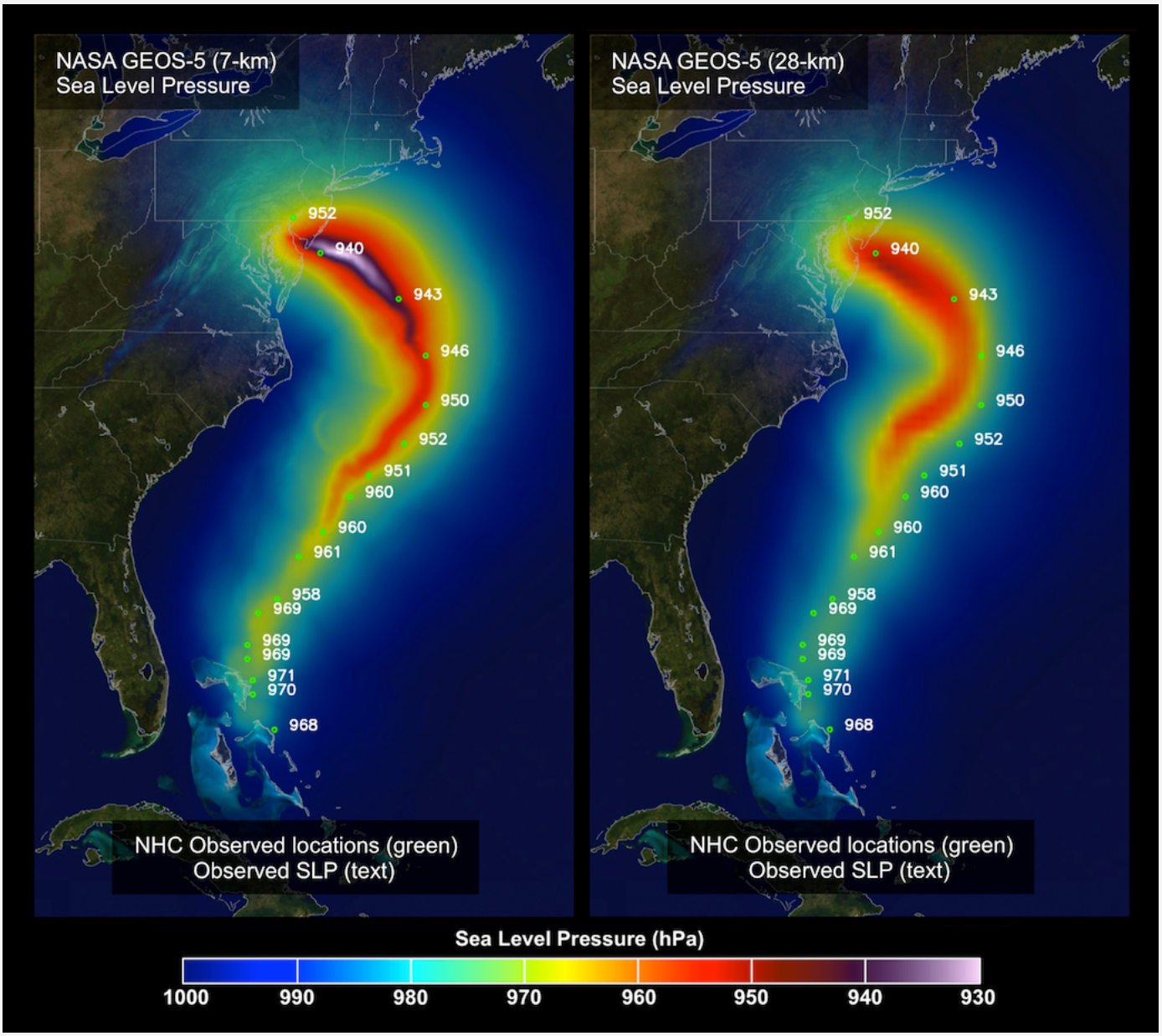
Global Modeling and Assimilation Office
Annual Report & Research Highlights
2012

