CMIP5 Model Output Requirements:
File Contents and Format, Data Structure and Metadata

Karl E. Taylor[1] and Charles Doutriaux

Program for Climate Model Diagnosis and Intercomparison (PCMDI)

7 January 2010

Changes made to this document after 7 January 2010 are summarized in the last section
below.

Overview

Past experience with model intercomparison projects, highlighted by the third phase of the Coupled Model Intercomparison Project (CMIP3), has demonstrated the exceptional value gained by archiving multi-model output in a structured and uniform way. The user community now expects to be able to extract data both efficiently and in a uniform way across all models. The effort to write the data in a uniform structure and format falls on the modeling centers contributing the data. Building on CMIP3, we are now preparing for the fifth phase of CMIP (CMIP5). In CMIP5 a more comprehensive suite of experiments is planned (see Taylor et al., 2009) and a more extensive list of model output is requested (see CMIP5 Requested Output). Here we provide the specifications for writing CMIP5 model output. It should be noted that these requirements also extend to the output from intercomparison experiments closely aligned with (or incorporated as part of) CMIP5, including AMIP, CFMIP, C4MIP, PMIP, and TAMIP.

A software library, CMOR2 (pronounced "see more two") has been written to facilitate writing model output that conforms to these requirements. This library is written in the C programming language, but can be accessed through interfaces from Fortran or Python programming languages. Documentation for this library explains how it can substantially reduce the burden placed on the modeling centers preparing CMIP5 model output. The library accesses the information contained in the excel spreadsheets that define the characteristics of the requested model output (after they have been reformatted into CMOR2-readable tables), so a group choosing to write its data through CMOR2 needs to supply only information specific to its own model; most of the metadata required by CMIP5 will be automatically provided by CMOR2. For those groups choosing not to use CMOR2, the following requirements must nevertheless be strictly adhered to. Although we have attempted to make the following specifications complete, the safest way to ensure that model output conforms to CMIIP5 requirements is to process it through CMOR2. In

---

[1] Send questions/comments/suggestions to taylor13@llnl.gov.

this document red text indicates that the user must be especially careful to adhere to the specifications, since it will be very difficult for anyone else to determine whether or not the information is correct. Compliance with the rest of the requirements is pretty much guaranteed if the data is written with CMOR2. CMOR2 is distributed with a Python-based checker (check_CMOR_compliant.py). Output that has not been written through CMOR2, but is thought to adhere to the requirements, may be passed through the CMOR2 checker to catch some errors.

The requirements for CMIP5 are similar to those required in CMIP3, but there are a few major changes:
- Model output may now be contributed on any native grid, even one that is not a Cartesian latitude-longitude grid.
- Filenames and directory structures are now mandated according to a defined template.
- A number of additional global attributes are now required.
- A few new variable attributes are now required (when appropriate).

The requirements for data contributed to the archive are listed in five sections below, the first specifying the general structure and format of the data, the second the directory structure and names of files and directories, the third the required and recommended "global attributes", the fourth the metadata required for describing the coordinates, and last the constraints imposed on the variables themselves.

Data format, data structure, and file content requirements:

- Data must be written in the netCDF-3 format and conform to the CF metadata standards. The output must be readable through the netCDF-3 API (application program interface) and conform to the netCDF "classic" data model. This means that if the data is written using the netCDF-4 API, the mode must be set to NC_CLASSIC_MODEL and *do **not** invoke chunking/compression/shuffling*.
- Each file must contain only a single output field from a single simulation (i.e., a single run). Each file will also include coordinate variables, attributes and other metadata as specified below. If the field is a function of time, more than one time sample (but not necessarily all time samples) may be included in a single file. Data representing a long time-series, typical of many coupled model simulations, will usually be split into several files, which should neither be too large (to be unwieldy) nor too small (as to create vexing I/O performance issues). Monthly data, for example, might be divided into multi-decade chunks. It is recommended that the same size chunks should be used for all variables found in the same table of the CMIP5 Requested Output. Note that when several tables are grouped together (e.g., under the single name "Amon"), each of the "sub-tables" should be considered as different tables when following the above recommendation. For example, 2-D and 3-D fields usually appear in separate "sub-tables" of "Amon", and one could use different "chunk" lengths in this case, without violating the recommendation. There may be cases in which 2-D and 3-D fields appear in the

2

same table, and there may be good reasons to choose different "chunk" lengths in this case (going against the above recommendation).
- Some atmospheric fields that are functions of the vertical coordinate must be interpolated to standard pressure levels (as specified in the CMIP5 Requested Output list of variables). Other fields (e.g., the 3-d cloud fraction) will reside on the original model levels. There are different metadata and attribute requirements specified below for these two types of "3-d" fields.
- Oceanic fields that are a function of the vertical coordinate should usually be reported on the native grid.

Structure and names of directories and names of files

The IPCC database will comprise output from many different models, dozens of experiments, and perhaps several ensemble members, which have been sampled (or averaged) in a number of different ways (e.g., monthly, daily, 3-hourly). The directory structure for all of the output is specified in the "CMIP5 Data Reference Syntax (DRS) and Controlled Vocabularies" document, subsequently referred to here as the "DRS document". The names of the directories must be drawn from the "controlled vocabulary" specified in the same document. Finally, the filenames themselves must strictly follow the template given in the DRS document.

The directory structure will be as follows (see the DRS document for definitions of the different elements):

<activity>/<product>/<institute>/<model>/<experiment>/<frequency>/<model ing realm>/<variable name>/<ensemble member>/

Here are two examples:
        /CMIP5/output/UKMO/HADCM3/decadal1990/day/atmos/tas/r3i2p1/
        /CMIP5/output/UKMO/HADCM3/rcp45/mon/ocean/uo/r1i1p1/

Note that <model> should be identical to model_id, one of the global attributes described in a subsequent section, except that the following characters, if they appear in model_id should be replaced by a hyphen (i.e., by '-'): _ ( ) . ; , [ ] : / * ? < > " ' { } & and/or a "space". If, after substitution, any hyphens are found at the end of the string, they should be removed.

The filenames will follow the template that is described more fully in the DRS document:

filename = *<variable name>_<MIP table>_<model>_<experiment>_<ensemble member>[_<temporal subset>]*.nc

Note that the *<temporal subset>* is omitted for variables that are time-independent (so-called "fixed" fields). For these "fixed" fields the ensemble member should invariably be set to r0i0p0, denoting that this field is valid for all "r", "i", and "p". For gridspec files (which are "fixed" fields) the template is also slightly different in that the variable name is

replaced by "gridspec" and a *modeling realm* identifier is added (see the DRS document for options).

Note that *<variable name>* and *<MIP table>* together uniquely define the variable (except in the case of gridspec files where the *modeling realm* qualifier is also necessary).

Here are two examples:
    tas_Amon_HADCM3_ historical_r1i1p1_185001-200512.nc
    gridspec_atmos_fx_IPSL-CM5_historical_r0i0p0.nc

Requirements for global attributes:

There are required attributes, optional standardized attributes, and the user may define any additional attributes thought to be useful.

- Required global attributes:
  - → branch_time = time in parent experiment when this simulation started (expressed in the units of the parent experiment). [See parent_experiment_id for more information about the "parent".] For example, if the child run were spun off from a control run at a time of "2000" in the control run, and the time units in the control run were "days since 500-01-01", then regardless of the units in the child experiment, the user would store branch_time=2000 (i.e., this time should be relative to the basetime of the control, not relative to a basetime of 0-01-01 and not relative to the basetime of the child). The branch_time should be set to 0.0 if not applicable (for example an AMIP run or a control run that was not initiated from another run).
  - → contact = name and contact information (e.g., email, address, phone number) of person who should be contacted for more information about the data.
  - → Conventions = 'CF-1.4'
  - → creation_date = a string representation of the date when the file was created in the format: "*YYYY-MM-DD*-T*HH*:*MM*:*SSZ*" with replacement of all but "T" and "Z" by the obvious date or time indicator (e.g., "2010-03-23-T05:56:23Z").
  - → experiment = a string providing a title for the experiment, as specified in the controlled vocabulary found in the table column labeled "Experiment Name" in Appendix 1.1 of the DRS document.
  - → experiment_id = a short string identifying the experiment, as specified in the controlled vocabulary found in the table column labeled "Short Name of Experiment" in Appendix 1.1 of the DRS document.
  - → forcing = a string containing a list of the "forcing" agents that should cause the climate to change in the experiment. A forcing agent will show some secular variation due to prescribed changes in concentration or emissions (or in the case of land-use, change in prescription of surface conditions). Sometimes the change will be due to emissions of a precursor species that relatively quickly becomes transformed into the forcing agent itself (e.g., transformation of SO2 emissions to sulfate aerosol. Changes in

composition resulting from the simulated climate change itself should not be counted as "forcing"; they are regarded as feedbacks. For a control run with no variation in radiative forcing or for any other experiment for which there are no externally imposed changes in radiative forcing agents, set this to "N/A".   Otherwise, the forcing should be expressed as a comma separated list of identifying strings that are part of the so-called DRS controlled vocabulary described in Appendix 1.2 of the DRS document. Within or following this machine-interpretable list may be text enclosed in parentheses providing further information. Use the terms in Appendix 1.2 that are most specific (i.e., avoid "Nat" and "Ant"). If, for example, only CO2, methane, direct effects of sulfate aerosols, tropospheric and stratospheric ozone, and solar irradiance varied, then specify "GHG, SD, Oz, Sl (GHG includes only CO2 and methane)".

