EU-CIRCLE

A pan-European framework for strengthening Critical Infrastructure resilience to climate change

D2.4 EU-CIRCLE climate hazards metadata and standards

Contractual Delivery Date: 02/2018	**Actual Delivery Date**: 07/2018
Type: OTHER	**Version**: V1.0

Dissemination Level Public Deliverable

Statement

This deliverable summarizes the metadata overview and specific EU-CIRCLE activities that contributed to the capabilities of CIRP to properly read and internally describe climate and climate hazard data. Deliverable D2.4 gives summary of the implemented frameworks related to (1) file-naming conventions and recommendations, (2) local variable and global file metadata conventions/schemes/standards, (3) input file formats and (4) relationships between common EU-CIRCLE data and metadata.

⚠ **EU-CIRCLE** is a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653824. Please see http://www.eu-circle.eu/ for more information.

Preparation Slip			
	Name	Partner	Date
From	I. Güttler	DHMZ	18/07/2018
Reviewer	N. Karatarakis	HNMS	17/07/2018
Reviewer			
For delivery	**Thanasis Sfetsos**	**NCSRD**	**19/07/2018**

Document Log			
Issue	Date	Comment	Author / Organization
0.1	08/02/2018	Table of contents (TOC).	Ivan Güttler (DHMZ)
0.2	01/03/2018	Metadata for G-FMIS_FWI, Return period, Probability of exceedance, Scenario Selection, G.FMIS_FIRESIM, FEPS, HYSPLIT Deposition and HYSPLIT Height.	Nadia Politi and NCSRD team
0.3	03/04/2018	Preparing the first version of the Executive summary.	Ivan Güttler (DHMZ)
0.4	05/04/2018	Extending the content of the Primary Climate Data section. Extending the content of the Secondary Climate Data section.	Ivan Güttler (DHMZ)
0.5	03/06/2018	Deliverable extensions based on the telco discussions in May 2018 between DHMZ, SATWAYS, ADIT and NCSRD: (1) adding information about CMIP5 variables, CDO indices. (2) including figures summarizing literature review and metadata vs. data relationship.	Ivan Güttler (DHMZ)
1.0	18/07/2018	Results of the internal review by HNMS included. Comments replied by NCSRD included.	Ivan Güttler (DHMZ)

Executive Summary

EU-CIRCLE DoA specifies the scope of the **Task 2.4** *EU-CIRCLE extended climate hazards standard* as follows:

> “Within Task 2.4, the systematic annotation of climate related metadata will be accomplished. This will greatly facilitate the processing and sharing of data files and related information. The work within EU-CIRCLE will be an extension of existing efforts and best practices already existing in the scientific community and includes (but is not limited to) Open Geospatial Consortium (OGC), CF (Climate and Forecast) metadata conventions, WMO-TD No. 1186, outcome of the ongoing FP7 CHARMe project, OASIS etc. Its extension will be mainly to facilitate the output of secondary climate hazards, which currently exhibit a persistent gap. The work will identify and assess existing related standards and data protocols in the literature that are suitable for the purposes of EU-CIRCLE and expand those annotations in a consistent manner to cover all sources of climate related hazards."

The results of the Task 2.4 are presented in this report as the **D2.4** *EU-CIRCLE climate hazards metadata and standards.* **D2.4** is final deliverable in series of EU-CIRCLE **WP2** *Climatic Data Capture and Processing* deliverables:

- **D2.1** *Report on Typology of Climate Related Hazards*
- **D2.2** *Report on Climate related hazards information collection mechanisms*
- **D2.3** *Tools for processing Climate Hazards Information*,

and builds upon the joint work of the project consortia during the project.

The aim of this report is to present in a consistent framework, the structure and content of metadata covering results of climate observations and modelling, secondary climate products and results of the impact models. The main recommendation in this report is to present the extension of climate modelling metadata approach as done by the CF (Climate and Forecast) metadata conventions and applied in recent CMIP5 (Coupled Model Intercomparison Project Phase 5; https://cmip.llnl.gov/cmip5/) and CORDEX (Coordinated Regional Climate Downscaling Experiment; http://cordex.org/) programmes. This is the EU-CIRCLE contribution to the standardization of the various metadata across wide range of currently fragmented practices and is based on the data and metadata needs encountered during the EU-CIRCLE case studies.

Finally, the work on Task 2.4 presented in this report is complemented by the inclusion of the file-naming and variable-naming structure into the CIRP platform, and the extension of the CIRP set of the metadata entries related to the climate data and climate hazards input fields and time-series. In summary, we suggest the use of the standardized formats (e.g. NetCDF and XML), and standardized set of file-naming and metadata entries (such as CF, CMIP5 and CORDEX) for the purpose of more systematic linkage of different steps when exploring the impact of weather and climate events on the critical infrastructure.

Contents

1 Introduction

This deliverable, D2.4, *EU-CIRCLE climate hazards metadata and standards,* presents approaches that were done in Task 2.4 *EU-CIRCLE extended climate hazards* from the WP2 *Climatic Data Capture and Processing* of the EU-CIRCLE project. The main results of D2.4 are common approaches in how to define metadata for the different types of climate data that are used in EU-CIRCLE project and in the CIRP. In general, metadata give information about the data, such as who made the data, what are the data dimensions, references to algorithms etc. Examples of relevant metadata attributes are given in this deliverable.

Introductory information to metadata vs. data relationship is presented in the three figures in this section:

Fig. 1.1 demonstrates three types of the input information that governed the creation of D2.4:

1. Literature and state-of-the-art review was made based on the available reports and common practices. For example, common convention for metadata in data produced by the climate models is the so called the CF convention, and is widely accepted in the climate community.
2. We have examined needs and activities in the EU-CIRCLE case studies relevant for the metadata of the climate data. For example, when working with the data from the climate models, EU-CIRCLE experts required details about the specific models, scenarios and units of several different climate variables.
3. Several telco and in person discussions have been done in order to acquire relevant information about the use of metadata and other technical aspects in different subdisciplines. For example, while NetCDF format is common way to share climate model data, some other modelling communities share their spatial data in XML or SHP formats.

The central part of Fig. 1.1 stresses out three EU-CIRCLE WPs, i.e. WP2, WP3 and WP4, where all three produce in the final step a deliverable related to the metadata in climate data, risk modelling and resilience framework respectively. These deliverables feed the WP8 and D8.11 *Final report on collected and harmonized data and metadata* and give inputs to the CIRP developers in terms of the expected data formats and metadata formats, and how to include them efficiently in the CIRP workflows.

Fig. 1.2 overviews eight major initiatives and programmes that consider various types of the metadata at the different level. The main goal of all these efforts is to have community accepted standards of practice that allow for the safe, no-error and transparent use of the underlying data. This includes OGC, OASIS and ISO at the highest level of generalization, then moves to CF, CORDEX and WMO-TD No. 1186 at the next level where community relevant (in this case meteorology and climatology) recommendations are given and ends with more specific activities in FP7 project CHARMe and "WMO Core Metadata" profile. Other ways of organizing this type of information are also possible. In addition, at this stage of the CIRP development, some levels and types of metadata considerations are more or less relevant. The amount of information on this subject is vast, and readers are invited to consult on the original source of data by following given links in Fig. 1.2.

Finally, we end this Introduction by approaching the subject of climate data and its metadata from the perspective of general CIRP user. During the scope of the EU-CIRCLE and discussion related to Task 2.4, we have recognized four main elements to consider, all presented in Fig 1.3: (1) *In which format data have to be in order to CIRP to read and process them in next steps?*, (2) *How should one name the finals in order to ease* *the file reading and processing?*, (3) *In which format should the metadata about input files be organized?*, (4) *What conventions or schemas or standards should users follow when preparing the metadata?* An example of possible answers to these four questions is also given in Fig. 1.1 in the Appendix 1 of this deliverable.

In summary, for different communities (e.g. meteorology & climatology, hydrology, fire protection, air quality, GIS), different approaches to metadata representation exist. From the perspective of the data & metadata user, one has interest to acquire all available information. However, from the perspective of the data & metadata producer, usually a balance must be made between the needs and capabilities when organizing data & metadata. This is why community driven efforts for the common conventions, schemas and standards help to optimize workflows of both data & metadata users and producers. We continue this deliverable by demonstrating how CIRP and EU-CIRCLE consortia approached to metadata related to (1) primary climate data, (2) secondary climate data, and (3) data produced by the impact models related to different climate hazards, such as floods and fires that impact state and operations of the critical infrastructure.

