-20, v1.1.1) [New versions will be released every 6 months; each update will extend the observations by 6 months to near present. The older datasets should be identical to this version except for insignificant changes in the data affecting the least significant figures, and of course the most recent data are missing.]

Further information/documentation: <https://pcmdi.llnl.gov/mips/amip> or for release notes see <https://github.com/PCMDI/amipbcs/releases/tag/v1.1.4>

Common characteristics of datasets in collection:

- Used in following expts.: AMIP
- Spatial domain: global
- Spatial resolution: 1x1 degree
- Temporal domain: 1870-01 through 2017-12
- Temporal resolution: monthly

Datasets:

- Ocean Grid-Cell Area data (areacello)
  - Data volume: 64KB; 1 file; 1 variable
- Sea area fraction (sftof)
  - Data volume: 64KB, 1 file; 1 variable
- Sea ice monthly-mean data (siconc)
  - Data volume: 31MB; 1 file; 1 variable
- Sea ice fraction mid-month boundary condition dataset (siconcbcs)
  - Data volume: 74MB; 1 file; 1 variable
- SST monthly-mean data (tos)
  - Data volume: 232MB; 1 file; 1 variable
- SST mid-month boundary condition dataset (tosbcs)
  - Data volume: 241MB; 1 file; 1 variable

Usage notes:



- Be sure to understand the difference between tosbcs and tos (and similarly, siconbcs and siconc) before using the data for AMIP simulations. (see <https://pcmdi.llnl.gov/mips/amip>)

Additional notes:

- These data sets will be updated approximately every 6 months as additional observational data become available (next update due October 2018)

## FORCING DATASETS FOR CMIP6 ENDORSED MIPS

### AerChemMIP Forcing

#### Future emissions (all greenhouse gases and air pollutants)

Contacts:

Steven Smith ([ssmith@pnnl.gov](mailto:ssmith@pnnl.gov))  
Matthew Gidden ([matthew.gidden@gmail.com](mailto:matthew.gidden@gmail.com))  
Keywan Riahi ([riahi@iiasa.ac.at](mailto:riahi@iiasa.ac.at))

Available at: <https://esgf-node.llnl.gov/search/input4mips>

Status: Preliminary version available

Latest version: x-x (2018-06-xx)

Latest version: <#> <2016-MM-DD>

Further information/documentation: <URL's and/or doi's>

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain:
- Spatial resolution:
- Temporal domain:
- Temporal resolution:

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable

Usage notes:

- <note 1>

Additional notes:

- <note 1>

### C4MIP Forcing

Historical Carbon Isotopes in Atmospheric CO<sub>2</sub>

Contact:

Heather Graven [h.graven@imperial.ac.uk](mailto:h.graven@imperial.ac.uk)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: ready for use

Version: version 1.1 for delta13co2 and version 2.0 for Delta14co2 (2017-07-21), replaced versions 1.0 (2017-04-28)

Further information/documentation:

- GMD manuscript: <https://www.geosci-model-dev.net/10/4405/2017/>
- C4MIP: <http://www.geosci-model-dev.net/9/2853/2016/gmd-9-2853-2016.html>
- OMIP: <http://www.geosci-model-dev-discuss.net/gmd-2016-155/>

Datasets:

- delta13co2-in-air\_input4MIPs\_atmosphericState\_C4MIP-OMIP\_ImperialCollege-1-1\_gm\_1850-2015.nc: Global-mean annual-mean  $\delta^{13}\text{C}$  in  $\text{CO}_2$ 
  - Data volume: 32 KB; 1 file; 1 variable
- Delta14co2-in-air\_input4MIPs\_atmosphericState\_C4MIP-OMIP\_ImperialCollege-2-0\_gz-NHT RSH\_1850-2015.nc: annual-mean  $\Delta^{14}\text{C}$  in atmospheric  $\text{CO}_2$  for Northern Hemisphere (30-90°N), Tropics (30°S-30°N), and Southern Hemisphere (30-90°S)
  - Data volume: 50 KB; 1 file; 1 variable in 3 sectors

Usage notes:

- These datasets provide historical  $\delta^{13}\text{C}$  and  $\Delta^{14}\text{C}$  in atmospheric  $\text{CO}_2$  to be used as boundary conditions for historical ocean and land model simulations of  $^{13}\text{C}$  and  $^{14}\text{C}$ .
- $\delta^{13}\text{C}$  is reported relative to the VPDB standard, referenced to the CSIRO measurement scale.
- $\Delta^{14}\text{C}$  is reported relative to the Modern standard, including corrections for age and fractionation following Stuiver and Polach 1977.
- Recent values of  $\Delta^{14}\text{C}$  are estimated from limited data (2013-15 for NH and SH, 2009-15 for Tropics).
- Version 1.1 for delta13co2 is the same as version 1.0 but the GMDD paper citation has been added to the file
- Version 2.0 for Delta14co2 updates recent data - see details in GMDD paper

## CFMIP Forcing

Contacts:

Mark Webb [mark.webb@metoffice.gov.uk](mailto:mark.webb@metoffice.gov.uk)

Available and documented in: supplementary information of the CFMIP paper at <http://www.geosci-model-dev.net/10/359/2017/gmd-10-359-2017.html> (externally hosted)

Status: released for CMIP6 use. [Please note that this is an exact copy of the forcing file provided previously for the amipFuture experiments in CMIP5](#)

Latest version: <#> <2017-01-25>

Further information/documentation:

<http://www.geosci-model-dev.net/10/359/2017/gmd-10-359-2017.html>

Common characteristics of datasets in collection:

- Used in following expts.: amip-future4K
- Spatial domain: Global
- Spatial resolution: 2 degrees latitude, 2.5 degrees longitude
- Temporal domain: 12 month climatology
- Temporal resolution: monthly

#### Datasets:

- Patterned SST warming anomalies from CMIP3 ensemble
  - Data volume: 631 kB (1 variable stored in a single file)

#### Usage notes:

- See: <http://www.geosci-model-dev.net/10/359/2017/gmd-10-359-2017.html>

## DAMIP Forcing

### Detection and attribution MIP

#### Contacts:

David Plummer [david.plummer@canada.ca](mailto:david.plummer@canada.ca)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

#### Latest version:

- hist-stratO3 - 1.0 (2017-08-14)
- hist-nat, hist-sol and hist-volc - 1.1 (2018-05-29)
  - Version 1.1 was created with the QBO term dropped from the datasets
  - The QBO term had been included in version 1.0 of these three datasets and should not be used Further information/documentation:

<http://blogs.reading.ac.uk/ccmi/forcing-databases-in-support-of-cmip6/>

#### Common characteristics of datasets in collection:

- Used in following expts.: DAMIP - hist-stratO3, hist-nat, hist-sol and hist-volc. See <https://www.geosci-model-dev.net/9/3685/2016/> for details.
- Spatial domain: global 3D
- Spatial resolution: approximately 2 degrees latitude x 2.5 degrees longitude; 66 pressure levels between 1000 and 0.0001 hPa
- Temporal domain: 1850-01 - 2014-12
- Temporal resolution: monthly
- One variable - ozone volume mixing ratio (vmro3) - in four files (185001-189912; 190001-194912; 195001-199912; 200001-201412)
- Total data volume for each dataset: 4.0 Gb

