Overview of the Reanalysis Project Dataset
prepared by Katherine K. Hirschboeck 1-22-03

The following summary information is from **The NCEP/NCAR 40-Year Reanalysis Project** (1996) by E. Kalnay and others. *Bulletin of the American Meteorological Society*, 77: 437-471

Summary of Key Concepts

NCEP = National Center for Environmental Prediction (formerly known as the National Meteorological Center or NMC) Visit: http://www.ncep.noaa.gov/ to find out what this center does.
NCAR = National Center for Atmospheric Research   Visit: http://www.ncar.ucar.edu/ncar/ to find out what this center does.
CDC = Climate Diagnostics Center Visit: http://www.cdc.noaa.gov/ to find out what this center does
CPC = Climate Prediction Center Visit: http://www.cpc.ncep.noaa.gov/index.html to find out what this center does.

Reanalysis Project. The purpose of the Project was to reanalyze historical atmospheric data using state-of-the-art models. The Project started in 1989 at NCEP with the goal of building a "Climate Data Assimilation System" (CDAS). (A data assimilation can be defined as a process that "makes similar" or alters by assimilation; a process of receiving new data in conformity with what is already available.) The motivation for the Project was to remove the apparent "climate changes" that resulted from the many adjustments in the numerical weather prediction operational systems which were made over time in an effort to improve forecasts. The Project has produced a research quality dataset suitable for many uses, including weather and short-term climate research, and it is an ongoing activity, i.e, there is currently a Reanalysis 2 in the works. The unique characteristics of the project are the length of the period covered (now 1948- present for most variables) and the assembly of a very comprehensive *observational* database. The observations include: balloon soundings (global rawindsonde data), surface marine data (from buoys), aircraft data, surface synoptic data (e.g., surface gauge measurements, etc.) and satellite measurements. These are all real observations, not output from a numerical model.

Model Resolution. The assimilated data are computed by the reanalysis model at individual gridpoints to make gridded fields extending horizontally over the whole globe at **28 different levels** in the atmosphere. (Some of these levels correspond to the "mandatory" pressure height level at which soundings are taken, e.g., 1000, 850, 700, 500, 250 mb, etc.) The horizontal resolution of the gridpoints is based on the **T62 model resolution** (T62 = "Triangular 62-waves truncation") which is a grid of 192 x 94 points, equivalent to an average horizontal resolution of a gridpoint every 210 km**.** The pressure level data are saved on a 2.5° latitude-longitude grid. *Note that the gridpoints for computed model output are more numerous and much closer together in the mid and high latitudes, and fewer and farther apart over the low latitudes.*

Reanalysis Output Fields. The gridded output fields computed for different variables have been classified into four classes, A, B, C, and D depending on the relative influence of (1) the observational data, and (2) the model, on the gridded variable. *[See **Appendix A** for a list of the A,B,C, & D variables]*

> **Class A** = the *most reliable* class of variables; "analysis variable is strongly influenced by observed data;" value is closest to a real observation Class A variables: **mean sea level pressure**, **geopotential height** (i.e. height of 500 mb surface, 700 mb surface, etc.), **air temperature**, **wind** (expressed as two vectors dimensions: zonal = **u wind** (west-east) and meridional = **v wind** (north-south), and **vorticity** (a measure of rotation)

Class B = the next most reliable class of variables; "although some observational data directly affect the value of the variable, the model also has a very strong influence on the output values." Class B variables: **surface pressure, surface temperature** (and near-surface **2-m temperature**) , **max and min temperature, vertical velocity**, near-surface wind (**u & v wind at 10 m**), **relative humidity, mean relative humidity, precipitable water content,** and **snow cover**

Class C = the *least reliable* class of variables; NO observations directly affect the variable and it is derived solely from the model computations; forced by the model's data assimilation process, not by any real data. Class C variables: **precipitation, snow depth, soil wetness and soil temperature, surface runoff, cloud fraction** (% high, middle, low), **cloud forcing, skin temperature** (temperature of the "skin," or surface, as in sea surface temperature), **surface wind stress, gravity wind drag,** and **latent and sensible heat fluxes** from surface or top of the atmosphere.

Class D = not really a variable but a mean field that is obtained from climatological values and does not depend on the model (e.g., surface roughness)

IMPORTANT: "the user should exercise caution in interpreting results of the reanalysis, especially for variables classified in categories B and C." (p 448)

Monitoring of Data.. Several statistical checks are performed on the model output using the monthly statistics of climatological means and daily standard deviations for each month of the year. These checks flag "outliers" for further diagnostics.

