A brief introduction to the netCDF Climate and Forecast (CF) metadata conventions

David Hassell
NCAS
University of Reading

Jonathan Gregory
NCAS
University of Reading
UK Met Office Hadley Centre

CF meeting, online, 2020-06-09

netCDF

- A Network Common Data Form (netCDF) file is a good medium for storing and sharing **self-describing array-orientated scientific data**, via standardised software libraries developed and maintained at Unidata.
    - In this context, “self-describing” means that a file contains, for each data array, an associated description of what it represents scientifically, i.e. **metadata**.
- NetCDF is very flexible and, by design, imposes very few rules on the metadata, so that interpreting the description of the data is generally left to the creators and users of the file.
    - There is a problem if people assume different interpretations ...

The CF conventions

CF-netCDF metadata conventions

http://cfconventions.org

- The CF-netCDF conventions (“CF conventions”) define wide-ranging rules for netCDF metadata for describing **geoscientific data** in netCDF files, with the aims of
  - providing a **definitive description** of what the data in each variable represents, and the spatiotemporal and other dimensional properties of the data
  - enabling users to **identify comparable data** held in different files
  - facilitating the **development of software** to extract, process, analyse, and display the data.

Geoscientific uses

- model-generated climate and forecast data for the atmosphere, surface and ocean
- observational datasets
- derived products
- satellite data
- quantities dependent on biological taxa
- etc.

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

netCDF

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
// global attributes:
                :Conventions = "CF-1.8" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

netCDF

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
// global attributes:
                :Conventions = "CF-1.8" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

netCDF

CF

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
// global attributes:
                :Conventions = "CF-1.8" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

CF

The CF conventions

```
$ ncdump file.nc
netcdf file {
dimensions:
        lat = 5 ;
        lon = 8 ;
        bounds2 = 2 ;
variables:
        double lat_bnds(lat, bounds2) ;
        double lat(lat) ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:bounds = "lat_bnds" ;
        double lon_bnds(lon, bounds2) ;
        double lon(lon) ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:bounds = "lon_bnds" ;
        double time ;
                time:units = "days since 2018-12-01" ;
                time:standard_name = "time" ;
        double x(lat, lon) ;
                x:project = "research" ;
                x:standard_name = "specific_humidity" ;
                x:units = "1" ;
                x:coordinates = "time" ;
                x:cell_methods = "area: mean" ;
// global attributes:
                :Conventions = "CF-1.8" ;
```

```
data:

 lat_bnds =
  -90, -60,
  -60, -30,
  -30, 30,
  30, 60,
  60, 90 ;

 lat = -75, -45, 0, 45, 75 ;

 lon_bnds =
  0, 45,
  45, 90,
  90, 135,
  135, 180,
  180, 225,
  225, 270,
  270, 315,
  315, 360 ;

 lon = 22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5 ;

 time = 31 ;

 x =
  0.007, 0.034, 0.003, 0.014, 0.018, 0.037, 0.024, 0.029,
  0.023, 0.036, 0.045, 0.062, 0.046, 0.073, 0.006, 0.066,
  0.11, 0.131, 0.124, 0.146, 0.087, 0.103, 0.057, 0.011,
  0.029, 0.059, 0.039, 0.07, 0.058, 0.072, 0.009, 0.017,
  0.006, 0.036, 0.019, 0.035, 0.018, 0.037, 0.034, 0.013 ;
}
```

CF-compliant

The CF conventions

General principles for the development of CF

- Data should be **self-describing**—no external tables needed to interpret it
- Development is driven by **use-cases** , i.e. only cater for things we know we need
- **Practicable** for both data-writers and data-readers
- Metadata **easily parsed** by humans as well as programs
- **Minimise redundancy** and any other possibilities for silly mistakes

The CF conventions

Other aspects of CF to be covered in this session

- Standard names
- Governance
- Change procedures

NCAS work on CF is additionally supported by the IS-ENES3 project funded by the European Commission under grant agreement N°824084