A Quick Guide to the Reanalysis Datasets

David Smith

smith.dav@husky.neu.edu

This documentation is mainly meant to be a collection of information about the four reanalysis datasets that are contained on the CMIP5 Community Server. These datasets have existing documentation and summaries that users are encouraged to read prior to use of data from the dataset, and the documentation will be linked in the entries below.

Datasets

NCEP/NCAR Reanalysis I

- *Timeframe: 1948-present*
- *Spatial Coverage: either 2.5°x 2.5° global grid (144x73), or a T62 Gaussian grid with 192x94 points (depending on whether the file directory with grid level name contains gauss)*
- *Frequencies: Monthly means, Daily means*
- *Updated: Monthly*
- *Pressure levels: 1000,925,850,700,600,500,400,300,250,200,150,100,70,50,30,20,10 (17 levels)*
  - *Omega only goes to 100mb, and the humidities only go to 300mb*
- *Citation Info: https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html#citations*
- *NCAR/UCAR ClimateDataGuide: https://climatedataguide.ucar.edu/climate-data/ncep-ncar-r1-overview*

Flux data (data provided on a Gaussian grid) can have one of several different meanings with regards to the reference time, depending on the variable. There is currently no 6-hour timescale data made available on the community server, but the daily averages are derived from the 6-hourly data, so it is still important to recognize how the variable is handled. Some variables are forecasts that are valid 6 hours after the reference time, making the daily mean a mean of 6-hour forecasts. Min/max temperature variables are 6-hour hindcasts, while most radiation variables are 6-hour averages. Review the documentation here: https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surfaceflux.html to check particular variables. Note that any data provided on a 2.5°x 2.5° global grid is instantaneous at the 6 hour frequency.

For non-Gaussian gridded data, data from **1948-1957** *is slightly different than after that year. The inputs available in that era in the model were instead at 3Z, 9Z, 15Z, and 21Z, so the data at times 0Z, 6Z, 12Z, and 18Z were forecasted to make the period 3 hours.* **Only the 0Z, 6Z, 12Z, AND 18Z forecasted data were used to make the daily and monthly mean files.**

NOAA-CIRES 20th Century Reanalysis, Version 2c

- *Timeframe: 1851-2014*
- *Spatial Coverage: either 2°x 2° global grid (180x91), or a T62 Gaussian grid with 192x94 points*
- *Frequencies: Monthly means, Daily means*
- *Updated: Irregularly (Timeframe does not expand)*

- *Pressure levels: 1000,950,900,850,800,750,700,650,600,550,500,450,400,350,300,250,200,150,100,70,50, 30,20,10 (24 levels)*
- *Citation Information:*
  https://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2c.pressure.html#citations
- *NCAR/UCAR ClimateDataGuide: https://climatedataguide.ucar.edu/climate-data/noaa-20th-century-reanalysis-version-2-and-2c*

*Similar to NCEP/NCAR Reanalysis I, data provided on the Gaussian grid is a derivative of forecasts rather than instantaneous analysis. All Gaussian variables for this dataset have their daily and monthly mean files derived from **3-hour forecasts** for the relevant day or month. For details about the model, boundary conditions, and other dataset details, consult*
https://www.esrl.noaa.gov/psd/data/gridded/data.20thC_ReanV2c.subsurface.html#detail.

ERA-Interim

- *Timeframe: 1979-present (3-month lag)*
- *Spatial Coverage: $0.75° x 0.75°$*
- *Frequencies: Monthly means, 6 hourly analysis, 12 hourly forecasts*
- *Updated: Monthly*
- *Pressure levels (Monthly):*
  1000,975,950,925,900,875,850,825,800,775,750,700,650,600,550,500,450,400,350,300,250,225,200,175, 150,125,100,70,50,30,20,10,7,4,3,2,1
- *Pressure levels (6-hourly): 975,950,850,750,500,250,200,100,50*
- *Citation Info: https://software.ecmwf.int/wiki/display/CKB/How+to+reference+or+acknowledge+ERA-Interim+data+in+a+publication*
- *NCAR/UCAR ClimateDataGuide: https://climatedataguide.ucar.edu/climate-data/era-interim*

For ERA-Interim, data frequencies and the meaning of the reference time are related to whether the variable is an analysis variable or a forecast variable. Analysis variables are available at one instantaneous value every 6 hours when daily files are made available, while forecast variables represent a 12- hour forecast valid up to the reference time. For example, the first point of a daily forecast variable's file will have a reference time of 12Z on January 1st, and represents the 12-hour forecast valid from 0Z-12Z. To check whether a variable is analysis or forecast, refer to the documentation here:
https://docs.google.com/spreadsheets/d/1CM7O7lbe2DpLHEixP78uV2c0Sm3jie96AIps0SfZ4zE

ERA-20C

- *Timeframe: 1900-2010*
- *Spatial Coverage: $0.75° x 0.75°$*
- *Frequencies: Monthly means, 6 hourly analysis, daily forecasts*
- *Updated: Irregularly (Timeframe does not expand)*

- *Pressure levels (Monthly):*
  1000,975,950,925,900,875,850,825,800,775,750,700,650,600,550,500,450,400,350,300,250,225,200,175,
  150,125,100,70,50,30,20,10,7,4,3,2,1
- *Pressure levels (6-hourly): 975,950,850,750,500,250,200,100,50*
- *Citation Info: https://reanalyses.org/atmosphere/era-20c-references*
- *NCAR/UCAR ClimateDataGuide: https://climatedataguide.ucar.edu/climate-data/era-20c-ecmwfs-atmospheric-reanalysis-20th-century-and-comparisons-noaas-20cr*

ERA-20C works essentially the same as ERA-Interim. The one major difference with the data files is in the way forecast variables work. Forecasts are made once per day starting at 6Z. Therefore, for ERA-20C forecast variables the first reference time will be 6Z on January 2nd, and will have the 24-hour forecast made from January 1st at 6Z.