#### Datasets:

- vmro3\_input4MIPs\_ozone\_DAMIP\_CCMI-hist-stratO3-1-0\_gn\_\*
  - modified CMIP6 specified ozone for the DAMIP hist-stratO3 experiment
  - constructed following DAMIP instructions, keeping the full timeseries of ozone above the tropopause but repeating the annual cycle of pre-industrial control ozone on levels below the tropopause, with the tropopause defined as the 100 ppbv ozone in the preindustrial control
- vmro3\_input4MIPs\_ozone\_DAMIP\_CCMI-hist-nat-1-1\_gn\_\*
  - modified CMIP6 specified ozone for the DAMIP hist-nat experiment
  - derived from the base CMIP6 specified ozone dataset using Multiple Linear Regression that included terms for a linear trend, solar, volcanic and QBO signals
  - The hist-nat dataset retained the solar and volcanic terms only
  - Note that all MLR-derived datasets (nat, sol, volc) fitted the MLR terms over 1850-1910 to estimate the responses under pre-industrial-like conditions, then extended these forward using the full timeseries of the proxies

- vmro3\_input4MIPs\_ozone\_DAMIP\_CCM1-hist-sol-1-1\_gn\_\*
  - modified CMIP6 specified ozone for the DAMIP hist-sol experiment
  - derived from the base CMIP6 specified ozone dataset using Multiple Linear Regression retaining the solar term only
- vmro3\_input4MIPs\_ozone\_DAMIP\_CCM1-hist-volc-1-1\_gn\_
  - Modified CMIP6 specified ozone for the DAMIP hist-volc experiment
  - Derived from the base CMIP6 specified ozone dataset using Multiple Linear Regression retaining the volcanic term only

## DCPP Forcing

SST anomalies used for restoring SSTs to observations.

Contacts:

Christophe Cassou [christophe.cassou@cerfacs.fr](mailto:christophe.cassou@cerfacs.fr)

Doug Smith [doug.smith@metoffice.gov.uk](mailto:doug.smith@metoffice.gov.uk)

George Boer [george.boer@canada.ca](mailto:george.boer@canada.ca)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 1.1 (2017-01-23)

Further information/documentation: <https://www.wcrp-climate.org/dcp-overview>

Common characteristics of datasets in collection:

- Used in following expts.: DCPD
- Spatial domain:
- Spatial resolution:
- Temporal domain:
- Temporal resolution:

Datasets:

- AMV index sea surface temperature (tos)
  - Data volume: 74KB; 1 file; 1 variable
  - AMV index obtained after filtering (10-yr zero-phase butterworth filter from Matlab)
- Mask for restoring to be used in dcppC-amv-ExTrop experiments
- IPV index sea surface temperature (tos)
  - Data volume: 74KB; 1 file; 1 variable
  - IPV index obtained after filtering (10-yr zero-phase butterworth filter from Matlab)
- Mask for restoring to be used in dcppC-amv-Nextrop experiments

## FAFMIP Forcing

Surface ocean flux anomalies from CMIP5 1pctCO2 simulations

Contacts:

Jonathan Gregory [j.m.gregory@reading.ac.uk](mailto:j.m.gregory@reading.ac.uk)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 2.1.0 (2018-01-20; replaced 2.0, 2015-08-21)

Further information/documentation: on the FAFMIP website (as above)

Common characteristics of datasets in collection:

- Used in following expts.: All FAFMIP experiments
- Spatial domain: Global ocean
- Spatial resolution: 1 degree in latitude and longitude
- Temporal domain: Climatological seasonal cycle
- Temporal resolution: Monthly
- Each file contains a single variable, is in CF-netCDF format and occupies 3.1 Mbyte

Datasets:

- Surface momentum flux perturbation
  - Two files: zonal and meridional components, tauu and tauv respectively
  - Used by faf-stress and faf-all
- Surface heat flux perturbation
  - One file: hfds
  - Used by faf-heat, faf-all and faf-passiveheat
- Surface water flux perturbation
  - One file: wfo
  - Used by faf-water and faf-all

Usage notes:

- Instructions for application of the perturbations and implementation of the experiments can be found on the FAFMIP website (as above)

## HighResMIP Forcing

Observed SST and sea-ice dataset 1950-2014

Future SST and sea-ice dataset 2015-2050

Spatially smoothed version of the observed SST dataset 1950-2014

Leaf area index dataset

Contacts:

Malcolm Roberts [malcolm.roberts@metoffice.gov.uk](mailto:malcolm.roberts@metoffice.gov.uk)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 2.2.0.0-r0 (2017-05-05)

Further information/documentation: Initially in HighResMIP protocol paper, Haarsma et al, doi:10.5194/gmd-2016-66, also on website <http://collab.knmi.nl/highresmip>

Common characteristics of datasets in collection:

- Used in following expts.: highresSST-present, highresSST-4xco2, highresSST-LAI, highresSST-smoothed, highresSST-p4K
- Spatial domain: Global
- Spatial resolution: 0.25 degree
- Temporal domain: 19500101-20141231
- Temporal resolution: day

Datasets:

- Observed SST and sea-ice concentration dataset 1950-2014
  - Data volume: 200 GB; 65 files; 2 variables; 1.5GB/variable

- 0.25 degree grid, daily data based on HadISST2.2

Usage notes:

- Used in highresSST-present, highresSST-4xco2, highresSST-LAI

Additional notes:

- Future SST and sea-ice concentration dataset 2015-2050
  - Data volume: 112 GB; 37 files; 2 variables; 1.5GB/variable
  - 0.25 degree grid, daily data based, using HadISST2.2 repeated period for variability, and CMIP5-RCP8.5 multi-model simulations for future trends

Usage notes:

- Used in highresSST-future

Additional notes: To be produced

- Spatially smoothed version of present day SST and sea-ice concentration dataset 1950-2014
  - Data volume: 112200 GB; 3765 files; 2 variables; 1.5GB/variable
  - 0.25 degree grid, daily data based on HadISST2.2, with a spatial filter applied to smooth out small-scale features

Usage notes:

- Used in highresSST-smoothed

Additional notes: To be produced

- Leaf area index dataset 1950-2014
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- Used in highresSST-LAI

Additional notes:

- SST dataset with 4K added to the highresSST-present dataset
  - No separate datasets, simply manipulate the highresSST-present dataset

Usage notes:

- Used in highresSST-p4K

Additional notes: Simply add a uniform 4K to all points

## LS3MIP Forcing

Pseudo-observations for prescribing land conditions; Princeton forcings; CRU-NCEP forcings; WFDEI forcings

Contacts:

Sonia Seneviratne [sonia.seneviratne@ethz.ch](mailto:sonia.seneviratne@ethz.ch)

Bart Van Den Hurk [bart.van.den.hurk@knmi.nl](mailto:bart.van.den.hurk@knmi.nl)

Available at: <URL or ftp address>

Status: status uncertain

Latest version: <#> <2016-MM-DD>

Further information/documentation: <URL's and/or doi's>

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain:
- Spatial resolution:
- Temporal domain:
- Temporal resolution:

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- <note 1>

Additional notes:

- <note 1>

## OMIP Forcing

Surface variables needed to compute surface fluxes in ocean–sea-ice models. There are two versions:

- Dataset1 (for omip1): CORE-2 forcing data (not available through input4MIPs)
- Dataset2 (for omip2): JRA55-do forcing data (available through input4MIPs)

Contacts:

Hiroyuki Tsujino [htsujino@mri-jma.go.jp](mailto:htsujino@mri-jma.go.jp)