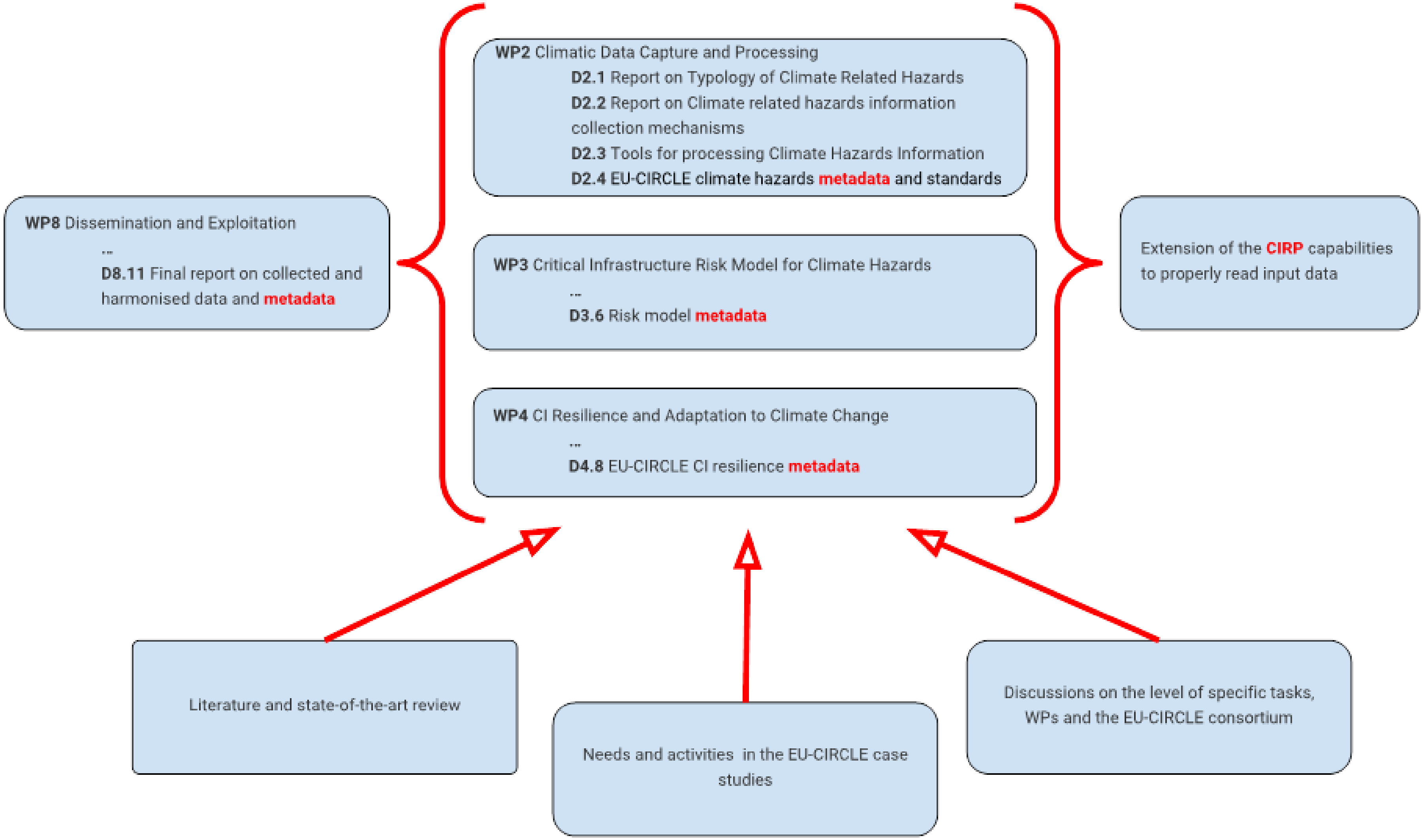

Figure 1.1 Relationship between deliverables from the EU-CIRCLE work packages WP2, WP3 and WP4. Deliverables D2.4, D3.6 and D4.8 give input to D8.11 related to metadata. Focus of this document is to document concepts implemented in CIRP, and related to the climate metadata.

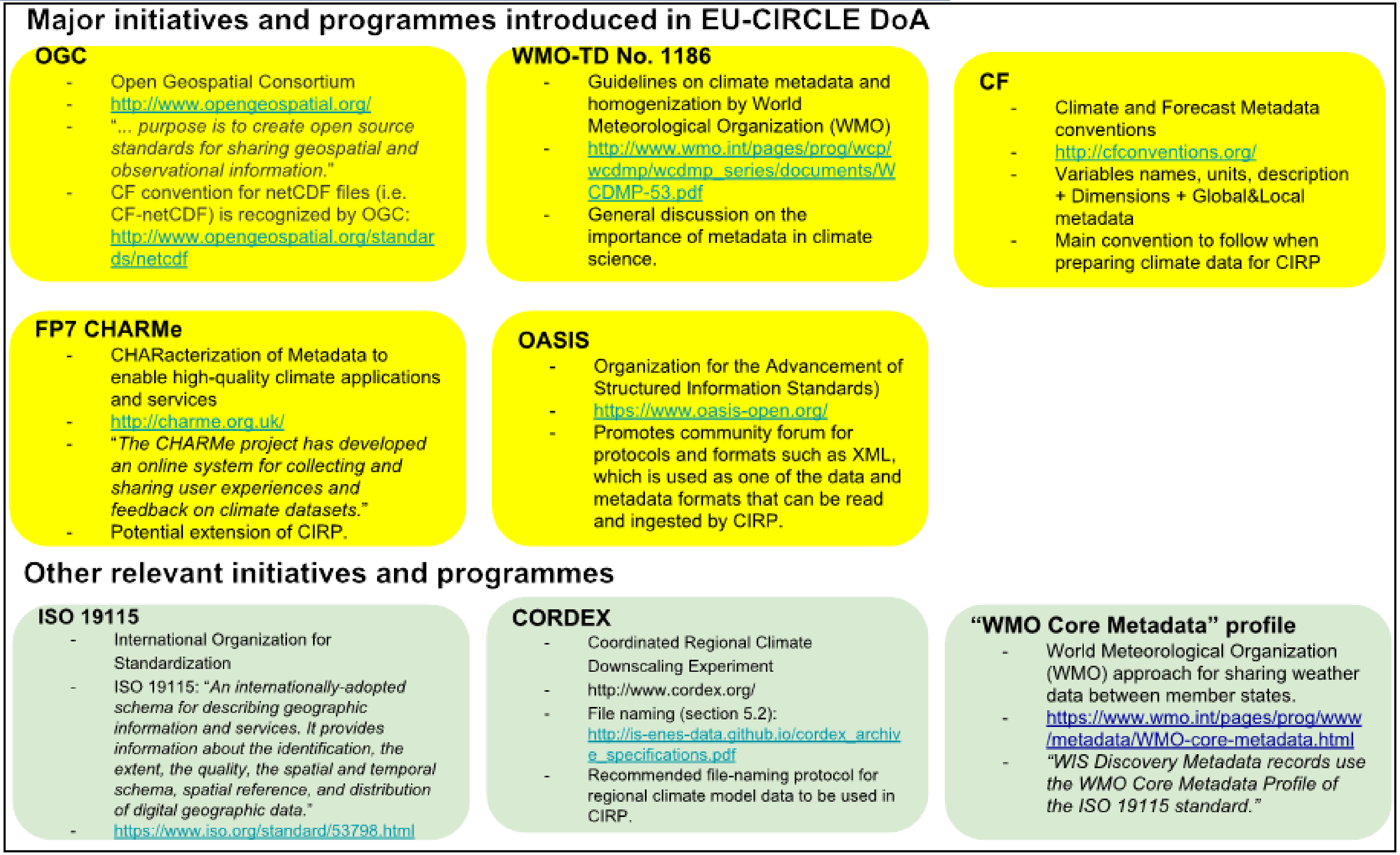

Figure 1.2 Summary of major initiatives and programmes relevant for the EU-CIRCLE metadata. In yellow: introduced in EU-CIRCLE Description of Action (DoA): OGC, OASIS, CF, WMO-TD No. 1186 report and EU FP7 project CHARMe. In light blue: additional relevant initiatives and programmes: ISO (ISO 19115), CORDEX, "WMO Core Metadata" profile. Information give: short description, URL and relevance to EU-CIRCLE and CIRP.