→ frequency = a string indicating the interval between individual time-samples in the atomic dataset. The following are the only options: "yr", "mon", "day", "6hr", "3hr", "subhr" (sampling frequency less than an hour), "monClim" (climatological monthly mean) or "fx" (fixed, i.e., time-independent).   The sampling frequency is specified at the top of each spreadsheet (cell G1) in CMIP5 Requested Output. For a few tables some variables within the table are sampled differently, as indicated by an entry in the "frequency column (T) of the spreadsheets.

→ initialization_method = an integer (≥1) referring to the initialization method used or different observational datasets used to initialize. If only a single method and dataset was used to initialize the model, then this argument should normally be given the value 1. For fields appearing in table "fx" in the CMIP5 Requested Output, set initialization_method=0 (violating the general rule that it should be a positive definite integer). See the DRS document for guidance on assigning initialization_method. Note that the initialization_method is used in constructing the "ensemble member" called for in the DRS document; it is the value of M in r<N>i<M>p<L>.

→ institute_id = a short acronym describing "institution" (e.g., 'GFDL') For CMIP5, the institute_id should be officially approved by the CMIP Panel (through PCMDI).

→ institution = character string identifying the institution that generated the data [e.g., 'GFDL (Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA']

→ model_id = a string containing an acronym that identifies the model used to generate the output.   For CMIP5, the model_id should be officially approved by the CMIP Panel (through PCMDI). It should be as short as possible, so that it can be used, for example, in labeling curves on multi-model plots (e.g., as might appear in the Fifth Assessment Report of the IPCC). The acronym may include the acronym of the modeling center and the model name/version separated by a hyphen (e.g., "IPSL-CM4"), but it may be o.k. to omit the modeling center. Please note that you might in the future want to submit results from a successor to the present model, so if appropriate, you may want to indicate a model version, but please keep it

simple e.g., CCSM4, not CCSM4.1.2. Full version information will appear in the "source" global attribute described below. The model_id, possibly modified as necessary to eliminate characters not permitted by the DRS, will be used to construct directory and filenames. For further information, see the earlier section describing the directory and filenames.

→ modeling_realm = a string that indicates the high level modeling component which is particularly relevant. For CMIP5, permitted values are: "atmos", "ocean", "land", "landIce", "seaIce", "aerosol" "atmosChem", or "ocnBgchem" (ocean biogeochemical). Note that sometimes a variable will be equally (or almost equally relevant) to two or more "realms", in which case a primary "realm" is assigned, but cross-referenced or aliased to the other relevant "realms".   The modeling realm(s) is (are) specified in the "realm" column (S) of the spreadsheets found in CMIP5 Requested Output.

→ parent_experiment_id = experiment_id indicating which experiment this simulation branched from. This should match the experiment_id of the parent unless the "parent" is irrelevant, in which case this should be set to "N/A". The experiment_id's can be found in the table column labeled "Short Name of Experiment" in Appendix 1.1 of the DRS document.

→ parent_experiment_rip = identifier indicating which member of an ensemble of parent experiment runs this simulation branched from. This identifier should be defined even when only a single parent experiment simulation was performed, but if parent_experiment_id="N/A", then parent_experiment_rip should also be set to "N/A". The "rip" value is constructed from the "realization", "initialization_method", and "physics_version" of the parent experiment, using the template "r<N>i<M>p<L>" to define the ensemble member. This template is described under "ensemble member" in the DRS document. When possible and not inappropriate, the child experiment should inherit the "rip" value from the parent.

→ physics_version = an integer (≥1) referring to the physics version used by the model If there is only one physics version of the model, then this argument should be normally given the value 1. Note that model versions that are substantially different should be given a different "model_id"; assigning a different "physics_version" should be reserved for closely-related model versions (e.g., as in a "perturbed physics" ensemble) or for the same model, but with different forcing or feedbacks active. In CMIP5, one would distinguish, for example, among runs forced by different combinations of "forcing" agents (as called for under the "historicalMisc" experiment – experiment 7.3) by assigning different values to physics_version. For fields appearing in table "fx" in the CMIP5 Requested Output, set physics_version=0 (violating the general rule that it should be a positive definite integer). Note that the physics_version is used in constructing the "ensemble member" called for by the DRS document; it is the value of L in r<N>i<M>p<L>.

→ product = "output", which indicates that the data you are writing is model output.

→ project_id = "CMIP5" for CMIP5. [For the "Transpose AMIP" project, it will be assigned "TAMIP".]

→ realization = an integer (≥1) distinguishing among members of an ensemble of simulations (e.g., 1, 2, 3, etc.). If only a single simulation was performed, then it is recommended that realization=1. For fields appearing in table "fx" in the CMIP5 Requested Output, set realization=0 (violating the general rule that it should be a positive definite integer). Note that if two different simulations were started from the same initial conditions, the same realization number should be used for both simulations. For example if a historical run with "natural forcing" only and another historical run that includes anthropogenic forcing were initiated from the same point in a control run, both should be assigned the same realization. Also, each so-called RCP (future scenario) simulation should normally be assigned the same realization integer as the historical run from which it was initiated. This will allow users to easily splice together the appropriate historical and future runs. A similar convention should be followed, when appropriate, with other simulations (e.g., the decadal simulations). Note that the realization can be used in constructing the "ensemble member" called for by the DRS document; it is the value of N in r<N>i<M>p<L>. [Note that for the "Transpose AMIP" project, the "realization" number is used to distinguish among the 16 members of each of 4 suites of runs (i.e., the 4 "seasons") generated from different observed conditions, spaced 30 hours apart. So, for example, the 16-member ensemble of runs initialized at 00Z on 15 Oct 2008, 06Z 16 Oct 2008, 12Z 17 Oct 2008, and so-on, would be assigned "r1", "r2", "r3", etc.]

→ source = character string fully identifying the model and version used to generate the output. The first portion of the string should be a copy of the global attribute "model_id". Additionally, this attribute must include the year (i.e., model vintage) when this model version was first used in a scientific application. Finally, it should include information concerning the component models. The following template should be followed in constructing this string: '<model_id> <year> atmosphere: <model_name> (<technical_name>, <resolution_and_levels>); ocean: <model_name> (<technical_name>, <resolution_and_levels>); sea ice: <model_name> (<technical_name>); land: <model_name> (<technical_name>)' For some models, it may not make much sense to include all these components, and nothing following "<year>" is absolutely mandatory. As an example, "source" might contain the string: 'CCSM2 2002 atmosphere: CAM2 (cam2_0_brnchT_itea_2, T42L26); ocean: POP (pop2_0_ver_1.4.3, 2x3L15); sea ice: CSIM4; land: CLM2.0'. For some models it might be appropriate to list only a single component, in which case the descriptor (e.g., 'atmosphere') may be omitted along with the other model components (e.g., for an aquaplanet experiment: 'CAM2 2002 (cam2_0_brnchT_itea_2,

→ table_id = should be assigned a character string identifying the CMIP5 Requested Output table where this variable appears. This string should be of the form "Table <table name>" as it appears in cell F1 of the spreadsheets in CMIP5 Requested Output>", followed by "(<date of table>", where the date is the date of the requested output table (e.g., "table_id=Table Amon (10 June 2010)").

→ tracking_id = a string that is almost certainly unique to this file and must be generated using the OSSP utility which supports a number of different DCE 1.1 variant UUID options. For CMIP5 version 4 (random number based) is required. Download the software from http://www.ossp.org/pkg/lib/uuid/. The tracking_id might look something like: `02d9e6d5-9467-382e-8f9b-9300a64ac3cd`.

- Optional global attributes

  → comment = A character string containing additional information about the data or methods used to produce it. The user might, for example, want to provide a description of how the initial conditions for a simulation were specified and how the model was spun-up (including the length of the spin-up period).

