Estimating the Analysis Uncertainty by an Ensemble of Analyses: Results and the “NCEP-NCAR 2006 Annual Analyses DVD”

Wesley Ebisuzaki (1), Steve Worley (2), Chi-Fan Shih (2)
(2) National Centers for Environmental Prediction, Camp Springs, MD, USA, (2) National Center for Atmospheric Research, Boulder, CO, USA
Email: wesley.ebisuzaki@noaa.gov

Abstract: Gridded atmospheric analyses such as the NCEP/NCAR Reanalysis don't come with error bars. Here we estimate the analysis uncertainty by using an ensemble of analyses. (We are assuming each analysis is an independent estimate and the spread of the ensemble is a measure of the uncertainty.) The uncertainty estimates show considerable flow dependence and influence by observations. The value for many is that the database is available and one maybe able to use it to estimate the uncertainty for a particular application.

Introduction: There are three main methods of trying to estimate the uncertainty of a product such as a gridded analyses.

Method 1: A comparison with independent data
Except for satellite data and rain gauge, independent data are “spotty” in space and time.

Method 2: A comparison with dependent data.
Good coverage over land by in-situ observations. An example of this approach is a comparison to raobs by S. Saha,

http://wwwt.emc.ncep.noaa.gov/gmb/ssaha/

Raobs have poor coverage over water and suffer from both instrument and representativeness errors.

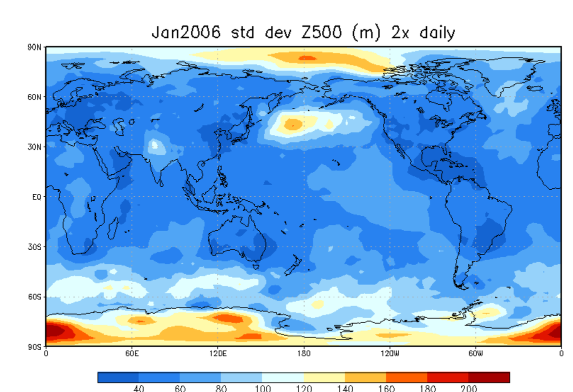
Representativeness error: in-situ observations are for a small “parcel” whereas the analyses are representative of a 250x250 km grid box for N/N Reanalysis. The difference between a perfect analysis and observation is the representativeness error. Typically this is larger than the instrument error.

Method 3: Using an ensemble of analyses
There are many atmospheric analyses and one could try to estimate the uncertainty by assuming that they are independent estimates. The uncertainty estimates could be an over-estimate because (i) numerical models may make similar errors, (ii) use many of the same (biased) observations. However, method 3 has the advantage that is is global, easy to use and can be applied to other than directly measured quantities such as the vorticity, the projection onto an EOF, the horizontal moisture flux, etc. Here, we will use method 3.

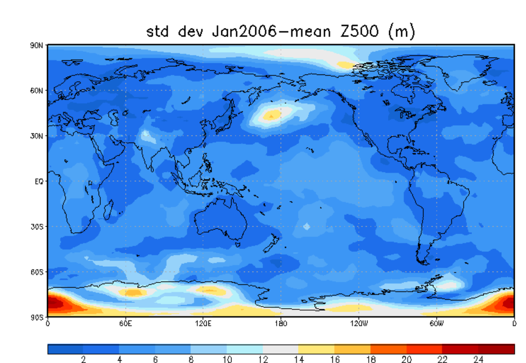
Method 3: 7 Ensemble Members

Data: 00Z/12Z for 2006, 2.5x2.5 degree grid

CMC: Environment Canada, Canadian Meteorological Centre, operational, global
ECMWF: European Centre for Medium-Range Weather Forecasting, operational, global
FNOC: Fleet Numerical Meteorology and Oceanography Center (NOGAPS), operational, global
GDAS: also known as FNL, made with GFS, NCEP, operational, global
R1: also known as NCEP/NCAR Reanalysis, global
R2: also known as NCEP/DOE Reanalysis, global
UKMET: UK Met Office, operational, global

