National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

A brief overview of the CF conventions

David Hassell

National Centre for Atmospheric Science

University of Reading

KNMI, 15 March 2017

data

`3.56, 6.78, ..., -0.32, 1.86`

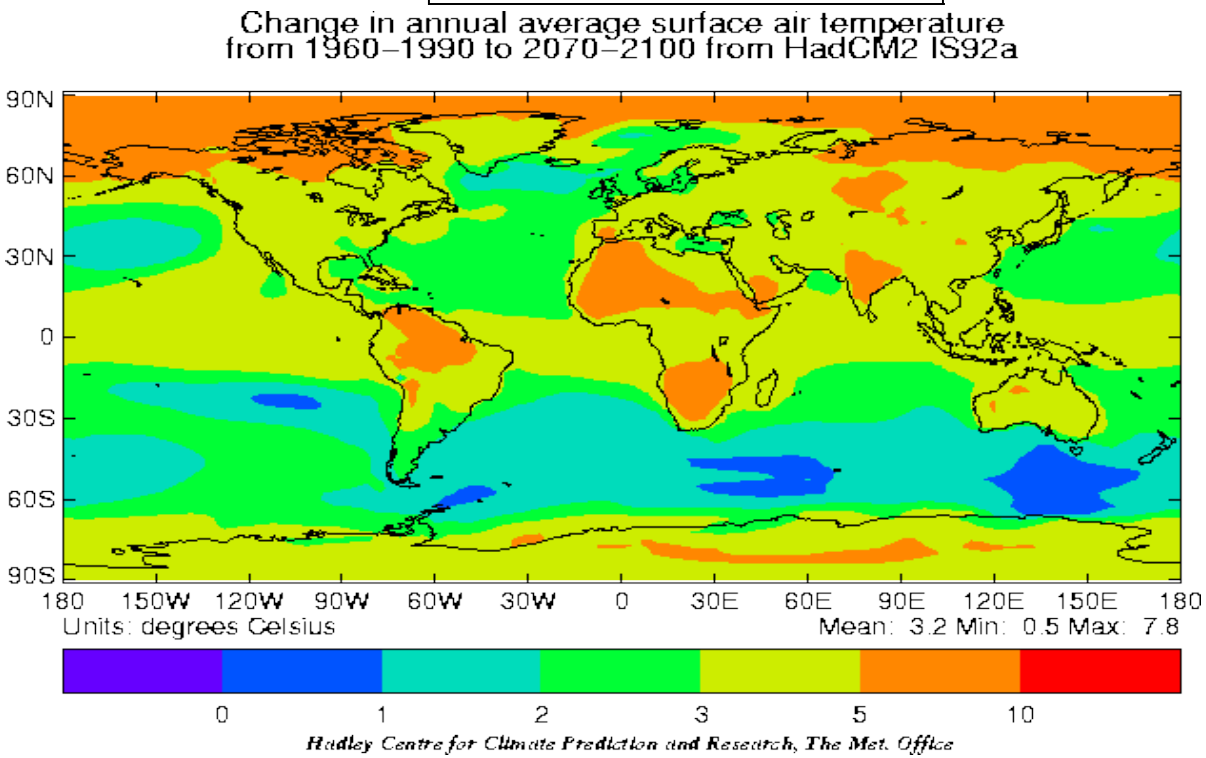
Data, metadata and the CF conventions

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

data

```
3.56, 6.78, ..., -0.32, 1.86
```

data + metadata

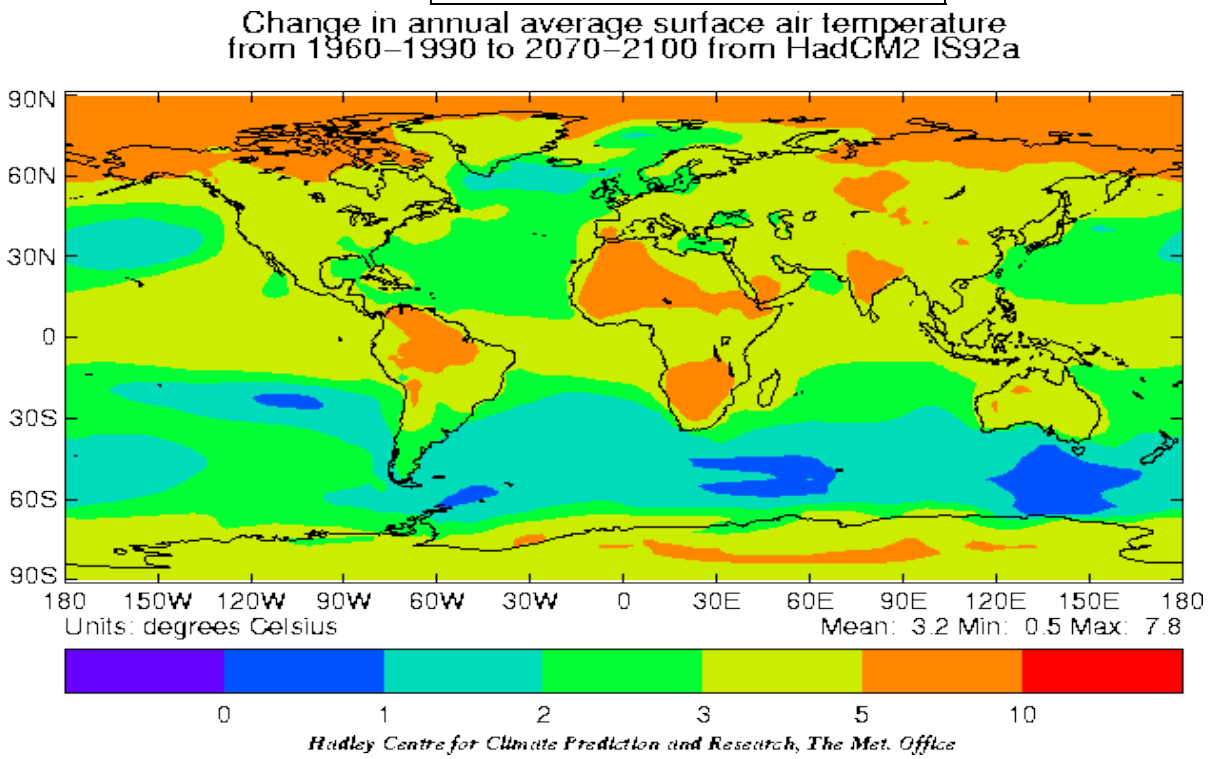


National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

data

```
3.56, 6.78, ..., -0.32, 1.86
```

data + metadata



- The metadata need to follow recognised conventions (CF) so that they can be understood by anyone or anything looking at them

CF conventions

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- A set of rules for storing self-describing geoscientific data in netCDF files

- CF describes a netCDF variable (V) containing scientific data that is a discretised within a domain (d)

$$V = V(d) = V(d(t, z, y, x, \ldots))$$

- The data variable is associated with metadata in the forms of “simple” attributes (e.g. the standard name) and other variables in the file (such as coordinate variables)

NetCDF variables and attributes defined by CF

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- **Data variable** Data variable (V) discretised within a domain (d)

$$V = V(d) = V(d(t, z, y, x, ...))$$

- **Dimension** Independent axis of a domain
- **Coordinate variable** Unique coordinates for a single dimension
- **Scalar coordinate variable** A coordinate for an unspecified dimension
- **Auxiliary coordinate variable** Alternative coordinates for any dimensions
- **Boundary variable** Cell vertices
- **Grid mapping variable** Horizontal coordinate system
- **Formula terms attribute** Vertical coordinate system
- **Cell measure variable** Cell areas or volumes
- **Ancillary data variable** Descriptive metadata that depends on the domain
- **Standard name attribute** Physical description of variables
- **Cell methods attribute** Variation within cells
- ...

Example CF-netCDF file

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

```
netcdf file {
dimensions:
    lat = 106 ;
    lon = 110 ;
    bounds = 2 ;
variables:
    double time ;
        time:standard_name = "time" ;
        time:units = "days since 2016-12-01" ;
        time:calendar = "Gregorian" ;
        time:bounds = "time_bounds" ;
    double lat(lat) ;
        y:standard_name = "latitude" ;
        y:units = "degrees" ;
        y:bounds = "lat_bounds" ;
    double lon(lon) ;
        x:standard_name = "longitude" ;
        x:units = "degrees" ;
        x:bounds = "lon_bounds" ;
    double time_bounds(bounds) ;
    double lat_bounds(lat, bounds) ;
    double lon_bounds(lon, bounds) ;
    double lon(y, x) ;
        lon:standard_name = "longitude" ;
        lon:units = "degrees_east" ;
    double lat(y, x) ;
        lat:standard_name = "latitude" ;
        lat:units = "degrees_north" ;
    double temp(lat, lon) ;
        temp:standard_name = "air_temperature" ;