NASA Goddard Space Flight Center June 2013



The image on the front cover shows the observed and the GEOS-5 forecast track and intensity (pressure in hPa) for Hurricane Sandy, starting October 26 at 0900 UTC. The left panel depicts results from a high-resolution GEOS-5 simulation conducted by GMAO's Bill Putman with the GEOS-5 AGCM on the cubed sphere grid at a nominal 7-km resolution, while the right shows results from the GEOS-5 operational production configuration at 28-km resolution. The forecasts were conducted on the Discover supercomputer at the NASA Center for Climate Simulation. More information is available at http://gmao.gsfc.nasa.gov/research/atmosphericassim/tracking_hurricanes.

Table of Contents

2012 in Review

Over the last year, the Global Modeling and Assimilation Office (GMAO) has continued to advance our GEOS-5-based systems, updating products for both weather and climate applications. We contributed hindcasts and forecasts to the North American Multi-Model Ensemble (NMME) of seasonal forecasts and the suite of decadal predictions to the Coupled Model Intercomparison Project (CMIP5). We also continued preparations for GEOS-5 support of NASA's Soil Moisture Active Passive (SMAP) mission and for generation of Level 4 SMAP soil moisture products.

Developments of the Gridpoint Statistical Interpolation (GSI) analysis have continued, in collaboration with NOAA/NCEP, providing improved high quality products from GEOS-5 for NASA instrument teams and other science groups who use GEOS-5 products. During this year, we transitioned systems to two production streams – one at 1/2° resolution to allow a re-processing with the updated system for NASA instrument teams, and one at 1/4° resolution for near-real-time analysis and forecasts to support regular upgrades and support of field campaigns with the best system available. The former production stream is referred to as FP-IT (and RP-IT for re-processing streams that cover January 2000 to December 2012); products are available through the Goddard Earth Sciences Data and Information Services Center (GES DISC). The latter production stream is referred to as FP. FP products, including 5-day forecasts, are openly available on the NASA Center for Climate Simulation (NCCS) data portal with several download options through the GMAO experimental forecast web page.

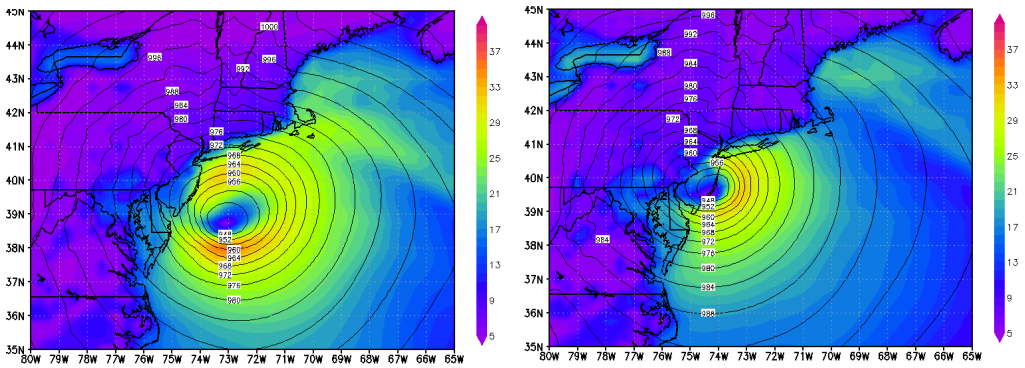
MERRA production has continued as an ongoing analysis of the atmospheric climate. Products, released only a few weeks behind real time, are distributed through the GES DISC. As of December 2012, over 1.5 PB of MERRA data have been distributed through the GES DISC.