Steve Griffies [stephen.griffies@noaa.gov](mailto:stephen.griffies@noaa.gov)

Gokhan Danabasoglu [gokhan@ucar.edu](mailto:gokhan@ucar.edu)

James Orr [James.Orr@lscce.ipsl.fr](mailto:James.Orr@lscce.ipsl.fr)

### **Dataset1 (CORE-2) for omip1:**

Available at: <http://data1.gfdl.noaa.gov/nomads/forms/core/COREv2.html>

Status: released for CMIP6 use

Latest version: CORE-2

Further information/documentation:

<[http://data1.gfdl.noaa.gov/nomads/forms/core/COREv2/doc\\_v2.html](http://data1.gfdl.noaa.gov/nomads/forms/core/COREv2/doc_v2.html)>

Common characteristics of datasets in collection:

- Used in following expts.: <omip1>, <omip1-spunup>
- Spatial domain: Global
- Spatial resolution: T62(atmospheric variables), 1°x1°(river discharge to the ocean)
- Temporal domain: 1948.01.01-2009.12.31
- Temporal resolution: 6hr (surface atmospheric variables),  
1day (downwelling radiations),  
mon (precipitation)  
mon (river discharge to the ocean)

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- <note 1>

Additional notes:

- <note 1>

## **Dataset2 (JRA55-do) for omip2:**

Available at: <https://esgf-node.llnl.gov/search/input4mips>

Status: released for CMIP6 use

Latest version: JRA55-do 1.3.1 (2018-04-12)

Further information/documentation: The description paper for this dataset (Tsuji et al. 2018) is available from the following:.

<<https://doi.org/10.1016/j.ocemod.2018.07.002>>.

Common characteristics of datasets in collection:

- Used in following expts.: <omip2>, <omip2-spunup>
- Spatial domain: Global
- Spatial resolution: TL319 (~55km) (atmospheric variables)  
0.25°x0.25°(river runoff to the ocean)
- Temporal domain: 1958.01.01-2018.02.01 (to be updated at least annually)
- Temporal resolution: 3hr (atmospheric variables)  
1 day (river runoff to the ocean)

Datasets:

- JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do)
  - Data volume: 757 GBytes; 61 files for each variable (one file for each year); 10 main variables and 11 supplementary variables; max ~1.2Gbytes/variable
  - Main variables (CMOR name): tas; huss; uas; vas; psl; rsds; rlds; prra; prsn; friver
  - Supplementary variables (CMOR names): ts; siconca; tos; siconc; sos; uo; vo; areacello; sftof; licalvf; areacellg
  - Atmospheric variables are based on the JRA-55 reanalysis product. Raw JRA-55 variables have been adjusted relative to reference datasets taken from satellites and other reanalysis products.
  - River runoff (river discharge to the ocean) dataset is a merged product taken from Suzuki et al. (2017) (river), Bamber et al. (2012) (Greenland), and Depoorter et al. (2013) (Antarctica).
  - Sea surface salinity is based on World Ocean Atlas 2013 version 2 (Zweng et al. 2013; Boyer et al. 2015).
  - Surface ocean current is based on GlobCurrent product (Rio et al. 2014).
  - Sea surface temperature and surface ice concentration are taken from COBESST (Ishii et al. 2005)

Usage notes:

- <note 1>

Additional notes:

-



## PMIP Forcing

Contacts:

Masa Kageyama [Masa.Kageyama@lsce.ipsl.fr](mailto:Masa.Kageyama@lsce.ipsl.fr)

Jean-Yves Peterschmitt [jean-yves.peterschmitt@lsce.ipsl.fr](mailto:jean-yves.peterschmitt@lsce.ipsl.fr)

Available at: <https://pmip4.lsce.ipsl.fr/doku.php>

Status: under review

Latest version: <#> <2017-06-30>

## RFMIP Forcing (of offline radiation codes)

Atmospheric condition imposed on offline clear-sky radiation codes

Contacts:

Robert Pincus [Robert.Pincus@colorado.edu](mailto:Robert.Pincus@colorado.edu)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>, also at

<https://www.earthsystemcog.org/projects/rfmip/resources/>

Status: Preliminary version available

Latest version: 0.4 (2017-01-23)

Further information/documentation: <https://www.earthsystemcog.org/projects/rfmip/resources/>

Common characteristics of datasets in collection:

- Used in following expts.: RFMIP-IRF
- Spatial domain: Quasi-global (100 samples or “sites”)
- Spatial resolution: NA
- Temporal domain: Time slices at present-day, pre-industrial, “future”, Last Glacial Maximum
- Temporal resolution: NA

Datasets:

- `multiple_input4MIPs_radiation_RFMIP_UColorado-RFMIP-0-4_none.nc`
  - Data volume: 1.9 Mb, 1 file, 65 variables (some scalars)
  - Usage notes at <https://www.earthsystemcog.org/projects/rfmip/resources/>

Usage notes:

- This dataset provides atmospheric conditions for radiative transfer calculations including temperature, pressure, and gas concentration profiles. It also includes boundary conditions including total solar irradiance, solar zenith angle, and surface temperature. These are to be provided to each model’s radiation code and the resulting spectrally-integrated fluxes reported at each level
- Users should use the data and compute results on the vertical grid provided

Additional notes:

- Reference results will be provided by RFMIP
- Current data set is almost complete but missing CMIP6 specification of “future” greenhouse gas concentrations to be taken from RCP8.5-equivalent in year 2100
- Format is non-standard and still being discussed with PCMDI

## ScenarioMIP Forcing

## Overall questions related to ScenarioMIP

### Contacts:

Detlef van Vuuren [Detlef.vanVuuren@pbl.nl](mailto:Detlef.vanVuuren@pbl.nl)  
Brian O'Neill [boneill@ucar.edu](mailto:boneill@ucar.edu)  
Claudia Tebaldi [tebaldi@ucar.edu](mailto:tebaldi@ucar.edu)

## Future emissions (all greenhouse gases and air pollutants)

### Contacts:

Steven Smith ([ssmith@pnnl.gov](mailto:ssmith@pnnl.gov))  
Matthew Gidden ([matthew.gidden@gmail.com](mailto:matthew.gidden@gmail.com))  
Keywan Riahi ([riahi@iiasa.ac.at](mailto:riahi@iiasa.ac.at))

Available at: <https://esgf-node.llnl.gov/search/input4mips>

Status: released for CMIP6 use

Latest version: 1-1 (2018-06-21 and 2018-07-12)

Further information/documentation: <URL's and/or doi's>

### Common characteristics of datasets in collection:

- Used in following expts.: AerChemMIP, ScenarioMIP
- Spatial domain: Global
- Spatial resolution: 0.5°
- Temporal domain: 2015 - 2100
- Temporal resolution: Monthly for each of the following years: 2015, 2020, 2030, 2040, 2050, 2060, 2070, 2080, 2090, 2100
- Bulk emission species (BC, OC, SO<sub>2</sub>, NO<sub>x</sub>, NH<sub>3</sub>, CO, NMVOC, CO<sub>2</sub>, CH<sub>4</sub>)
- 4 variables for each bulk emission
  - em AIR anthro (aviation emissions)
  - em anthro (all other anthropogenic emission sectors)
  - em openburning (total open burning emissions)
  - openburning share (share of open burning emissions for 4 aggregate land-types)

### Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable

### Usage notes:

- The anthropogenic files are identical in format and definition to the historical data with the exception of the years covered.
- Checksum files for future emissions are posted at <http://www.globalchange.umd.edu/ceds/ceds-cmip6-data/> under "Auxiliary Gridded Emissions Information"
- The open burning emission files, including speciated NMVOCs from open burning, differ from the historical files as follows
  - Resolution is 0.5 instead of 0.25
  - Variable names contain "openburning" so as to distinguish between open burning and anthro sectors and differ in overall format so as to be consistent in format with the anthropogenic emission files

- “Sector” shares for future openburning emissions are provided for only the four aggregate land classifications available from the future scenario data (peat - held constant, forest, grassland, agricultural waste burning on fields)
- Other metadata differ slightly so as to be consistent with anthro emission files as noted below:

<b>Historical Open Burning</b>	<b>Future Open Burning</b>
Dimensions include ‘latitude’ and ‘longitude’	Dimensions include ‘lat’ and ‘lon’
latitude/longitude ‘long_name’ variable title case	lat/lon ‘long_name’ variable lower case
<i>(not present)</i>	lat and time dimensions have ‘realtopology’ variable
latitude:units = ‘degree_north’	lat:units = ‘degrees_north’
longitude:units = ‘degree_east’	lon:units = ‘degrees_east’
time ‘calendar’ variable = ‘365_day’	time ‘calendar’ variable = ‘noleap’
emission variable = ‘Higher_Alkenes’	emission variable = ‘NMVOC_Higher_Alkenes_em_speciated_VOC_op enburning’
emission ‘long_name’ variable lower case	emission ‘long_name’ variable title case
lat_bnds/lon_bnds/time_bnds variables are floats	lat_bnds/lon_bnds/time_bnds variables are doubles
grid is 0.25x0.25 degree latitude x longitude	grid is 0.5x0.5 degree latitude x longitude
<i>(not present)</i>	contains ‘molecular_weight’ global attribute
<i>(not present)</i>	contains ‘molecular_weight_unit’ global attribute
<i>(not present)</i>	contains ‘product’ global attribute
contains ‘dataset_version_number’ global attribute	contains ‘source_version’ global attribute
<i>(not present)</i>	contains ‘table_id’ global attribute
<i>(not present)</i>	contains ‘VOC_name’ global attribute
<i>(not present)</i>	contains ‘data_usage_tips’ global attribute
<i>(not present)</i>	contains ‘reporting_unit’ global attribute

Additional notes:

- 

### **Future concentrations greenhouse gases**

Contacts:

Malte Meinshausen ([malte.meinshausen@unimelb.edu.au](mailto:malte.meinshausen@unimelb.edu.au))  
Zebedee Nicholls ([zebedee.nicholls@climate-energy-college.org](mailto:zebedee.nicholls@climate-energy-college.org))  
IAM team members

Available at: [https://esgf-node.llnl.gov/search/input4mips/?source\\_version=1.2.0](https://esgf-node.llnl.gov/search/input4mips/?source_version=1.2.0) (and soon also at <http://climatecollege.unimelb.edu.au/cmip6>)

Status: released for CMIP6 use

Latest version: 1-2-0 (2018-06-11)

Further information/documentation:

<http://climatecollege.unimelb.edu.au/cmip6>

Common characteristics of datasets in collection:

- Used in following expts.: ScenarioMIP all and AerChemMIP ssp370-lowNTCF
- Spatial domain: global
- Spatial resolution: 0.5 degree latitudinal resolution (gr-0p5x360deg files), 15 degree latitudinal resolution (gn-15x360deg) and also as NH, SH and global means (gr1-GMNHSH files)
- Temporal domain: Jan 2015 to Dec 2500
- Temporal resolution: monthly (also annual for NH, SH and global mean files, see above)

Datasets:

- <http://climatecollege.unimelb.edu.au/cmip6>
- mole-fraction-of-XXX-in-air\_input4MIPs\_GHGConcentrations\_TARGETMIP\_UoM-SCENARIOID-1-2-0\_gr1-GMNHSH\_201501-250012.nc: Hemispheric- and global-mean monthly-mean surface mole fraction of greenhouse gas XXX
  - Data volume: 120MB; 414 files; 46 variables by 9 scenarios; 290kB/variable
- mole-fraction-of-XXX-in-air\_input4MIPs\_GHGConcentrations\_TARGETMIP\_UoM-SCENARIOID-1-2-0\_gn-15x360deg\_201501-250012.nc: Zonal mean, monthly-mean surface mole fraction of greenhouse gas XXX in 15-degree latitudinal bands
  - Data volume: 169MB; 414 files; 46 variables by 9 scenarios; 408kB/variable
- mole-fraction-of-XXX-in-air\_input4MIPs\_GHGConcentrations\_TARGETMIP\_UoM-SCENARIOID-1-2-0\_gr-0p5x360deg\_201501-250012.nc: Zonal mean, monthly-mean surface mole fraction of greenhouse gas XXX in 0.5-degree latitudinal bands
  - Data volume: 3.3GB; 414 files; 46 variables by 9 scenarios; 8.0MB/variable
  - Note that this finely-interpolated dataset is a mean-preserving interpolation of the 15-degree data, provided for convenience to interpolate to the respective model grid. The underlying observational data assimilation and projection procedure cannot provide information on GHG surface concentrations on such a fine latitude resolution.

Usage notes:

- For greenhouse gases, XXX, provided see [Historical Concentrations](#)
- The 9 scenarios for which data is provided are (written in the format ‘TARGETMIP, SCENARIOID’):
  - ScenarioMIP, IMAGE-ssp119
  - ScenarioMIP, IMAGE-ssp126
  - ScenarioMIP, MESSAGE-GLOBIOM-ssp245
  - ScenarioMIP, AIM-ssp370
    - i. Note: this scenario is sponsored by both ScenarioMIP and AirChemMIP but only ScenarioMIP appears in the filename
  - AerChemMIP, AIM-ssp370-lowNTCF
  - ScenarioMIP, GCAM4-ssp434

- ScenarioMIP, GCAM4-ssp460
- ScenarioMIP, REMIND-MAGPIE-ssp534-over
- ScenarioMIP, REMIND-MAGPIE-ssp585

Additional notes:

- See [Historical Concentrations](#) (noting that the vertical extension was developed with the historical datasets so we make no guarantee about the validity of its use for projections)

Known issues:

- No major issues known

Minor issues will be highlighted in manuscript (TBC)

## Future Land Use

Contacts:

George Hurtt [gchurtt@umd.edu](mailto:gchurtt@umd.edu)

Ritvik Sahajpal [ritvik@umd.edu](mailto:ritvik@umd.edu)

Louise Chini [lchini@umd.edu](mailto:lchini@umd.edu)

Detlef van Vuuren [detlef.vanvuuren@pbl.nl](mailto:detlef.vanvuuren@pbl.nl)

Kate Calvin

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 2.1f (2017-10-05, 2017-11-29 and 2018-01-19)

Further information/documentation:

<https://cmip.ucar.edu/lumip>

<http://luh.umd.edu>

Common characteristics of datasets in collection:

- Used in following expts.: ScenarioMIP and All LUMIP experiments
- Domain: global (land only)
- Spatial resolution: 0.25 x 0.25 degree
- Temporal domain: 2015-2100
- Temporal resolution: annual

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- <note 1>

Additional notes:

- <note 1>

## Future Ozone

2015-2100

Contacts:

Michaela Hegglin [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk)

Available at:

Status: With the mid-June release of the greenhouse gas concentration and emissions data for the scenarios, release of the future ozone dataset is projected for September, 2018. Note that the ozone datasets will cover the four Tier 1 SSPs - SSP1-26, SSP2-45, SSP3-70 and SSP5-85.

