# NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA)

Michael G. Bosilovich, Siegfried D. Schubert, Michele Rienecker, Ricardo Todling, Max Suarez, Julio Bacmeister, Ron Gelaro, Gi-Kong Kim, Ivanka Stajner, Junye Chen

Global Modeling and Assimilation Office, NASA GSFC, Greenbelt MD, USA

## 1. Objectives

Retrospective analyses (or reanalyses) synthesize temporally and spatially irregular observations to provide a spatially complete gridded meteorological database spanning the historical data record. The first and second rounds of reanalyses (Kalnay et al., 1996; Uppala et al., 2005) facilitated tremendous research in weather and climate in the community. These research efforts feed back into the reanalysis system by identifying future development pathways. This leads to improvements in our models, data assimilation techniques and the quality of the observations, which should then result in improvements in subsequent retrospective analysis.

There are substantial differences between the existing reanalyses. In part, these differences are an important measure of the uncertainty in any reanalysis product, due to deficiencies and differences in the assimilating models and how the models interact with the assimilated data. Such differences can be exacerbated by observing system changes that are often manifest in reanalysis time series by abrupt variations or discontinuities. These impacts from observing system changes must be distinguished from real climate variations. As an example, consider the 25 year Japanese Reanalysis (JRA25, http://www.jreap.org/) in which total column water vapor from Special Sensor Microwave/Imager (SSM/I) are assimilated, beginning in June 1987. While there is some evidence that the SSM/I assimilation has improved the quality of the reanalysis precipitation, a discontinuity is apparent in the time series (Figure 1). On the other hand, in 1995 there is a shift in the precipitation anomalies that grows in time. This change also appears in the NCEP/NCAR reanalysis, NCEP-2 reanalysis and ERA-40 reanalysis (not shown) and may be related to a possible phase change in the Pan Pacific Decadal variability (Chen, 2005).

The Global Modeling and Assimilation Office (GMAO) at the NASA Goddard Space Flight Center is developing a new data assimilation system to synthesize the large volume of EOS era and other observations (e.g. Bosilovich et al., 2007), with the goal of improving our understanding and predictions of the Earth's environment and climate. Using this system for reanalysis allows us to better address such questions as: What is the impact of new data on the climate time series? How do we place new satellite observations into the context of the long-term climate record?

To this end NASA is producing the Modern Era Retrospective-analysis for Research and Applications (MERRA). The MERRA project supports NASA's Earth science interests by:

1) utilizing the NASA global data assimilation system to produce long-term (1979-present) synthesis that places the current suite of research satellite observations in a climate data context;

2) providing the science and applications communities with state-of-the-art global analyses, with emphasis on improved estimates of the hydrological cycle on a broad range of weather and climate time scales.

In this brief article, we will summarize our plans for MERRA.

## 2. Model and Data Assimilation

The data assimilation system (DAS) for MERRA consists of the Goddard Earth Observing System version 5 (GEOS-5) atmospheric model coupled to the Grid-point Statistical Interpolation (GSI) analysis scheme being developed by the Environmental Modeling Center of the National Centers for Environmental Prediction (NCEP/EMC) and GMAO.

The GSI analysis solver was developed at NCEP to support inhomogeneous and anisotropic 3D background error covariances (e.g., Wu et al., 2002; Derber *et al*. 2003; Purser *et al*. 2003). The data streams currently assimilated by the DAS are shown in Table 1. Jack Woolen (NCEP) has provided an updated conventional observation data set for reanalyses.

The GEOS-5 atmospheric model is a weather-and-climate capable model using the finite-volume dynamical core (Lin, 2004). In developing GEOS-5, attention has focused on the representation of moist processes (see http://gmao.gsfc.nasa.gov/systems/geos5/). The tropical precipitation morphology has been analyzed in Bacmeister et al. (2006). The moist physics package uses a single phase prognostic condensate and a prognostic cloud fraction. Two separate cloud types are distinguished by their source: *anvil cloud* originates in detraining convection, and *large-scale cloud* originates in a PDF-based condensation calculation. Ice and liquid phases for each cloud type are considered. Once created, condensate and fraction from the anvil and statistical cloud types experience the same loss processes: evaporation of condensate and fraction, autoconversion of liquid or mixed phase condensate, sedimentation of frozen condensate, and accretion of condensate by falling precipitation.

