

CORDEX-CMIP6 Archiving Specifications for Dynamical Downscaling

26th May 2026

DOI: [10.5281/zenodo.20274448](https://doi.org/10.5281/zenodo.20274448)

Grigory Nikulin¹, Lars Bunttemeyer², Jesús Fernández³, Seth McGinnis⁴ and Jason P. Evans⁵

1. Swedish Meteorological and Hydrological Institute (SMHI), Sweden
2. Climate Service Center Germany (GERICS), Germany
3. Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Spain
4. National Center for Atmospheric Research (NCAR), USA
5. Climate Change Research Centre, University of New South Wales (UNSW), Australia

This document provides Data Reference Syntax (DRS) elements necessary for post-processing CORDEX-CMIP6 simulations and publishing them on the Earth System Grid Federation (ESGF). The document includes file and directory naming conventions, global attributes and ESGF Search Facets Mapping. Known issues are collected at <https://github.com/WCRP-CORDEX/archive-specifications/issues>. Newer versions of this document may exist. Find the latest version under <https://doi.org/10.5281/zenodo.10961068>.

1. DRS elements

The DRS element values must consist of the characters a-z, A-Z, 0-9 and '-' (dash). No other character is allowed. The terms in brackets following the DRS element names in the list below indicate whether the values must be taken from a controlled vocabulary ('CV'), i.e. a fixed list of values, must be registered with CORDEX ('CV to register'), or must follow a predefined structure ('structured form'). Note that most elements must have the same value as the mandatory NetCDF global attribute.

`variable_id` (CV) is the short name of the variable. The name is taken from the [CORDEX-CMIP6 Variable List](#) or [CORDEX-CMIP6 CMOR tables](#).

`domain_id` (CV) is the name assigned to each of the CORDEX regions and includes a flag for resolution as listed in the [CORDEX-CMIP6 domain id CV](#).

`driving_source_id` (CV) is an identifier of the driving data. The name consists of a model identifier. For reanalysis driven runs this is the name of the reanalysis data (ERA5). For runs driven by CMIP6 model data this is the associated CMIP6 `source_id`, which can be found in the [CORDEX-CMIP6 driving source id CV](#).

`driving_experiment_id` (CV) is either "evaluation" for the ERA5-driven experiment, the value of the CMIP6 `experiment_id` from the ScenarioMIP activity, or "historical" for the historical experiment from CMIP. The values for `driving_experiment_id` can be found in the [CORDEX-CMIP6 driving experiment id CV](#).

`driving_variant_label` (CV) identifies the ensemble member of the CMIP6 simulation that produced the forcing data. It must have the same value as the `variant_label` DRS element of the driving CMIP6 simulation (ripf, see [CMIP6 DRS Appendix 1: Global Attributes for Labeling Experiments](#)). For the evaluation experiment driven by ERA5 it is "r1i1p1f1". For all invariant fields (frequency = fx, e.g. orog or sftlf),

`driving_variant_label` must be the same as `variant_label` of the driving CMIP6 simulation (e.g. r1i1p1f1, r2i1p1f1, etc.); r0i0p0f0 is not allowed.

`institution_id` (CV to register) is an identifier for the institution that is responsible for generating and providing CORDEX simulations. All CORDEX institutions must be registered to publish their simulations on ESGF. See instructions on [how to register an institution](#) and the current state of the [CORDEX-CMIP6 institution id CV](#).

`source_id` (CV to register) is an identifier (acronym) of the CORDEX RCM. All CORDEX RCMs must be registered to publish their simulations on ESGF. See instructions on [how to register a model](#) and the current state of the [CORDEX-CMIP6 source id CV](#). Different configurations of the same RCM such as different combinations of parameterization schemes or changes in parameters for existing schemes must be reflected in `source_id` by a free text suffix (e.g. RCM123, RCM123A for Africa or RCM123T for the tropics). RCM simulations with spectral nudging must use the "SN" suffix in `source_id` (e.g. RCM123-SN) while RCM simulations with Newtonian/dynamical nudging must use the "NN" suffix (e.g. RCM123-NN).