There were several important upgrades that affect both the FP and FP-IT systems. The GEOS-5 atmospheric general circulation model (AGCM) now uses a cubed-sphere grid scheme, which not only avoids computational problems at the poles on the lat-lon grid but also helps increase processing speeds. The GEOS-5.9.1 system used for FP-IT and the GEOS-5.11.0 system used for FP (as of the end of 2012) are both capable of assimilating new observation types, including GPS radio occultation, Aura OMI ozone, and brightness temperature data from GOES-13, S-NPP ATMS, Metop-A, Metop-B, and NOAA-19. The systems now use OSTIA Sea Surface Temperature (SST) and sea-ice products that have higher horizontal and temporal resolutions than used with earlier systems. Both systems also include aerosol distributions generated from the GEOS aerosol assimilation system (GAAS), which runs coupled to the GEOS-5 atmospheric data assimilation system (ADAS). GAAS assimilates MODIS aerosol optical depth from EOS-Aqua and Terra.

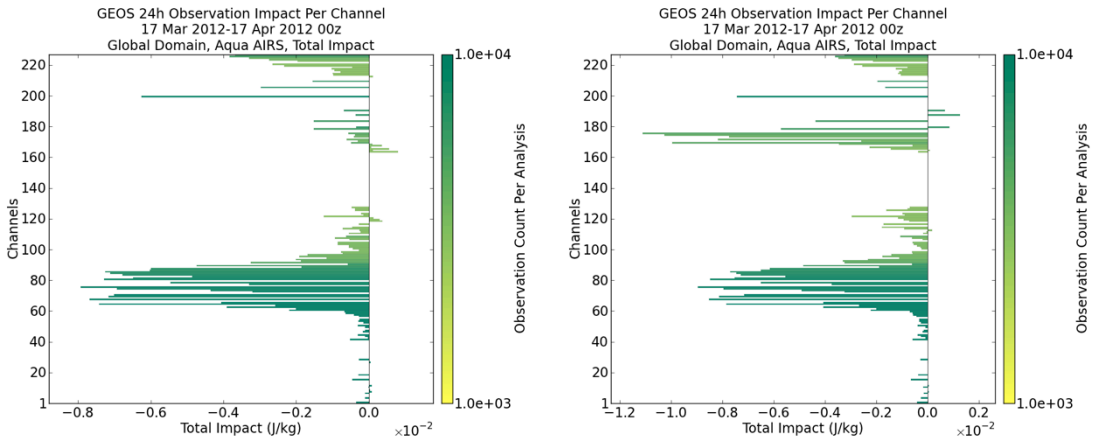
In addition to the switch to the cubed-sphere grid scheme, several atmospheric model developments have been included in the most recent GEOS-5 1/4° production systems. Of note is the updated snow model over glaciated land surfaces, implemented as a unified snow model for the land surface component. Along with this change, the albedo calculation over Arctic sea ice was updated to be seasonally varying. These developments reduced high latitude biases in surface temperature.

A significant system development has been that of the moist adjoint needed for assimilation of all-sky radiances and for the conventional 4DVar system. This has facilitated the upgrade of the GEOS-5 Observation Impact Monitoring System to include the effects of moist physical processes. The upgrade provides more accurate estimates of observation impact in GEOS-5, especially for moist-sensitive observations from AIRS, MHS, TMI, and some wind observations.

Developments of the hybrid ensemble-variational atmospheric assimilation have continued, with collaborations with NOAA/NCEP and ESRL. The system has been brought to sufficient maturity that it is now on the schedule for parallel production tests, with the goal of operational implementation by the end of 2013. In the process of system development, tests have been conducted of a *filter-free* ensemble scheme, wherein an ensemble of analyses is generated simply by adding appropriate perturbations to the central, hybrid, variational analysis. Experiments show the filter-free approach to be just as reliable as those from the Ensemble Kalman Filter based hybrid strategy.