Developments of GEOS-5 were guided by a realistic representation of tracer transports and stratospheric dynamics. The ozone analysis of the DAS is input to the radiation package along with an aerosol climatology. GEOS-5 is coupled to a catchment-based hydrologic model (Koster et al., 2000) and a sophisticated multi-layer snow model (Stieglitz et al., 2001).

## 3. MERRA Validation and Production

MERRA will use a ½ degree resolution model and analysis, with 72 levels to 0.01 hPa. A 2-degree resolution version of the MERRA system (called Sweeper) is being prepared and is planned to run for the period 1972 through 2005. From the software engineering point of view, the Sweeper will progress quickly through the input data, identifying problems that would otherwise slow the MERRA production. The Sweeper will be spilt into 4 streams, 3 of which will provide 5 years of analysis spin up to the land surface states.

Full resolution validation experiments will be run for 2001 and 2004, prior to initiating MERRA production. The water cycle, specifically global precipitation and evaporation, as well as the dynamical circulation and radiative processes are being evaluated. An interdisciplinary user group, external to the GMAO, will have access to the validation data and review the validation results. MERRA production will be accomplished through three streams (Figure 2). The streams will be initialized by the Sweeper run, and then spun up one more year to address the downscaling of the coarse grid data. When one stream catches another, there will be some overlap in the processed data. We intend to continue some of this overlap to evaluate the uncertainty and sensitivity of the analysis system to initial conditions.

In 1987, SSM/I becomes available and will strongly affect the water cycle. Subsequent advanced data products from EOS-era satellites, such as AIRS, TMI and QuikSCAT, likewise contribute to the analysis data and impact the climate record. To help assess the effect of new data on the reanalysis, an additional analysis will continue with the Reduced Observing System Baseline (ROSB) data stream. The ROSB will use only the non-satellite conventional data and the TOVS data for the entire period. With the ROSB we will assess the impact of the modern observing system on forecasts and analyses (e.g. water and energy budgets) and begin to understand how to incorporate new instruments and data into the long term climate time series.

## 4. Data Products and Portals

While disk technology and internet bandwidth still limit data distribution, they are improving rapidly with time. So, our product distribution plans are ambitious. The three dimensional data will be available at the analysis time, every six hours. Two dimensional and vertically integrated fields will be available at hourly intervals. This will allow for detailed analysis into the variations of the diurnal temperature range and precipitation frequency statistics. Full budget information will be available in these data products, including the analysis increment. In addition, post processing of the data products to monthly means and monthly mean diurnal cycle (for the surface data) will be done. A reduced spatial resolution (2.5 degrees) data set will also be prepared at 6-hourly and monthly intervals.

Current estimates indicate that 100 TB of data will be available from online drives. We plan to use the Nomads system (Rutledge et al., 2006; http://nomad3.ncep.noaa.gov/), a flexible system that includes a quick look viewer, ftp, http and OpenDAP access to the data. The form of the eventual system is still being finalized but will include several additional portals into the online data holdings.

## 5. Summary

The GMAO is producing the Modern Era Retrospective-analysis for Research and Applications (MERRA) to support NASA's Earth science interests by placing the current suite of research satellite observations in a climate context, and by providing the science and applications communities with state-of-the-art global analyses.

The MERRA system is currently in the last stages of assembly and is entering final validation before production. The latest information and updates on MERRA can be found at http://gmao.gsfc.nasa.gov/merra/. Production should begin in the summer of 2006. It is anticipated that it will take 1 ½ years to complete the time series. The terminus of the time series will likely be dictated by the lifetime of the computing platform; however, we expect that the time series will eventually include at least all of 2007. We expect to start making data available to the community prior to completion of the full time series.

Acknowledgements. The MERRA project is supported by NASA's Research, Education, Applications Solutions Network (REASoN) and Modeling, Analysis and Prediction (MAP) programs. The super computing support is provided by the NASA Center for Computational Sciences (NCCS).

## 6. References

Bacmeister, J.T., M.J. Suarez, and F.R. Robertson, 2006. Rain re-evaporation, boundary-layer/convection interactions and Pacific rainfall patterns in an AGCM, (accepted J. Atmos. Sci.)

Bosilovich, M. G., J.D. Radakovich, A. da Silva, R. Todling, and F. Verter, 2007: Skin Temperature Analysis and Bias Correction in a Coupled Land-Atmosphere Data Assimilation System. Submitted to JMSJ.