`version_realization` (structured form) is a combination that identifies i) versions of CORDEX datasets (simulations) related to technical, configuration, or postprocessing errors and ii) realizations with different initial conditions for RCMs. This DRS element has the form "vN-rM". "N" in the version part "vN" is 1 for the first release of dataset (v1) and sequential numbers (2, 3, 4, etc.) for any rerun or re-processing of the dataset (v2, v3, v4, etc.). The later version always supersedes the earlier version. "M" in the realization part "rM" is sequential numbers (1, 2, 3 etc.) that reflect multiple RCM simulations with perturbed initial conditions (r1, r2, r3, etc.) driven by the same GCM and the same GCM member. The version and realization parts are separated by a dash "-" (e.g. v1-r1, v1-r2, v1-r3). The version part of this DRS element should not be confused with the ESGF-related DRS element version that has the form "vYYYYMMDD" and is only included in the ESGF directory structure (see [4. ESGF Directory Structure](#)).

`frequency` (CV) is the output frequency indicator: 1hr - 1 hourly, 3hr - 3 hourly, 6hr - 6 hourly, day - daily, mon - monthly, and fx - invariant fields; see the [CORDEX-CMIP6 frequency CV](#).

`StartTime` and `EndTime` (structured form) indicate the time span of the file content. The format is YYYY[MM[DD[hhmm]]], i.e. the year is represented by 4 digits, while the month, day, hour, and minutes are represented by exactly 2 digits, if they are present at all (monthly output - YYYYMM, daily - YYYYMMDD, sub-daily - YYYYMMDDhhmm). The `StartTime` and `EndTime` of sub-daily instantaneous and average data are based on the time values of the first and last record in the file. The two dates are separated by a dash. All time stamps refer to UTC. Constant fields (`frequency=fx`) do not have the `StartTime - EndTime` element in their file names.

`activity_id` (CV) - an identifier of different CORDEX activities such as dynamical downscaling (DD) and empirical-statistical downscaling (ESD), see the [CORDEX-CMIP6 activity id CV](#). Regarding the activity described in this document (dynamical downscaling on standard CORDEX domains at continental scale; CORDEX-Domain), "DD" is the only option. However, other values are possible for other activities under the CORDEX-CMIP6 project, such as Flagship Pilot Studies (FPS), which will provide their own archiving specifications.

`activity_id` can also be a space-separated list of activities if a simulation can belong to several of them. This is useful for faceted searches. In this case, the first entry in the list will be the only value used in the DRS path (see [Section 4](#)).

`project_id` (CV) - project identifier ("CORDEX-CMIP6" is the only option)

2. Global attributes

Table 1: CORDEX-CMIP6 global attribute description (see examples in section [13.6](#)) and comparison with CORDEX-CMIP5

| CORDEX-CMIP6 global attribute | description | examples | corresponding attribute in CORDEX-CMIP5 | form | when required? |
|-------------------------------|--|--|---|-----------|----------------|
| <code>activity_id</code> | an identifier of different CORDEX activities as dynamical downscaling (DD) or empirical- statistical downscaling (ESD). Used in faceted searches, part of DRS. | "DD" is the only option, (see reference CV) | - | CV | always |
| <code>comment</code> | additional information about the data or methods used to produce the simulation | - | comment | free form | never |

| | | | | | |
|-----------------------|--|--|---------------|-----------------|--------|
| contact | contact information of the institution that is responsible for CORDEX simulations (avoid personal contact information) | - | contact | free form | always |
| Conventions | Climate and Forecast (CF) metadata conventions version | "CF-1.11" is the only option. | Conventions | CV | always |
| creation_date | date when the file was created in format YYYY-MM-DDTHH:MM:SSZ | "2023-01-15T14:30:23Z" | creation_date | structured form | always |
| domain | name of the CORDEX region | "Africa", "South-East Asia" (see reference CV) | - | CV | always |
| domain_id | an identifier assigned to each CORDEX region including a flag for resolution. Used in faceted searches, part of DRS. | "AFR-25", "SEA-25" (see reference CV) | CORDEX_domain | CV | always |
| driving_experiment | short description of the reanalysis-driven experiment or the driving CMIP6 experiment | "reanalysis simulation of the recent past", "all-forcing simulation of the recent past", "gap-filling scenario reaching 7.0 based on SSP3" (see reference CV) | experiment | CV | always |
| driving_experiment_id | an identifier of the reanalysis-driven experiment or the driving CMIP6 experiment. Used in faceted searches, part of DRS. | "evaluation", "historical", "ssp370" (see reference CV) | experiment_id | CV | always |

