DOCUMENTATION OF OPERATIONAL NCEP CFS DATA FILES

Suranjana Saha
Global Climate and Weather Modeling Branch,
Environmental Modeling Center,
NCEP/NWS/NOAA/DOC Washington, DC, U. S. A.
Suranjana.Saha@noaa.gov

Revised: November 14, 2008

The NCEP Climate Forecast System (CFS) was developed at the Environmental Modeling Center at NCEP. It is a fully coupled model representing the interaction between the Earth's oceans, land and atmosphere. A description of the CFS is given in Saha et al, 2006. Please reference this article when using the data. The CFS became operational at NCEP in August 2004. There are four members per day from initial conditions for the atmosphere and ocean, which are 1 day old. The atmospheric initial conditions are obtained from NCEP Reanalysis-2 and the ocean initial condition is obtained from NCEP GODAS (Global Ocean Data Assimilation). The integrations are complete for the first partial month + 9 full months into the future. The raw (i.e. not corrected for bias) monthly mean data is available at the official NWS TOC data server every day. The raw daily CFS output forecast data is now available on a local NCEP ftp server. These data will remain at these sites for 7 days only.

References:
S. Saha, S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, P. Xie, **2006** : *The NCEP Climate Forecast System*. Journal of Climate, Vol. 19, No. 15, pages 3483-3517.

DAILY OUTPUT FROM CFS OPERATIONAL FORECASTS

Please download the daily CFS output data from the following ftp server:
ftp://ftpprd.ncep.noaa.gov/pub/data1/nccf/com/cfs/prod/forecast.yyyymmdd/daily_grib_memb
where yyyy is the initial year, mm is the initial month, dd is the initial day and memb is the member.
Example:
ftp://ftpprd.ncep.noaa.gov/pub/data1/nccf/com/cfs/prod/forecast.20060911/daily_grib_01/z500.01.2006091100.daily
is the file containing the 500 hPa geopotential daily output forecast data from the first CFS member initialized from the 11 September 2006. There are 2 CFS members run every day. *This data will remain on the site for only 7 days.*

Please visit the CFS website (http://cfs.ncep.noaa.gov) for documentation and download of the retrospective CFS daily forecasts of the variables available at this operational site. *It is highly recommended that some kind of calibration (bias correction) be done on the operational forecasts (using the retrospective data) before they are used in any predictive environment.*
For this purpose, we have constructed forecast and observed daily climatologies for selected variables, for both the mean and standard deviation. These datasets can be accessed by 2 methods:
- by an anonymous ftp login into cfs.ncep.noaa.gov,
  then go to the following directory: pub/raid1/cfs/daily/climatology/$var/$type,
  where var is variable and type is meanf (forecast mean) or sd (standard deviation).
The observed climatologies are available for the following pameters: mslp, wnd200,

wnd850, sfvp200, sfvp850, z200, z500, z700, z850 and z1000. They are in
pub/raid1/cfs/daily/climatology/$var/R2.

- by downloading the data from a web browser : http://cfs.ncep.noaa.gov/dailyclim

Please match the *month* (mm) and *day* (dd) of the initial condition of the real time
operational forecast with its counterpart in the daily climatology.

In our earlier example,

http://cfs.ncep.noaa.gov/dailyclim/z500/meanf/z500.09.11.mean.clim.daily

is the forecast mean climatology, that needs to be subtracted from the real time forecast to
get a bias-corrected forecast anomaly.

http://cfs.ncep.noaa.gov/dailyclim/z500/sdf/z500.09.11.sd.clim.daily

is the forecast standard deviation climatology, that the bias corrected forecast (obtained in
the previous step) needs to be divided by, in order to get the standardized forecast
anomaly.

http://cfs.ncep.noaa.gov/dailyclim/z500/R2/z500.R2.mean.data

is the observed mean climatology, that needs to be subtracted from the real time
observation to get a bias-corrected observation anomaly.

http://cfs.ncep.noaa.gov/dailyclim/z500/R2/z500.R2.sd.data

is the observed standard deviation climatology, that the bias corrected observation
(obtained in the previous step) needs to be divided by, in order to get the standardized
observation anomaly.