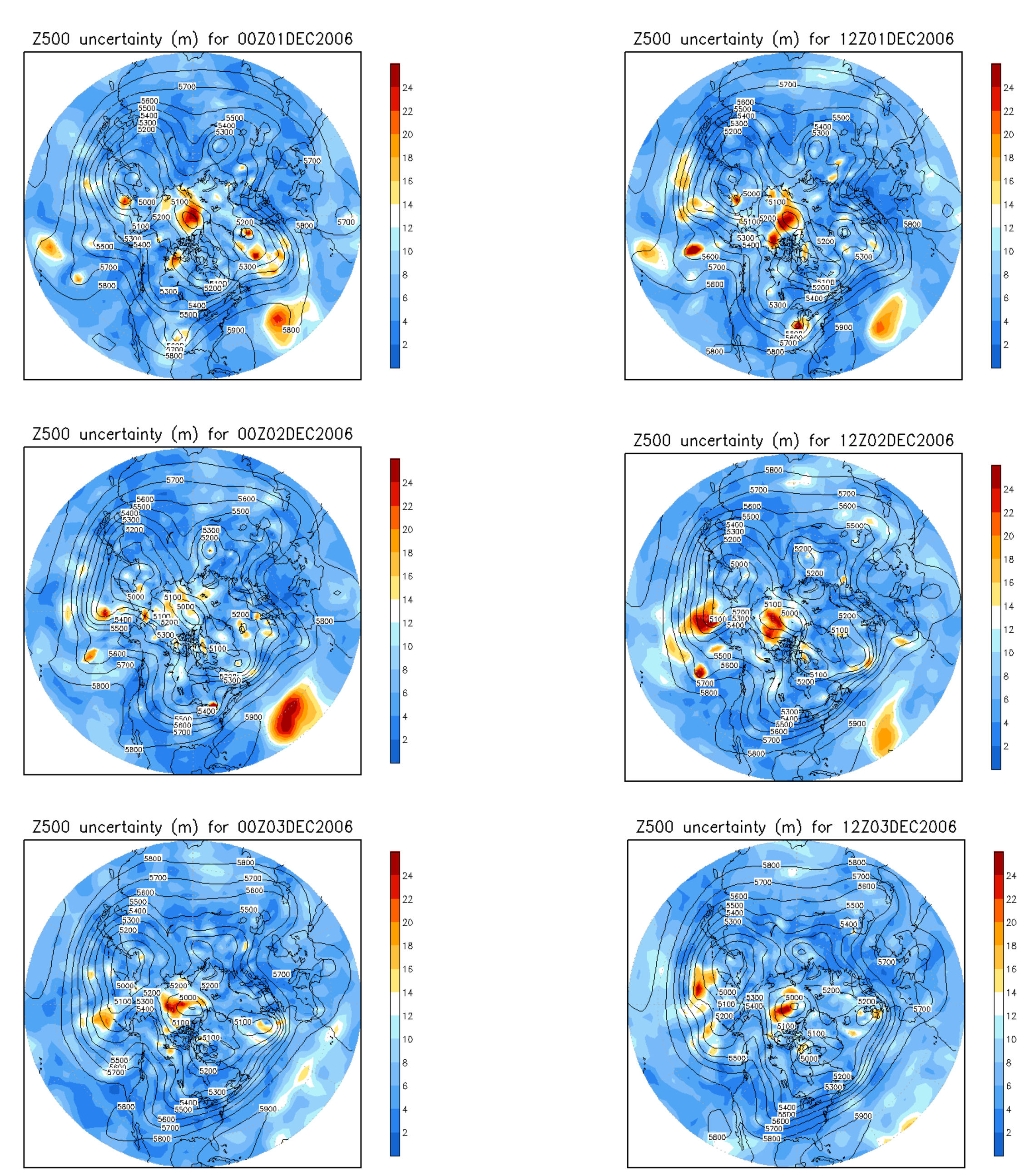


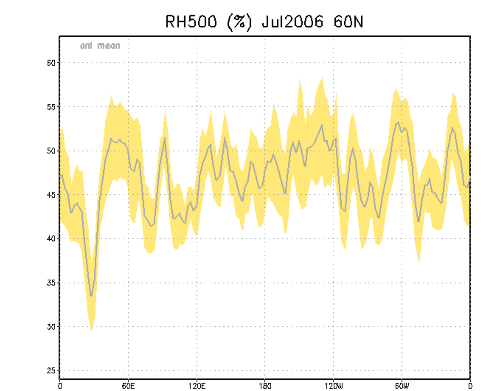
The average uncertainty (one standard deviation) for an individual analysis in Jan 2006, 500 hPa geopotential height (Z500).