| | | | | | |
|------------------------|--|---|---|-----------|--------|
| driving_institution_id | an identifier of the institution that is responsible for the driving reanalysis or CMIP6 simulation | "ECMWF" (see reference CV) | part of driving_model_id | CV | always |
| driving_source_id | reanalysis or CMIP6 model identifier Used in faceted searches, part of DRS. | "ERA5" (reference CV) | part of driving_model_id | CV | always |
| driving_variant_label | variant_label of the CMIP6 simulation or driving reanalysis. Used in faceted searches, part of DRS. | "r1i1p1f1", "r2i1p1f1" | driving_model_ensemble_member (ensemble_member in CMIP5) | | always |
| frequency | sampling frequency Used in faceted searches, part of DRS. | day, mon, 6hr, 3hr, 1hr, fx (see reference CV) | frequency | CV | always |
| grid | provides information about the horizontal grid and regridding procedure | see note 1 | - | free form | always |
| history | a timestamped trail for modifications to the original data, as suggested by the CF conventions | - | history | free form | never |

| | | | | | |
|----------------|---|--|------------------|----------------|--------|
| institution_id | an identifier of the institution that is responsible for CORDEX simulations Used in faceted searches, part of DRS. | (see reference CV) | institute_id | CV to register | always |
| license | provides information about the license | " https://cordex.org/data-access/cordex-cmip6-data/cordex-cmip6-terms-of-use " is the only option | - | CV | always |
| mip_era | determines what cycle of CMIP defines experiment and data specifications | "CMIP6" is the only option. | - | CV | always |
| product | product type Used in faceted searches. | "model-output" is the only option | product | CV | always |
| project_id | project identifier Used in faceted searches, part of DRS. | "CORDEX-CMIP6" is the only option. | project_id | CV | always |
| references | published or web-based references that describe the data, model or methods used | - | references | free form | never |
| source | full model name/version, matching the "label_extended" registered in the CV | see "label_extended" in reference CV | - | CV to register | always |
| source_id | model identifier (acronym) Used in faceted searches, part of DRS. | see "source_id" in reference CV | part of model_id | CV to register | always |

| | | | | | |
|--------------------------|--|--|----------------|-----------------|-----------------------|
| source_type | model configuration Used in faceted searches | "ARCM", "AORCM", "AGCM", "AOGCM" (see reference CV) | - | CV | always |
| tracking_id | unique file identifier | see note 3 | tracking_id | structured form | always |
| variable_id | variable identifier Used in faceted searches, part of DRS. | "tas", "pr" (see CORDEX-CMIP6 CMOR Tables) | - | CV | always |
| version_realization | identifies versions of CORDEX datasets and RCM realizations Used in faceted searches, part of DRS. | "v1-r1", "v2-r1", "v1-r2" | rcm_version_id | structured form | always |
| version_realization_info | description of version_realization if it is not "v1-r1" | see note 2 | - | free form | never but recommended |

Table 1 notes:

1. The "grid" global attribute can be used to describe the horizontal grid and regridding procedure. There is no standard form used to record this information, but it is suggested that when appropriate the following be indicated: brief description of native grid and resolution, and if data have been regridded, regridding procedure and description of target grid (see note 10 in [CMIP6 DRS](#)). Here are some examples:

```

grid = "Lambert conic conformal with 25 km grid spacing"
grid = "Rotated-pole latitude-longitude with 0.22 degree grid spacing"
grid = "Rotated-pole latitude-longitude with 0.11 degree grid spacing, interpolated by 2nd order conservative remapping from the original unstructured icosahedral ICON grid R13B05 (~12.1 km)"
grid = "NEMO ORCA tripolar grid with a 1/12 degree (6-8km) grid spacing (no grid_mapping); Mediterranean Sea only"

```

2. The `version_realization_info` global attribute provides information on how new reruns (e.g. v2, v3, etc.) and/or realizations (e.g. r2, r3, etc.) are generated; recommended if the `version_realization` is not v1-r1.

- `tracking_id` must be of the form `<PID>/<uuid>` where PID is a Persistent Identifier (`hdl:21.14103` for CORDEX-CMIP6) and `uuid` is a Universally Unique Identifier e.g. `hdl:21.14103/187fcd6c-7cc6-11ee-9481-7824afb1963b`. The `tracking_id` should be unique for each CORDEX-CMIP6 file published in ESGF. The `<uuid>` should be generated using the OSSP utility which supports a number of different DCE 1.1 variant UUID options. For CORDEX-CMIP6, version 4 (random number based) is required. Download the software from [OSSP uuid](#). (see note 15 in [CMIP6 DRS](#)).