The 2012 Atlantic hurricane season was extremely active and the GEOS-5.7.2 system used for NRT analysis and forecasts performed very well. Hurricane Sandy had a devastating impact, being the second most costly hurricane in US history. The GEOS-5 forecast initialized on October 25 1200 UTC predicted that the trough that was keeping the hurricane offshore would develop a northwest-to-southeast tilt and alter the hurricane's track, pulling it in towards the northeast coast. The model predicted Sandy would make landfall as an intense storm with an exceptionally large radius of tropical storm force winds. In addition to forecasting Sandy's track and intensity rather well, especially in the one-to-three-day range, GEOS-5 also forecast the structure of its strong winds well. The above figures show hourly-averaged sea level pressure (hPa), and 10m winds (m/s) of Sandy for 1900-2000 UTC October 29 (left) and 2300 October 29 - 0000 UTC October 30 (right), as forecast by GEOS-5 from 1200 UTC October 26. The compact field of high winds is forecast, as is the movement of the wind maximum from directly south of the storm's center at the earlier forecast time to east of the storm center closer to the time of the storm's landfall on October 30.



An important component of the GEOS-5 atmospheric data assimilation system is the adjoint-based observation impact tool, used to routinely monitor the observation network. In order to improve the analysis of the impact from moisture-sensitive observations, moist physics have recently been included in an updated version of the GEOS-5 adjoint model. This required the development of a linearization of both the convection and large-scale precipitation schemes used in GEOS-5. A moist energy norm, which combines errors in wind, temperature, surface pressure and humidity, is used to measure observation impact in the updated system. For moisture-sensitive instruments, such as AIRS and MHS, the reported impact is larger in the moist adjoint system. The total impact of AIRS observations increases by approximately 40% and the total impact of MHS observations increases by nearly a factor of 5. The above figures show the total impact per channel for AIRS. In the moist adjoint system, the moisture-sensitive channels (~160–200) are seen to have a much larger impact. Temperature-sensitive channels 50–120 also have a larger impact and the negative impact of certain channels is reduced.

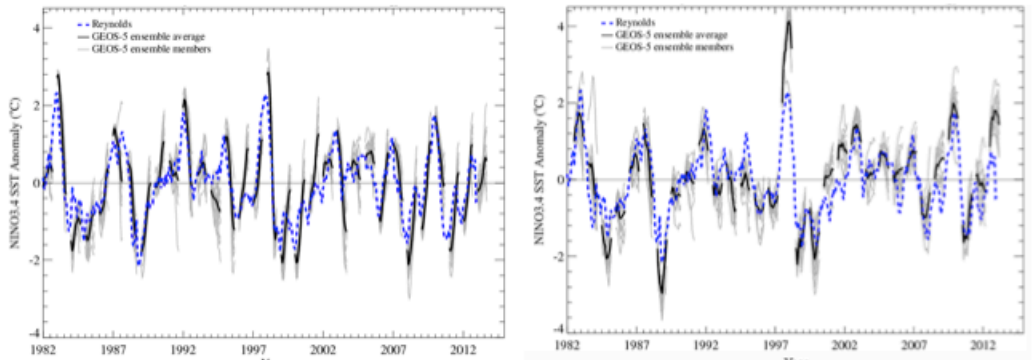
One of the recommendations from the 2010 Technical Workshop on Reanalyses was for reanalysis groups to collaborate in the examination of data utilization. To make progress along these lines, GMAO has made gridded information available from the MERRA observation feedback files (innovations and analysis residuals). A common grid, consistent with MERRA's native model grid, is used for all observations and the resulting output files are easily accessible for the multiple decades of the reanalysis. Collectively, this processing of the reanalysis observations is called Gridded Innovations and Observations (GIO) and is available online with the MERRA collections at the GES DISC.

The GMAO was well represented at the 4th WCRP International Conference on Reanalyses. In addition to a summary of MERRA, presentations were given on MERR*Aero*, MERRA-Land, on the MERRA-based ocean reanalysis, and on model and climate diagnostics using reanalyses.