Reliability of the Reanalysis. The daily atmospheric and surface fields are, "for some variables, close to a best estimate of the evolving state of the atmosphere." The dynamics of the analysis computations allow it to "transport information from data-rich to data-poor regions, so that even in relatively data-void areas the reanalysis can estimate the evolution of the atmosphere over both synoptic and climatological timescales. **A researcher using the reanalysis should be aware, however, that the different outputs are not uniformly reliable."** (p 453).

The Reanalysis Project Data are most reliable **in the middle latitudes** and **over continents**. This is due to the spatial distribution of observations that go into the Class A variables (see sounding map below) and the resolution and arrangement of the gridpoints at which the T62 model computes and produces output.

Map of locations of Raobs soundings for the globe:
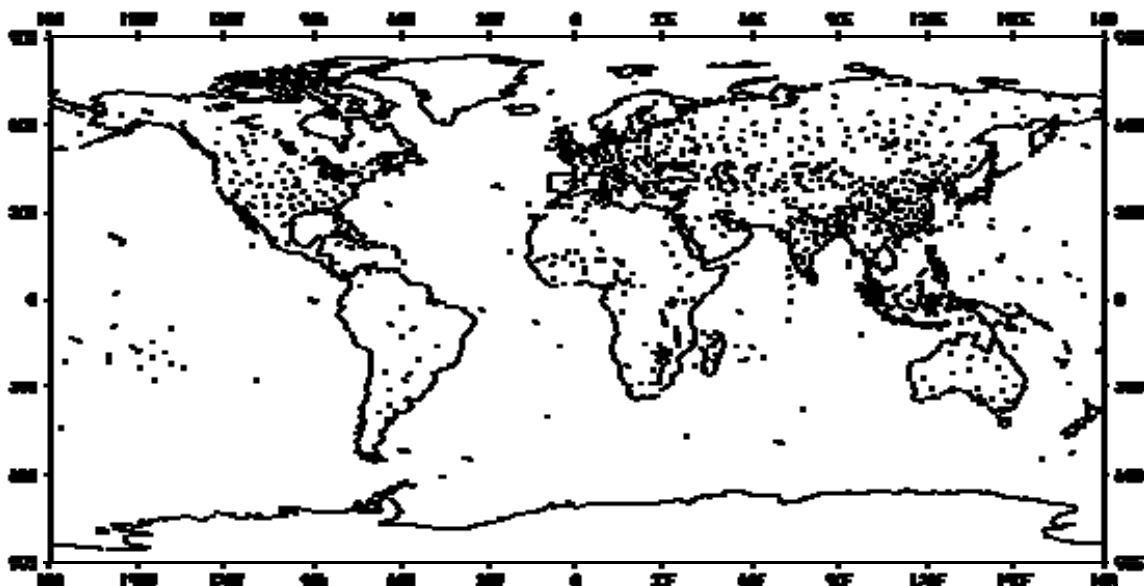
HANDBOOK FOR USING THE NCEP / NCAR REANALYSIS SITE
TO PRODUCE MAPS

Variable, years and levels available

Data are available from Jan 1958 to the present for most variables and is updated the first or 2nd week of the month with the previous month's data. Some variables are only available for certain time periods, etc.

Simple Black & White Composite Plots

Go to CLIMATE LINKS & go to CDC NECEP/NCAR Reanalaysis links. A good place to start is CLIMATIC GRIDDED DATA - **Visualize NCEP Data**
When you get to this page:
Select a **statistic** (e.g. Monthly Longterm Mean or Individual Monthly Mean)
Select a **variable** to plot (e.g. Geopotential Height) *(then click "Continue Image Specification)*
On next page check to see if what you've selected is what you want; keep lat and lon selected for the two dimensions, then click "Continue Image Specification"
On next page select the following for your Image Specification for a B&W plot:
* lat & lon, 0 N - 90N 0 - 360E for whole N. Hemisphere
* pick a level (e.g. 700 mb)
* pick beginning and ending periods (e.g. for winter season pick Dec - Feb)
*Select: **create plot**, under Plot Output Options select**: plot on WHITE background,** *uncheck* **Color Plot,** *check* **Polar stereographic**
* Click on "Create Plot or Subset of Data"
* wait patiently for plot to come out!

Plotting regions

To plot over the dateline, use values from 0 to 720. For example, to plot 180W eastward to 180W, use 180 to 540. Be sure that the western most longitude is less than the eastern., For example, to plot 100W to 70W, use -100 to -70 or 260 to 290 and NOT 100 to 70. There are 6 custom projections:

Northern Hemisphere: 0-90N, 0-360W in using a polar stereographic projection
Globe: 90S-90N 0E-360W
United States: 20N-65N; 235-285 polar stereographic projection
Tropics: 60S-60N
Tropical Pacific: 35N-35S 100E to 60W
Indio-Pacific (20S-20N,60E-160W)
4 Corner Region US (xxS-xxN,xxE-xxW)

To rotate polar stereographic plots (e.g., so N. America is at bottom of plot)

To rotate map, choose "custom" map projection and then choose either northern or southern polar stereographic projection. For lat range, enter 0 to 90 (for northern hemisphere) or -90 to 0 for southern. For longitude, the center of the longitude range input will be at the **bottom** of the plot. To center along 0E, choose -180 to 180, for example. To center at 90E, choose -90 to 270. You can plot sectors as well. The longitudes -90 to 90 in the NH will plot the half hemisphere from the US across the Atlantic to Europe.