3. File naming

`file_name=`

```
<variable_id>_<domain_id>_<driving_source_id>_<driving_experiment_id>_<driving_variant_label>_<institution_id>_<source_id>_<version_realization>_<frequency>[_<StartTime>-<EndTime>].nc
```

Examples:

```
tas_AFR-25_ERA5_evaluation_r1i1p1f1_INST_RCM123_v1-r1_mon_201101-202012.nc  
tas_AFR-25_GCM_historical_r1i1p1f1_INST_RCM123_v1-r1_mon_201101-201412.nc  
tas_AFR-25_GCM_ssp370_r1i1p1f1_INST_RCM123_v1-r1_mon_201501-202012.nc  
orog_AFR-25_GCM_ssp370_r1i1p1f1_INST_RCM123_v1-r1_fx.nc
```

In contrast to CORDEX-CMIP5:

- the institution that is responsible for CORDEX simulations (`institution_id`) and model acronym (`source_id`) are 2 different DRS elements, i.e. separated by the underscore "_" in the file name
- the institution that is responsible for the driving CMIP6 simulation (`driving_institution_id`) is not a part of DRS and not included in the file name and ESGF directory structure.

4. ESGF directory structure

`directory_structure=`

```
<project_id>/<activity_id>/<domain_id>/<institution_id>/<driving_source_id>/<driving_experiment_id>/<driving_variant_label>/<source_id>/<version_realization>/<frequency>/<variable_id>/<version>/
```

The `version` DRS element indicates an approximate date of model output files or publication on ESGF and has the form "vYYYYMMDD" (e.g., "v20231206"). This is the only DRS element that is not stored as a global attribute. Note that files contained in a single `version` subdirectory at the end of the directory path should represent all the available time-samples reported from the simulation; a time-series can be split across several files, but all the files must be found in the same subdirectory. This implies that `version` will not generally be the actual date that all files in the subdirectory were written or published (see also Directory structure template in [CMIP6 DRS](#)).

Examples:

```
/CORDEX-CMIP6/DD/AFR-25/INST/ERA5/evaluation/r1i1p1f1/RCM123/v1-r1/mon/tas/v20240319  
/CORDEX-CMIP6/DD/AFR-25/INST/GCM/historical/r1i1p1f1/RCM123/v1-r1/mon/tas/v20240319  
/CORDEX-CMIP6/DD/AFR-25/INST/GCM/ssp370/r1i1p1f1/RCM123/v1-r1/mon/tas/v20240319  
/CORDEX-CMIP6/DD/AFR-25/INST/GCM/ssp370/r1i1p1f1/RCM123/v1-r1/fx/orog/v20240319
```

5. File format

Data files must be in NetCDF format, version 4, using the NetCDF 4 classic data model. It is recommended that data should be compressed by using "deflate level" 1 and with "shuffle" turned on. Data files must conform to the [CF Conventions 1.11](#).

Each file may contain only one output field (target variable) from a single simulation. It must include a target variable with attributes defined in the [CORDEX-CMIP6 CMOR tables](#) and coordinate variables. The entire time series of a target variable is to be distributed over several files as described in section [8 Time period for each data file](#).

All output fields must be single precision (type NC_FLOAT), while all coordinate variables (time and space) must be double precision (type NC_DOUBLE). All missing data must be assigned the single precision floating point value of 1.e20.

File compliance to these archive specifications can and should be checked by means of the [WCRP Compliance Checker Plugins](#) `wcrp_cordex_cmip6` and `cf:1.11`. Data files rising high severity errors in these checks will not be published on ESGF.

It is strongly recommend to "repack" all files as a last postprocessing step, in order to have a cloud-optimized internal structure. The [ncrepack-cordex](#) tool is provided to rearrange the internal metadata and chunking pattern. CORDEX data not properly repacked will limit the ability of ESGF data servers to provide remote access to them.

6. CORDEX domains and horizontal coordinates

The CORDEX domains are defined in the [CORDEX domain tables](#). A domain must lie fully inside the RCM interior computational domain, i.e. in the area left once the relaxation zone is excluded. It is strongly recommended that RCMs using the rotated-pole coordinate system exactly follow the CORDEX grid definition provided. The rotated-pole coordinate system is always defined in terms of rotation of the North Pole in accordance with [CF-1.11](#). All variables from a model component must be provided on the same grid (i.e., variables on staggered grids must be regridded to the standard grid). Zonal and meridional winds must be provided as real north- and eastward winds if the RCM uses a coordinate system/projection that does not coincide with real north- and eastward directions (e.g. the rotated pole, Lambert Conformal, etc.).