For further information on the construction of these climatologies and a list of the
variables, please see: http://cfs.ncep.noaa.gov/cfs.daily.climatology.doc

MONTHLY OUTPUT FROM CFS OPERATIONAL FORECASTS

CLIMATOLOGY DATA

The CFS has been run in retrospective mode for 25 years (1981-2005) with a 15member
ensemble for each calendar month. This amounts to over 4000+ integrations of the CFS,
each for 9-month forecasts. For documentation and download of these retrospective
forecasts, please visit the CFS website at: http://cfs.ncep.noaa.gov. From this
retrospective data, 10 sets (5 for monthly means and 5 for seasonal means) of
climatologies over the period 1981-2002 have been prepared, namely:

1.   CFS monthly and seasonal mean uncorrected forecast climatologies
      (modclim)which can be used to correct the model bias in the real time forecasts
      and to obtain corrected forecast anomalies;
2.   Reanalysis-2 and GODAS (proxy for observations) monthly and seasonal mean
      climatologies (obsclim);
3.   CFS monthly and seasonal mean forecast standard deviations (modstd), which can
      be used to obtain standardized forecast anomalies after bias correction of real time
      forecasts;
4.   Reanalysis-2 and GODAS (proxy for observations) monthly and seasonal mean

standard deviations (obsstd), and
5.   CFS monthly and seasonal mean uncorrected forecast root-mean-square-error
        (rmseclim), from which a corrected rmse can be computed using the terms above.

The above set of files, interpolated to the initial day and valid for use along with the real
time forecasts, will be available on the NWS server every day and will stay there for 7
days only.

CFS operational forecast data can be found at the NWS/TOC, the "Telecommunication
Operations Center". (http://www.weather.gov/tg/)

The website for real time data retrieval is at:
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.fcst/

And the climatological data is at:
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/

DOWNLOADING CFS DATA

ALL DATA IS IN OPERATIONAL GRIB FORMAT

Documentation on GRIB can be found at: http://www.nco.ncep.noaa.gov/pmb/docs/
And GRIB codes can be downloaded at: http://www.nco.ncep.noaa.gov/pmb/codes

There are 3 types of forecast and climatology files:

a.   Monthly mean pressure level data (press) of winds, temperature, geopotential,
        pressure vertical velocity, etc., at 17 standard pressure levels (1000, 925, 850,
        700, 600, 500, 400, 300 ,250, 200, 150, 100, 70, 50, 30, 20 and 10 hPa).
b.   Monthly mean surface and single level flux quantities (sflux), such as 2-meter
        temperature, precipitation, evaporation, etc.
c.   Monthly mean ocean data (ocean) of temperature, zonal and meridional velocity,
        vertical velocity, etc., at 40 levels in the ocean.

Be aware the initial condition is specified only in the parent directory, and care needs to
be taken when downloading files, since common names appear in different directories.

A *typical forecast file* has the following naming convention:
dd.yyyymm_pa.member(nn)_t1.type, where,
type stands for the type of file, i.e. press, sflux or ocean,
nn is the ensemble member
yyyymm is the verification year and month.
(dd is just a place holder and does not mean anything)

Example:

ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.fcst/RD.20040812/PA.month avg/PT.grid_DF.gr1/dd.200410_pa.member01_tl.ocean, is the ocean forecast file from initial condition of 12 August 2004, first member, valid at 2-month forecast lead, for October 2004. (Note: The initial condition is specified as RD.20040812 in a parent directory)

For Climatology files :

Lead-0 forecast from initial days 1-21 of a month contains the partial mean of the initial month, example :
Monthly Mean Lead-0 forecast from 12 August is the partial mean of 12 Aug to 31 Aug.
Seasonal Mean Lead-0 forecast from 12 August is the mean of 12 Aug to 31 Oct (ASO).