  → history = A character string containing an audit trail for modifications to the original data. Each modification is typically preceded by a "timestamp". The "history" attribute provided here will be a global one that should not depend on which variable is contained in the file. A variable-specific "history" can also be included as an attribute attached to the output variable.

  → references = A character string containing a list of published or web-based references that describe the data or the methods used to produce it. Typically, the user should provide references describing the model formulation here.

  → title = "*<model_id>* model output prepared for CMIP5 *<experiment>*" where *<model_id>* should be replaced by the contributing model's acronym or name (see description above) and *<experiment>* should be replaced by one of the experiment id's found in the table column labeled "Experiment Name" in Appendix 1.1 of the DRS document. A sample title is: 'IPSL-CM5 model output prepared for CMIP5 historical"

Requirements for coordinate variables:

- All coordinate variables must be written as double precision floating point numbers (netCDF type NC_DOUBLE).   [Note that simple "index" dimensions are not considered to be "coordinate variables".]
- The names for all coordinate variables are specified in the "output dimension name" column (C) of the "dims" table in CMIP5 Requested Output.

- The units for all coordinate variables are specified in the "units" column (H) of the "dims" table in CMIP5 Requested Output.
- For all time coordinates the units must be "days since \<basetime\>", where \<basetime\> must be specified by the user, typically in the form year-month-day (e.g., "days since 1800-01-01"). Follow these rules:
  - → The same 'base time' should apply to all time samples in a single simulation (i.e., when creating a series of files containing model output for a single run, don't change the units or 'base time' from one file to the next; in the above example 1800-01-01 is the 'base time').
  - → For simulations meant to represent a particular historical period, set the 'base time' to the time at the beginning of the simulation. A historical run initialized with forcing for year 1850 would, for example, have units of "days since 1850-01-01". The 'base time' used in the control run is irrelevant as is the time in the control run when the historical run was initialized.
  - → For the future scenario runs, retain the same 'base time' as used in the historical run from which it was initiated.
  - → For simulations not tied to a particular date, the 'base time' is arbitrary, but it should correspond to the beginning of the run. Thus, in these simulations, 'base time' is whatever time one decides to assign to the beginning of the run.
  - → All values of the time coordinate should be positive, which will be ensured by following the above rules.
- For time-mean data, a time coordinate value must be defined as the mid-point of the interval over which the average is computed. (More generally, this same rule applies whenever time-bounds are included: the time coordinate value should be the mean of the two time bounds.)
- The "bounds" are required for certain coordinate variables as indicated by the "bounds?" column (K) of the dims table in CMIP5 Requested Output. The bounds variable should be double precision.
- For grids other than Cartesian latitude-longitude grids, the output field can be a function of 1, 2, 3, or more horizontal "dimensions", depending on the natural logical structure of the model's grid. For these grids the dimensions will usually be simple index dimensions, and when they are, the dimension names should be drawn from the list: 'i', 'j', 'k', 'l', 'm', or in the case of the CMIP5 cfSites table the index dimension should be named 'site', and in the case of the so-called "offline" variables in the CMIP5 cf3hr table, the index dimension should be named 'loc'. In some cases when there are two horizontal dimensions, these dimensions might represent the earth according to a well-know map projection involving actual spatial dimensions, in which case the names for these dimensions should be 'x' and 'y'. In all of the above cases a "coordinates" attribute is required and should be attached to the output variable. This attribute will include the string "lat lon", indicating that variables lat and lon contain the latitude and longitude coordinates. Those variables must have the standard names "latitude" and "longitude" respectively and must have the units "degrees_north" and "degrees_east", respectively. In addition (except for the case of 'site' and 'loc')

these coordinate variables must include "bounds" attributes pointing to 'lat_vertices" and "lon_vertices", respectively, which should be stored arrays named lat_vertices(lat,vertices) and lon_vertices(lon,vertices), respectively. The length of the "vertices" dimension should be just long enough to accommodate the maximum number of vertices needed to describe any cell in the grid. If a cell does not have the maximum number of vertices, the remaining values should be set to 1.e20. Finally, it should be noted that in the case of the "offline" variables in the CMIP5 cf3hr table, the longitude and latitude associated with each "location" depend on time, so longitude and latitude will be a function of location and time. See example 6 below where data is stored which is not on a Cartesian latitude-longitude grid .

- In addition for grids other than Cartesian latitude-longitude grids, a file recording all the grid configuration information called for by "gridspec" must be created for each modeling realm (atmosphere, ocean, land surface, etc.).

- For a coordinate representing model atmospheric level, the user must include in the file all the information needed to positively and uniquely indicate the vertical location of the data. Usually in this case the "formula_terms" attribute must be defined, and additional variables or parameters (and their attributes) will need to be stored in the file (e.g., surface pressure and pressure at the top of the model for an atmospheric-sigma coordinate system, along with their units and standard names, and possibly grid description information). The names of the variables and parameters needed to describe each of the known vertical coordinate systems are provided in Appendix 1 below. See example 5 below, and section 4.3.2 of the CF-conventions and Appendix D of the CF-conventions.

- Variables that are stored as a function of "basin", "line", or "type" (which are simple index dimensions) will require auxiliary coordinate variables named "region", "passage", and "type_description", respectively, which are pointed to by a "coordinates" attribute attached to the variable. These will be netCDF type NC_CHAR and will be dimensioned region(basin, strlen), passage(line, strlen), and type_description(type, strlen), respectively, where strlen is the maximum string length. For each of these auxiliary coordinate variables, a long_name attribute and in the case of "region", the standard_name attribute should be stored as specified in the dims table of CMIP5 Requested Output. The values and order of the labels stored in these auxiliary variables should be consistent with the lists found in the "requested" column (S) of the dims table in CMIP5 Requested Output. The values and order of the labels stored in "type_description" are not standardized, but when possible should be selected from the CF conventions "Area Type Table".

- Required attributes for coordinate variables:
  → axis = "X", "Y", "Z", or "T" as and if appropriate (see section 4 of the CF-conventions). For a dimension that represents an axis other than a spatial or temporal dimension, this attribute should be omitted. Also omit this attribute for grids that are not Cartesian latitude-longitude grids unless the grid is logically represented as two-dimensional with these two dimensions being actual coordinates (not simple index dimensions), in which case the "X" and "Y" axis attributes should be assigned as one would like to see a

plotted figure of the field in that two-dimensional space. See example 6 below.

→ bounds = a character string containing the name of the variable where the cell bounds are stored. Whether or not bounds are required for a given coordinate is specified in the "bounds?" column (K) of the dims table in CMIP5 Requested Output. Usually, the bounds name should be the same as the coordinate name, but with the suffix "_bnds" appended. For grids other than those that are Cartesian in latitude and longitude, however, the bounds name should be the same as the coordinate name, but with the suffix "_vertices" appended. For Cartesian latitude-longitude grids, the "bounds" variable should be a function of the coordinate and have a second dimension (varying most quickly) of length 2 in the normal case. For example for a lon coordinate the bounds would be lon_bnds(lon,2).[2] For other grids an example is lon_vertices(lon,vertices).

→ calendar = (but for time coordinates only), one of the options described in section 4.4.1 of the CF-conventions. Note that if a model has a gregorian calendar and the base time is defined to be prior to 1582, then the user should specify 'proleptic_gregorian', rather than 'gregorian' for the calendar attribute. If none of the calendars defined by the CF-conventions applies, then the user may set calendar to a non-standard value. For paleoclimate simulations, calendar might, for example, be given the value "21 kyr B.P." For non-standard calendars include the month_lengths, and, as needed, leap_year, and leap_month attributes.

→ formula_terms = (but for dimensionless vertical coordinates only) a character string, as described in section 4.3.2 of the CF-conventions and as illustrated in example 5 below. Further information can be found in Appendix D of the CF-conventions.

→ positive = (but for vertical coordinates only) "down" or "up", as appropriate. See section 4.3 of the CF-conventions.

→ standard_name = a character string containing the standard name specified in the dims table of CMIP5 Requested Output.

→ units = a character string containing the units specified in the dims table of CMIP5 Requested Output. Note that in the case of a dimensionless vertical coordinate, this attribute may be omitted or set to "1".

Requirements for output variables:

▪ All output fields must be written consistent with the data type (e.g., float or integer) specified in the "type" column (P) of the CMIP5 Requested Output tables. Almost all variables will be written as single precision floating point numbers (netCDF type NC_FLOAT).

---

[2] Note that the ordering of array indices specified here assumes the C language convention (opposite the FORTRAN convention) in which the last coordinate varies most rapidly.