The domain acronym (`domain_id`) is 'domain name'-'resolution', where 'resolution' is the nearest grid spacing in km of the 3 resolutions used in CORDEX-CMIP5 and CORDEX-CMIP6 (50, 25 and 12 km). Changing 'resolution' from degrees, which are related only to the rotated coordinate system (CORDEX-CMIP5), to the more common

kilometres allows us to unify the terminology used in CORDEX, making it easily understandable by all users. For example, "AFR-25" means the CORDEX-Africa domain with 25 km resolution in a projected coordinate system and 0.22° resolution in the rotated pole coordinate system. The resolution flag indicates the resolution of the atmospheric component of CORDEX models. All variables from a simulation must be provided with the same `domain_id`.

For each model native grid, there is also a corresponding regular latitude-longitude grid. These grids have roughly the same resolution as the native grids used by the RCMs (50 km ↔ 1/2°, 25 km ↔ 1/4°, 12 km ↔ 1/8°) with grid cell boundaries (not centers) on integer degrees of latitude and longitude. The domain acronyms for the regular grids are the same as those for the corresponding model native grid with the letter 'i' appended to the resolution (e.g. "AFR-25i" is the 0.25° lat-lon grid for the CORDEX-Africa domain). Unlike model native grids, regular i-grids are always matching across models and model components. We recommend but do not require providing the most important variables for impacts on the regular grid in addition to the native grid.

Data must be provided for the CORDEX domain only, i.e. the relaxation zones must be removed before the data is delivered. Names of the CORDEX domains are provided in [CORDEX-CMIP6 domain id CV](#).

Data files for models with native projected coordinate systems must contain a full description:

- the 1-dimensional coordinate variables (e.g. rlon and rlat for the rotated pole coordinate system or x and y for the Lambert Conformal Conic (LCC) projection),
- grid mapping variable `crs`¹ describing the coordinate reference system and
- the variable attribute - `grid_mapping = "crs"`

in accordance with [CF-1.11](#) (see examples [13.1](#) and [13.2](#)). The grid mapping variable `crs` is of arbitrary type (e.g. char or int) since it contains no data. The shape and size of the Earth used for the model grid must be specified. For a spherical earth this is done via the `crs` attribute `earth_radius`. If a model grid specifies an ellipsoid for the shape of the earth then see [CF-1.11 Appendix F](#).

For model components with non-projected coordinate systems, the `grid_mapping` variable and attribute can be skipped, and a text description of the grid must be provided in the `grid` attribute (see examples in Table 1, [note 1](#) above). The specific text "(no grid_mapping)" should be added to the `grid` description, as in the example above, in order to let automatic quality checkers know about this exception. The 1-dimensional coordinate variables may also be skipped in these non-projected coordinate systems.

The 2-dimensional geographic latitudes and longitudes of the model grid cells (lon and lat) must be also provided as auxiliary coordinates. Longitude coordinates must be strictly monotonically increasing, except for domains that include a pole or cross both the 0° and 180° meridians (e.g. ANT, ARC). Longitude coordinates must also be confined to the range -180 to 360 and have absolute values as small as possible given the other two constraints (e.g., store 170 E to 170 W as 170 to 190, but store 150 W to 130 W as -150 to -130, not 210 to 230).

For models with native unstructured grids, it is up to the regional CORDEX communities to decide whether data must be remapped to one of the regular lat-lon domain grids (e.g., AFR-25i) or to the most common native RCM grid used for a specific CORDEX domain.

7. Time coordinate

The units of the time coordinate is `days since 1950-01-01`² for all files. The earlier reference date `days since 1850-01-01` is also allowed if a RCM group downscales a longer period that includes the pre-1950 era. All time dependent variables must have an attribute `cell_methods` specifying the time aggregation, as provided in the [CORDEX-CMIP6 CMOR tables](#).

The time value of the instantaneous data is [0Z, 6Z, 12Z, 18Z], [0Z, 3Z, 6Z, 9Z, 12Z, 15Z, 18Z, 21Z], and [0Z, 1Z, 2Z, 3Z, ..., 20Z, 21Z, 22Z, 23Z] of each day for the 6-, 3-, and 1-hourly data respectively.

