CORDEX archive specifications for river variables
for Med-CORDEX

This following is based on the "CORDEX archive specifications for ocean variables" document produced for Med-CORDEX and adapted where needed

Text underlined in green is related to the new specificities adapted to the river component

1. Introduction
same as in CORDEX_archive_specifications_3march2014.pdf

2. File format
same as in CORDEX_archive_specifications_3march2014.pdf

2.1 Netcdf attributes
same as in CORDEX_archive_specifications_3march2014.pdf

3. Grids

3.1 CORDEX domains
The domain acronym has to be 'domain'-'resolution', for example 'domain'-44, 'domain'-22, 'domain'-11, etc. corresponding to the chosen grid spacing. Names of the domains are provided in Table 1. For native coordinate systems data files must contain required grid information in the variable attribute - grid_mapping and in the coordinate variables in accordance with the CF-1.4 or later convention [3]. Examples can be found in Appendix A. It is strongly recommended to provide the geographic latitudes and longitudes of the model grid cell positions as well, in order to facilitate the data analyses. Note that longitudes have to have absolute values as small as possible and monotonic. All variables have to be provided on their native computational grid.
Do we have to define a range of possible resolutions?

Examples to define the River model resolution:
The model resolution is given in degrees.
River grid at an horizontal resolution of 0.22° = 110km x 0.22° = 25km(1/4°)
Examples:
MED-09 = 0.09° = 10km
To choose your MED-XX, take your model mean resolution (ex:12km) and make:
12km/110km = 0.109° = 0.11 → MED-11

3.2 Regular geographic grid

not yet defined for CORDEX-River domains

3.3 Vertical coordinates
All requested variables are two-dimensional.

4. Time coordinates
same as in CORDEX_archive_specifications_3march2014.pdf

5 Data Reference System
same as in CORDEX_archive_specifications_3march2014.pdf

5.1 DRS elements
The Data Reference Syntax (DRS) element values have to 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 are prescribed ('single value'), have to be taken from a fixed list of values ('CV'), have to be registered with the CORDEX ('CV to register'), or can be chosen freely. Note that most elements must have the same value as a mandatory NetCDF attribute. The attribute name is included in the brackets after the element name in these cases.

The elements are (in alphabetical order):

activity (single value; project_id) has to have the value 'CORDEX'.

CMIP5EnsembleMember (CV; driving_model_ensemble_member) identifies the ensemble member of the CMIP5 experiment that produced the forcing data. It has to have the same value in CORDEX as in CMIP5. For evaluation runs it has to be r1i1p1. Invariant fields (frequency=fx) may have the value r0i0p0 or that of the corresponding GCMEnsembleMember attribute.

CMIP5ExperimentName (CV; driving_experiment_name) is either evaluation or the value of the CMIP5 experiment_id of the data used.

Domain (CV; CORDEX_domain) is the name assigned to each of the CORDEX regions and includes a flag for resolution as listed in Table 1. Please refer also to section 3.1.

GCMModelName (CV; driving_model_id) is an identifier of the driving data. The name consists of an institute identifier and a model identifier. For reanalysis driven runs these are ECMWF and a name for the reanalysis data (ERAINT). For runs driven by CMIP5 model data these are the associated CMIP5 institute_id and the CMIP5 model_id. The two parts of the name are separated by a '-' (dash). Note that dashes in either of the two parts are allowed. The CV list of GCMModelName can be found at

http://cordex.dmi.dk/joomla/images/CORDEX/GCMModelName.txt.

Institution (CV; institute_id) is an identifier for the institution that is responsible for the scientific aspects of the CORDEX simulation (RCM configuration, experiments ...). The CV for Institution has to be coordinated in the worldwide CORDEX community. The actual state of the CV is found at

http://cordex.dmi.dk/joomla/images/CORDEX/RCMModelName.txt

New names should be added for the coupled configurations such as ITU or ENEA for example

product (single value) has to have the value output.

RCMModelName (CV to register; model_id) is an identifier of the CORDEX RCM. It consists of the Institution identifier (see above) and a model acronym, connected by a dash (e.g. CNRM-RCSM4 or ENEA-PROTHEUS or ICTP-RegCM-ES). The CV of the RCMModelName has to be coordinated in the worldwide CORDEX community. The actual state of the CV is found at
http://cordex.dmi.dk/joomla/images/CORDEX/RCMModelName.txt
New names should be added for the coupled configurations such as CNRM-RCSM4 or ENEA-PROTHEUS for example

RCMVersionID (free string; rcm_version_id) identifies reruns with perturbed parameters or smaller RCM release upgrades, i.e. equivalent simulations. Major upgrades and improvements should be reflected in the RCMModelName.

Frequency (CV; frequency) is the output frequency indicator: 3hr=3 hourly, 6hr=6 hourly, day=daily, mon=monthly, sem=seasonal, and fx=invariant fields.