Latest version: <??> (2016-XX-XX)

Further information/documentation: [http://www.met.reading.ac.uk/ccmi/?page\\_id=375](http://www.met.reading.ac.uk/ccmi/?page_id=375) (externally hosted)

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain:
- Spatial resolution:
- Temporal domain:
- Temporal resolution:

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- <note 1>

Additional notes:

- <note 1>

## Future Nitrogen Deposition

2015-2100

Contacts:

Michaela Hegglin [m.i.hegglin@reading.ac.uk](mailto:m.i.hegglin@reading.ac.uk)

Available at:

Status: status uncertain: Awaiting greenhouse gas concentration data as input to ozone simulations. Once these are delivered, the data should become available within 2.5 months

Latest version: <??> (2016-XX-XX)

Further information/documentation: [http://www.met.reading.ac.uk/ccmi/?page\\_id=375](http://www.met.reading.ac.uk/ccmi/?page_id=375) (externally hosted)

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain:
- Spatial resolution:
- Temporal domain:
- Temporal resolution:

Datasets:

- <dataset 1>
  - Data volume: <total bytes>; <# of> files; <# of> variables; <bytes>/variable
  - <additional dataset-specific characteristic 1>

Usage notes:

- <note 1>

Additional notes:

- <note 1>

## Future Solar

Contacts:

Katja Matthes [kmatthes@geomar.de](mailto:kmatthes@geomar.de)

Bernd Funke [bernd@iaa.es](mailto:bernd@iaa.es)

Available at: <https://esgf-node.llnl.gov/search/input4mips/>

Status: released for CMIP6 use

Latest version: 3.2 (2016-10-24; replaced v3.1)

Further information/documentation: <http://solarisheppa.geomar.de/cmip6>

Common characteristics of datasets in collection:

- Used in following expts.: <experiment\_id 1>, <experiment\_id 2>, ...
- Spatial domain: (only variables describing particle-induced ion-pair production rate) global, zonal mean (in geomagnetic coordinates)
- Spatial resolution: variable
- Temporal domain: 1850-01-01 – 2299-12-31

Datasets:

- solarforcing\_ref\_day\_3.1.nc
  - Data volume: 3.1 GB; 1 file; 22 variables
  - Temporal resolution: day
- solarforcing\_ref\_mon\_3.1.nc
  - Data volume: 77 MB; 1 file; 16 variables
  - Temporal resolution: mon

Usage notes:

- The solar forcing is provided for radiative properties, i.e., total solar irradiance (TSI) and solar spectral irradiance (SSI), and F10.7cm radio flux, as well as particle forcing, i.e., geomagnetic indices Ap and Kp, and ionisation rates to account for effects of solar protons, electrons and galactic cosmic rays.

Additional notes:

- Particle forcing is provided only in the daily-resolved dataset.

## VolMIP Forcing

Contacts:

Davide Zanchettin [davidoff@unive.it](mailto:davidoff@unive.it)

Matt Toohey [mtoohey@geomar.de](mailto:mtoohey@geomar.de)

Claudia Timmreck [claudia.timmreck@mpimet.mpg.de](mailto:claudia.timmreck@mpimet.mpg.de)  
Myriam Khodri [myriam.khodri@locean-ipsl.upmc.fr](mailto:myriam.khodri@locean-ipsl.upmc.fr)  
Beiping Luo [beiping.luo@env.ethz.ch](mailto:beiping.luo@env.ethz.ch)

The stratospheric aerosol forcing data set to be used for the volc-pinatubo-XXX experiments corresponds to the CMIP6 stratospheric aerosol forcing data set for the historical experiment, see [Stratospheric Aerosols](#)

Available at: [ftp://iacftp.ethz.ch/pub\\_read/luo/CMIP6/<YOUR\\_MODEL>](ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/<YOUR_MODEL>)

Status: ready for use: version 1.0 2016-06-02

Further information/documentation: [ftp://iacftp.ethz.ch/pub\\_read/luo/CMIP6/](ftp://iacftp.ethz.ch/pub_read/luo/CMIP6/)

Common characteristics of datasets in collection:

- Used in following expts.: volc-pinatubo-full, volc-pinatubo-strat, volc-pinatubo-surf, volc-pinatubo-ini, volc-pinatubo-slab
- Spatial domain: global YZ (latitude, height)
- Spatial resolution: 90S and 90N and from 5 km to 39.5 km in 0.5 km resolution
- Temporal domain: 1990-1994
- Temporal resolution: monthly

Datasets:

- YOUR\_MODEL/CMIP\_YOUR\_MODEL\_radiation.nc
  - Data volume: ~ 5 MB/band

Usage notes:

Only a subset is needed covering the period 1990-1994

Volcanic forcing dataset for the volc-long/cluster experiments

Available at: <http://www.geosci-model-dev-discuss.net/gmd-2016-83/> (Supplementary)

Status: accepted for GMD: version 1.0 2016-05-19

Further information/documentation: <URL's and/or doi's>

Common characteristics of datasets in collection:

- Used in following expts.: volc-long-eq, volc-long-hlN, volc-long-hlS, volc-cluster-ctrl, volc-cluster-mill, volc-cluster-21C
- Spatial domain: global YZ (latitude, height)
- Spatial resolution: model specific
- Temporal domain: whole integration, dependent on experiment
- Temporal resolution: monthly

Usage notes:

- The Easy Volcanic Aerosol (EVA) module must be used to generate model-specific volcanic forcing input datasets following the VolMIP protocol. The EVA module is available as supplementary material at <http://www.geosci-model-dev-discuss.net/gmd-2016-83>



# Forcing Dataset Specifications

Data providers should attempt to generate datasets that follow the input4MIPs specifications which follow closely specifications for the obs4MIPs project (see <https://www.earthsystemcog.org/projects/obs4mips/HowtoContribute>). Data should conform to the CF-Conventions (<http://cfconventions.org/>) and follow the general outlines as specified in the CMIP5 model output specifications found at: [https://pcmdi.llnl.gov/mips/cmip5/docs/CMIP5\\_output\\_metadata\\_requirements.pdf](https://pcmdi.llnl.gov/mips/cmip5/docs/CMIP5_output_metadata_requirements.pdf)

## Proposed directory structure for input4MIPs:

input4MIPs/<mip\_era>/<target\_mip>/<institution\_id>/<source\_id>/<realm>/<frequency>/<variable\_id>/<grid\_label>/<version>/files.nc

*Directory e.g. input4MIPs/CMIP6/CMIP/PCMDI/PCMDI-AMIP-1-1-4/ocean/mon/tos/gn/v20180427/*

## Proposed filename for input4MIPs:

<variable\_id>\_input4MIPs\_<dataset\_category>\_<target\_mip>\_<source\_id>\_<grid\_label>[\_<time\_range>].nc

In the case when there are multiple variables in a file and the “multiple” variable\_id is being used, to prevent filename clashes, an additional <variable\_id>-[multiple\_file\_identifier] can be appended.