Contour range/interval option

A desired contour interval and range can be input instead of the default being used. Different plots can be easily compared (and the resulting gifs could be animated). For this option to work, the interval AND the range must be input. There must be at least 2 and less than 33 contours. The contour interval must be positive and the range must go from low to high.

Year range

Default years are 1958-1999 with the exception of seasons that would begin in 1957 (e.g. DJF) and variables that do not extend over the whole time period. For the former, 1959-1999 is used and for the latter the ending year varies. You can use your own range of years up to the latest possible data. Note that there is no data before January 1958. You **should** get an error message if you ask for data that is not there. For example, the season DJF 1958 contains December 1957 and therefore **will not** be calculated. The teleconnection values and the monthly means are usually updated within the next month.

How correlations are calculated

A seasonal mean is calculated for each year for both the index time series and the variable that is being correlated with. A correlation is then calculated from these seasonal means (The correlation is NOT calculated from the monthly values unless the season is one month long). The formula for the correlation is as follows:

$$r = sum(xy)/sqrt[sum(x**2)*(y**2)]$$

where x and y have the means removed.

The time indexes are not defined for some seasons (e.g. PNA in summer). If you choose that season for the correlation, you will get back a blank plot. Likewise, any missing values for the index will not be used the the calculation. Please look at the time series before interpreting the results.

Instructions for crossections

1. Choose a variable with multiple levels (temperature, geopotential heights, omega, humidity, wind, or potential temp).
2. From 'plot region' choose latitude by height or longitude by height.
3. For latitude by height plots, choose a latitudinal extent and a range of longitudes (or a single longitude) over which to average.
4. For longitude by height plots, choose a longitudinal extent and a range of latitudes (or a single latitude) over which to average.
5. Choose level range (bottom to top). Note that humidity variables only go through 300mb and omega only goes to 100mb.

More Guidelines For Plotting Maps On The Reanalysis Webpages

Lat begin box: *always* start with the **southernmost** latitude

Lon begin box: *always* start with the **westernmost** latitude

You can use a minus sign (-) to indicate longitude **degrees W** and latitude **degrees S.** (Positive values are used for degrees E and degrees N)

You can also enter w*est longitudes* in positive degrees going EAST from **0° E in a full circle of 360°** as in the chart below.

For **Polar Stereographic Projections**, the beginning longitude will be plotted at the *top* of the circular projection; see the chart below for beginning and ending longitudes to center a given location at the *bottom* of the circular projection.

Equivalent ways to express the same west longitude:			
10 W	-10	350 E	350
20 W	-20	340 E	340
30 W	-30	330 E	330
40 W	-40	320 E	320
50 W	-50	310 E	310
60 W	-60	300 E	300
70 W	-70	290 E	290
80 W	-80	280 E	280
90 W	-90	270 E	270
100 W	-100	260 E	260
110 W	-110	250 E	250
120 W	-120	240 E	240
130 W	-130	230 E	230
140 W	-140	220 E	220
150 W	-150	210 E	210
160 W	-160	200 E	200
170 W	-170	190 E	190
180 W	-180	180 E	180

| Beginning & Ending Longitudes (polar stereographic or cylindrical full-globe maps) to center on:	Begin	End
0° E (Greenwich, Prime Meridian)	-180	180
100 W (center lon. of North America)	-280	80
25E (center lon. of Europe)	-155	205
90E (center lon. of Asia)	-90	270
60 W (center lon. of South America	-240	120
20E (center lon. of Africa)	-160	200
135E (center lon. of Australia)	-45	315
170W (center lon. of Pacific Ocean	-350	10
40W (center lon of N. Atlantic Ocean)	-220	140
20W (center lon. of S. Atlantic Ocean)	-200	160
75 E (center lon of Indian Ocean)	-105	255
90 E (center lon of Antarctica)	-90	270

For Maps Centered on Specific Continents or Regions *(may need to be adjusted somewhat)*