- The variable names (in the netCDF file) must be the names specified in the "output variable name" column (F) of the CMIP5 Requested Output tables (e.g., surface air temperature data will be stored in a variable named "tas").
- The units required for the output fields are given in the "unformatted units" column (I) of the CMIP5 Requested Output tables.
- The positive direction of vertical fluxes must be consistent with that specified in the "positive" column (O) of the CMIP5 Requested Output table and is also indicated by the standard_name (e.g., "surface_upward_latent_heat_flux" and "surface_downward_eastward_stress").
- For each variable, the ordering of the required dimensions is given in the "CMOR dimensions" column (Q) of the CMIP5 Requested Output tables, but note that the *names* of these dimension, as they will appear in the netCDF files, are given in the "output dimension name" column (C) of the "dims" table of CMIP5 Requested Output. Note that the ordering given in column Q of the tables follows the Fortran language convention in which the first coordinate varies most rapidly, whereas in the examples below (at the end), the dimensions are ordered consistent with the C language convention in which the last coordinate varies most rapidly. Note that for grids that cannot be expressed as the intersections of points along constant longitude lines and constant latitude lines (i.e., not Cartesian latitude-longitude), the longitude and latitude dimensions should be replaced with a 1-, 2-, … n-dimensional set of axes, as appropriate, depending on the logical structure of the non-Cartesian grid. Finally note that some of the dimensions appearing in the "CMOR dimensions" column are so-called scalar dimensions and the variable should *not* include this dimension explicitly. Rather, as described below, a "coordinates" attribute should be included and should list this dimension as specified in the CF conventions.
- For Cartesian latitude-longitude grids, the data must be ordered with longitude increasing from west to east, starting with the first grid point greater than or equal to 0 degrees east. All coordinate locations must be unique (e.g., don't include both 0 and 360 degrees east).
- For Cartesian latitude-longitude grids, the data must be ordered with latitude increasing from south to north.
- If there is a vertical coordinate, data must be stored starting with the grid point nearest the surface. Thus, for the atmosphere the lowest layer is stored first, whereas for the ocean and land surface the top layer is stored first.
- If there is a time dimension, data must be stored with time increasing.
- If there is a "basin" coordinate or a "line" coordinate used to distinguish ocean basins, channels, or passages, the data should be stored in the order consistent with the specifications for these coordinates given in the previous section.
- All "missing data"[3] must be assigned the single precision floating point value of 1.e20

---

[3] As an example of "missing data", consider a globally gridded dataset of ocean surface salinity. The data representing land points on this dataset would be missing. Likewise, atmospheric data on portions of pressure surfaces that are underground are usually considered missing.

▪ Required attributes for variables:

→ associated_files = a string listing the base URL for CMIP5 and the location of the model's gridspec file, followed, as appropriate, by the name of the file containing the grid cell areas and/or grid cell volumes. For CMIP5 this string is: "baseURL:http://cmip-pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:<gridspec file name> [areacella:<atmos. cell area file name>] [areacello:<ocean cell area file name] [volcello:<ocean cell volume file name>]", where cell area and cell volume are only sometimes required. Here is an example:

→   associated_files=

→   "baseURL: http://cmip-pcmdi.llnl.gov/CMIP5/dataLocation

→   gridspecFile: gridspec_ocean_fx_IPSL-CM5_historical_r0i0p0.nc

→   areacello: areacello_fx_IPSL-CM5_historical_r0i0p0.nc

→   volcello: volcello_fx_IPSL-CM5_historical_r0i0p0.nc"

→   Note that the "cell_measures" column (U) of the CMIP5 Requested Output tables indicates whether or not areacella, areacello, and/or volcello should be included as associated_files. For each model version and experiment, only one area field is requested by CMIP5 for the atmosphere and only one area and one volume field are requested for the ocean. These cell areas should be defined such that exact global integrals of energy fluxes at the surface and "top of the atmosphere" can be computed. It is understood that for some staggered grids, the meshes for horizontal velocities might be offset from the radiation points, so in these cases exact global integrals of momentum would require areas not requested by CMIP5. In the associated_files attribute, include references to areacella, areacello, and volcello only for variables carried on the same mesh as the areas and volumes (i.e., only when it is appropriate to do so).

→ cell_measures = the string that appears in the "cell measures" column (U) of the CMIP5 Requested Output tables, or it is omitted if an entry in the table is blank. For some staggered grids, the area and/or volume measures called for in the column would not be correct (as discussed in the "associated_files" description above), so in these cases the cell_measures attribute should be omitted.

→ cell_methods = the string that appears in the "cell methods" column (J) of the CMIP5 Requested Output tables, or it is omitted if an entry in the table is blank.

→ coordinates = a character string that must be included for the following types of variables:

– When an output field has a coordinate that is single-valued (e.g., surface wind may have a vertical coordinate value of 10 meters), the "coordinates" attribute must be defined and must include the name of the scalar coordinate (as specified in column C of the "dims" table of CMIP5 Requested Output). The scalar value for the coordinate must be stored in the variable pointed to by that attribute. The dimensions appearing in the "CMOR dimensions" column (Q) of the CMIP5 Requested Output tables that are actually

13

scalar dimensions that need special treatment are identified by the presence of an entry in the "value" column (Q) of the "dims" table (i.e., if a value is found in that column, then the dimension is a scalar dimension). Note that the scalar "dimension" named "type" is a character string assigned the value appearing in column (Q) of the table. Units should not be assigned to these "variables". See section 5.7 of the CF-conventions for more information. Note that the attributes required to be attached to scalar dimensions are the same as the attributes required to be attached to a normal dimension. See example 3 below.

- For fields that are a function of "line", "basin", or "type" the "coordinates" attribute should be defined as "region", "passage", and "type_description", respectively. The coordinate variable found in the coordinates attribute must be defined in the file and be assigned the values listed in the "requested" column (S) of the "dims" table found in the CMIP5 Requested Output. See example 4 below and section 6.1.1 of the CF-conventions for more information.

- For fields stored on grids other than a Cartesian latitude-longitude grid, the "coordinates" attribute should include the string "lon lat", and the longitude and latitude values for each grid cell should be stored in the variables "lon" and "lat", respectively, as called for by section 5.2 of the CF-conventions and as illustrated in example 6 below.

→ _FillValue and missing_value = 1.e20 (but required only for variable that might have "missing values"). This is the value to be assigned to portions of the variable field that are "missing". Note that defining both of these attributes and giving them the same value seems to be redundant, but some software that reads netCDF files expects the missing_value attribute, while other software expects the _FillValue attribute, so including both is required here.

→ flag_values and flag_meanings should be defined (at least for the "Region Selection Index" in the "fx" table), with the values taken from the "flag values" and "flag meanings" columns of the CMIP5 Requested Output tables (columns V and W).

→ grid_mapping = (but only when appropriate) the name of the variable containing the grid_mapping information describing the grid for fields stored on certain grids (*not* Cartesian latitude-longitude grids), as called for in section 5.6 of the CF-conventions and as illustrated in example 6 below. Note that the grid_mapping variable must also be included in the netCDF file along with the all the appropriate attributes associated with it, as described in Appendix F of the CF-conventions.

→ long_name = a character string that might serve as a title for the variable on plots (e.g., "Surface Air Temperature"). It is strongly recommended that this string be copied from the "long name" column (B) of the CMIP5 Requested Output tables.

→ standard_name = a character string containing a CF standard name that can be found in the "standard name" column (G) of the [CMIP5 Requested Output](#) tables. Although usually required, for a few fields the standard_name does not appear in the table, and therefore it should be omitted as an attribute for these variables.

→ units = a character string containing the units given in the "unformatted units" column (I) of the [CMIP5 Requested Output](#) tables, which must be in a form recognizable by [UNIDATA's Udunits package](#) version 2 (e.g., "W m-2".)

- ▪ Optional attributes for variables:

  → comment = a character string providing further information concerning the variable (e.g., if the variable is mrso (soil_moisture_content), the comment might read "includes subsurface water, both frozen and liquid, but not surface water, snow or ice").

  → history = a character string containing an audit trail for modifications to the original data (e.g., indicate what calculations were performed in post-processing, such as interpolating to standard pressure levels or changing the units). Each modification is typically preceded by a "timestamp". Note that this history attribute is variable-specific, whereas the global history attribute defined above (see "Optional attributes" under "Requirements for global attributes") provides information concerning the model simulation itself or refers to processing procedures common to all variables.