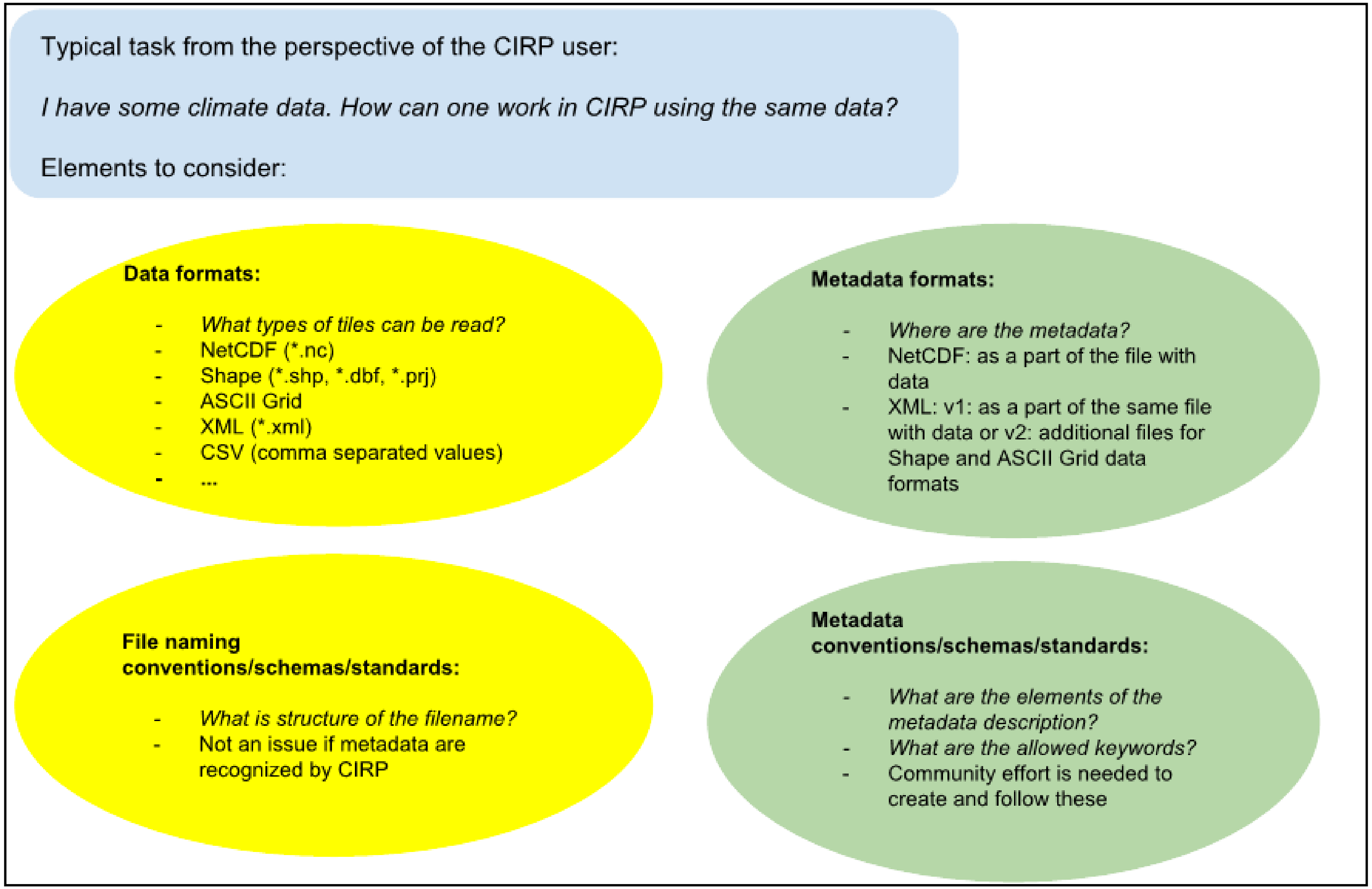

Figure 1.3 Main technical and implementation elements to consider from the perspective of the CIRP user when preparing and applying climate and climate hazard data and metadata. In yellow: elements to consider related to the data or file format, and specific way to name files. In light green: elements to consider related to the location of the metadata, and content of the metadata description. Specific example is given in Fig. 1.1.

2 Primary climate data

Primary climate data in this section refers to the following types of information: (1) results of the station observations, (2) interpolated station data to the regular grid, (3) remote sensing (e.g. by satellites) data available at the climate periods, (4) results of the global and regional climate models (cf. EU-CIRCLE D2.1). All types of data can be produced by the research communities and national and international bodies (often following community driven efforts for the common formats and conventions), and user generated data (often needs additional work to format specific data before sharing). Common format to share these data is NetCDF. Table 1.1 summarizes main elements in the NetCDF file related to variable dimensions, its local metadata (there may be several variables in the same file; e.g. air-temperature, precipitation amount, wind speed) and global metadata (describing details about the algorithms, infrastructure or models that produced the data).

Table 1.1 Main components of the NetCDF header under the CF metadata framework. Description of the components is given by **.**

Dimensions	These state the dimensions of the structure that the data will reside in and must always have different names. Dimensions can be spatial, temporal or any other variable. CF convention recommends that the dimensions be listed in the order TZYX, where T is time, Z is height or depth, Y is latitude and X is longitude.
Variables	**Data type:** • floating point ('float'), • double precision ('double') • integer ('int') numbers **Units** **Content description:** • standard_name • long_name **Missing data values:** describe the dummy values for missing data within the file.
Global Attributes	**Type Format:** CF-1.4 **title:** a succinct description of the data set **institute:** the organisation where the data were produced **source:** how the data were produced, e.g. model type, run number and circumstances **history:** an audit trail of data set processing **references:** web page or reference to the report and/or scientific publication **comment:** other useful information not covered elsewhere that adds value **author:** the person(s) who generated the data

2.1 Metadata for primary climate data

In Table 1.1.1 an overview is given of the typical types of information that describe dataset from the CORDEX initiative that brings together different research groups working on regional climate model development and application. Information is given at the level of (1) filename and (2) metadata content.

Table 1.1.1 Example of information that can be extracted from the results of the CORDEX regional climate models simulations in NetCDF format. Similar structure can be applied to other types of primary climate data and can be extended to (1) secondary climate data & impact models' (or climate hazard) data, and (2) other file formats.

Climate model data and metadata in NetCDF format	
Partner	**DHMZ**
Source of data	Results of the regional climate models simulations saved in the NetCDF format using CF and CORDEX metadata description.
References to relevant formats and protocols	NetCDF: https://www.unidata.ucar.edu/software/netcdf/ CF: http://cfconventions.org CORDEX: http://www.cordex.org/experiment-guidelines/experiment-protocol-rcms/
Typical filenames and content	*For example:* tas_EUR-11_CNRM-CERFACS-CNRM-CM5_rcp45_r1i1p1_SMHI-RCA4_v1_day_20060101-20101231.nc Files of this type contain a variable of interest (e.g. tas: near-surface air temperature) and various additional quantities that describe of the spatial and temporal characteristics of the data.
Description of the typical filenames	*For example:* **tas**: variable saved: near-surface air temperature **EUR-11**: domain: European domain at the 0.11° (~12.5 km) horizontal resolution. **CNRM-CERFACS**: institution that performed global climate model (GCM) simulation **CNRM-CM5**: specific GCM **rcp45**: greenhouse gases concentration scenario **r1i1p1**: additional details of the GCM simulation related to the model physics and initialization

	method. **SMHI**: institution that performed regional climate model (RCM) simulation **RCA4_v1**: specific RCM **day**: frequency of data: daily data **20060101**: first simulated time event in this file YYYYMMDD **20101231**: last simulated time event in this file YYYYMMDD
Dimensions	*For example:* dimensions: bnds = 2 ; time = UNLIMITED ; // (1826 currently) rlon = 424 ; rlat = 412 ;
Variables and **local attributes** saved in the respective file [NetCDF specific]	*For example:* variables: double **height** ; height:axis = "Z" ; height:long_name = "height" ; height:positive = "up" ; height:standard_name = "height" ; height:units = "m" ; double **time_bnds**(time, bnds) ; double **rlon**(rlon) ; rlon:standard_name = "grid_longitude" ; rlon:long_name = "longitude in rotated pole grid" ; rlon:units = "degrees" ; rlon:axis = "X" ; double **rlat**(rlat) ; rlat:standard_name = "grid_latitude" ; rlat:long_name = "latitude in rotated pole grid" ; rlat:units = "degrees" ; rlat:axis = "Y" ; char **rotated_pole** ;

	rotated_pole:grid_mapping_name= "rotated_latitude_longitude" ; rotated_pole:grid_north_pole_latitude = 39.25 ; rotated_pole:grid_north_pole_longitude = -162. ; double **time**(time) ; time:standard_name = "time" ; time:units = "days since 1949-12-01 00:00:00" ; time:calendar = "standard" ; time:long_name = "time" ; time:bounds = "time_bnds" ; time:axis = "T" ; float **tas**(time, rlat, rlon) ; tas:grid_mapping = "rotated_pole" ; tas:_FillValue = 1.e+20f ; tas:missing_value = 1.e+20f ; tas:standard_name = "air_temperature" ; tas:long_name= "Near-Surface Air Temperature" ; tas:units = "K" ; tas:coordinates = "lon lat height" ; tas:cell_methods = "time: mean" ; 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" ;
Global attributes saved in the respective file [NetCDF specific]	*For example:* global attributes: :Conventions = "CF-1.4" ; :contact = "rossby.cordex@smhi.se" ; :creation_date = "2013-07-03-T23:48:30Z" ; :experiment = "RCP4.5" ; :experiment_id = "rcp45" ; :driving_experiment = "CNRM-CERFACS-CNRM-CM5, rcp45, r1i1p1" ; :driving_model_id = "CNRM-CERFACS-CNRM-CM5" ;

	:driving_model_ensemble_member = "r1i1p1" ; :driving_experiment_name = "rcp45" ; :frequency = "day" ; :institution = "Swedish Meteorological and Hydrological Institute, Rossby Centre" ; :institute_id = "SMHI" ; :model_id = "SMHI-RCA4" ; :rcm_version_id = "v1" ; :project_id = "CORDEX" ; :CORDEX_domain = "EUR-11" ; :product = "output" ; :references="http://www.smhi.se/en/Research/Research-departments/climate-research-rossby-centre" ; :tracking_id = "0a9ffd2d-8fbb-4573-ba8b-8344ebfee6f9" ; :rossby_comment = "201144: CORDEX Europe 0.11 deg | RCA4 v1 | CNRM-CERFACS-CNRM-CM5 | r1i1p1 | rcp45 | L40" ; :rossby_run_id = "201144" ; :rossby_grib_path= "/nobackup/rossby16/rossby/joint_exp/cordex/201144/raw/"

In the previous table, example of the EURO-CORDEX NetCDF file metadata content was given for a specific variable (i.e. near-surface air temperature or tas), with the following metadata describing the actual 2D + time field. We extract this information in Table 1.1.2. In the scope of the EU-CIRCLE case studies some other primary climate data are also based on the regional climate models' simulation. These are summarized on the level of variable local metadata in Table 1.1.3.

Table 1.1.2 Local metadata extracted from Table 1.1.1, related to the variable Near-Surface Air Temperature. Local metadata contain (but are not limited to) grid mapping (*What is projection of the model grid?*), missing values (*How are missing values marked?* Covered by FillValue and missing_value), standard (following CF convention) and long variable name, variable units, coordinates, and the way the data are processed (in this table cell_methods defines the data are temporal mean from the data available at the higher frequency).