Continent / Region	Projection *	"Center" Lat/Lon	Longitude		Latitude	
			Begin	End	Begin	End
North America	polar	45 N / 100 W	-180	-20	10	90
United States (contiguous)	polar	35 N / 100 W	-150	-50	20	60
Central America	cylindrical	20 N / 100 W	-120	-65	5	35
Europe	polar	45 N / 25 E	-25	50	30	75
Asia	polar	40 N / 90 E	35	175	0	80
South America	cylindrical	20 S / 60 W	-95	-25	-60	20
Africa	cylindrical	0 N / 20 E	-25	65	-40	45
Oceania	cylindrical	25 S / 135 E	100	180	5	100
Antarctica	polar	90 S / 90 E	-180	180	-90	-55
N. Pacific Ocean	cylindrical	25 N / 170 W	-260	-80	-10	60
S. Pacific Ocean	cylindrical	25 S / 170 W	-240	-60	-70	10
N. Atlantic Ocean	cylindrical	25 N / 40 W	-100	20	-10	80
S. Atlantic Ocean	cylindrical	25 S / 20 W	-70	30	-70	15
Indian Ocean	cylindrical	20 S / 75 E	20	130	-60	30

* polar = **polar stereographic projection** (for single hemisphere maps of areas poleward of ~20°N or S latitude)

  cylindrical = **cylindrical equidistant projection** (default; for mapping low latitudes & in both N & S hemispheres)

OROGRAPHY FOR WESTERN UNITED STATES
"SURFACE LEVEL" USED IN THE NCEP/NCAR REANALYSIS

Plot for: 20 – 60 N 130-90 W
height in meters
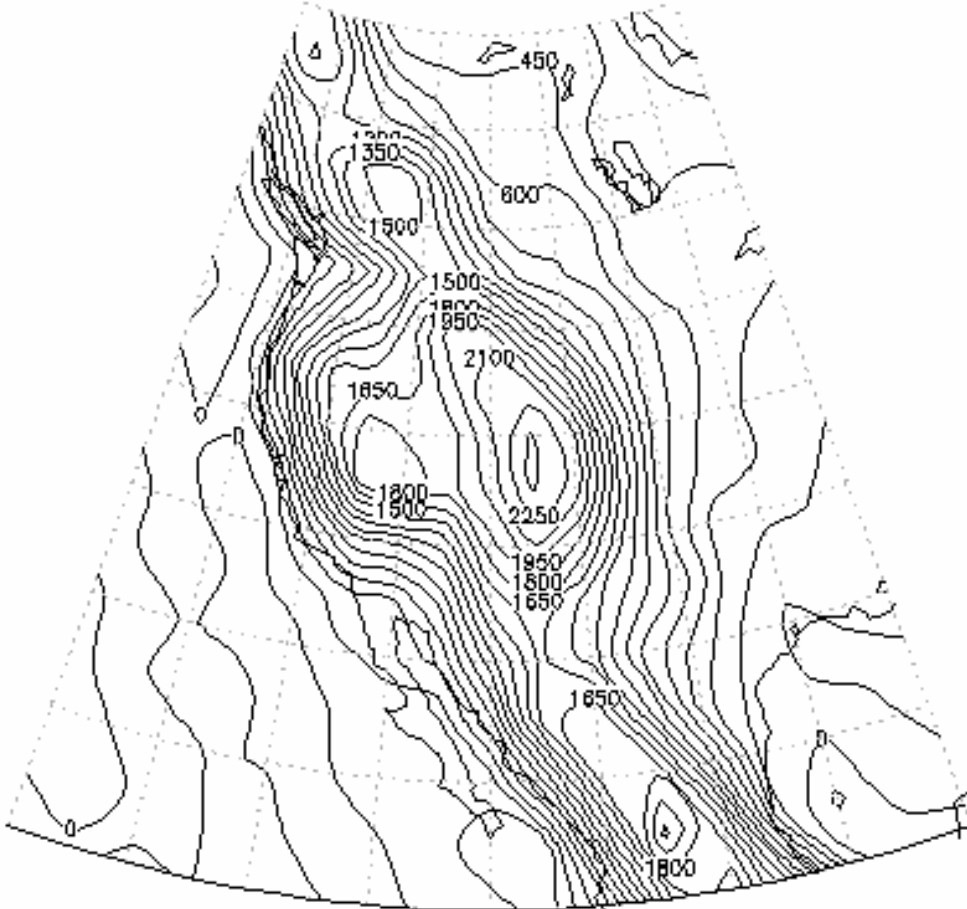
How to make a plot of the OROGRAPHY in the Reanalysis model:

- **Go to: NCEP Reanalysis Surface Level**
  http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.html#surface

- In the Surface Data archive parameter table, click on **Geopotential hgt:**

Geopotential hgt (time invariant) *	hgt.sfc	m	1.0

- Then click on **"make a plot or subset"** and proceed to make a plot as usual

(Surface Level or Averages Surface Level – doesn't mater which)