  → original_name = the name of the variable as it is commonly known at the user's modeling institute. If the variable being written was computed in some simple way from two or more original fields (e.g., subtracting the upwelling and downwelling fluxes to get a net flux, or multiplying by a scalar or changing the sign), then it is recommended that this be indicated in the "original_name" (e.g., "irup − irdown", where "irup" and "irdown" are the names of the original fields that were subtracted). If more complicated processing is required, this attribute would likely be omitted and the information would more naturally be included in a "history" attribute for this variable, described above.

  → original_units = a character string indicating the original units of the data before it has been processed to be consistent with the units appearing in the [CMIP5 Requested Output](#) tables.

THINGS TO DO: Need to add Appendix 1 and 2 describing more fully the grid_mapping information and dimensionless vertical coordinate information.

Recent Changes to this Document:

8 March 2010: Initial document written, drawing from its predecessor: [Requirements for IPCC Standard Output Contributed to the PCMDI Archive](#).

19 March 2010: Revised version created, including input from Charles Doutriaux.

25 April 2010: 1) Replaced all occurrences of "xline" with "oline". Explicitly noted that chunking/compression/shuffling should be turned "off". 2) Added the following to description of the global attribute "forcing": Use the terms in Appendix 1.2 that are most specific (i.e., avoid "Nat" and "Ant"). If, for example, only CO_2, methane, direct effects of sulfate aerosols, tropospheric and stratospheric ozone, and solar irradiance varied, then specify "GHG, SD, Oz, Sl (GHG includes only CO_2 and methane)". 3) Replaced the examples at the end of this document.

28 April 2010: Modified how the "associated_files" should be specified and made the examples consistent with this revised requirement.

8 November 2010: 1) Modified the gridspec filename to include "realm" and modified the associated_files attribute to correctly point to it. 2) Attempted to clarify what is meant by the "branch_time". 3) Defined a new required global attribute: "parent_experiment_rip". 4) Expanded descriptions of the "physics_version" global attribute so that it can be used to distinguish among different "forcing" combinations in the historicalMisc runs (experiment 7.3). 5) Noted that for "Transpose AMIP" the "project_id" global attribute will be assigned "TAMIP". 6) Noted that for TAMIP, the "realization" global attribute can be used to distinguish among different start times of runs within a single "season". 7) Noted that the basetime should be defined such that all time coordinate values are non-negative. 8) Corrected location pointed to be baseURL defined in the "associated_files" attribute. 9) Described cases when it would be incorrect to include areacella, areacello, and/or volcello in the "associated_files" attribute. 10) Indicated some relatively rare cases where for some fields it would be incorrect to include the "cell_measures" attribute. 11) Moved grid_mapping description (and expanded it) indicating that it is required for certain grids. 12) Corrected minor errors in the examples (below).

4 January 2011: 1) description of the naming convention for gridspec files was simplified. 2) minor corrections made to character type dimension requirements. 3) some links to webpages were updated. 4) shading added in Appendix 1

22 May 2014: 1) changed all occurrences of "typeDescription" to "type_description", to be consistent with CMOR output. 2) corrected date of previous revisions from 2010 to 2011. 3) Corrected sentence in the 5th bulleted item within the "Requirements for Output Variables" section above to read: "Note that the ordering given in column Q of the tables follows the Fortran language convention in which the first coordinate varies most rapidly, whereas in the examples below (at the end), the dimensions are ordered consistent with the C language convention in which the last coordinate varies most rapidly."

Appendix 1
Dimensionless Vertical Coordinate Attributes

Coming Soon???

Example 1: Usual Treatment of a 2-D Field
(surface latent heat flux; a function of longitude, latitude, month)

```
netcdf hfls_Amon_GICCM1_abrupt4xCO2_r1i1p1_198001-198002 {
dimensions:
      time = UNLIMITED ; // (2 currently)
      lat = 3 ;
      lon = 4 ;
      bnds = 2 ;
variables:
      double time(time) ;
            time:bounds = "time_bnds" ;
            time:units = "days since 1980-01-01" ;
            time:calendar = "standard" ;
            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 hfls(time, lat, lon) ;
            hfls:standard_name = "surface_upward_latent_heat_flux" ;
            hfls:long_name = "Surface Upward Latent Heat Flux" ;
            hfls:comment = "comment from CMIP5 table: includes both
                  evaporation and sublimation" ;
            hfls:units = "W m-2" ;
            hfls:original_name = "LATENT" ;
            hfls:history = "2010-04-21T21:05:23Z altered by CMOR:
                  Changed sign. Inverted axis: lat." ;
            hfls:cell_methods = "time: mean" ;
            hfls:cell_measures = "area: areacella" ;
            hfls:associated_files = "baseURL: http://cmip-
                  pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                  gridspec_atmos_fx_GICCM1_abrupt4xCO2_r0i0p0.nc
                  areacella: areacella_fx_GICCM1_abrupt4xCO2_r0i0p0.nc"
                  ;
            hfls:_FillValue = 1.e+20f ;
            hfls:missing_value = 1.e+20f ;

// global attributes:
            :institution = "GICC (Generic International Climate Center,
                  Geneva, Switzerland)" ;
```

```
            :institute_id = "GICC" ;
            :experiment_id = "abrupt4xCO2" ;
            :source = "GICCM1 (2002): atmosphere:  GICAM3
                  (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                  (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                  GILSM2.5" ;
            :model_id = "GICCM1" ;
            :forcing = "GHG (CO2 only)" ;
            :parent_experiment_id = "piControl" ;
            :parent_experiment_rip = "r1i1p1"
            :branch_time = 365. ;
            :contact = "Rusty Koder (koder@middle_earth.net)" ;
            :history = "Output from archive/giccm_03_std_2xCO2_2256.
                  2010-04-21T21:05:23Z CMOR rewrote data to comply with
                  CF standards and CMIP5 requirements." ;
            :references = "Model described by Koder and Tolkien (J.
                  Geophys. Res., 2001, 576-591).  Also see
                  http://www.GICC.su/giccm/doc/index.html  2XCO2
                  simulation described in Dorkey et al. \'(Clim. Dyn.,
                  2003, 323-357.)" ;
            :initialization_method = 1 ;
            :physics_version = 1 ;
            :tracking_id = "c35a58c3-8978-4628-9103-9a7d61002e07" ;
            :product = "output" ;
            :experiment = "abrupt 4XCO2" ;
            :frequency = "mon" ;
            :creation_date = "2010-04-21T21:05:23Z" ;
            :Conventions = "CF-1.4" ;
            :project_id = "CMIP5" ;
            :table_id = "Table Amon (02 April 2010)
                  235242a4410e9b772307262dc755a17b" ;
            :title = "GICCM1 model output prepared for CMIP5 abrupt
                  4XCO2" ;
            :modeling_realm = "atmos" ;
            :realization = 1 ;
data:

 time = 15.5, 45.5 ;

 time_bnds =
  0, 31,
  31, 60 ;

 lat = 10, 20, 30 ;

 lat_bnds =
  5, 15,
  15, 25,
  25, 35 ;

 lon = 0, 90, 180, 270 ;

 lon_bnds =
  -45, 45,
  45, 135,
  135, 225,
  225, 315 ;
```

```
 hfls =
  120, 116, 112, 108,
  104, 100, 96, 92,
  88, 84, 80, 76,
  119, 115, 111, 107,
  103, 99, 95, 91,
  87, 83, 79, 75 ;
}
```

