

MATLAB Introduction part 4
File Input/Output
November 17

File Formats

Type	Advantage	Disadvantage
Text (ASCII)	Portable (especially between software), human readable	Inefficient (both in file size and read/write speed)
MATLAB native (mat-file)	By far easiest for read/write within MATLAB. Portable within MATLAB, self-descriptive	Not easily portable outside of MATLAB
Binary	Fast, simple	Not easily portable
NetCDF, HDF5	Fast, portable, easy with compound data. Portable, self-descriptive (need library is installed)	Complex, some learning curve (interface is primitive)

MATLAB file examples

- Get KGRB.txt from IDL examples
 - Interactive examples...

NetCDF

- Very portable (need library installed)
- MATLAB has a built-in reader
- Self-describing, can contain useful metadata inside dataset attributes:
 - Engineering units for data arrays
 - Valid ranges
 - Any other arbitrary descriptor (attributes can be anything)
- Primary use in UCAR/NCAR (NCEP reanalysis, CESM model runs, etc), ACRF

NetCDF

```
$ ncdump -h rhum.mon.ltm.nc
netcdf rhum.mon.ltm {
dimensions:
    lon = 144 ;
    lat = 73 ;
    level = 8 ;
    time = UNLIMITED ; // (12 currently)
variables:
    float level(level) ;
        level:units = "millibar" ;
        level:long_name = "Level" ;
        level:positive = "down" ;
        level:GRIB_id = 100s ;
        level:GRIB_name = "hPa" ;
        level:actual_range = 1000.f, 10.f ;
        level:axis = "Z" ;
    float lat(lat) ;
        lat:units = "degrees_north" ;
        lat:actual_range = 90.f, -90.f ;
    .....
```

HDF

- Similar to NetCDF (self-describing, metadata in attributes, etc)
- Adds support for structures
- Internal “group” structure – just like directories/subdirectories in a computer's filesystem
- Commonly used for satellite data products from NASA (MODIS, AIRS, other A-train)
- NetCDF is more used in atm. sci., HDF is used more science-wide.