Chen, J., 2005: Understanding the observed tropical and mid-latitude radiative energy budget in the context of long-term climate variations. Ph.D. Dissertation, Columbia University.

Derber, J. C., R. J. Purser, W.-S. Wu, R. Treadon, M. Pondeca, D. Parrish, and D. Kleist, 2003: Flow-dependent Jb in a global grid-point 3D-Var. Proc. ECMWF annual seminar on recent developments in data assimilation for atmosphere and ocean. Reading, UK, 8-12 Sept. 2003.

Kalnay E. and co-authors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437-431.

Koster, R. D., M. J. Suárez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a GCM, Part 1, Model Structure. J. Geophys. Res., 105, 24809-24822.

Lin, S.-J., 2004: A vertically Lagrangian Finite-Volume Dynamical Core for Global Models. Mon. Wea. Rev., 132, 2293-2307.

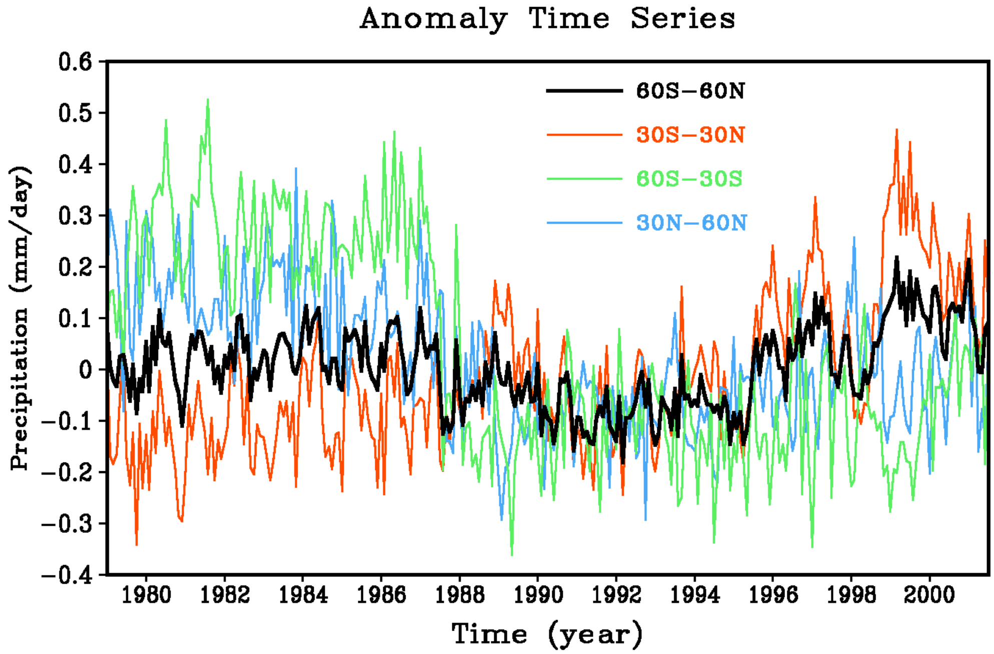
Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. Mon. Wea. Rev., 131, pp. 1524-1535.

Rutledge, G. K., J. Alpert and W. Ebisuzaki, 2006: Nomads, BAMS, 87, 327-341.
Stieglitz, M., A. Ducharne, R. D. Koster, and M. J. Suarez, 2001: The Impact of Detailed Snow Physics on the Simulation of Snow Cover and Subsurface Thermodynamics at Continental Scales. J. Hydromet., 2, 228-242.

Uppala SM, and co-authors, 2005: The ERA-40 re-analysis, QJRMS, 131 Part B, 2961-3012.

Wu, W.-S., R.J. Purser and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Wea. Rev., 130, 2905-2916.