Seasonal forecasts, initialized with MERRA and the MERRA-driven ocean and sea-ice reanalysis have continued in 2012 and were contributed each month to the NMME. Three new ensemble members were added from 1993 onwards, using satellite altimeter data to initialize the ocean model state.

Developments for the GEOS-5 AOGCM have included the integration of a new version of the ocean model, MOM5, developed at NOAA/GFDL. Our focus with the updated component is to improve the simulation of the Atlantic Meridional Overturning Circulation (AMOC). Initial tests have found that the new physical parameterizations used in MOM5 have some positive effect on AMOC strength, but the transport is still under-represented. Sensitivity tests underway indicate that the transport is sensitive to vertical mixing parameterization as well as air-sea fluxes.

One of the priorities this year has been the preparation for SMAP Level 4 Soil Moisture and Carbon products. In addition to the technical preparations for interface with the JPL SMAP product team, research developments have continued to facilitate direct assimilation of radiance data. For this a microwave radiative transfer model (RTM) has been coupled to the GEOS-5 Catchment land surface model and the relevant RTM parameters have been calibrated to obtain climatologically unbiased brightness temperature estimates.



Niño-3.4 SST hindcasts from January (left) and July (right) initial conditions using the GEOS-5 AOGCM. Anomaly correlations for SST hindcasts from 1982 to 2010 are above 60% over much of the ocean for lead times of up to three months, and are above 70% for a large portion of the equatorial Pacific. Skill stays high in the central and eastern equatorial Pacific for lead times of up to five months.

Hybrid Data Assimilation without Ensemble Filtering

Ricardo Todling and Amal El Akkraoui

It is now generally accepted that a practical, feasible way to introduce flow dependence in the background error covariances needed for either sequential or variational data assimilation procedures is to rely on an ensemble of short-range forecasts. At the time of this writing, multiple works have shown that combining the time-varying background error covariance derived from an ensemble with the typical, stationary, climatological background error covariance leads to non-trivial improvements to the resulting, so-called, hybrid data assimilation system. Most operational weather centers use three- or four-dimensional variational techniques and have implemented hybrid approaches in these contexts. With the variational component capable of accepting a hybrid formulation of its background error covariance, what remains to be specified is a methodology to generate the required ensemble of forecasts. Presently, the GMAO follows NCEP and uses the ensemble Kalman filter (EnKF; Whitaker et al. 2008) for this purpose. The small number of ensemble members used in practice requires some care to render adequate spread from the ensemble of forecasts. It is thus necessary to tweak the ensemble of analyses and: (i) apply multiplicative inflation to compensate for sampling errors; (ii) apply additive inflation to represent model uncertainty; and (iii) re-center the ensemble of analyses around the hybrid variational analysis to prevent possible divergence between the two assimilation systems.

During the process of implementation and testing of the EnKF to allow a hybrid strategy to be adopted for the GEOS atmospheric data assimilation system, we have found the latter two steps above to have a significant role. Indeed so much so, that the EnKF step seemed unnecessary. This led us to develop a so-called *filter-free* ensemble scheme where the EnKF is not used, but rather an ensemble of analyses is generated by simply adding perturbations to the central, hybrid, variational analysis – that is, steps (ii) and (iii) are what constitute the ensemble analysis strategy. The additive perturbations used in the procedure correspond to samples of the scaled, 48-minus-24-hour forecast differences, similar to those used to generate the climatological background error covariance of the traditional assimilation approach. These perturbations are also those used in step (ii) when the EnKF is exercised. A couple of highlights comparing the EnKF-based hybrid with the filter-free strategy are given below.

Figure 1 shows root-mean-square of observation-minus-background zonal wind radiosonde residuals, as a function of pressure level, over the month of April 2012 for three regions of the globe, and three experiments, namely: the control, traditional 3DVAR case (blue); and two hybrid 3DVAR strategies, one relying on the EnKF (red) and another using the filter-free ensemble (green). That hybrid assimilation brings improvement over traditional 3DVAR is evident in all regions. More surprising is the similarity of the results from the two hybrid strategies; the EnKF brings only very minor improvement over the filter-free approach in the Northern and Southern Hemispheres, and results from both hybrid strategies are indistinguishable in the tropics.