StartTime and **EndTime** (build rule) indicate the time span of the file content. The format is YYYY[MM[DD[HH[MM]]]], 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. In accordance with CMIP5, only those digits have to be included that are necessary to indicate the file content. The StartTime and EndTime of instantaneous data are based on the time values of the first and last record in the file, while averaged data use the left and right time_bnds values respectively. It is also allowed to use time values of the first and last records in NetCDF files for averaged data, however. The two dates are separated by a dash. All time stamps refer to UTC. Examples are found in Appendix C. See also section 5.4 for more details.

VariableName (CV) is the name of the target variable in the NetCDF files. The CV is found in column B of sheet 'all' of VR [1]p.

Most of the DRS elements correspond to a mandatory NetCDF global attribute which must have the same value. These pairs of DRS element and global NetCDF attribute are listed in section C of sheet 'Global attributes' of VR [1].

5.2 DRS file naming:
same as in CORDEX_archive_specifications_3march2014.pdf

+ see Appendix C

5.3 DRS Directory structure
same as in CORDEX_archive_specifications_3march2014.pdf

5.4 Time periods for each data file
same as in CORDEX_archive_specifications_3march2014.pdf

The time spans that have to be included into a single file is one year independently from the aggregation (hourly, daily and monthly)

6. Variables to output:
same as in CORDEX_archive_specifications_3march2014.pdf

7. Storage in the Med-CORDEX database:
When files are ready and names checked with the Med-CORDEX coordinators (S. Somot) and the database manager (E. Lombardi), you can upload your files on the Med-CORDEX/HyMeX database (www.medcordex.eu) with your data producer login and password.
Ask Emanuele Lombardi <emanuele.lombardi@enea.it> if you need help.
Then your river dataset will be searchable activating the river realm in the searching tools of the database: https://www.medcordex.eu/search/index.php

8. Metadata for the river model runs:
Metadata are key elements of a reliable dataset. It allows the users to know what are the main characteristics of the model and simulations and also, when needed, the known issues of the dataset. In HyMeX/Med-CORDEX, we decided to write one metedata sheet per simulation. Those sheets remains simple and can be full-filled on-line in the HyMeX database:
http://mistrals.sedoo.fr/HyMeX/Model-Data/
Ask Sophie Cloché <Sophie.Bouffies-Cloche@ipsl.jussieu.fr> or Karim Ramage <Karim.Ramage@ipsl.polytechnique.fr>
if you need help.

You can find example of already completed metadata sheets in:
http://mistrals.sedoo.fr/HyMeX/Plateform-search/?datsType=2
and an example for a coupled model here:
http://mistrals.sedoo.fr/?editDatsId=1312&datsId=1312&project_name=HyMeX
The metadata files are clickable links in the list of runs: for example for the coupled runs of Med-CORDEX
https://www.medcordex.eu/Tabelle_RUNs/Listofruns_Med-CORDEX_phase1_RCSM.pdf

Acknowledgments:
same as in CORDEX_archive_specifications_3march2014.pdf

References:
same as in CORDEX_archive_specifications_3march2014.pdf

Table 1

Area	Short name	Resolution	limits
Mediterranean river basins	MED-05, MED-11, MED-22	Model dependant	Common seasonal cycle for Nile river

Appendix A Examples of Netcdf headers

Examples of global attributes :

:Conventions = "CF" ;
:production = "CHyM hydrological model" ;
:institution = "ICTP(International Centre for Theoretical Physics), Italy, Trieste" ;
:contact = "coppolae@ictp.it" ;
:experiment_id = "evaluation" ;
:driving_model_id = "ECMWF-ERAINT" ;
:frequency = "day" ;
:creation_date = "2017-05-12 09:21:13" ;

Examples of coordinates attributes:
float latitude(y, x) ;
    latitude:units = "degrees_north" ;
    latitude:standard_name = "latitude" ;

float longitude(y, x) ;
    longitude:units = "degrees_east" ;
    longitude:standard_name = "longitude" ;

Examples of time attributes:

float time(time) ;
    time:standard_name = "time" ;
    time:units = "seconds since 1961-01-01 00:00:00" ;

Examples of 2D variable attributes : should at least contain : zos:units = "m3s-1" ;

rivo:_FillValue = 1.e+20f ;
rivo:standard_name = "water_flux_to_downstream" ;
rivo:valid_min = -1.e+20f ;
rivo:valid_max = 1.e+20f ;
rivo:cell_method = "time:mean" ;
rivo:coordinates = "latitude longitude time" ;

Appendix B List of CORDEX variables

CMIP6	1D/2D/3D files	Variable standard name	Variable name	Storage Frequency MM,DD,HH	Units	CORE/Tier1/Tier2	notes
y	2D	water_flux_to_down stream	rivo	MM,DD,HH	m3 s-1	CORE/Tier1/Tier2	
y	2D	water_flux_from_up stream	rivi	MM,DD,HH	m3 s-1	CORE/Tier1/Tier2	
NO	2D	total_runoff_flux	mrrot	MM,DD,HH	kg m-2 s-1	CORE/Tier1/Tier2	
y	2D	surface_runoff_flux	mrros	MM,DD,HH	kg m-2 s-1	CORE/Tier1/Tier2	
NO	2D	River Basin Numbers	basins	static	#		

A static field with the identification of the basins could be beneficial for analysis purpose

Appendix C Examples of file names

to be completed