Example 2: Usual Treatment of a 3-D Field on Pressure Levels
(a function of longitude, latitude, pressure, month)
Shading indicates information of particular interest for this type of field.

```
netcdf ta_Amon_GICCM1_abrupt4xCO2_r1i1p1_198001-198002 {
dimensions:
        time = UNLIMITED ; // (2 currently)
        plev = 17 ;
        lat = 3 ;
        lon = 4 ;
        bnds = 2 ;
variables:
        double time(time) ;
                time:bounds = "time_bnds" ;
                time:units = "days since 1980-01-01" ;
                time:calendar = "standard" ;
                time:axis = "T" ;
                time:long_name = "time" ;
                time:standard_name = "time" ;
        double time_bnds(time, bnds) ;
        double plev(plev) ;
                plev:units = "Pa" ;
                plev:axis = "Z" ;
                plev:positive = "down" ;
                plev:long_name = "pressure" ;
                plev:standard_name = "air_pressure" ;
        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 ta(time, plev, lat, lon) ;
                ta:standard_name = "air_temperature" ;
                ta:long_name = "Air Temperature" ;
                ta:units = "K" ;
                ta:original_name = "T" ;
                ta:history = "2010-04-21T21:05:23Z altered by CMOR: Inverted
                        axis: lat." ;
                ta:comment = "no future" ;
                ta:cell_methods = "time: mean" ;
                ta:cell_measures = "area: areacella" ;
                ta:associated_files = "baseURL: http://cmip-
                        pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                        gridspec_atmos_fx_GICCM1_abrupt4xCO2_r0i0p0.nc
```

```
                areacella: areacella_fx_GICCM1_abrupt4xCO2_r0i0p0.nc"
                    ;
            ta:_FillValue = 1.e+20f ;
            ta:missing_value = 1.e+20f ;


// global attributes:
            :institution = "GICC (Generic International Climate Center,
                Geneva, Switzerland)" ;
            :institute_id = "GICC" ;
            :experiment_id = "abrupt4xCO2" ;
            :source = "GICCM1 (2002): atmosphere:  GICAM3
                (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                GILSM2.5" ;
            :model_id = "GICCM1" ;
            :forcing = "GHG (CO2 only)" ;
            :parent_experiment_id = "piControl" ;
            :parent_experiment_rip = "r1i1p1"
            :branch_time = 365. ;
            :contact = "Rusty Koder (koder@middle_earth.net)" ;
            :history = "Output from archive/giccm_03_std_2xCO2_2256.
                2010-04-21T21:05:23Z CMOR rewrote data to comply with
                CF standards and CMIP5 requirements." ;
            :references = "Model described by Koder and Tolkien (J.
                Geophys. Res., 2001, 576-591).  Also see
                http://www.GICC.su/giccm/doc/index.html  2XCO2
                simulation described in Dorkey et al. \'(Clim. Dyn.,
                2003, 323-357.)" ;
            :initialization_method = 1 ;
            :physics_version = 1 ;
            :tracking_id = "93b039b7-77d1-41db-9ce9-95bcb33c27f0" ;
            :product = "output" ;
            :experiment = "abrupt 4XCO2" ;
            :frequency = "mon" ;
            :creation_date = "2010-04-21T21:05:23Z" ;
            :Conventions = "CF-1.4" ;
            :project_id = "CMIP5" ;
            :table_id = "Table Amon (02 April 2010)
                235242a4410e9b772307262dc755a17b" ;
            :title = "GICCM1 model output prepared for CMIP5 abrupt
                4XCO2" ;
            :modeling_realm = "atmos" ;
            :realization = 1 ;
data:

 time = 15.5, 45.5 ;

 time_bnds =
  0, 31,
  31, 60 ;

 plev = 100000, 92500, 85000, 70000, 60000, 50000, 40000, 30000, 25000,
    20000, 15000, 10000, 7000, 5000, 3000, 2000, 1000 ;

 lat = 10, 20, 30 ;
```

```
lat_bnds =
 5, 15,
 15, 25,
 25, 35 ;

lon = 0, 90, 180, 270 ;

lon_bnds =
 -45, 45,
 45, 135,
 135, 225,
 225, 315 ;

ta =
 288.5, 290.5, 292.5, 1.e20,
 280.5, 282.5, 284.5, 286.5,
 272.5, 274.5, 276.5, 278.5
          .
          .
          .

 166, 168, 170, 172,
 158, 160, 162, 164,
 150, 152, 154, 156,
 164, 166, 168, 170 ;
}
```

Example 3: Treatment of a Scalar Dimension
(near-surface air temperature)
Shading indicates information of particular interest for this type of field.

```
netcdf tas_Amon_GICCM1_abrupt4xCO2_r1i1p1_198001-198002 {
dimensions:
      time = UNLIMITED ; // (2 currently)
      lat = 3 ;
      lon = 4 ;
      bnds = 2 ;
variables:
      double time(time) ;
            time:bounds = "time_bnds" ;
            time:units = "days since 1980-01-01" ;
            time:calendar = "standard" ;
            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) ;
      double height ;
            height:units = "m" ;
            height:axis = "Z" ;
            height:positive = "up" ;
            height:long_name = "height" ;
            height:standard_name = "height" ;
      float tas(time, lat, lon) ;
            tas:standard_name = "air_temperature" ;
            tas:long_name = "Near-Surface Air Temperature" ;
            tas:comment = "comment from CMIP5 table: near-surface
                  (usually, 2 meter) air temperature." ;
            tas:units = "K" ;
            tas:original_name = "TS" ;
            tas:history = "2010-04-21T21:05:23Z altered by CMOR: Treated
                  scalar dimension: \'height\'.  Inverted axis: lat." ;
            tas:cell_methods = "time: mean" ;
            tas:cell_measures = "area: areacella" ;
            tas:coordinates = "height" ;
            tas:associated_files = "baseURL: http://cmip-
                  pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                  gridspec_atmos_fx_GICCM1_abrupt4xCO2_r0i0p0.nc
```

```
                    areacella: areacella_fx_GICCM1_abrupt4xCO2_r0i0p0.nc"
                        ;

// global attributes:
                :institution = "GICC (Generic International Climate Center,
                    Geneva, Switzerland)" ;
                :institute_id = "GICC" ;
                :experiment_id = "abrupt4xCO2" ;
                :source = "GICCM1 (2002): atmosphere:  GICAM3
                    (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                    (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                    GILSM2.5" ;
                :model_id = "GICCM1" ;
                :forcing = "GHG (CO2 only)" ;
                :parent_experiment_id = "piControl" ;
                :parent_experiment_rip = "r1i1p1"
                :branch_time = 365. ;
                :contact = "Rusty Koder (koder@middle_earth.net)" ;
                :history = "Output from archive/giccm_03_std_2xCO2_2256.
                    2010-04-21T21:05:23Z CMOR rewrote data to comply with
                    CF standards and CMIP5 requirements." ;
                :references = "Model described by Koder and Tolkien (J.
                    Geophys. Res., 2001, 576-591).  Also see
                    http://www.GICC.su/giccm/doc/index.html  2XCO2
                    simulation described in Dorkey et al. \'(Clim. Dyn.,
                    2003, 323-357.)" ;
                :initialization_method = 1 ;
                :physics_version = 1 ;
                :tracking_id = "fec0cec2-debf-4adf-b69c-382b0c0e6a20" ;
                :product = "output" ;
                :experiment = "abrupt 4XCO2" ;
                :frequency = "mon" ;
                :creation_date = "2010-04-21T21:05:23Z" ;
                :Conventions = "CF-1.4" ;
                :project_id = "CMIP5" ;
                :table_id = "Table Amon (02 April 2010)" ;
                :title = "GICCM1 model output prepared for CMIP5 abrupt
                    4XCO2" ;
                :modeling_realm = "atmos" ;
                :realization = 1 ;
data:

 time = 15.5, 45.5 ;

 time_bnds =
  0, 31,
  31, 60 ;

 lat = 10, 20, 30 ;

 lat_bnds =
  5, 15,
  15, 25,
  25, 35 ;

 lon = 0, 90, 180, 270 ;
```

```
 lon_bnds =
  -45, 45,
  45, 135,
  135, 225,
  225, 315 ;

 height = 2 ;

 tas =
  230, 238, 246, 254,
  262, 270, 278, 286,
  294, 302, 310, 318,
  232, 240, 248, 256,
  264, 272, 280, 288,
  296, 304, 312, 320 ;
}
```

Example 4: Treatment of Auxiliary Coordinates
(northward ocean heat transport; a function of latitude, ocean basin, month)
Shading indicates information of particular interest for this type of field.

```
netcdf htovgyre_Omon_GICCM1_abrupt4xCO2_r1i1p1_198001-198002 {
dimensions:
        time = UNLIMITED ; // (2 currently)
        basin = 3 ;
        lat = 3 ;
        bnds = 2 ;
        strlen = 21 ;
variables:
        double time(time) ;
                time:bounds = "time_bnds" ;
                time:units = "days since 1980-01-01" ;
                time:calendar = "standard" ;
                time:axis = "T" ;
                time:long_name = "time" ;
                time:standard_name = "time" ;
        double time_bnds(time, bnds) ;
        char region(basin, strlen) ;
                region:long_name = "ocean basin" ;
                region:standard_name = "region" ;
        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) ;
        float htovgyre(time, basin, lat) ;
                htovgyre:standard_name =
                    "northward_ocean_heat_transport_due_to_gyre" ;
                htovgyre:long_name = "Northward Ocean Heat Transport due to
                    Gyre" ;
                htovgyre:comment = "comment from CMIP5 table: function of Y,
                    basin" ;
                htovgyre:units = "W" ;
                htovgyre:original_name = "LATENT" ;
                htovgyre:history = "2010-04-21T21:05:23Z altered by CMOR:
                    Converted type from \'d\' to \'f\'. Inverted axis:
                    lat." ;
                htovgyre:cell_methods = "time: mean longitude: mean" ;
                htovgyre:associated_files = "baseURL: http://cmip-
                    pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                    gridspec_ocean_fx_GICCM1_abrupt4xCO2_r0i0p0.nc" ;
                htovgyre:coordinates = "region" ;

// global attributes:
                :institution = "GICC (Generic International Climate Center,
                    Geneva, Switzerland)" ;
                :institute_id = "GICC" ;
                :experiment_id = "abrupt4xCO2" ;
```