Variable: **tas** (Near-Surface Air Temperature)
float **tas**(time, rlat, rlon) ; tas:grid_mapping = "rotated_pole" ; tas:_FillValue = 1.e+20f ; tas:missing_value = 1.e+20f ; tas:standard_name = "air_temperature" ; tas:long_name = "Near-Surface Air Temperature" ; tas:units = "K" ; tas:coordinates = "lon lat height" ; tas:cell_methods = "time: mean" ;

Assuming one analyses the results of the same regional climate model (in this example SMHI-RCA4), forced by the same boundary conditions (in this example CNRM-CERFACS-CNRM-CM) under the same scenario of the greenhouse gases concentrations (in this example RCP4.5) the content of the metadata would be generally the same in terms of the dimensions and most of the global attributes (also assuming data are available at the same temporal frequency), while in local attributes major difference is found for the metadata concerning the variable in interest. In this sense, following examples of other variables are listed (under the CF, CMIP5 and CORDEX metadata conventions):

Table 1.1.3 Same as Table 1.1.2 for a set of variables. For the specific source of model data, differences between variables include different standard_name, long_name, units and cell_methods (e.g. if model produces data at the sub-daily frequency one can define daily mean, minimum, maximum etc.). Also for some surface variables there is no height dimension, while near-surface variables for temperature and humidity are usually defined for the level of 2m, and for wind speed for the level of 10m.

Variable: **hurs** (Near-Surface Relative Humidity)
float **hurs**(time, rlat, rlon) ; hurs:grid_mapping = "rotated_pole" ; hurs:missing_value = 1.e+20f ; hurs:_FillValue = 1.e+20f ; hurs:standard_name = "relative_humidity" ; hurs:long_name = "Near-Surface Relative Humidity" ; hurs:units = "%" ; hurs:coordinates = "lon lat height" ; hurs:cell_methods = "time: mean" ;
Variable: **mrso** (Total Soil Moisture Content)
float **mrso**(time, rlat, rlon) ; mrso:grid_mapping = "rotated_pole" ; mrso:_FillValue = 1.e+20f ; mrso:missing_value = 1.e+20f ; mrso:standard_name = "soil_moisture_content" ; mrso:long_name = "Total Soil Moisture Content" ; mrso:units = "kg m-2" ; mrso:coordinates = "lon lat" ; mrso:cell_methods = "time: mean" ;
Variable: **pr** (Precipitation)
float **pr**(time, rlat, rlon) ; pr:grid_mapping = "rotated_pole" ; pr:_FillValue = 1.e+20f ; pr:missing_value = 1.e+20f ; pr:standard_name = "precipitation_flux" ; pr:long_name = "Precipitation" ; pr:units = "kg m-2 s-1" ;

```
pr:coordinates = "lon lat" ;
pr:cell_methods = "time: mean" ;
```

Variable: **prhmax** (Daily Maximum Hourly Precipitation Rate)

```
float prhmax(time, rlat, rlon) ;
    prhmax:grid_mapping = "rotated_pole" ;
    prhmax:_FillValue = 1.e+20f ;
    prhmax:missing_value = 1.e+20f ;
    prhmax:standard_name = "precipitation_flux" ;
    prhmax:long_name = "Daily Maximum Hourly Precipitation Rate" ;
    prhmax:units = "kg m-2 s-1" ;
    prhmax:coordinates = "lon lat" ;
    prhmax:cell_methods = "time: maximum" ;
```

Variable: **psl** (Sea Level Pressure)

```
float psl(time, rlat, rlon) ;
    psl:grid_mapping = "rotated_pole" ;
    psl:_FillValue = 1.e+20f ;
    psl:missing_value = 1.e+20f ;
    psl:standard_name = "air_pressure_at_sea_level" ;
    psl:long_name = "Sea Level Pressure" ;
    psl:units = "Pa" ;
    psl:coordinates = "lon lat" ;
    psl:cell_methods = "time: mean" ;
```

Variable: **sfcwind** (Near-surface Wind Speed)

```
float sfcWind(time, rlat, rlon) ;
    sfcWind:grid_mapping = "rotated_pole" ;
    sfcWind:_FillValue = 1.e+20f ;
    sfcWind:missing_value = 1.e+20f ;
    sfcWind:standard_name = "wind_speed" ;
    sfcWind:long_name = "Near-Surface Wind Speed" ;
    sfcWind:units = "m s-1" ;
    sfcWind:coordinates = "lon lat height" ;
    sfcWind:cell_methods = "time: point" ;
```

Variable: **sfcWindmax** (Daily Maximum Near-Surface Wind Speed)

```
float sfcWindmax(time, rlat, rlon) ;
    sfcWindmax:standard_name = "wind_speed" ;
    sfcWindmax:long_name = "Daily Maximum Near-Surface Wind Speed" ;
    sfcWindmax:units = "m s-1" ;
    sfcWindmax:cell_methods = "time: maximum" ;
    sfcWindmax:coordinates = "height lat lon" ;
    sfcWindmax:missing_value = 1.e+20f ;
    sfcWindmax:_FillValue = 1.e+20f ;
```

sfcWindmax:grid_mapping = "rotated_pole" ;
Variable: **snd** (Snow Depth)
float **snd**(time, rlat, rlon) ; snd:standard_name = "surface_snow_thickness" ; snd:long_name = "Snow Depth" ; snd:units = "m" ; snd:coordinates = "lon lat" ; snd:_FillValue = 1.e+20f ; snd:missing_value = 1.e+20f ; snd:cell_methods = "time: mean" ; snd:grid_mapping = "rotated_pole" ;
Variable: **tasmax** (Daily Maximum Near-Surface Air Temperature)
float **tasmax**(time, rlat, rlon) ; tasmax:grid_mapping = "rotated_pole" ; tasmax:_FillValue = 1.e+20f ; tasmax:missing_value = 1.e+20f ; tasmax:standard_name = "air_temperature" ; tasmax:long_name = "Daily Maximum Near-Surface Air Temperature" ; tasmax:units = "K" ; tasmax:coordinates = "lon lat height" ; tasmax:cell_methods = "time: maximum" ;
Variable: **tasmin** (Daily Minimum Near-Surface Air Temperature)
float **tasmin**(time, rlat, rlon) ; tasmin:grid_mapping = "rotated_pole" ; tasmin:_FillValue = 1.e+20f ; tasmin:missing_value = 1.e+20f ; tasmin:standard_name = "air_temperature" ; tasmin:long_name = "Daily Minimum Near-Surface Air Temperature" ; tasmin:units = "K" ; tasmin:coordinates = "lon lat height" ; tasmin:cell_methods = "time: minimum" ;
Variable: **uas** (Eastward Near-Surface Wind)
float **uas**(time, rlat, rlon) ; uas:grid_mapping = "rotated_pole" ; uas:_FillValue = 1.e+20f ; uas:missing_value = 1.e+20f ; uas:standard_name = "eastward_wind" ; uas:long_name = "Eastward Near-Surface Wind" ; uas:units = "m s-1" ; uas:coordinates = "lon lat height" ; uas:cell_methods = "time: mean" ;

Variable: **vas** (Northward Near-Surface Wind)
float **vas**(time, rlat, rlon) ; vas:grid_mapping = "rotated_pole" ; vas:_FillValue = 1.e+20f ; vas:missing_value = 1.e+20f ; vas:standard_name = "northward_wind" ; vas:long_name = "Northward Near-Surface Wind" ; vas:units = "m s-1" ; vas:coordinates = "lon lat height" ; vas:cell_methods = "time: mean" ;

When primary climate data are not given in NetCDF format, in the EU-CIRCLE case studies other data format are valid too. Table 1.1.4 lists EU-CIRCLE recommendation of the metadata information needed with the CSV data format.

Table 1.1.4 EU-CIRCLE recommendation for the metadata of EURO-CORDEX results in the CSV format. Specific example includes climate model data processed from NetCDF format to CSV format before using in the R and GIS algorithms (cf. EU-CIRCLE D2.2 and D2.3).

Climate model data output: CSV format	
Partner	**NCSRD**
Headers	Data type : double precision ('double') Header description with units: • TIME , • LAT , LON , • T_00 [C] , • WS_00[m/s] , • T_06[C] , • WS_06[m/s] , • T_12[C] , • R_D[mm] , • WS_12[m/s] , • RH_D[%] , • SM_D[kg/m2] , • T_18[C] , • WS_18[m/s] Missing data values : -9999
External Attributes	Type Format —csv title: scenario[rcpXY]_startyear[YYYY]_endyear[YYYY]_pointID[idXXXXX].csv (for example rcp26_2006_2010_id142771.csv) coordinate system : LAT, LON climatic model : any climatic scenario : any institute : NSCRD source : post processing of NetCDF EURO-CORDEX climate data with NCL scripts and FORTRAN code Comment: author — Stelios Karozis – EU-CIRCLE project.

2.2 File-naming recommendation for the primary climate data

If the description of the data using available metadata is well set, the need for the specific file-naming convention is less relevant for the application in CIRP. Nevertheless, based on the experience in other initiatives, some good-practice approaches may be followed. For the example of the regional model data given in Table 1.1.1,

tas_EUR-11_CNRM-CERFACS-CNRM-CM5_rcp45_r1i1p1_SMHI-RCA4_v1_day_20060101-20101231.nc

the CORDEX instructions where followed, while the processed regional climate data are summarized in Table 1.1.4,

rcp26_2006_2010_id142771.ccsv

with elements between underscore ("_") separately defined. We recommend similar structure of the file names for primary climate data in other formats too. This enables efficient listing, searching and filtering procedures before reading in CIRP.