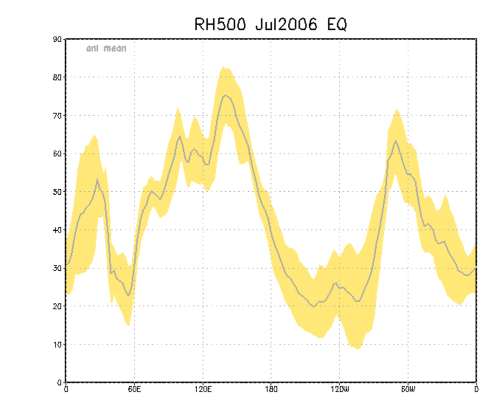
The same as left but for the monthly mean Z500. Note the differences in scales and lower uncertainty over land areas with strong observational networks.

The monthly mean of the uncertainty hides much of the day-to-day variability. This variability has synoptic scales and is often associated with the flow pattern (ridges, jets). Below are the Z500 uncertainty for 00Z 12/1/06 to 12Z 12/3/06

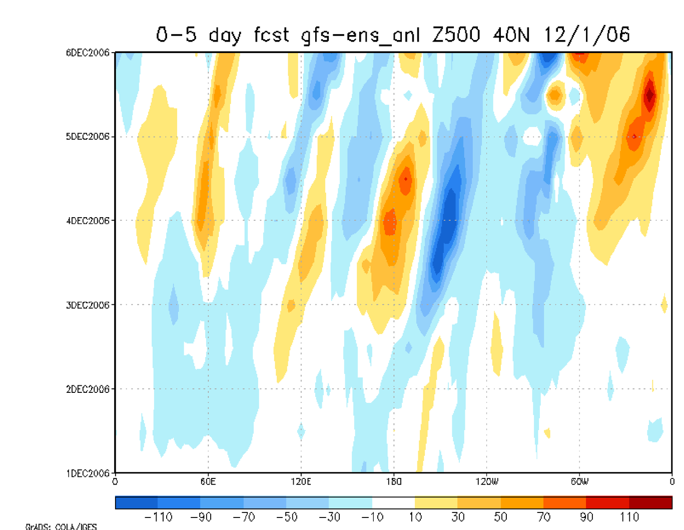




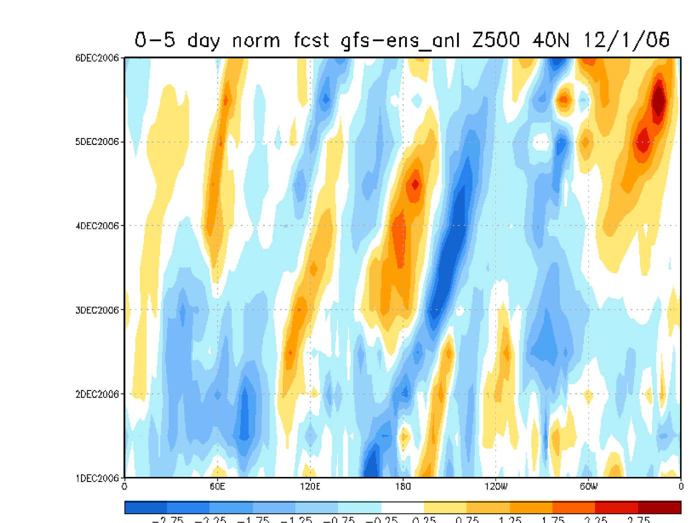
Mean July 2006 relative humidity at 500 hPA (RH500), Grey is the ensemble mean and the shaded region is the uncertainty of the monthly mean (1 std. dev.). All for 60N. RH uncertainty is often 10%.



Like the previous plot except for the equator. The uncertainty is larger than at 60N.



0-5 day Z500 GFS forecast – average analyses, 40N starting 12/1/06. As expected the forecast errors increase with forecast lead and errors propagate eastward.



Like previous plot except the variance at each forecast hour has been normalized. Note that the 5 day forecast errors seem to have connections to the difference between GFS initial conditions (GDAS) and ensemble mean analysis.

Results:

The uncertainty shows synoptic-scale variability (time and space).

Some fields have small relative errors (Z), other have quite large errors (RH). The errors, as expected, are smaller in observation-rich regions.

At least for the GFS, some 5-day forecast errors appear to have precursors in the difference of the GFS initial conditions and the ensemble-mean analysis.

For individual analyses, one often find features for which R1 and R2 form a cluster and the operational analyses form another cluster. This may be caused by the different satellite data (retrievals vs radiance assimilation) or the low resolution used by R1 and R2.

The database (2006) is available on DVD. Contact Chi-Fan Shih <chifan@ucar.edu> for details. This DVD is the continuation of the NCEP/NCAR Reanalysis CD-ROM series 1950-2006.

NCAR