Variables that are representative of an interval (averages, maxima, minima, sum) must use the midpoints of the time intervals (bounds) as time coordinate values. Therefore, these variables have the time values 0.5Z, 1.5Z, 2.5Z, ..., 23.5Z (1-hourly); 1.5Z, 4.5Z, 7.5Z, ... 22.5Z (3-hourly); 3Z, 9Z, 15Z, 21Z (6-hourly); 12Z (daily); Jan 16 12Z, Feb 15 0Z; (monthly, non leap years); Jan 16 12Z, Feb 12Z (monthly, leap years); Jan 16 0Z, Feb 16 0Z (monthly, 360 day); etc.

Furthermore, interval variables must have a `time_bnds` field of dimensions `(ntimes, 2)`, where `ntimes` is the dimension of the time coordinate (see an example in section 13.4). Intervals for daily and monthly data should start and end at 00:00:00 UTC of the appropriate day. Intervals for sub-daily data should start and end at 00:00:00 UTC or an integer multiple of the frequency (1, 3, or 6 hours) from that point.

The time variable must have a calendar attribute. Use of the proleptic-gregorian or standard calendar is strongly recommended when possible. Other calendars (`360_day` and `365_day`) inherited from the driving models are also allowed. If the driving model uses the gregorian calendar (which is deprecated since CF-1.9), use the standard calendar.

8. Time period for each data file

The time spans to be included in a single file depend on the aggregation, which is 1-hourly, 6-hourly, daily, monthly, or invariant:

- 1-hourly or 6-hourly: one year,
- daily: 5 years, or less if some of the years in the 5-year period are unavailable,
- monthly: 10 years, or less if some of the years in the 10-year period are unavailable,
- invariant: single file.

Files should always contain full years if the data are available, but are allowed to be shorter if it is not.

Files with monthly data start with years that end with '1' or the first year of the experiment; they end with '0' or the last year of the experiment. Daily data files start with years that end with '1' or '6' or the first experiment year; the last year they contain ends with '5' or '0' or is the last experiment year. For example, the ERA5-driven evaluation experiment for 1979-2021 with 1979 as a spin-up:

| monthly | daily | subdaily |
|-----------|-----------|-----------|
| 1980-1980 | 1980-1980 | 1980-1980 |
| 1981-1990 | 1981-1985 | 1981-1981 |
| 1991-2000 | 1986-1990 | 1982-1982 |
| 2001-2010 | ... | ... |
| 2011-2020 | 2016-2020 | 2020-2020 |
| 2021-2021 | 2021-2021 | 2021-2021 |

9. License

All CORDEX modeling groups choose a license for their CMIP6-driven simulations depending on institutional and/or funding agency policies. This information is necessary to register a RCM in the [CORDEX-CMIP6 source id CV](#). It is strongly recommended to use the Creative Commons Attribution 4.0 International (CC BY 4.0) license, as is currently used in [CMIP6](#). Note, that any kind of "non-commercial" license will significantly limit the use of the data in downstream climate mitigation and adaptation applications.

The global attribute license has the only option "<https://cordex.org/data-access/cordex-cmip6-data/cordex-cmip6-terms-of-use>" leading to the CORDEX-CMIP6 Terms of Use page.

10. Registration

All institutions (modelling groups) that contribute or plan to contribute to CORDEX-CMIP6 must

1. register their institution and model following the instructions on [how to register institutions and models](#) and
2. provide information about their planned simulations in the [CORDEX-CMIP6 downscaling plan](#).

The modelling groups will not be able to publish their CORDEX-CMIP6 simulations on ESGF without first registering their institution and model.

11. ESGF Search Facets Mapping

| ESGF Search Facet | CORDEX-CMIP6 DRS element or global attribute | Values |
|-------------------|--|--|
| Project | project_id | "CORDEX-CMIP6" is the only option (CV) |
| Activity | activity_id | "DD" is the only option (CV) |
| Product | product | "model-output" is the only option (CV) |

| | | |
|---------------------|-----------------------|---|
| Domain ID | domain_id | AFR-25, EUR-12, AFR-25i, etc. (CV) |
| Driving Source ID | driving_source_id | ERA5 or CMIP6 models (CV) |
| Source ID | source_id | CORDEX models (CV) |
| Institution ID | institution_id | CORDEX institutions (CV) |
| Source Type | source_type | ARCM, AORCM, AGCM, AOGCM (CV) |
| Experiment ID | driving_experiment_id | evaluation, historical, ssp370, etc. (CV) |
| Variant Label | driving_variant_label | r1i1p1f1, r2i1p1f1, etc. (structured form) |
| Version-Realization | version_realization | v1-r1, v2-r1, v1-r2, etc. (structured form) |
| Frequency | frequency | mon, day, 6hr, 1hr, fx (CV) |
| Variable | variable_id | tas, pr, tasmax, etc. (CMOR tables) |

12. User support

In case of any questions or doubts please create an issue in <https://github.com/WCRP-CORDEX/archive-specifications>.