3 Secondary climate data

In this section we provide an overview of the metadata need for the secondary climate data. For example, one starts with primary climate data such as daily precipitation fields (either observed or modelled), and needs some specific climate indices. The question is: *How should one define metadata of the final result?* Solution to this issue depends on the structure and results of the applied algorithm. We will demonstrate a few examples that can be generalized where needed. Additionally, we will end this Section by summarizing examples of related metadata when primary climate data where used to produce secondary climate data, such as the Fire Weather Index, or the estimates of the characteristics of the probability distribution function for climate variables used in the EU-CIRCLE case studies.

3.1 Metadata for secondary climate data

In Table 2.1.1 an example is given using CDO tool (cf. EU-CIRCLE D2.3 and https://code.mpimet.mpg.de/projects/cdo/wiki), where primary climate data (e.g. near-surface air temperature) is used as an input to determine tropical nights climate index. Similar examples for heavy precipitation days index and consecutive strong wind days index are given in Tables 2.1.2 and 2.1.3.

Table 2.1.1 Application of CDO to create secondary climate data and related metadata: example of tropical nights climate index. Creation of the metadata is handled by the CDO tool. Extension of the additional metadata can be made using the NCO or CDO tools, or most of the existing software that can process files in the NetCDF format.

Metadata of the input	Local metadata as for variable tas in Table 1.1.2; Dimensions and global metadata as in Table 1.1.1.
Algorithm	Bash command in Unix/Linux operating system: cdo **eca_tr** -daymean tas_FR-11_CNRM-CERFACS-CNRM-CM5_historical_r1i1p1_SMHI-RCA4_v1_3hr_197001010000-197012312100.nc eca_tr_tas_FR-11_CNRM-CERFACS-CNRM-CM5_historical_r1i1p1_SMHI-RCA4_v1_3hr_197001010000-197012312100.nc
Metadata of the output	float **tropical_nights_index_per_time_period**(time, rlat, rlon) ; **tropical_nights_index_per_time_period**:long_name = "Tropical nights index is the number of days where minimum of temperature is above 20 degree Celsius. The time period should be defined by the bounds of the time coordinate." ; **tropical_nights_index_per_time_period**:units = "No." ; **tropical_nights_index_per_time_period**:grid_mapping = "rotated_pole" ; **tropical_nights_index_per_time_period**:coordinates = "height lat lon" ; **tropical_nights_index_per_time_period**:_FillValue = 1.e+20f ; **tropical_nights_index_per_time_period**:missing_value = 1.e+20f ;

Comment	The computation of the number of tropical nights in specific period is trivial using the CDO tool (cf. D2.3 and https://code.mpimet.mpg.de/projects/cdo). Local metadata for the final variable are automatically generated by the CDO software. Similar entries in the local metadata section can be generated by other tools. Other local metadata, dimension description and global metadata are inherited from the input file. In this specific example, one first computes **daily mean** from the input with the sub-daily frequency. Details of the climate indices computation in CDO: https://code.mpimet.mpg.de/projects/cdo/embedded/cdo_eca.pdf

Table 2.1.2 Application of CDO to create secondary climate data and related metadata: example of heavy precipitation days index. Creation of the metadata is handled by the CDO tool. Extension of the additional metadata can be made using the NCO or CDO tools, or most of the existing software that can process files in the NetCDF format.

Metadata of the input	Local metadata as for variable **pr** in Table **1.1.3**; Dimensions and global metadata as in Table **1.1.1.**
Algorithm	Bash command in Unix/Linux operating system: cdo **eca_r10mm** -mulc,86400 pr_FR-11_CNRM-CERFACS-CNRM-CM5_historical_r1i1p1_SMHI-RCA4_v1_day_19810101-19851231.nc eca_r10mm_pr_FR-11_CNRM-CERFACS-CNRM-CM5_historical_r1i1p1_SMHI-RCA4_v1_day_19810101-1985123.nc
Metadata of the output	float **heavy_precipitation_days_index_per_time_period**(time, rlat, rlon) ; **heavy_precipitation_days_index_per_time_period**:long_name = "Heavy precipitation days is the number of days per time period with daily precipitation sum exceeding 10 mm. The time period should be defined by the bounds of the time coordinate." ; **heavy_precipitation_days_index_per_time_period**:units = "No." ; **heavy_precipitation_days_index_per_time_period**:grid_mapping = "rotated_pole" ; **heavy_precipitation_days_index_per_time_period**:coordinates = "lat lon" ; **heavy_precipitation_days_index_per_time_period**:_FillValue = 1.e+20f ; **heavy_precipitation_days_index_per_time_period**:missing_value = 1.e+20f ;
Comment	The computation of the number of heavy precipitation days in specific period is trivial using the CDO tool (cf. D2.3 and https://code.mpimet.mpg.de/projects/cdo). Local metadata for the final variable are automatically generated by the CDO software. Similar entries in the local metadata section can be generated by other tools. Other local metadata, dimension description and global metadata are inherited from the input file. In this specific example, one first **converts** input field from **$kg/m^2/s$** units to **mm** units.

	Details of the climate indices computation in CDO: https://code.mpimet.mpg.de/projects/cdo/embedded/cdo_eca.pdf

Table 2.1.3 Application of CDO to create secondary climate data and related metadata: example of strong wind days climate index. Creation of the metadata is handled by the CDO tool. Extension of the additional metadata can be made using the NCO or CDO tools, or most of the existing software that can process files in the NetCDF format.

Metadata of the input	Local metadata as for variable **sfcwind** in Table **1.1.3**; Dimensions and global metadata as in Table **1.1.1.**
Algorithm	Bash command in Unix/Linux operating system: cdo **strwin** -daymean sfcWind_FR-11_CNRM-CERFACS-CNRM-CM5_rcp45_r1i1p1_SMHI-RCA4_v1_3hr_200701010000-200712312100.nc strwin_sfcWind_FR-11_CNRM-CERFACS-CNRM-CM5_rcp45_r1i1p1_SMHI-RCA4_v1_3hr_200701010000-200712312100.nc
Metadata of the output	**float strong_wind_days_index_per_time_period**(time, rlat, rlon) ; strong_wind_days_index_per_time_period:long_name = "Strong wind days index is the number of days per time period where maximum wind speed is above 10 m/s. The time period should be defined by the bounds of the time coordinate." ; strong_wind_days_index_per_time_period:units = "No." ; strong_wind_days_index_per_time_period:grid_mapping = "rotated_pole" ; strong_wind_days_index_per_time_period:coordinates = "height lat lon" ; strong_wind_days_index_per_time_period:_FillValue = 1.e+20f ; strong_wind_days_index_per_time_period:missing_value = 1.e+20f ; **float consecutive_strong_wind_days_index_per_time_period**(time, rlat, rlon) ; consecutive_strong_wind_days_index_per_time_period:long_name = "Greatest number of consecutive strong wind days per time period. The time period should be defined by the bounds of the time coordinate." ; consecutive_strong_wind_days_index_per_time_period:units = "No." ; consecutive_strong_wind_days_index_per_time_period:grid_mapping = "rotated_pole" ; consecutive_strong_wind_days_index_per_time_period:coordinates = "height lat lon" ; consecutive_strong_wind_days_index_per_time_period:_FillValue = 1.e+20f ; consecutive_strong_wind_days_index_per_time_period:missing_value = 1.e+20f ;
Comment	The computation of the number of strong wind days in specific period is trivial using the CDO tool (cf. D2.3 and https://code.mpimet.mpg.de/projects/cdo). Local metadata for the final variable are automatically generated by the CDO software. Similar entries in the local metadata section can be generated by other tools. Other local metadata, dimension description and global metadata are inherited from the input file. In this specific example, one first compute **daily mean** from the input with the sub-daily frequency. Also, two output variables are computed **consecutive_strong_wind_days_index_per_time_period** and

	strong_wind_days_index_per_time_period. Details of the climate indices computation in CDO: https://code.mpimet.mpg.de/projects/cdo/embedded/cdo_eca.pdf

In summary, we recommend the secondary metadata for the typical climate indices to follow implementation in the CDO tool. Following list of the indices is currently available in the CDO and can be implemented in other tools:

Table 2.1.4 Climate indices as implemented in the CDO tool (source: https://code.mpimet.mpg.de/projects/cdo/embedded/cdo_eca.pdf)