27

```
            :source = "GICCM1 (2002): atmosphere:  GICAM3
                    (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                    (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                    GILSM2.5" ;
            :model_id = "GICCM1" ;
            :forcing = "GHG (CO2 only)" ;
            :parent_experiment_id = "piControl" ;
            :parent_experiment_rip = "r1i1p1"
            :branch_time = 365. ;
            :contact = "Rusty Koder (koder@middle_earth.net)" ;
            :history = "Output from archive/giccm_03_std_2xCO2_2256.
                    2010-04-21T21:05:23Z CMOR rewrote data to comply with
                    CF standards and CMIP5 requirements." ;
            :references = "Model described by Koder and Tolkien (J.
                    Geophys. Res., 2001, 576-591).  Also see
                    http://www.GICC.su/giccm/doc/index.html  2XCO2
                    simulation described in Dorkey et al. \'(Clim. Dyn.,
                    2003, 323-357.)" ;
            :initialization_method = 1 ;
            :physics_version = 1 ;
            :tracking_id = "e8408167-9bc7-40dd-8ddb-05f3dda8a97c" ;
            :product = "output" ;
            :experiment = "abrupt 4XCO2" ;
            :frequency = "mon" ;
            :creation_date = "2010-04-21T21:05:23Z" ;
            :Conventions = "CF-1.4" ;
            :project_id = "CMIP5" ;
            :table_id = "Table Omon (02 April 2010)" ;
            :title = "GICCM1 model output prepared for CMIP5 abrupt
                    4XCO2" ;
            :modeling_realm = "ocean" ;
            :realization = 1 ;
data:

 time = 15.5, 45.5 ;

 time_bnds =
  0, 31,
  31, 60 ;

 region =
  "atlantic_arctic_ocean",
  "indian_pacific_ocean",
  "global_ocean" ;

 lat = 10, 20, 30 ;

 lat_bnds =
  5, 15,
  15, 25,
  25, 35 ;

 htovgyre =
  -80, -84, -88,
  -100, -104, -76,
  -120, -92, -96,
  -79, -83, -87,
```

```
  -99, -103, -75,
  -107, -111, -115 ;
}
```

Example 5: Treatment of a 3-D Field on Model Levels
(cloud fraction; a function of longitude, latitude, model level, month)
Shading indicates information of particular interest for this type of field.

```
netcdf cl_Amon_GICCM1_abrupt4xCO2_r1i1p1_198001-198002 {
dimensions:
     time = UNLIMITED ; // (2 currently)
     lev = 5 ;
     lat = 3 ;
     lon = 4 ;
     bnds = 2 ;
variables:
     double time(time) ;
             time:bounds = "time_bnds" ;
             time:units = "days since 1980-01-01" ;
             time:calendar = "standard" ;
             time:axis = "T" ;
             time:long_name = "time" ;
             time:standard_name = "time" ;
     double time_bnds(time, bnds) ;
     double lev(lev) ;
             lev:bounds = "lev_bnds" ;
             lev:units = "1" ;
             lev:axis = "Z" ;
             lev:positive = "down" ;
             lev:long_name = "hybrid sigma pressure coordinate" ;
             lev:standard_name =
                     "atmosphere_hybrid_sigma_pressure_coordinate" ;
             lev:formula = "p = a*p0 + b*ps" ;
             lev:formula_terms = "p0: p0 a: a b: b ps: ps" ;
     double lev_bnds(lev, bnds) ;
             lev_bnds:formula = "p = a*p0 + b*ps" ;
             lev_bnds:standard_name =
                     "atmosphere_hybrid_sigma_pressure_coordinate" ;
             lev_bnds:units = "1" ;
             lev_bnds:formula_terms = "p0: p0 a: a_bnds b: b_bnds ps: ps"
                     ;
     float p0 ;
             p0:long_name = "vertical coordinate formula term: reference
                     pressure" ;
             p0:units = "Pa" ;
     double a(lev) ;
             a:long_name = "vertical coordinate formula term: a(k)" ;
     double b(lev) ;
             b:long_name = "vertical coordinate formula term: b(k)" ;
     float ps(time, lat, lon) ;
             ps:standard_name = "surface_air_pressure" ;
             ps:long_name = "Surface Air Pressure" ;
             ps:comment = "not, in general, the same as mean sea-level
                     pressure" ;
             ps:units = "Pa" ;
             ps:original_units = "hPa" ;
```

```
        ps:history = "2010-04-21T21:05:23Z altered by CMOR:
                Converted units from \'hPa\' to \'Pa\'." ;
        ps:cell_methods = "time: mean" ;
    double a_bnds(lev, bnds) ;
        a_bnds:long_name = "vertical coordinate formula term:
                a(k+1/2)" ;
    double b_bnds(lev, bnds) ;
        b_bnds:long_name = "vertical coordinate formula term:
                b(k+1/2)" ;
    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 cl(time, lev, lat, lon) ;
        cl:standard_name = "cloud_area_fraction_in_atmosphere_layer"
                ;
        cl:long_name = "Cloud Area Fraction" ;
        cl:comment = "comment from CMIP5 table: Report on model
                layers (not standard pressures).  Include both large-
                scale and convective cloud." ;
        cl:units = "%" ;
        cl:original_name = "CLOUD" ;
        cl:history = "2010-04-21T21:05:23Z altered by CMOR: Inverted
                axis: lev. Inverted axis: lat." ;
        cl:cell_methods = "time: mean" ;
        cl:cell_measures = "area: areacella" ;
        cl:associated_files = "baseURL: http://cmip-
                pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                gridspec_atmos_fx_GICCM1_abrupt4xCO2_r0i0p0.nc
                areacella: areacella_fx_GICCM1_abrupt4xCO2_r0i0p0.nc"
                ;
// global attributes:
        :institution = "GICC (Generic International Climate Center,
                Geneva, Switzerland)" ;
        :institute_id = "GICC" ;
        :experiment_id = "abrupt4xCO2" ;
        :source = "GICCM1 (2002): atmosphere:  GICAM3
                (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                GILSM2.5" ;
        :model_id = "GICCM1" ;
        :forcing = "GHG (CO2 only)" ;
        :parent_experiment_id = "piControl" ;
        :parent_experiment_rip = "r1i1p1"
        :branch_time = 365. ;
        :contact = "Rusty Koder (koder@middle_earth.net)" ;
```

```
            :history = "Output from archive/giccm_03_std_2xCO2_2256.
                2010-04-21T21:05:23Z CMOR rewrote data to comply with
                CF standards and CMIP5 requirements." ;
            :references = "Model described by Koder and Tolkien (J.
                Geophys. Res., 2001, 576-591).  Also see
                http://www.GICC.su/giccm/doc/index.html  2XCO2
                simulation described in Dorkey et al. \'(Clim. Dyn.,
                2003, 323-357.)" ;
            :initialization_method = 1 ;
            :physics_version = 1 ;
            :tracking_id = "438535ef-7d52-451f-a665-595a284da7b4" ;
            :product = "output" ;
            :experiment = "abrupt 4XCO2" ;
            :frequency = "mon" ;
            :creation_date = "2010-04-21T21:05:23Z" ;
            :Conventions = "CF-1.4" ;
            :project_id = "CMIP5" ;
            :table_id = "Table Amon (02 April 2010)" ;
            :title = "GICCM1 model output prepared for CMIP5 abrupt
                4XCO2" ;
            :modeling_realm = "atmos" ;
            :realization = 1 ;
data:

 time = 15.5, 45.5 ;

 time_bnds =
  0, 31,
  31, 60 ;

 lev = 0.92, 0.72, 0.5, 0.3, 0.1 ;

 lev_bnds =
  1, 0.83,
  0.83, 0.61,
  0.61, 0.4,
  0.4, 0.2,
  0.2, 0 ;

 p0 = 100000 ;

 a = 0.12, 0.22, 0.3, 0.2, 0.1 ;

 b = 0.8, 0.5, 0.2, 0.1, 0 ;

 ps =
  97000, 97400, 97800, 98200,
  98600, 99000, 99400, 99800,
  100200, 100600, 101000, 101400,
  97100, 97500, 97900, 98300,
  98700, 99100, 99500, 99900,
  100300, 100700, 101100, 101500 ;

 a_bnds =
  0.06, 0.18,
  0.18, 0.26,
  0.26, 0.25,
```

```
  0.25, 0.15,
  0.15, 0 ;

 b_bnds =
  0.94, 0.65,
0.65, 0.35,
0.35, 0.15,
0.15, 0.05,
0.05, 0 ;

 lat = 10, 20, 30 ;

 lat_bnds =
  5, 15,
  15, 25,
  25, 35 ;

 lon = 0, 90, 180, 270 ;

 lon_bnds =
  -45, 45,
  45, 135,
  135, 225,
  225, 315 ;

 cl =
  72.8, 73.2, 73.6, 74,
  71.6, 72, 72.4, 72.4,
  70.4, 70.8, 70.8, 71.2,
  67.6, 69.2, 69.6, 70,
  66, 66.4, 66.8, 67.2,
  64.8, 65.2, 65.6, 66,
  63.6, 64, 64.4, 64.4,
  60.8, 61.2, 62.8, 63.2,
  59.6, 59.6, 60, 60.4,
  58, 58.4, 58.8, 59.2,
  56.8, 57.2, 57.6, 58,
  54, 54.4, 54.8, 56.4,
  52.8, 53.2, 53.2, 53.6,
  51.6, 51.6, 52, 52.4,
  50, 50.4, 50.8, 51.2,
  72.9, 73.3, 73.7, 74.1,
  71.7, 72.1, 72.5, 72.5,
  70.5, 70.9, 70.9, 71.3,
  67.7, 69.3, 69.7, 70.1,
  66.1, 66.5, 66.9, 67.3,
  64.9, 65.3, 65.7, 66.1,
  63.7, 64.1, 64.5, 64.5,
  60.9, 61.3, 62.9, 63.3,
  59.7, 59.7, 60.1, 60.5,
  58.1, 58.5, 58.9, 59.3,
  56.9, 57.3, 57.7, 58.1,
  54.1, 54.5, 54.9, 56.5,
  52.9, 53.3, 53.3, 53.7,
  51.7, 51.7, 52.1, 52.5,
  50.1, 50.5, 50.9, 51.3 ;
}
```