Acknowledgments

We thank all who provided their valuable comments and suggestions on the CORDEX-CMIP6 archiving specifications for dynamical downscaling.

13. Examples

13.1 Rotated Pole Coordinate System

```
char crs ;
  crs:grid_mapping_name = "rotated_latitude_longitude" ;
  crs:grid_north_pole_latitude = 39.25 ;
  crs:grid_north_pole_longitude = -162. ;
  crs:earth_radius = 6371229. ;

double rlon(rlon) ;
  rlon:standard_name = "grid_longitude" ;
  rlon:long_name = "longitude in rotated pole grid" ;
  rlon:units = "degrees" ;

double rlat(rlat) ;
  rlat:standard_name = "grid_latitude" ;
  rlat:long_name = "latitude in rotated pole grid" ;
  rlat:units = "degrees" ;

double lon(rlat, rlon) ;
  lon:standard_name = "longitude" ;
  lon:long_name = "longitude" ;
  lon:units = "degrees_east" ;

double lat(rlat, rlon) ;
  lat:standard_name = "latitude" ;
  lat:long_name = "latitude" ;
  lat:units = "degrees_north" ;

float pr(time, rlat, rlon) ;
  pr:standard_name = "precipitation_flux" ;
  pr:long_name = "Precipitation" ;
  pr:units = "kg m-2 s-1" ;
  pr:coordinates = "lon lat" ;
  pr:_FillValue = 1.e+20f ;
  pr:missing_value = 1.e+20f ;
  pr:cell_methods = "time: mean" ;
  pr:grid_mapping = "crs" ;
```

13.2 Lambert Conformal Conic projection

```
int crs ;
  crs:grid_mapping_name = "lambert_conformal_conic" ;
  crs:standard_parallel = 49.5 ;
  crs:longitude_of_central_meridian = 10.5 ;
  crs:latitude_of_projection_origin = 49.5 ;
  crs:false_easting = 2925000. ;
  crs:false_northing = 2925000. ;
  crs:earth_radius = 6371229. ;

double x(x) ;
  x:standard_name = "projection_x_coordinate" ;
  x:long_name = "x coordinate of projection" ;
  x:units = "m" ;

double y(y) ;
  y:standard_name = "projection_y_coordinate" ;
  y:long_name = "y coordinate of projection" ;
  y:units = "m" ;

double lon(y, x) ;
  lon:standard_name = "longitude" ;
  lon:long_name = "longitude" ;
  lon:units = "degrees_east" ;

double lat(y, x) ;
  lat:standard_name = "latitude" ;
  lat:long_name = "latitude" ;
  lat:units = "degrees_north" ;

float pr(time, y, x) ;
  pr:standard_name = "precipitation_flux" ;
  pr:long_name = "Precipitation" ;
  pr:units = "kg m-2 s-1" ;
  pr:coordinates = "lon lat" ;
  pr:_FillValue = 1.e+20f ;
  pr:missing_value = 1.e+20f ;
  pr:cell_methods = "time: mean" ;
  pr:grid_mapping = "crs" ;
```

13.3 Vertical coordinate depth

a surface model with 4 layers: layer 1: 0-7cm, layer 2: 7-28cm, layer 3: 28-100cm, layer 4: 100-289cm

```
dimensions:  
  bnds = 2 ;  
  depth = 4 ;  
  
variables:  
  
  double depth(depth) ;  
    depth:standard_name = "depth" ;  
    depth:long_name = "depth" ;  
    depth:units = "m" ;  
    depth:bounds = "depth_bnds" ;  
    depth:positive = "down" ;  
  
  double depth_bnds(depth, bnds) ;  
  
data:  
  
  depth = 0.035, 0.175, 0.64, 1.945 ;  
  
  depth_bnds =  
    0, 0.7,  
    0.7, 0.28,  
    0.28, 1,  
    1, 2.89 ;
```