These results are corroborated further when the quality of mid-range forecasts are compared. Figure 2 shows 500 hPa geopotential heights forecast anomaly correlations for the same three experiments, up to 5 days, in the Northern (left) and Southern (right) Hemisphere Extra-Tropics. Again here, both hybrid strategies amount to improvements over the control, traditional 3DVAR, approach; the EnKF-based hybrid (red) is only marginally better than the filter-free approach – differences from the control are statistically significant in the Southern Hemisphere, but not statistically significant when comparing the two hybrid strategies with each other (not shown).

Additional diagnostic evaluations show the filter-free approach to provide improvements that are just as reliable as those from the EnKF-based hybrid strategy.

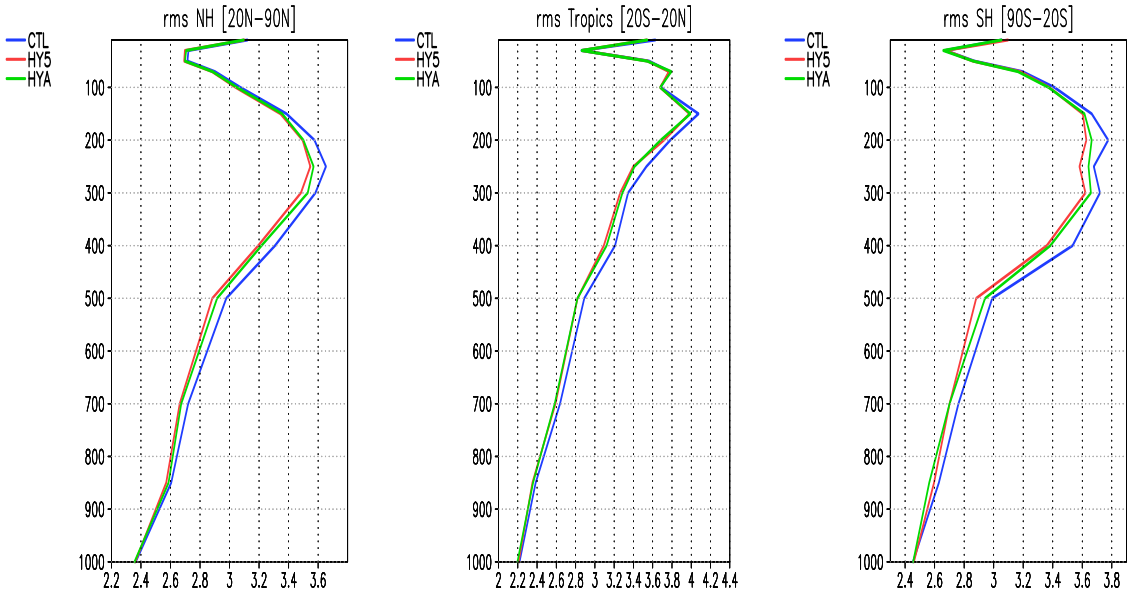


Figure 1. Root-mean-square of observation-minus-background zonal wind radiosonde residuals, as a function of pressure, over the month of April 2012 for three regions: Northern Hemisphere (left), Tropics (middle) and Southern Hemisphere (right). Curves are for control, traditional 3DVAR (blue), EnKF-based hybrid (red), and filter-free hybrid (green).