Example 6: Treatment of Data *Not* on a Cartesian Latitude-Longitude Grid
(surface latent heat flux; a function of x, y, month)
Shading indicates information of particular interest for this type of field.

```
netcdf hfls_Amon_GICCM1_amip_r1i1p1_198001-198002 {
dimensions:
        time = UNLIMITED ; // (2 currently)
        y = 4 ;
        x = 3 ;
        bnds = 2 ;
        vertices = 4 ;
variables:
        double time(time) ;
                time:bounds = "time_bnds" ;
                time:units = "days since 1980-01-01" ;
                time:calendar = "standard" ;
                time:axis = "T" ;
                time:long_name = "time" ;
                time:standard_name = "time" ;
        double time_bnds(time, bnds) ;
        double y(y) ;
                y:units = "m" ;
                y:axis = "Y" ;
                y:long_name = "y coordinate of projection" ;
                y:standard_name = "projection_y_coordinate" ;
        double x(x) ;
                x:units = "m" ;
                x:axis = "X" ;
                x:long_name = "x coordinate of projection" ;
                x:standard_name = "projection_x_coordinate" ;
        int lambert_conformal_conic ;
                lambert_conformal_conic:grid_mapping_name =
                    "lambert_conformal_conic" ;
                lambert_conformal_conic:standard_parallel = -20. ;
                lambert_conformal_conic:longitude_of_central_meridian = 175.
                    ;
                lambert_conformal_conic:latitude_of_projection_origin = 13.
                    ;
                lambert_conformal_conic:false_easting = 8. ;
                lambert_conformal_conic:false_northing = 0. ;
        float lat(y, x) ;
                lat:standard_name = "latitude" ;
                lat:long_name = "latitude coordinate" ;
                lat:units = "degrees_north" ;
                lat:bounds = "lat_vertices" ;
        float lon(y, x) ;
                lon:standard_name = "longitude" ;
                lon:long_name = "longitude coordinate" ;
                lon:units = "degrees_east" ;
                lon:bounds = "lon_vertices" ;
        float lat_vertices(y, x, vertices) ;
                lat_vertices:units = "degrees_north" ;
        float lon_vertices(y, x, vertices) ;
```

```
            lon_vertices:units = "degrees_east" ;
    float hfls(time, y, x) ;
            hfls:standard_name = "surface_upward_latent_heat_flux" ;
            hfls:long_name = "Surface Upward Latent Heat Flux" ;
            hfls:comment = "comment from CMIP5 table: includes both
                evaporation and sublimation" ;
            hfls:units = "W m-2" ;
            hfls:original_name = "HFLS" ;
            hfls:history = "2010-04-21T21:05:24Z altered by CMOR:
                Changed sign. Converted type from \'d\' to \'f\'." ;
            hfls:cell_methods = "time: mean" ;
            hfls:associated_files = "baseURL: http://cmip-
                pcmdi.llnl.gov/CMIP5/dataLocation gridspecFile:
                gridspec_atmos_fx_GICCM1_abrupt4xCO2_r0i0p0.nc
                areacella: areacella_fx_GICCM1_abrupt4xCO2_r0i0p0.nc"
                ;
            hfls:grid_mapping = "lambert_conformal_conic" ;
            hfls:coordinates = "lat lon" ;

// global attributes:
            :institution = "GICC (Generic International Climate Center,
                Geneva, Switzerland)" ;
            :institute_id = "GICC" ;
            :experiment_id = "amip" ;
            :source = "GICCM1 (2002): atmosphere:  GICAM3
                (gicam_0_brnchT_itea_2, T63L32); ocean: MOM
                (mom3_ver_3.5.2, 2x3L15); sea ice: GISIM4; land:
                GILSM2.5" ;
            :model_id = "GICCM1" ;
            :forcing = "GHG, SD, BC, Sl, Vl (GHG includes only CO2 and
                methane)" ;
            :parent_experiment_id = "N/A" ;
            :parent_experiment_rip = "N/A"
            :branch_time = 365. ;
            :contact = "Rusty Koder (koder@middle_earth.net)" ;
            :history = "Output from archive/giccm_03_std_2xCO2_2256.
                2010-04-21T21:05:24Z CMOR rewrote data to comply with
                CF standards and CMIP5 requirements." ;
            :references = "Model described by Koder and Tolkien (J.
                Geophys. Res., 2001, 576-591).  Also see
                http://www.GICC.su/giccm/doc/index.html  2XCO2
                simulation described in Dorkey et al. \'(Clim. Dyn.,
                2003, 323-357.)" ;
            :initialization_method = 1 ;
            :physics_version = 1 ;
            :tracking_id = "e3c0d67a-efa8-4466-9eae-8d69ae5ad538" ;
            :product = "output" ;
            :experiment = "AMIP" ;
            :frequency = "mon" ;
            :creation_date = "2010-04-21T21:05:24Z" ;
            :Conventions = "CF-1.4" ;
            :project_id = "CMIP5" ;
            :table_id = "Table Amon (02 April 2010)" ;
            :title = "GICCM1 model output prepared for CMIP5 AMIP" ;
            :modeling_realm = "atmos" ;
            :realization = 1 ;
data:
```

```
time = 15.5, 45.5 ;

time_bnds =
 0, 31,
 31, 60 ;

y = 0, 1, 2, 3 ;

x = 0, 1, 2 ;

lambert_conformal_conic = _ ;

lat =
 10, 0, -10,
 20, 10, 0,
 30, 20, 10,
 40, 30, 20 ;

lon =
 290, 300, 310,
 300, 310, 320,
 310, 320, 330,
 320, 330, 340 ;

lat_vertices =
 10, 0, 10, 20,
 0, -10, 0, 10,
 -10, -20, -10, 0,
 20, 10, 20, 30,
 10, 0, 10, 20,
 0, -10, 0, 10,
 30, 20, 30, 40,
 20, 10, 20, 30,
 10, 0, 10, 20,
 40, 30, 40, 50,
 30, 20, 30, 40,
 20, 10, 20, 30 ;

lon_vertices =
 280, 290, 300, 290,
 290, 300, 310, 300,
 300, 310, 320, 310,
 290, 300, 310, 300,
 300, 310, 320, 310,
 310, 320, 330, 320,
 300, 310, 320, 310,
 310, 320, 330, 320,
 320, 330, 340, 330,
 310, 320, 330, 320,
 320, 330, 340, 330,
 330, 340, 350, 340 ;

hfls =
 72, 68, 64,
 88, 84, 80,
 104, 100, 96,
```

```
120, 116, 112,
71, 67, 63,
87, 83, 79,
103, 99, 95,
119, 115, 111 ;
```