7. Figures



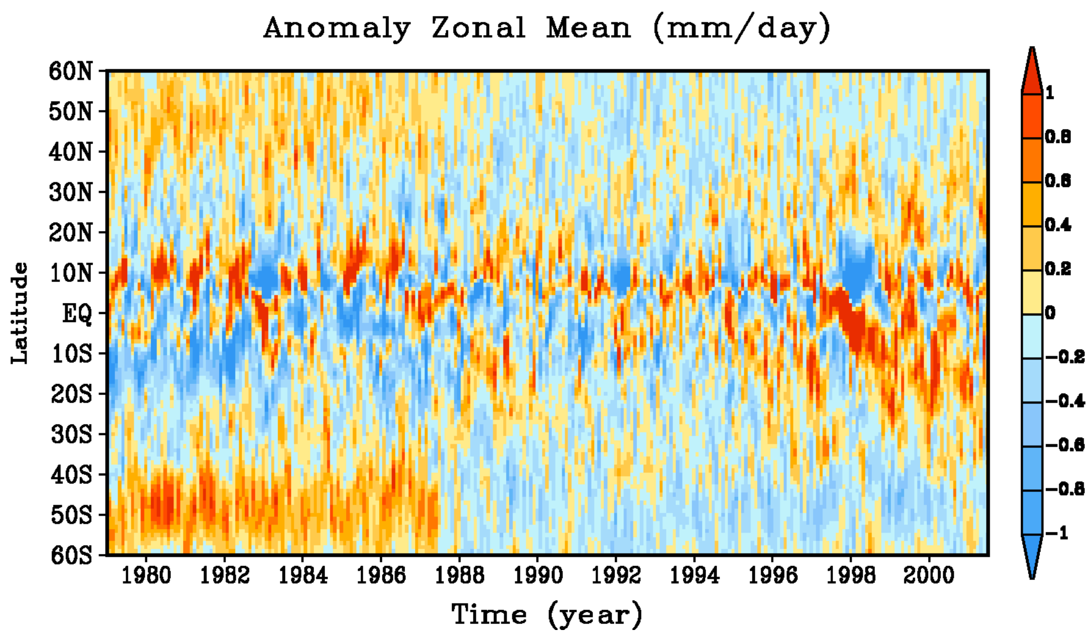




Figure 1. The time variation of monthly precipitation in JRA-25 reanalysis. Upper plot: the regional mean of the time series of anomalies, which defined as the deviation from the mean annual cycle. Black: 60S-60N; red: 30S-30N; green: 60S-30S; blue: 30N-60N. Lower plot: the latitude-time distribution of the zonal mean of the precipitation anomalies.

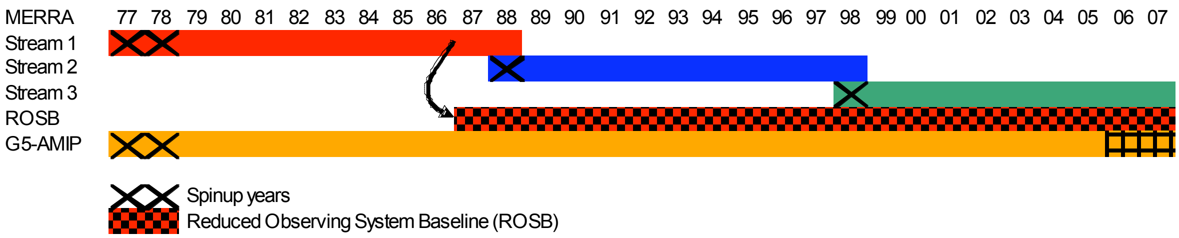



Figure 2. Configuration of MERRA production computing streams. The highest priority streams (1-3) include the full resolution output and complete input data. The Reduced Observing System Baseline (ROSB) will continue from Stream 1 but withhold new and advanced observing systems. G5-AMIP is a model-only integration to define the model's climate.

Table 1 Input observation data sources and parameters.

| |
|---|
| **Conventional Data** |
| Radiosondes |
| Pibal winds |
| Wind profiles |
| Conventional aircraft reports, ASDAR, MDCARS |
| NEXRAD radar winds |
| Dropsondes |
| Surface land observations |
| Surface ship and buoy observations |
| |
| **Satellite Retrievals** |
| GMS, METEOSAT, cloud drift IR and visible winds |
| MODIS clear sky and water vapor winds |
| GOES cloud drift IR winds |
| GOES water vapor cloud top winds |
| SSM/I rain rate and wind speed |
| TMI rain rate |
| QuikSCAT wind speed and direction |
| SBUV2 ozone (Version 8 retrievals) |
| |
| **Satellite Radiances** |
| TOVS 1b Radiances |
| DMSP SSM/I radiances |
| GOES sounder $T_B$ |
| Aqua/AIRS radiances (150 channels) |
| Aqua/AMSU-A radiances |