### File e.g.

*tos\_input4MIPs\_SSTsAndSeaIce\_CMIP\_PCMDI-AMIP-1-1-4\_gn\_187001-201712.nc (ncdump example below)*  
*multiple-states\_input4MIPs\_landState\_ScenarioMIP\_UofMD-AIM-ssp370-2-1-f\_gn\_2015-2100.nc*

<dataset\_category> = aerosolProperties

atmosphericState

emissions

GHGConcentrations

landState

ozone

radiation

SSTsAndSeaIce

solar

surfaceAir

surfaceFluxes

<target\_mip> =

e.g. CMIP (DECK and historical), CFMIP, DCP, FAFMIP, ScenarioMIP etc  
See identified MIPs in the earlier pages. In the case that the data will be used by multiple MIPs, include these in a comma separated string e.g. “C4MIP, OMIP”

<frequency> =

e.g. “3hr”, “6hr”, “day”, “mon”, “monC”, “yr”, “yrC”, “fx”

<realm> =

e.g. “aerosol”, “atmos”, “land”, “landIce”, “ocean”, “ocnBgchem”, “seaIce”

<source\_id> =

e.g. “PCMDI-AMIP-1-1-4” (should include an institution and version identifier)  
if multiple scenarios are included in the <target\_mip> indicate this here, e.g. CEDS-ssp245-SO2-BIO-1-2-0; The identical value for source\_id is used in the directory structure, filename and global\_attributes

<grid\_label> =

e.g. “gn”, “gr” (see Appendix 1 & 2 in the following pages)

<time\_range> =

If monthly, time\_range = “YYYYMM-YYYYMM”,

If daily = “YYYYMMDD-YYYYMMDD”

<multiple\_file\_identifier>

For time-invariant forcing fields, this element is omitted (noted in [ ] above)  
If your dataset is comprised of multiple files that contain use “multiple” as the file variable(s) and that cover the same temporal period (and are not delimited by a single variable identified in the file name), this identifier can be appended to provide further granularity between files. This is an optional identifier (noted in [ ] above)

## Documentation and miscellaneous supporting files:

We also provide the ability for documentation (in pdf format) and additional supporting files to be hosted alongside the contributed data. These files will use the same directory structure as the contributed data, with documentation/none replacing the <variable\_id>/<grid\_label> pair noted above:

input4MIPs/CMIP6/<target\_mip>/<institution\_id>/<source\_id>/<realm>/<frequency>/none/none/<version>/files.\*  
E.g. input4MIPs/CMIP6/CMIP/PCMDI/PCMDI-AMIP-1-1-4/ocean/mon/none/none/<version>/doc.pdf

## Dataset formats and organization:

Data should be netcdf files (netcdf4 using deflation/in-file compression, and not tarballs with multiple differing formats) and it is recommended that all time-slices for a single forcing dataset reside in the same file (e.g., don't split "historical" data for 1850-present across several files). In the case of the larger datasets, data should be subset in time providing multiple files of approximate size <1GB.

It is also recommended that only a single variable be placed in each file. A number of software packages that integrate with ESGF have been built around this standard ESGF/CMIP5 data format and if ESGF data does not conform to a single variable per file these packages will not function correctly.

Please see the AMIP Boundary Forcing dataset for example files

<https://esgf-node.llnl.gov/search/input4mips/> (a ncdump output of the "tos" variable is included below).

## Dataset standards compliance:

Data contributors providing data in netcdf format can validate their datasets using

<http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl>

## Incoming data for review (can be uploaded - please do not upload multi-GB files without contacting us):

<ftp://ftp.llnl.gov/> (login to ftp.llnl.gov first, then "cd incoming" and "mkdir <yourdirectory>")

Username: anonymous

Password: [me@email.com](mailto:me@email.com) <user's email address>

## Example file:

Here is an example of a forcing data set that others should follow:

KEY: yellow means "absolutely essential" ; blue means "only value permitted"; green are notes

```
netcdf tos_input4MIPs_SSTsAndSeaIce_CMIP_PCMDI-AMIP-1-1-4_gn_187001-201712 {
dimensions:
    time = UNLIMITED ; // (1776 currently)
    lat = 180 ;
    lon = 360 ;
    bnds = 2 ;
variables:
    double time(time) ;
        time:bounds = "time_bnds" ;
        time:units = "days since 1870-01-01" ;
        time:calendar = "gregorian" ;
        time:axis = "T" ;
        time:long_name = "time" ;
        time:standard_name = "time" ;
    double time_bnds(time, bnds) ;
    double lat(lat) ;
        lat:bounds = "lat_bnds" ;
        lat:units = "degrees_north" ;
        lat:axis = "Y" ;
        lat:long_name = "latitude" ;
        lat:standard_name = "latitude" ;
    double lat_bnds(lat, bnds) ;
```

```

double lon(lon) ;
    lon:bounds = "lon_bnds" ;
    lon:units = "degrees_east" ;
    lon:axis = "X" ;
    lon:long_name = "longitude" ;
    lon:standard_name = "longitude" ;
double lon_bnds(lon, bnds) ;
float tos(time, lat, lon) ;
    tos:standard_name = "sea_surface_temperature" ; please set standard name if available (see CF Standard Name Table)
    tos:long_name = "Sea Surface Temperature" ;
    tos:units = "degC" ; must be included in udunits library (see here)
    tos:cell_methods = "time: mean" ;
    tos:cell_measures = "area: areacello" ;
    tos:missing_value = 1.e+20f ;
    tos:_FillValue = 1.e+20f ;

// global attributes:
:Conventions = "CF-1.7 CMIP-6.2" ; "CF-1.6" is also acceptable
:activity_id = "input4MIPs" ;
:comment = "Based on Hurrell SST/sea ice consistency criteria applied to merged HadISST (1870-01 1981-10) & NCEP-OI2 (1981-11 to
2017-12)" ;
:contact = "PCMDI (pcmdi-cmip@llnl.gov)" ;
:creation_date = "2018-04-27T18:53:31Z" ;
:data_specs_version = "01.00.23" ;
:dataset_category = "SSTsAndSeaIce" ;
:external_variables = "areacello" ;
:frequency = "mon" ;
:further_info_url = "https://pcmdi.llnl.gov/mips/amip" ;
:grid = "1x1 degree longitude x latitude" ; see appendix 1 below
:grid_label = "gn" ; see appendix 1 below
:history = "2018-04-27T18:53:31Z; CMOR rewrote data to be consistent with input4MIPs, CMIP6, and CF-1.7 CMIP-6.2 standards;\n",
"File processed: 27-04-2018 18:52:06 PM UTC; San Francisco, CA, USA" ;
:institution = "Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA
94550, USA" ;
:institution_id = "PCMDI" ;
:mip_era = "CMIP6" ;
:nominal_resolution = "1x1 degree" ; see appendix 1 below
:product = "observations" ;
:realm = "ocean" ;
:references = "Taylor, K.E., D. Williamson and F. Zwiers, 2000: The sea surface temperature and sea ice concentration boundary
conditions for AMIP II simulations. PCMDI Report 60, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National
Laboratory, 25 pp. Available online: https://pcmdi.llnl.gov/report/pdf/60.pdf ;
:region = "global_ocean" ;
:release_year = "2018" ;
:source = "PCMDI-AMIP 1.1.4: Merged SST based on UK MetOffice HadISST and NCEP OI2" ;
:source_description = "Sea surface temperature and sea-ice datasets produced by PCMDI (LLNL) for the AMIP (DECK) experiment of
CMIP6" ;
:source_id = "PCMDI-AMIP-1-1-4" ;
:source_type = "satellite_blended" ;
:source_version = "1.1.4" ; This should be all numeric, following the MAJOR.MINOR.PATCH convention
:table_id = "input4MIPs_Omon" ;
:table_info = "Creation Date:(25 April 2018) MD5:0e0d2efd376ad34d4e628cf785da9" ;
:target_mip = "CMIP" ; allows search to ascertain whether data is associated with "historical", "future" or other experiments
:title = "PCMDI-AMIP 1.1.4 dataset prepared for input4MIPs" ;
:tracking_id = "hdl:21.14100/73bdc5e3-f46a-43dc-9c28-2d523f93f362" ;
:variable_id = "tos" ;
:license = "AMIP boundary condition data produced by PCMDI is licensed under a Creative Commons
Attribution-[NonCommercial-]ShareAlike 4.0 International License (https://creativecommons.org/licenses). Consult
https://pcmdi.llnl.gov/CMIP6/TermsOfUse for terms of use governing input4MIPs output, including citation requirements and proper
acknowledgment. Further information about this data, including some limitations, can be found via the further_info_url (recorded as a global
attribute in this file). The data producers and data providers make no warranty, either express or implied, including, but not limited to,
warranties of merchantability and fitness for a particular purpose. All liabilities arising from the supply of the information (including any liability
arising in negligence) are excluded to the fullest extent permitted by law" ;
:cmor_version = "3.3.2" ;
}