CDO command	Short description
eca cdd	Consecutive dry days index per time period
eca cfd	Consecutive frost days index per time period
eca csu	Consecutive summer days index per time period
eca cwd	Consecutive wet days index per time period
eca cwdi	Cold wave duration index w.r.t. mean of reference period
eca cwf	Cold-spell days index w.r.t. 10th percentile of reference period
eca etr	Intra-period extreme temperature range
eca fd	Frost days index per time period
eca gsl	Growing season length index
eca hd	Heating degree days per time period
eca hwdi	Heat wave duration index w.r.t. mean of reference period
eca hwf	Warm spell days index w.r.t. 90th percentile of reference period
eca id	Ice days index per time period
eca r75p	Moderate wet days w.r.t. 75th percentile of reference period
eca r75ptot	Precipitation percent due to R75p days
eca r90p	Wet days w.r.t. 90th percentile of reference period
eca r90ptot	Precipitation percent due to R90p days
eca r95p	Very wet days w.r.t. 95th percentile of reference period
eca r95ptot	Precipitation percent due to R95p days
eca r99p	Extremely wet days w.r.t. 99th percentile of reference period
eca r99ptot	Precipitation percent due to R99p days
eca pd	Precipitation days index per time period
eca r10mm	Heavy precipitation days index per time period
eca r20mm	Very heavy precipitation days index per time period
eca rr1	Wet days index per time period
eca rx1day	Highest one day precipitation amount per time period

eca rx5day	Highest five-day precipitation amount per time period
eca sdii	Simple daily intensity index per time period
eca su	Summer days index per time period
eca tg10p	Cold days percent w.r.t. 10th percentile of reference period
eca tg90p	Warm days percent w.r.t. 90th percentile of reference period
eca tn10p	Cold nights percent w.r.t. 10th percentile of reference period
eca tn90p	Warm nights percent w.r.t. 90th percentile of reference period
eca tr	Tropical nights index per time period
eca tx10p	Very cold days percent w.r.t. 10th percentile of reference period
eca tx90p	Very warm days percent w.r.t. 90th percentile of reference period

After summarizing recommended approach to define metadata secondary climate data as implemented in CDO tool, we demonstrate metadata defined in the EU-CIRCLE case studies related to the estimation of the Fire Weather Index. Table 2.1.5 describes metadata for the variable *Fire Weather Index*, while Tables 2.1.6. and 2.1.7 give additional information about related variables *Drought Code* and *Fine Fuel Moisture Code.*

Table 2.1.5 Metadata for secondary climate data: Fire Weather Index (FWI). Data and metadata are generated using the GIS software.

Secondary climate data: FWI	
Partner	**NSCRD**
Dimensions	Number of columnsx Number of rowsx xllcornerx yllcornerx cellsizex
Variables	Data type:double precision ('double') Units : no dimension Content description: FWI values Missing data values: -9999
External Global Attributes	Type Format: ArcInfo ASCII GRID title: Fire weather Index coordinate system: WGS84 world projected data creation date: dd/mm/yyyy data reference date: dd/mm/yyyy climatic model: x climatic scenario: x description of the data set: daily Fire Weather Index values (FWI) institute: NSCRD source: G-FMIS_FWI software history: Climatic data inputs from CORDEX references: Van Wagner, C.E.; Pickett, T.L. 1985.Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Canadian Forest Service, Ottawa, ON. Forestry Technical Report 33. Comment: author: Vassiliki Varela – EU-CIRCLE project
Associated file	Projection file *.prj

Table 2.1.6 Metadata for secondary climate data: Drought Code (DC). Data and metadata are generated using the GIS software.

Secondary climate data: DC	
Partner	**NSCRD**
Dimensions	Number of columnsx Number of rowsx xllcornerx yllcornerx cellsizex
Variables	Data type: double precision ('double') Units: no dimension Content description: DC values Missing data values: -9999
External Global Attributes	Type Format: ArcInfo ASCII GRID title: Drought Code coordinate system: WGS84 world projected data creation date: dd/mm/yyyy data reference date: dd/mm/yyyy climatic model: x climatic scenario: x description of the data set: daily Drought Code (DC) institute: NSCRD source: G-FMIS_FWI software history: Climatic data inputs from CORDEX references: Van Wagner, C.E.; Pickett, T.L. 1985.Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Canadian Forest Service, Ottawa, ON. Forestry Technical Report 33. Comment : author: Vassiliki Varela – EU-CIRCLE project
Associated file	Projection file *.prj

Table 2.1.7 Metadata for secondary climate data: Fine Fuel Moisture Code. Data and metadata are generated using the GIS software.

Secondary climate data: Fine Fuel Moisture Code	
Partner	**NSCRD**
Dimensions	Number of columnsx Number of rowsx xllcornerx yllcornerx cellsizex
Variables	Data type: double precision ('double') Units: no dimension Content description: Fine Fuel Moisture Code values Missing data values: -9999
External Global Attributes	Type Format: ArcInfo ASCII GRID title: Fine Fuel Moisture Code coordinate system: WGS84 world projected data creation date:dd/mm/yyyy data reference date: dd/mm/yyyy climatic model: x climatic scenario: x description of the data set: daily Fine Fuel Moisture Code (FFMC) institute: NSCRD source: G-FMIS_FWI software history: Climatic data inputs from CORDEX references: Van Wagner, C.E.; Pickett, T.L. 1985.Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Canadian Forest Service, Ottawa, ON. Forestry Technical Report 33. Comment: - author: Vassiliki Varela – EU-CIRCLE project
Associated file	Projection file *.prj

Last demonstration of the implementation of the metadata definition in EU-CIRCLE related to the secondary data is about the estimates of the (1) return period, (2) probability of exceedance and (3) dates with extreme events (Table 2.1.8).

Table 2.1.8 Metadata for secondary climate data: return period, probability and scenario. Data and metadata are generated using the R software.

Partner	**NCSRD**
Return period metadata	
Variables	Attribute name: RP (return period) Header description: 20, 25, 50, 100, 500 Data type: double precision ('double') Units: same units with the examined climate variable
External Global Attributes	Type Format: csv title: RP.csv climatic scenario: x description of the dataset: return period of 20, 25, 50, 100, 500 years for specific climate variable institute :NSCR"D" source: R scripts based on extreme value theory history: input of post processing climate data from CORDEX (see Metadata of EURO-CORDEX csv files) author: Nadia Politi – EU-CIRCLE project
Probability metadata	
Variables	Attribute name: prob (probability of exceedance) Data type : double precision ('double') Units : no dimension
External Global Attributes	Type Format: csv title: prob.csv climatic scenario: x description of the dataset: probability of exceeding a value institute :NSCR"D" source: R scripts based on extreme value theory history: input of post processing climate data from CORDEX (see Metadata of EURO-CORDEX csv files) author: Nadia Politi – EU-CIRCLE project

Scenarios metadata	
Variables	Attribute name: scenarios Header description: date Data type: date [YYMMDD] Units: no dimension
External Global Attributes	Type Format: csv title : scenarios.csv climatic scenario : x description of the data set : dates with extreme events, called scenarios, selected through PCA and cluster analysis institute: NSCR"D" source: R scripts – scenarios selection tool history: input of post processing climate data from CORDEX (see Metadata of EURO-CORDEX csv files) author: Nadia Politi – EU-CIRCLE project

3.2 File-naming recommendation for the secondary climate data

Similar as in the Subsection 1.2, if the description of the data using available metadata is well set, the need for the specific file-naming convention is less relevant for the application in CIRP. However, some general suggestions may be followed.

In the example from Table 2.1.1, three elements where considered:

1. input data: RCM results in the file named:

tas_EUR-11_CNRM-CERFACS-CNRM-CM5_rcp45_r1i1p1_SMHI-RCA4_v1_day_20060101-20101231.nc

2. data processing: cdo **eca_tr**
3. output data: secondary climate results in the file named:

eca_tr_tas_FR-11_CNRM-CERFACS-CNRM-CM5_historical_r1i1p1_SMHI-RCA4_v1_3hr_197001010000-197012312100.nc,

where output file inherits global metadata, dimensions and file-naming structure from the input file. Recommendation is given to follow similar approach when creating secondary climate data from (1) other types of primary climate data, (2) using different input and output format, and (3) applying different analytical and programming tools and approaches.

4 Impact models' data

In this section we summarize several recommendations for preparing the results of users impact models related to different climate hazards or any events that will modulate under the effects of the climate change. As for the previous sections, we end this section by demonstrating possible approaches to define metadata and file-naming rules.

4.1 Metadata for impact models' data

Impact models in this subsection are models that simulate or model events such as fires, smoke plumes and floods in a great detail. This leads to different tools and set of community rules in various research areas. We start again from the perspective of the CIRP users whose aim is to link the results of the impact model and CIRP.

In Tables 3.1.1, 3.1.2 and 3.1.3, we extracted the content of the CF convention for the global climate models. In these type of models fires and floods are simulated on a coarse model grid, but metadata content related to standard names, units and description is more generalized and can find application in the way one specifies metadata needed in CIRP.

Table 3.1.1 List of variables related to the **flood** characteristics. Recommended CF standard names, units and descriptions (source: http://cfconventions.org/Data/cf-standard-names/54/build/cf-standard-name-table.html).