13.4 Time coordinate

for a daily variable representative of an interval (e.g. pr, tasmax)

```
dimensions:  
  
  time = UNLIMITED ;  
  bnds = 2 ;  
  
variables:  
  
  double time(time);  
    time:long_name = "time";  
    time:units = "days since 1950-01-01";  
    time:calendar = "standard"  
    time:bounds = "time_bnds";  
  
  double time_bnds(time, bnds) ;  
  
data:  
  
  time = 0.5, 1.5, 2.5, ... ;  
  
  time_bnds =  
    0, 1,  
    1, 2,  
    2, 3,  
    ... ;
```

13.5 Scalar coordinate variables

for variables on individual levels (e.g. at 2 meters, or at 850 hPa), do not include an extra dimension of size 1. Use a [CF scalar coordinate variable](#) instead.

```
variables:

    double height ;
        height:units = "m" ;
        height:axis = "Z" ;
        height:positive = "up" ;
        height:long_name = "height" ;
        height:standard_name = "height" ;

    float tas(time, rlat, rlon) ;
        [...]
        tas:coordinates = "height lat lon" ;
data:

    height = 2. ;
```

13.6 Global attributes

```
// global attributes:
:Conventions = "CF-1.11" ;
:activity_id = "DD" ;
:comment = "optional" ;
:contact = "cordex-data@iircm.org" ;
:creation_date = "2023-11-19T18:01:15Z" ;
:domain = "Africa" ;
:domain_id = "AFR-25" ;
:driving_experiment = "reanalysis simulation of the recent past" ;
:driving_experiment_id = "evaluation" ;
:driving_institution_id = "ECMWF" ;
:driving_source_id = "ERA5" ;
:driving_variant_label = "r1i1p1f1" ;
:frequency = "mon" ;
:grid = "Lambert conic conformal with 25 km grid spacing"
:institution = "Interdisciplinary Institute of Regional Climate Modeling" ;
:institution_id = "IIRCM" ;
:license = "https://cordex.org/data-access/cordex-cmip6-data/cordex-cmip6-terms-of-use" ;
:mip_era = "CMIP6" ;
:product = "model-output" ;
:project_id = "CORDEX-CMIP6" ;
:source = "Interdisciplinary Regional Climate Model version 1" ;
:source_id = "InterRCM1" ;
:source_type = "ARCM" ;
:tracking_id = "hdl:21.14103/187fcd6c-7cc6-11ee-9481-7824afb1963b"
:variable_id = "tas" ;
:version_realization = "v1-r1" ;
```

14. Version history

| Version | Date | Comment |
|---------|------------|--|
| v3 | 2026-05-26 | <ul style="list-style-type: none"> • First version official in mkdocs at https://wcrp-cordex.github.io/archive-specifications/CORDEX-CMIP6_archiving_specifications_DD. • Update user support target repository. • Fix inconsistency in <code>driving_experiment</code> attribute for the evaluation experiment (#26). • Fix case in <code>long_name</code> of LCC projected coordinates (#27). • Skip mandatory nature of <code>grid_mapping</code> in non-projected grids (#29). • Add quality checker info and repacking recommendation (#38). • Allow for different grids for different model components of the same simulation (#40). • Allow for multiple <code>activity_id</code>s (#45). • Some smaller updates and rewording(#30 #49). |
| v2 | 2025-03-21 | <ul style="list-style-type: none"> • Change of <code>project_id</code> (CORDEX to CORDEX-CMIP6) due to ESGF publication requirements (#22) • <code>mip_era</code> is excluded from the ESGF directory structure and search facets mapping, must be presented as the global attribute only (#24) • Fixed example for <code>creation_date</code> global attribute (#19) • New example (13.5) illustrating scalar coordinate variables for height (#18) • Specify time coordinate origin, but not specific formatting (#5) • Allow grid mapping variable to match the <code>grid_mapping_name</code> (#17) |
| v1 | 2024-04-10 | Initial release on April, 10th, 2024. |

1. The grid mapping variable is also allowed to be called after the CF standard `grid_mapping_name` used. This is the default hardcoded behaviour in the CMOR library as of version 3.9.0 (2024-08-28). Of course, the main variable corresponding `grid_mapping` attribute should then match this grid mapping variable name. ←
2. The format of the reference date can be either "days since 1950-01-01" (preferred), "days since 1950-01-01 00:00:00", or "days since 1950-01-01T00:00:00Z" ←