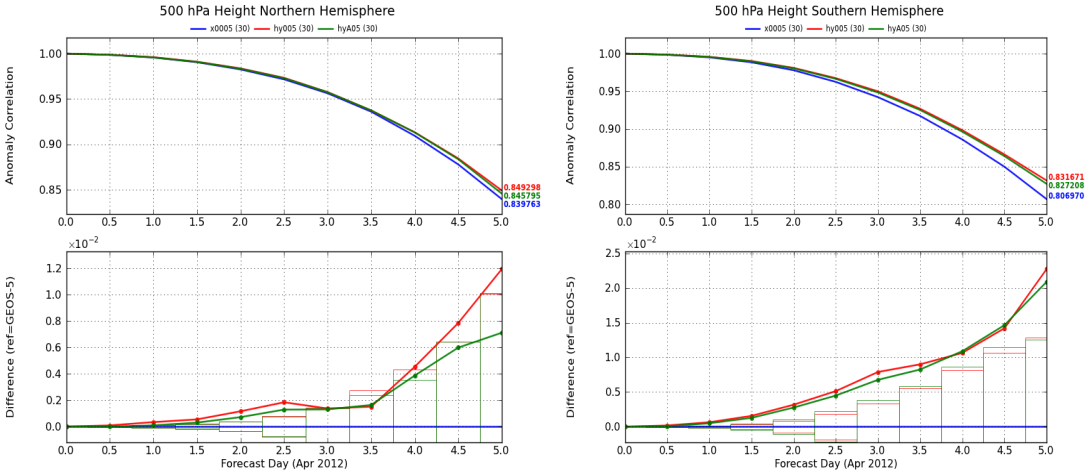


Figure 2. Anomaly correlation for the 500 hPa geopotential height for the Northern (left) and Southern (right) Hemispheres with three configurations of GEOS DAS: traditional 3DVAR (blue); EnKF hybrid 3DVAR; and filter-free hybrid 3DVAR. Significance plots appear beneath anomaly correlations with significance boxes color according to experiment designation; results are statistically significant when the curve appears outside, and above, the corresponding box.

Reference

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463-482.

Publication

Todling, R., and A. El Akkraoui, 2013: Hybrid data assimilation without ensemble filtering. *(Draft ms).*

Impacts of EOS-Aura/MLS Temperature Retrievals in GEOS-5

Lawrence Coy, Jianjun Jin, Steven Pawson, and Krzysztof Wargan

Accurate determination of global temperatures near the stratopause (~55 km) is problematic in modern global data assimilation systems such as GEOS-5 since these altitudes are above direct measurement by routine weather balloon soundings. Nadir viewing satellite infrared and microwave radiometers, such as AMSU-A (Advanced Microwave Sounding Unit-A), have high-peaking channels that are a potential source of information about the stratopause temperature. However, the depth of the nadir viewing weighting functions over the relatively sharp stratopause temperature peak and the uncertainty in the modeled stratopause temperature has limited the usefulness of the AMSU-A stratospheric channels. Comparisons of GEOS-5 products with temperatures retrieved from the EOS-Aura Microwave Limb Sounder (MLS) reveal cold biases of about 5 K near the stratopause (near 1 hPa, or 50 km altitude) and warm biases of 5-15 K in the lower mesosphere.