CF Standard name	Units	CF Description
flood_water_speed	$m\ s^{-1}$	"Speed is the magnitude of velocity. Flood water is water that covers land which is normally not covered by water."
northward_flood_water_velocity	$m\ s^{-1}$	"A velocity is a vector quantity. "Northward" indicates a vector component which is positive when directed northward (negative southward). Flood water is water that covers land which is normally not covered by water."
eastward_flood_water_velocity	$m\ s^{-1}$	"A velocity is a vector quantity. "Eastward" indicates a vector component which is positive when directed eastward (negative westward). Flood water is water that covers land which is normally not covered by water."
flood_water_duration_above_threshold	s	"The quantity with standard name flood_water_duration_above_threshold is the time

		elapsed between the instant when the flood depth first rises above a given threshold until the time falls below the same threshold for the last time at a given point in space. If a threshold is supplied, it should be specified by associating a coordinate variable or scalar coordinate variable with the data variable and giving the coordinate variable a standard name of flood_water_thickness. The values of the coordinate variable are the threshold values for the corresponding sub-arrays of the data variable. If no threshold is specified, its value is taken to be zero. Flood water is water that covers land which is normally not covered by water.”
flood_water_thickness	m	“The flood_water_thickness is the vertical distance between the surface of the flood water and the surface of the solid ground, as measured at a given point in space. The standard name ground_level_altitude is used for a data variable giving the geometric height of the ground surface above the geoid. "Flood water" is water that covers land which is normally not covered by water.”
time_of_maximum_flood_depth	s	“The quantity with standard name time_of_maximum_flood_depth is the time elapsed between the breaking of a levee (origin of flood water simulation) and the instant when the flood depth reaches its maximum during the simulation for a given point in space. Flood water is water that covers land which is normally not covered by water.“
time_when_flood_water_falls_below_threshold	s	“The quantity with standard name time_when_flood_water_falls_below_threshold is the time elapsed between the breaking of a levee (origin of flood water simulation) and the instant when the depth falls below a given threshold for the last time, having already risen to its maximum depth, at a given point in space. If a threshold is supplied, it should be specified by associating a coordinate variable or scalar coordinate variable with the data variable and giving the coordinate variable a standard name of flood_water_thickness. The values of the coordinate variable are the threshold values for the corresponding

		sub-arrays of the data variable. If no threshold is specified, its value is taken to be zero. Flood water is water that covers land which is normally not covered by water."
time_when_flood_water_rises_above_threshold	s	"The quantity with standard name time_when_flood_water_rises_above_threshold is the time elapsed between the breaking of a levee (origin of flood water simulation) and the instant when the depth first rises above a given threshold at a given point in space. If a threshold is supplied, it should be specified by associating a coordinate variable or scalar coordinate variable with the data variable and giving the coordinate variable a standard name of flood_water_thickness. The values of the coordinate variable are the threshold values for the corresponding sub-arrays of the data variable. If no threshold is specified, its value is taken to be zero. Flood water is water that covers land which is normally not covered by water".

Table 3.1.2 List of variables related to the **fire** characteristics. Recommended CF standard names, units and descriptions (source: http://cfconventions.org/Data/cf-standard-names/54/build/cf-standard-name-table.html).

CF Standard name	Units	CF Description
fire_area	m^2	“"X_area" means the horizontal area occupied by X within the grid cell. The extent of an individual grid cell is defined by the horizontal coordinates and any associated coordinate bounds or by a string valued auxiliary coordinate variable with a standard name of "region". "Fire area" means the area of detected biomass fire.”
fire_radiative_power	W	“The product of the irradiance (the power per unit area) of a biomass fire and the corresponding fire area. A data variable containing the area affected by fire should be given the standard name fire_area.”
fire_temperature	K	“The overall temperature of a fire area due to contributions from smouldering and flaming biomass. A data variable containing the area affected by fire should be given the standard name fire_area.”

Table 3.1.3 Examples of variables related to the **smoke** characteristics. Recommended CF standard names, units and descriptions (source: http://cfconventions.org/Data/cf-standard-names/54/build/cf-standard-name-table.html). Similar definitions and standard names are summarized for the purpose of this Table.

Standard name	Units	Description
surface_upward_mass_flux_of_**carbon_dioxide**_expressed_as_**carbon**_due_to_emission_**from_fires**	kg m^{-2} s^{-1}	"The surface called "surface" means the lower boundary of the atmosphere. "Upward" indicates a vector component which is positive when directed upward (negative downward). In accordance with common usage in geophysical disciplines, "flux" implies per unit area, called "flux density" in physics. The phrase "expressed_as" is used in the construction A_expressed_as_B, where B is a chemical constituent of A. It means that the quantity indicated by the standard name is calculated solely with respect to the B contained in A, neglecting all other chemical constituents of A. The chemical formula for carbon dioxide is CO_2. The specification of a physical process by the phrase "due_to_" process means that the quantity named is a single term in a sum of terms which together compose the general quantity named by omitting the phrase. "Emission" means emission from a primary source located anywhere within the atmosphere, including at the lower boundary (i.e. the surface of the earth). "Emission" is a process entirely distinct from "re-emission" which is used in some standard names. The term "fires" means all biomass fires, whether naturally occurring or ignited by humans. The quantity with standard name surface_upward_mass_flux_of_carbon_dioxide_expressed_as_carbon_due_to_emission_from_fires is the sum of the quantities with standard names surface_upward_mass_flux_of_carbon_dioxide_expressed_as_carbon_due_to_emission_from_vegetation_in_fires and surface_upward_mass_flux_of_carbon_dioxide_expressed_as_carbon_due_to_emission_from_litter_in_fires."
Similar variables as in the upper cell include: 1. *from_fires_excluding_anthropogenic_land_use_change 2. *from_litter_in_fires 3. *from_natural_fires 4. *from_vegetation_in_fires	kg m^{-2} s^{-1}	Equivalent description as for surface_upward_mass_flux_of_carbon_dioxide_expressed_as_carbon_due_to_emission_from_fires.

surface_upward_mass_flux_of_**nitrogen_compounds**_expressed_as_**nitrogen**_due_to_emission_from_fires	kg m^{-2} s^{-1}	“The surface called "surface" means the lower boundary of the atmosphere. "Upward" indicates a vector component which is positive when directed upward (negative downward). In accordance with common usage in geophysical disciplines, "flux" implies per unit area, called "flux density" in physics. The phrase "expressed_as" is used in the construction A_expressed_as_B, where B is a chemical constituent of A. It means that the quantity indicated by the standard name is calculated solely with respect to the B contained in A, neglecting all other chemical constituents of A. "Nitrogen compounds" summarizes all chemical species containing nitrogen atoms. The list of individual species that are included in this quantity can vary between models. Where possible, the data variable should be accompanied by a complete description of the species represented, for example, by using a comment attribute. The specification of a physical process by the phrase "due_to_" process means that the quantity named is a single term in a sum of terms which together compose the general quantity named by omitting the phrase. "Emission" means emission from a primary source located anywhere within the atmosphere, including at the lower boundary (i.e. the surface of the earth). "Emission" is a process entirely distinct from "re-emission" which is used in some standard names. The term "fires" means all biomass fires, whether naturally occurring or ignited by humans.”
tendency_of_atmosphere_mass_content_of_**alcohols**_due_to_emission_from_forest_fires	kg m^{-2} s^{-1}	“"tendency_of_X" means derivative of X with respect to time. "Content" indicates a quantity per unit area. The "atmosphere content" of a quantity refers to the vertical integral from the surface to the top of the atmosphere. For the content between specified levels in the atmosphere, standard names including "content_of_atmosphere_layer" are used. The mass is the total mass of the molecules. The specification of a physical process by the phrase "due_to_" process means that the quantity named is a single term in a sum of terms which together compose the general quantity named by omitting the phrase. "Emission" means emission from a primary source located anywhere within the atmosphere, including at the lower boundary

		(i.e. earth's surface). "Emission" is a process entirely distinct from "re-emission" which is used in some standard names. Alcohols include all organic compounds with an alcoholic (OH) group. In standard names "alcohols" is the term used to describe the group of chemical species that are represented within a given model. The list of individual species that are included in a quantity having a group chemical standard name can vary between models. Where possible, the data variable should be accompanied by a complete description of the species represented, for example, by using a comment attribute. The "forest fires" sector comprises the burning (natural and human-induced) of living or dead vegetation in forests. "Forest fires" is the term used in standard names to describe a collection of emission sources. A variable which has this value for the standard_name attribute should be accompanied by a comment attribute which lists the source categories and provides a reference to the categorization scheme, for example, "IPCC (Intergovernmental Panel on Climate Change) source category 5 as defined in the 2006 IPCC guidelines for national greenhouse gas inventories". "
Similar variables as upper in cells and includes: **1.** ***alcohols**_due_to_emission_savanna_and_grassland_fires **2.** **2.1** ammonia **2.2** benzene **2.3** butane **2.4** carbon-monoxide **2.5** chlorinated hydrocarbons **2.6** dimethyl sulphide **2.7** elemental (or black) carbon dry aerosol **2.8** ethane **2.9** ethene **2.10** ethers **2.11** ethyne **2.12** formaldehyde **2.13** isoprene **2.14** ketones **2.15** methane **2.16** molecular hydrogene **2.17** nitrogene dioxide **2.18** nitrogene monoxide	kg m^{-2} s^{-1}	Equivalent description as for tendency_of_atmosphere_mass_content_of_alcohols_due_to_emission_from_forest_fires

2.19 NMVOC **2.20** NOx **2.21** organic acids **2.22** particulate organic matter dry aerosol **2.23** pentane **2.24** propane **2.25** propene **2.26** sulfur dioxide **2.27** terpenes **2.28** toluene **2.29** xylene from (A) forest and from (B) savanna&grassland fires.		

While in Tables 3.1.1-3.1.3 we summarized recommended approach for documenting metadata from different impact models in general, in the following part of the subsection we present how were metadata from the impact models applied in the EU-CIRCLE section defined: (1) in Table 3.1.4 metadata for the results of the G.FMIS FIRESIM model are presented, (2) in Table 3.1.5 metadata from the results of the FEPS model are presented, (3) while in Tables 3.1.6 and 3.1.7 metadata for two variables from the HySPLIT models are summarized. Similar approach is recommended for future use of the CIRP related to the impact models' or climate hazard model outputs.

Table 3.1.4 Impact model data: G.FMIS_FIRESIM metadata (cf. EU-CIRCLE D2.3).

