

# An introduction to the CF-netCDF metadata conventions and the CF data model

**David Hassell**

NCAS  
University of Reading

**Jonathan Gregory**

NCAS  
University of Reading  
UK Met Office Hadley Centre

2020 ESIP summer meeting, online, 2020-07-22



## 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.
  - self-describing = the dataset also contains the **metadata** for each data array
- 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 ...

## 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
  - enabling users to **identify comparable data** held in different files
  - facilitating the **development of software** to extract, process, analyse, and display the data.

# 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

## Understanding and interpreting CF

- Read the latest release of the CF conventions (all 183 pages!)
- The meaning is often obfuscated by netCDF encoding details, e.g.

*In the second format, it is a blank-separated list of words "<gridMappingVariable>: <coordinatesVariable> [<coordinatesVariable> ...] [<gridMappingVariable>: <coordinatesVariable>...]", which identifies one or more grid mapping variables, and with each grid mapping associates one or more coordinatesVariables, i.e. coordinate variables or auxiliary coordinate variables.*

## Understanding and interpreting CF

- Read the latest release of the CF conventions (all 183 pages!)
- The meaning is often obfuscated by netCDF encoding details, e.g.

*In the second format, it is a blank-separated list of words "<gridMappingVariable>: <coordinatesVariable> [<coordinatesVariable> ...] [<gridMappingVariable>: <coordinatesVariable>...]", which identifies one or more grid mapping variables, and with each grid mapping associates one or more coordinatesVariables, i.e. coordinate variables or auxiliary coordinate variables.*

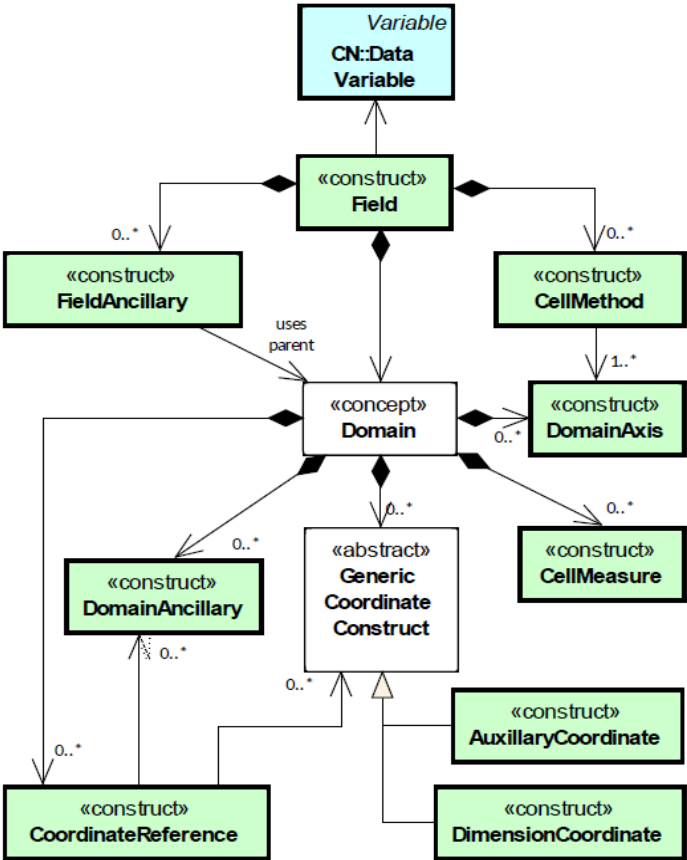
- The recent inclusion of the CF data model helps to understand CF

- **Definition:** The CF data model identifies distinct elements of CF and describes how they relate to each other and the real world
- **Benefits**
  - Allows for a better understanding of CF
  - Guides enhancements to CF
  - Encourages the creation of CF-complete software

