CF conventions for RAPID inputs/outputs

By Cédric H. David (cedric.david@jpl.nasa.gov),

06 Jan 2016, modified 07 Jan 2016, 02 Feb 2016, 19 Feb 2016

Introduction

The largest inputs and outputs of RAPID are stored in netCDF files with two dimensions: river reach ID, and time. The lack of metadata in these netCDF files has hindered many useful applications RAPID.

Previous data format

The format for RAPID outputs (valid until tag 20151119, included) is the following:

Water inflow from land and subsurface upstream of each river reach (input)

```
dimensions:
      Time = UNLIMITED ;  // (11680 currently)
      COMID = 5175 ;
variables:
      float m3_riv(Time, COMID) ;
```

River discharge downstream of each river reach (output)

```
dimensions:
      Time = UNLIMITED ;  // (11680 currently)
      COMID = 5175 ;
variables:
      float Qout(Time, COMID) ;
      int COMID(COMID) ;
```

Volume of water inside each river reach (output)

```
dimensions:
      Time = UNLIMITED ;  // (11680 currently)
      COMID = 5175 ;
variables:
      float V(Time, COMID) ;
      int COMID(COMID) ;
```

Updated data format

The proposed format for RAPID outputs is presented in an example below (red font for changes, grey font for new static additions, blue font is populated from namelist, green font is populated from metadata in water inflow file, brown font is automatically at runtime):

Water inflow from land and subsurface upstream of each river reach (input)

```
dimensions:
      time = UNLIMITED ;  // (11680 currently)
      rivid = 5175 ;
      nv = 2 ;
variables:
      float m3_riv(time, rivid) ;
        m3_riv:long_name = "accumulated external water volume inflow upstream of each river reach" ;
        m3_riv:units = "m3" ;
        m3_riv:coordinates = "lon lat" ;
        m3_riv:grid_mapping = "crs" ;
        m3_riv:cell_methods = "time: sum" ;
      int rivid(rivid) ;
        rivid:long_name = "unique identifier for each river reach" ;
        rivid:units = "1" ;
        rivid:cf_role = "timeseries_id" ;
      int time(time) ;
        time:standard_name = "time";
        time:long_name = "time" ;
        time:units = "seconds since 1970-01-01 00:00:00 +00:00" ;
        time:axis = "T";
        time:calendar = "gregorian";
        time:bounds = "time_bnds";
      int time_bnds(time, nv) ;
      double lon(rivid) ;
        lon:standard_name = "longitude";
        lon:long_name = "longitude of a point related to each river reach";
        lon:units = "degrees_east";
        lon:axis = "X";
      double lat(rivid) ;
        lat:standard_name = "latitude";
        lat:long_name = "latitude of a point related to each river reach" ;
        lat:units = "degrees_north" ;
        lon:axis = "Y";
      int crs ;
        crs:grid_mapping_name = "latitude_longitude";
        crs:semi_major_axis = "6378137";
        crs:inverse_flattening = "298.257222101";

// global attributes:
        :Conventions = "CF-1.6";
        :title = "RAPID data corresponding to the San Antonio and Guadalupe Basins, TX";
        :institution = "Jet Propulsion Laboratory, California Institute of Technology";
        :source = "RRR: v1.x.y, runoff: NLDAS2_VIC0125_3H_20100101_20131231.nc";
        :history = "date_created: 2016-01-06T14:11:31-08:00";
        :references = "https://github.com/c-h-david/rrr/";
        :comment = "None";
        :featureType = "timeSeries";
```

River discharge downstream of each river reach (output)

```
dimensions:
      time = UNLIMITED ;  // (11680 currently)
      rivid = 5175 ;
      nv = 2 ;
variables:
      float Qout(time, rivid) ;
        Qout:long_name = "average river water discharge downstream of each river reach" ;
        Qout:units = "m3 s-1" ;
        Qout:coordinates = "lon lat" ;
        Qout:grid_mapping = "crs" ;
        Qout:cell_methods = "time: mean" ;
      int rivid(rivid) ;
        rivid:long_name = "unique identifier for each river reach" ;
        rivid:units = "1" ;
        rivid:cf_role = "timeseries_id" ;
      int time(time) ;
        time:standard_name = "time";
        time:long_name = "time" ;
        time:units = "seconds since 1970-01-01 00:00:00 +00:00" ;
        time:calendar = "gregorian";
        time:axis = "T";
        time:bounds = "time_bnds";
      int time_bnds(time, nv) ;
      double lon(rivid) ;
        lon:standard_name = "longitude";
        lon:long_name = "longitude of a point related to each river reach";
        lon:units = "degrees_east";
        lon:axis = "X";
      double lat(rivid) ;
        lat:standard_name = "latitude";
        lat:long_name = "latitude of a point related to each river reach" ;
        lat:units = "degrees_north" ;
        lon:axis = "Y";
      int crs ;
        crs:grid_mapping_name = "latitude_longitude";
        crs:semi_major_axis = "6378137";
        crs:inverse_flattening = "298.257222101";

// global attributes:
        :Conventions = "CF-1.6";
        :title = "RAPID data corresponding to the San Antonio and Guadalupe Basins, TX";
        :institution = "Jet Propulsion Laboratory, California Institute of Technology";
        :source = "RAPID: v1.5.0, water inflow: m3_riv_San_Guad_20100101_20131231_utc.nc";
        :history = "date_created: 2016-01-07T07:23:01-08:00";
        :references = "https://github.com/c-h-david/rapid/, http://dx.doi.org/10.1175/2011JHM1345.1";
        :comment = "None";
        :featureType = "timeSeries";
```

Volume of water inside each river reach (output)

```
dimensions:
      time = UNLIMITED ;  // (11680 currently)
      rivid = 5175 ;
      nv = 2 ;
variables:
      float V(time, rivid) ;
        V:long_name = "average river water volume inside of each river reach" ;
        V:units = "m3" ;
        V:coordinates = "lon lat" ;
        V:grid_mapping = "crs" ;
        V:cell_methods = "time: mean" ;
      int rivid(rivid) ;
        rivid:long_name = "unique identifier for each river reach" ;
        rivid:units = "1" ;
        rivid:cf_role = "timeseries_id" ;
      int time(time) ;
        time:standard_name = "time";
        time:long_name = "time" ;
        time:units = "seconds since 1970-01-01 00:00:00 +00:00" ;
        time:calendar = "gregorian";
        time:axis = "T";
        time:bounds = "time_bnds";
      int time_bnds(time, nv) ;
      double lon(rivid) ;
        lon:standard_name = "longitude";
        lon:long_name = "longitude of a point related to each river reach";
        lon:units = "degrees_east";
        lon:axis = "X";
      double lat(rivid) ;
        lat:standard_name = "latitude";
        lat:long_name = "latitude of a point related to each river reach" ;
        lat:units = "degrees_north" ;
        lon:axis = "Y";
      int crs ;
        crs:grid_mapping_name = "latitude_longitude";
        crs:semi_major_axis = "6378137";
        crs:inverse_flattening = "298.257222101";

// global attributes:
        :Conventions = "CF-1.6";
        :title = "RAPID data corresponding to the San Antonio and Guadalupe Basins, TX";
        :institution = "Jet Propulsion Laboratory, California Institute of Technology";
        :source = "RAPID: v1.5.0, water inflow: m3_riv_San_Guad_20100101_20131231_utc.nc";
        :history = "date_created: 2016-01-07T07:23:01-08:00";
        :references = "https://github.com/c-h-david/rapid/, http://dx.doi.org/10.1175/2011JHM1345.1";
        :comment = "None";
        :featureType = "timeSeries";
```