Lead-0 forecast from initial days after the 21st of a month contains the full mean of the next month, example :
Monthly Mean Lead-0 forecast from 22 August is a mean of the full month of September.
Seasonal Mean Lead-0 forecast from 22 August is the mean of 1 Sep to 30 Nov (SON).

A typical monthly mean climatology file has the following naming convention:
dd.xxxx.mm_pa.climtype_t1.type where
mm stands for the verification month
climtype stands for modclim, obsclim, modstd, obsstd, rmseclim (see section B)
type stands for the type of file,i.e. press, sflux or ocean (see section C)
(dd.xxxx is just a place holder and does not mean anything)

Example:

ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.mont havg/PT.grid_DF.gr1/dd.xxxx10_pa.modclim_tl.ocean, is the monthly ocean forecast climatology file which applies to the monthly forecast at 2month lead, valid for October, from initial condition of 12 August. (Note again : The initial condition is specified as RD.20040812 in a parent directory)

A typical seasonal mean climatology file has the following naming convention:

dd.xxxx.mm_pa.sclimtype_t1.type where,
mm stands for the first month of the 3-month verification season
climtype stands for modclim, obsclim, modstd, obsstd, rmseclim (see section B)
type stands for the type of file,i.e. press, sflux or ocean (see section C)
(dd.xxxx is just a place holder and does not mean anything)

Example:

ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.seasa vg/PT.grid_DF.gr1/dd.xxxx10_pa.smodclim_t1.ocean,
is the seasonal mean ocean forecast standard deviation climatology file which applies to the seasonal mean forecast at 2-month lead, valid for Oct-Nov-Dec, from initial condition

of 12 August. (Note again : The initial condition is specified as RD.20040812 in a parent directory)

WHAT TO DOWNLOAD

First, download the real time forecast file you need.

Then, match the appropriate set of 5 climatology files (modclim, obsclim, modstd, obsstd and rmseclim) to the real time forecast file, checking for matching month and day of forecast initial condition (in the parent directory) and then matching the verification month in the name of the file. (If seasonal means are being used, first average the 3 months of real time forecasts)

Example:

If the forecast file you download is:
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.fcst/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.200410_pa.member01_tl.ocean (F)

Then the set of 5 climatology files that apply to this forecast file are:

model climatology :
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.xxxx10_pa.modclim_tl.ocean (A)

observed climatology :
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.xxxx10_pa.obsclim_tl.ocean (B)

rmse climatology :
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.xxxx10_pa.rmseclim_tl.ocean (C)

model standard deviation :
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.xxxx10_pa.modstd_tl.ocean (D)

observed standard deviation :
ftp://tgftp.nws.noaa.gov/SL.us008001/ST.opnl/MT.cfs_MR.clim/RD.20040812/PA.monthavg/PT.grid_DF.gr1/dd.xxxx10_pa.modstd_tl.ocean (E)

HOW TO APPLY THE CLIMATOLOGY DATA TO REAL TIME FORECASTS

(Note: A, B, C, D, E and F are defined above)

1. To obtain a corrected forecast full field, first subtract the forecast climatological value (A) from the forecast grid (F), and then add the observed climatological value (B) to it.

2. To obtain a corrected forecast anomaly field, subtract out the forecast climatological value (A) from the forecast grid (F).

3. To obtain a corrected forecast standardized anomaly field, divide the corrected forecast anomaly field (obtained in item 2) by the forecast standard deviation (D).

4. To obtain a corrected forecast root-mean-square-error, use the expression: C' = sqrt (C*C - (B-A)*(B-A)).

5. To obtain the corrected anomaly correlation (AC), use the expression AC = 0.5 * (D*D + E*E – C*C + (A-B)*(A-B)) / (D*E)

To obtain cross-validated scores, multiply C', D and E by N / (N-1), where N is the number of years in the climatology, presently set at 22.