        temp:units = "K" ;
        temp:cell_methods = "time: mean" ;
        temp:coordinates = "time" ;
    double wind(lat, lon) ;
        wind:standard_name = "eastward_wind" ;
        wind:units = "m s-1" ;
        wind:cell_methods = "time: maximum" ;
        wind:coordinates = "time" ;
```

The standard name attribute

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- For the systematic identification of the physical quantity contained in variables
- Permissible values are listed in the standard name table (**`http://cfconventions.org/standard-names.html`**), which includes precise definitions
- They answer the question, "*What does this mean?*", rather than the question, "*What do you call this?*"
  - "precipitable water" has the standard name of `atmosphere_mass_content_of_water_vapor`
  - there is no name for plain "potential temperature", since we have to distinguish it in air (`air_potential_temperature`) and in sea water (`sea_water_potential_temperature`)
- CF also has a "long name" attribute which is not standardised

Standard name: units

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Each standard name implies particular physical dimensions (mass, length, time and other dimensions corresponding to SI base units) expressed as a "canonical unit"
  - `large_scale_rainfall_amount` (kg m^{-2})
  - `large_scale_rainfall_flux` (kg m^{-2} s^{-1})
  - `large_scale_rainfall_rate` (m s^{-1})
  - `number_of_days_with_air_temperature_below_threshold` (1)
  - `spell_length_of_days_with_air_temperature_below_threshold` (day)

Standard name: additional constraints

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- Some standard names require the existence of additional metadata
  - `downwelling_radiance_per_unit_wavelength_in_air` requires there to be a coordinate variable with the standard name of `radiation_wavelength`
  - `number_of_days_with_air_temperature_above_threshold` must have a coordinate variable with the standard name of `air_temperature`

Variation within cells

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- CF describes variation within cells by use of the “cell methods” attribute
  - `time: maximum`
  - `area: mean`
  - `time: mean (interval: 1 day)`
  - `time: mean (El Nino years)`
  - `area: mean time: maximum`

More complex cell methods: climatological statistics

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

- A series of data points representing sets of time intervals which are not contiguous
  - Corresponding portions of the annual cycle in a set of years, e.g. decadal averages for January:

```
time: mean within years time: mean over years
```

  - Corresponding portions of days, e.g. the average diurnal cycle in April 1997:

```
time: range within days time: mean over days
```

  - Both at once, e.g. the average winter daily minimum temperature from the years 1961 to 1990:

```
time: minimum within days time: mean over days
time: mean over years
```

  - Number of frost days in winter

```
time: minimum within days time: sum over days
```