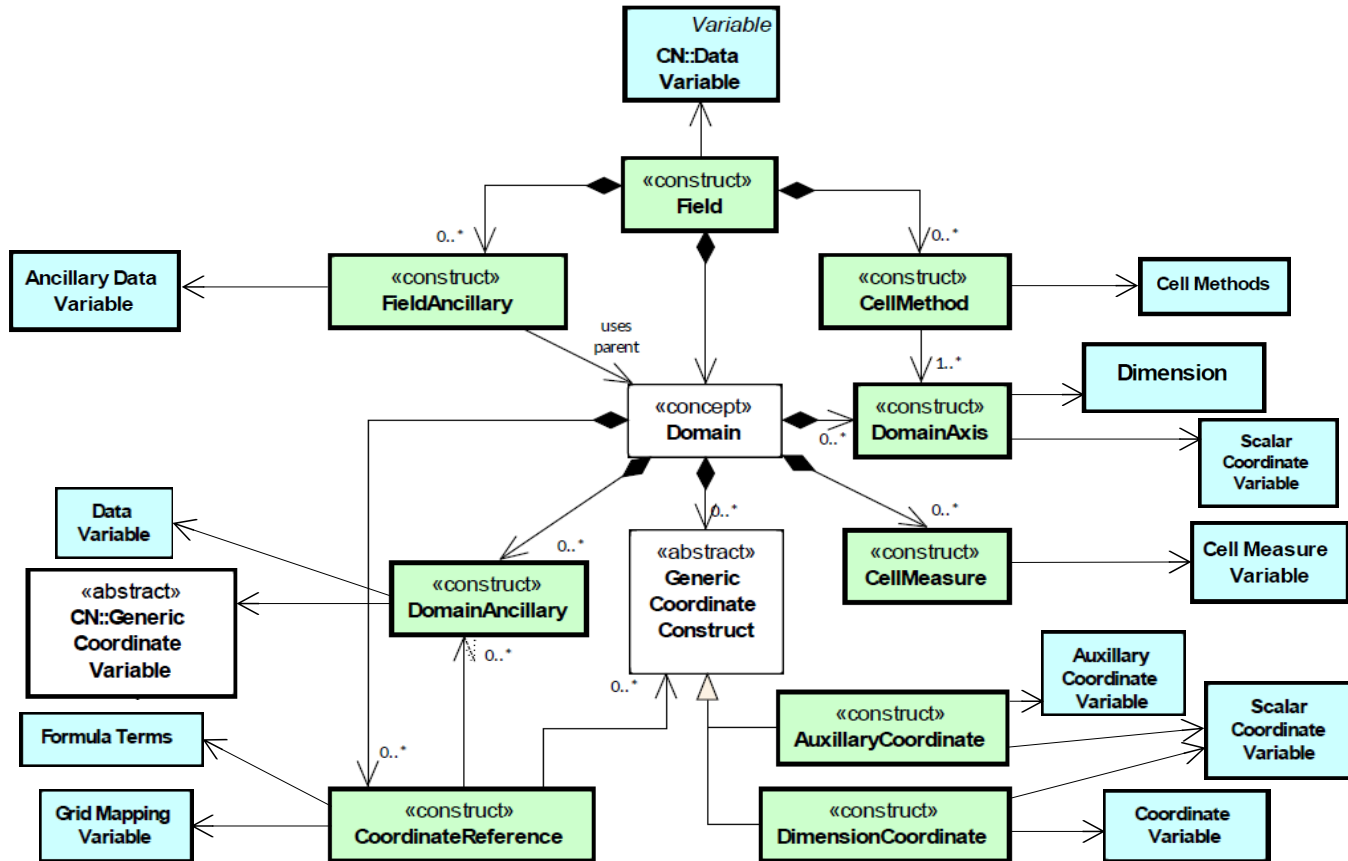
## Design criteria for the CF data model

- **Necessary**
  - The data model should not introduce additional elements not presently needed or used by CF
- **Sufficient**
  - The data model should be composed of a minimal set of elements that are sufficient for accommodating all aspects of the CF
- **Logical**
  - The data model should be a logical data model, i.e. **independent of any encoding**
  - Sets the groundwork for applying CF to other file formats, if the need arises

# The CF data model



# The CF data model



## Find out more

- **A data model of the Climate and Forecast metadata conventions (CF-1.6) with a software implementation (cf-python 2.1)**
  - Geoscientific Model Development (2017)
  - <https://doi.org/10.5194/gmd-10-4619-2017>
  
- **CF-1.9**
  - Draft version at <https://cfconventions.org/cf-conventions/cf-conventions.html>