```

## Appendix 1: Guidance for Defining “grid” and “grid\_label” Attributes

[This appendix draws from and is consistent with the [“CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV’s”](#) google doc dated (14 September 2016 (v6.0.0))

The “grid” global attribute can be used to describe the horizontal grid and any regridding procedure applied. There is no standard form used to record this information, but it is suggested that when appropriate the following be indicated: brief description of original (“native”) grid and resolution, and if data have been regridded, regridding procedure and description of target grid. Here are some examples:

- grid = “data provided on T63 gaussian grid (64x128 latxlon), which is the highest resolution currently available.”
- grid = “data regridded to a CMIP6 standard 1x1 degree latxlon grid from the original T63 grid using an area-average preserving method.”
- grid = “data regridded via bilinear interpolation to a 3x3 deg latxlon grid from the original T63 gaussian grid (64x128 latxlon)”

Data providers may choose to report their output on the original (“native”) grid and/or regrid it to one or more target grids. To distinguish between output reported on different grids, a “grid\_label” attribute is defined.

The rules for assigning grid labels should make clear which grid is considered best -- the so-called “primary” grid. If output is reported on the native grid, this is always deemed the “primary” grid. If output is not reported on the native grid, then modeling groups should regrid the data to some primary grid of its choosing. For the “primary” grid the following labels apply:

- grid\_label = “gn” (output is reported on the native grid)
- grid\_label = “gr” (output is not reported on the native grid, but instead is regridded to the “primary grid” of choice)
- grid\_label = “gm” (global mean output is reported, so data are not gridded)

If besides the “primary” grid, output is regridded to an additional grid, then for this output:

grid\_label = “gr[i]” (a “secondary” grid), where <i> should be replaced by a positive integer less than 10, which distinguishes this output from other regridded output.

Note that:

- If model output is reported on a native grid, then if regridded output is also reported, it must not be labeled “gr”, but instead should be of the form gr[i] (e.g., gr1, gr2, ...).
- If model output is reported on a native grid, then if regridded output is also reported, it must not be labeled “gr”, but instead should be of the form gr[i] (e.g., gr1, gr2, ...).
- The grid label provides no information about the grid other than to indicate whether or not the data have been regridded (from the native grid) and whether or not the grid is considered to be a “primary” grid by the data provider.
- Output for different variables may be reported on different grids, so “gn” and “gr” may not uniquely define a grid even within a single model.
- If a variable is reported on more than one grid, one of these grids must be labeled “gn” or “gr” (as appropriate), and others must be labeled as secondary grids (“gr1”, “gr2”, etc.).
- The grid labels (“gr1”, “gr2”, etc.) must be defined consistently for all variables reported from a single model on the same secondary grid. Thus, if “gr3” indicates output on a 2x2 degree grid for one variable, then the same label should be used for other variables output on this grid (unless for

another variable output is considered to be primary, in which case the output would be labeled “gn” or “gr”, as appropriate).

- Output on the same grid but from different models will usually not be labeled with the same grid\_label, since data providers independently assign the labels.
- A “nominal\_resolution” attribute must be defined as specified in Appendix 2, and a brief description of the grid should be recorded in the “grid” global attribute (see note 10).
- For zonal mean output, a “z” should be appended to the grid label that would apply before performing the zonal mean (e.g., “gnz”, “grz”, “gr2z”). The “gnz” label would likely only be appropriate when zonally-averaging data on a native cartesian latxlon grid.
- For global mean output or basin mean output, the grid label should refer to the grid used in computing the area mean (presumably the native grid would be best for this purpose, so “gn”).
- For “site” data the label “gn” should be used presuming data are obtained from the single native grid cell located nearest each site.
- For “transport through a straight” (and the like), the grid label should reflect the grid relied on in calculating the transport (presumably the native grid would be best for this purpose, so “gn”).

## Appendix 2: Algorithms for Defining the “nominal\_resolution” Attribute

[This appendix was copied from the [“CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV’s”](#) google doc dated (23 March 2017 (v6.2.2))]

There are various ways grid resolution might be defined, but in CMIP6 this should be done in the same way by all models. If the following procedure seems inappropriate for a model, the modeling group may request an exception from the WGCM Infrastructure Panel (WIP).

Algorithm for defining the nominal\_resolution global attribute:

1. For each grid cell, calculate the distance (in km) between each pair of cell vertices and select the maximum distance (“ $d^{\max}$ ”). For latxlon grid cells, for example,  $d^{\max}$  would be the diagonal distance.
2. Calculate the mean over all cells of  $d^{\max}$ , weighting each by the grid-cell’s area ( $A$ ). This defines the “mean resolution” (dmax). The formula is:

$$\overline{d^{\max}} = \frac{\sum d_i^{\max} A_i}{\sum A_i}$$

where the sums are over all grid cells except for the following cases:

- For a global ocean grid, only sum over the ocean grid cells.
- For a sea ice model calculated on the ocean grid, include *all* ocean cells, whether or not they contain sea ice.
- For land surface models calculated on the atmospheric grid include all grid cells (not just those over land).
- For a land surface model calculated on its own grid, include all land grid cells .
- For data reported on a sub-domain of the globe (e.g., northern high latitudes only), include only those grid cells in the domain.
- For data reported at individual sites, calculate as if every grid cell contained one site (i.e., include all grid cells).