Partner	NCSRD
Dimensions	xllcorner x yllcorner x xurcorner x yurcorner x
Variables	Attribute name: STEP Step.Contentdescription: Fire step Step.definition: x minutes Step.Datatype: integer Step.Units: no dimension Attribute name: FLIN Flin.Contentdescription: Fireline Intensity Flin.Datatype: double precision ('double') Flin.Units: kW/m Attribute name: ROS ros.Contentdescription: Fire Rate of Spread ros.Data type: double precision ('double') ros.Units: m/min
External Global Attributes	Type Format: point shape file title: forest fire simulation steps data creation date: dd/mm/yyyy: hh/ss coordinate system: WGS84 world projected simulation start date: dd/mm/yyyy:hh/ss simulation end date: dd/mm/yyyy:hh/ss description of the data set: forest fire simulation under specific meteorological and fuel moisture conditions institute: NSCR"D" source: G-FMIS_FIRESIM software history: input of fuel moisture values and wind measurements

	author: Vassiliki Varela – EU-CIRCLE project
Associated file	Projection file *.prj

Table 3.1.5 Impact model data: Fire Emission Production Simulator (FEPS) metadata. (cf. EU-CIRCLE D2.3).

Partner	NCSRD
Dimensions	lonllcorner x latllcorner x lonurcorner x laturcorner x
Variables	column name: DATE Date.Contentdescription: date of incident Date.Datatype : dd/mm/yyyy Date.Units : date Attribute name: HOUR Hour.Content description: hour of incident Hour.Datatype: double precision ('double') Hour.Units hour Attribute name: ACRES Acres.Content description: affected area in Acres Acres.Data type: double precision ('double') Acres.Units: acres Attribute name: PM2.5 PM2.5.Content description: PM2.5 concentration PM2.5.Data type : double precision ('double') PM2.5. Units :micrograms Attribute name: PLUME_RISE Plume_rise.Content description: average height of plume Plume_rise.Data type: double precision ('double') Plume_rise.Units: meters
External Global Attributes	Type Format : ASCII file title: fire emissions input file for Hysplit data creation date :dd/mm/yyyy: hh/ss coordinate system: WGS84 geographic

	simulation start date: dd/mm/yyyy:hh/ss simulation end date: dd/mm/yyyy:hh/ss description of the data set: smoke/emissions estimation under specific meteorological and fuel conditions institute: NSCR”D” source: FEPS software history: input of fuel and meteo conditions author :Vassiliki Varela – EU-CIRCLE project

Table 3.1.6. Impact model data: HySPLIT Deposition-Ground Level metadata (cf. EU-CIRCLE D2.3).

Partner	NSCRD
Variables	Data type : double precision ('double') Units : μg /m^2 Header description: • YEAR, MO, DA, HR, • LAT, LON, • PARTICLE, • GROUND LEVEL Missing data values: x
External Attributes	Type Format: ASCII / CSV title: cdep[TIMEAVERAGE]_[YYYYMMDD].txt e.g.cdep24HR_20480601.txt concentration, deposition, time interval of average, date reference coordinate system: LAT, LON climatic model: any climatic scenario: any institute: NSCRD source: HySPLIT model references: Stein, A.F., R.R. Draxler, G.D. Rolph, B.J. Stunder, M.D. Cohen, and F. Ngan, 2015: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bull. Amer. Meteor. Soc.*, 96, 2059–2077, https://doi.org/10.1175/BAMS-D-14-00110.1 Comment: Author: Stelios Karozis – EU-CIRCLE project

Table 3.1.7 Impact model data: HySPLIT Concentration-Height metadata (cf. EU-CIRCLE D2.3).

Partner	**NSCRD**
Headers	Data type: double precision ('double') Units: μg/m^3 Header description: • YEAR, MO, DA, HR, • LAT, LON, • PARTICLE, • HEIGHTS Missing data values: x
External Attributes	Type Format — ASCII / CSV title : cdep[TIMEAVERAGE]_[YYYYMMDD].txt e.g.cht24HR_20480601.txt concentration, height, time interval of average, date reference coordinate system: LAT, LON climatic model: any climatic scenario: any institute: NSCRD source: HySPLIT references: Stein, A.F., R.R. Draxler, G.D. Rolph, B.J. Stunder, M.D. Cohen, and F. Ngan, 2015: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bull. Amer. Meteor. Soc.*, 96, 2059–2077, https://doi.org/10.1175/BAMS-D-14-00110.1 Comment: x Author: Stelios Karozis – EU-CIRCLE project

4.2 File-naming recommendation for the impact models' data

Because of the large diversity of the impact models whose results are or could be potentially read or ingested by CIRP, CIRP users are advised to use simple but informative approaches when defining their file names. As for the case of the filenames containing primary and secondary climate data, the filenames of the impact models may contain:

1. short variable name,
2. main forcing datasets,
3. spatial and temporal resolution,
4. spatial location and temporal range,

5. algorithm or model that produced the final results and,
6. the institute or group that created the dataset and files.

Since variety of data formats are expected (with a subset of the most common that are already implemented in the CIRP), XML files containing metadata attributes should be prepared, giving more details about:

1. variable units,
2. standard and long name of the variable,
3. details of the spatial and temporal dimensions,
4. alpha-numerical codes for different flags such as missing values and,
5. any specificity relevant for the efficient analysis such as the averaging operators.

Appendix 1 Example: preparing climate data for the CIRP

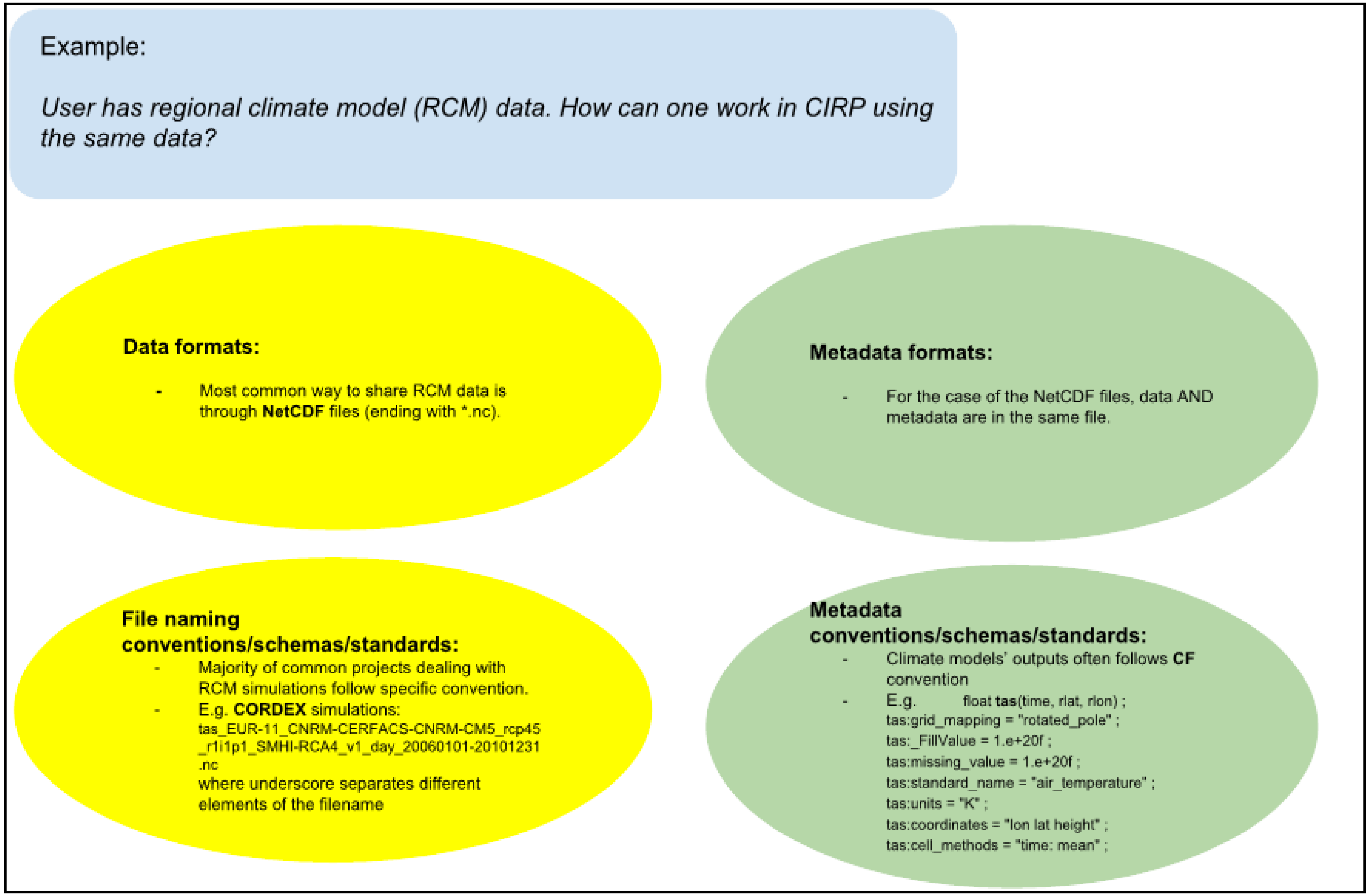

Figure S1.1 Example using CORDEX regional climate model data. General framework presented in Fig. 1.3. Same technical answers have to considered when working with e.g. climate observations (sometimes in NetCDF format, but larger diversity exists), secondary climate information (existing in various formats and produced by different research groups and institutions) and impact models' (or climate hazards) results (existing in various formats and produced by different research and application communities).