3. Except in the case of a CMIP6 “standard grid” (see item 4 below), define the global attribute “nominal\_resolution” according to:

If	dmax < 0.72 km,	nominal_resolution = “0.5 km”
else If	dmax < 1.6 km,	nominal_resolution = “1 km”
else if	dmax < 3.6 km,	nominal_resolution = “2.5 km”
else if	dmax < 7.2 km,	nominal_resolution = “5 km”
else if	dmax < 16 km,	nominal_resolution = “10 km”
else if	dmax < 36 km,	nominal_resolution = “25 km”
else if	dmax < 72 km,	nominal_resolution = “50 km”
else if	dmax < 160 km,	nominal_resolution = “100 km”
else if	dmax < 360 km,	nominal_resolution = “250 km”
else if	dmax < 720 km,	nominal_resolution = “500 km”
else if	dmax < 1600 km,	nominal_resolution = “1000 km”
else if	dmax < 3600 km,	nominal_resolution = “2500 km”
else if	dmax < 7200 km,	nominal_resolution = “5000 km”
else		nominal_resolution = “10000 km”

The different nominal\_resolution values are approximately spaced logarithmically and the bounds on each are logarithmically approximately half-way between the values.

For a regular latlon grid it is possible to calculate the approximate mean resolution analytically:

$$\overline{d}^{max} = \frac{r_{earth}\Delta\phi}{2} \left[ 1 + \frac{\Delta\phi^2 + \Delta\lambda^2}{\Delta\phi \Delta\lambda} \tan^{-1} \left( \frac{\Delta\lambda}{\Delta\phi} \right) \right]$$

where  $r_{earth}$  is the radius of the Earth (in km), and  $\Delta\phi$  and  $\Delta\lambda$  are the latitude and longitude angular dimensions of each cell (measured in radians). If these dimensions are identical, then the mean resolution is:

$$\overline{d}^{max} = \frac{r_{earth}\Delta\phi}{2} \left[ 1 + \frac{\pi}{2} \right] \approx 1.285 r_{earth} \Delta\phi$$

For a 0.5x0.5 degree grid, the mean resolution, according to this formula is 71.5 km (given 6371 km as the Earth’s radius).

4. Note that for so-called “standard CMIP6 grids,” nominal\_resolution will be assigned a string defined by the WGCM Infrastructure Panel (WIP), rather than applying the above algorithm. Currently the only WIP-assigned standard grid is a 1x1 degree latitude by longitude grid with 180 latitudes and 360 longitudes of equal (angular) width, and with one longitude centered at 0.5 degree east longitude (consistent with the World Ocean Atlas). For this CMIP6 standard grid, nominal\_resolution = “1x1 degree”, rather than “100 km” (which would be the result of the above algorithm). Defining the standard grid resolution in this way makes it easy for users to download only data that have been

regridded to the standard grid, since “1x1 degree” will be selectable from the “nominal\_resolution” search facet on ESGF.

### Appendix 3: Relevant web links

<https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6> - CMIP6 WCRP webpage

<http://goo.gl/r8up31> - Short link to this document

<http://goo.gl/v1drZl> - Short link to the Global Attributes, DRS document

### Appendix 4: Document version information

The document version number consists of 3 integers separated by “.”. The first integer is “6”, indicating the document applies to CMIP6. The second integer will be incremented if errors are found in the DECK/Historical forcing data that had been previously released and approved by the CMIP panel. The third digit will be incremented whenever an amendment or update is made to this document, or when the download details for an existing contributing dataset (to the current 6.x release) are amended. The third digit will also be incremented if the version number of a contributed satellite MIP dataset changes, with these changes most commonly related to preparing data to conform to the input4MIPs specifications required for publishing on ESGF. **Note that red font is used to highlight those versions where 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 the original 6.0.0 version released on the 20th December 2016. Two key datasets from the original release have been revised:

- [Anthropogenic SLCF \(Short Lived Climate Forcings\), CO2, and CH4 Emissions data v2017-05-01 \(Steven Smith\)](#)
  - updated from 2016-06 and 2016-07 versions (corrected an error with earlier versions)
  - fixed two issues with the gridded historical data (see the [README](#) file for further details)
  - added CO2 and CH4 gridded products
- [AMIP Boundary forcing v1.1.2 \(PCMDI\)](#) - further details are available in the [release notes](#)
  - updated from the 1.1.1 version (extended temporal coverage from 2016-06 to 2016-12)
  - added the sftof variable
  - the older SST and sea ice data should be the same in this version except for insignificant changes in the data affecting the least significant figures

For all groups that have downloaded earlier versions, please replace those versions with the updated datasets

**6.1.1 (22nd May 2017)** - Corrected [Anthropogenic SLCF \(Short Lived Climate Forcings\), CO2, and CH4 emissions data v2017-05-18 \(Steven Smith\)](#) made available through input4MIPs

**6.1.2 (6th June 2017)** - [HighResMIP \(Malcolm Roberts\)](#) details updated; Minor type clarifications made

**6.1.3 (6th June 2017)** - [Aerosol Optical Properties and Relative Change in Cloud Droplet Number Concentration \(Stephanie Fiedler\)](#) details updated; Minor type clarifications made

**6.1.3 (6th June 2017)** - Corrected frequency <invariant> to <fx> consistent with the [CMIP6\\_CVs](#)

**6.1.4 (18th July 2017)** - [Aerosol Optical Properties and Relative Change in Cloud Droplet Number Concentration \(Stephanie Fiedler\)](#) made available through input4MIPs

**6.1.5 (25th July 2017)** - Informational text about CMIP Panel releases added to the first page following suggestions by Veronika Eyring

**6.1.6 (28th August 2017)** - [DAMIP Forcing \(David Plummer\)](#), [C4MIP Forcing \(Heather Graven\)](#), and [Anthropogenic SLCF \(Short Lived Climate Forcing\), CO2, and CH4 emissions \(Leyang Feng\)](#) datasets and monthly climatologies (Anthropogenic SLCF, Leyang Feng) made available through input4MIPs; DAMIP forcing information added to current document, was previously omitted

**6.2.0 (11th September 2017)** - [Anthropogenic SLCF \(Short Lived Climate Forcings\), corrected aircraft emissions data v2017-08-30 \(Steven Smith\)](#) made available through input4MIPs. A new v3 release of the [Stratospheric aerosol data \(Beiping Luo\)](#) was also made available on the 15th September 2017

**6.2.0 (26th September 2017)** - The input4MIPs filename template was revised to remove the “\_[\_multiple\_file\_identifier].nc” example which was breaking filename consistency across the project. Rather we now suggest that this identifier is appended to the <variable\_id>. See example [above](#)

**6.2.1 (6th October 2017)** - [Anthropogenic SLCF \(Short Lived Climate Forcings\), corrected SO2 aircraft emissions v2017-10-05 \(Steven Smith\)](#) data made available through input4MIPs (v2017-08-30 SO2 aircraft emissions data were deprecated). [Stratospheric aerosol v3.0.0 \(Beiping Luo\)](#) data made available through input4MIPs. [Future Land Use v2.1f for ScenarioMIP \(George Hurtt\)](#) data made available through input4MIPs (IMAGE-ssp126, AIM-ssp370 and MAGPIE-ssp585 scenarios)

**6.2.2 (1st November 2017)** - [AMIP Boundary forcing v1.1.3 \(PCMDI\)](#) made available through input4MIPs; Updated example file ncdump output following the tos v1.1.3 file

**6.2.3 (29th November 2017)** - [Future Land Use v2.1f for ScenarioMIP \(George Hurtt\)](#) data made available through input4MIPs (GCAM-ssp434 and GCAM-ssp460 scenarios)

**6.2.4 (19th January 2018)** - updated details for [forcing dataset specifications](#)

**6.2.5 (22nd January 2018)** - complete input4MIPs project republication on the ESGF (cleaned up existing data, identifiers and protocols to enable project replication); [Future Land Use v2.1f for ScenarioMIP \(George](#)