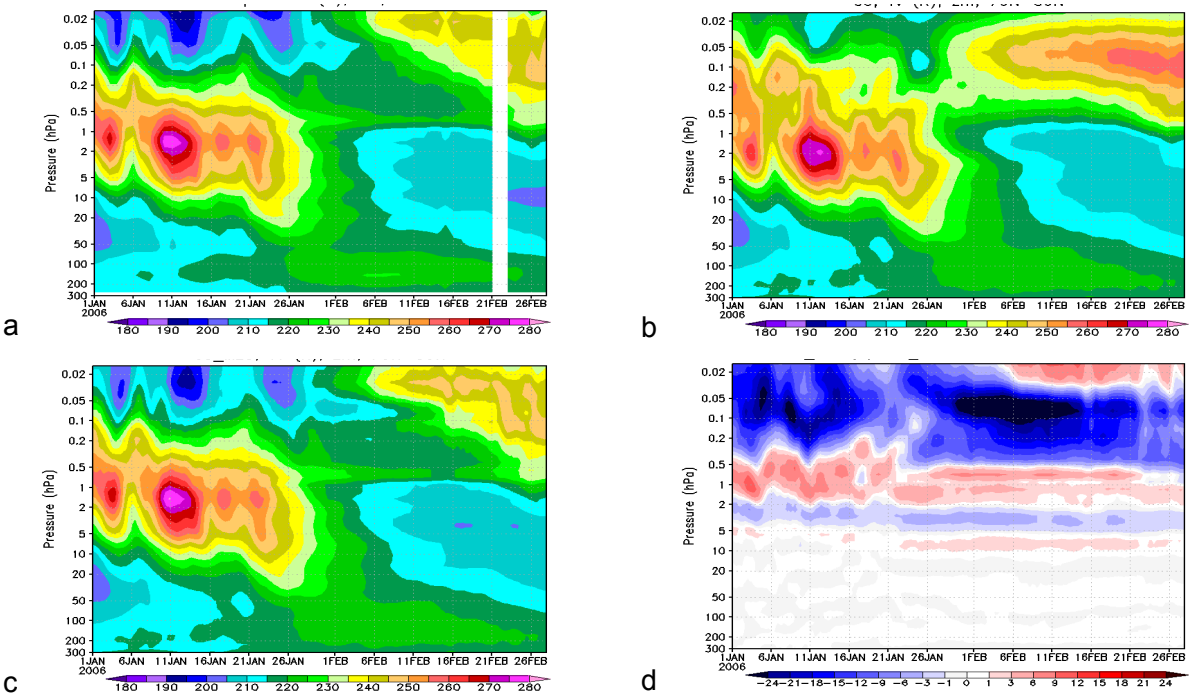


Figure 1. Time series of zonal-averaged temperature between 70°N and 80°N, for Jan. 1 – Feb. 28, 2006. (a) EOS-Aura MLS retrievals; (b) GEOS-5 with AMSU-A radiance observations assimilated; (c) GEOS-5 with AMSU-A radiances and MLS assimilated; (d) difference of GEOS-5 with and without MLS observations.

The limb-viewing MLS instrument has a vertical resolution of about 5 km near the stratopause and provides complementary information to the nadir viewing AMSU-A instruments. To better constrain the upper stratosphere and mesosphere, MLS temperatures (level-2 products) have been assimilated into GEOS-5 at pressures lower than 5 hPa, which is the peak level of AMSU-A Channel 13. The assimilation experiments were conducted for the period May 2005 through April 2007. The resultant temperature analyses are significantly improved in the upper stratosphere and lower mesosphere. Figure 1 shows high-latitude (70°-80°N) temperature time series between 300 and 0.015 hPa through January and February 2006, spanning the major Stratospheric Sudden Warming in early February. Manney et al. (2008) showed that MLS (and other) observations reveal that the thermal structure is disturbed by the sudden warming, with the polar stratopause reforming in late February at a higher altitude than before the warming (Fig. 1a). This feature is not captured as well by meteorological analyses with standard observations (Fig. 1b) wherein the stratopause is reformed lower than observed and with a stronger, more coherent structure. Adding the MLS observations into GEOS-5 provides a means of forcing the analyses to this observed state (Fig. 1c), resulting in very large corrections to the temperature analyses (Fig. 1d).

Variational bias correction (VBC) was not applied to AMSU-A channel-14 radiances in MERRA: in the absence of other observations, the model bias propagates into the resultant product, leading to a degradation of the stratopause temperature (Rienecker et al., 2011). This issue is revisited here in the context of the MLS observations. Four experiments, using two combinations of with/without MLS data and with/without VBC on AMSU-A channel 14 are examined. Figure 2 shows the global distribution of the AMSU-A channel 14 observation minus forecast (O-F) radiance temperatures at 18 UTC 10 January 2006 for the four experiments. The assimilations that include MLS temperature profiles (Fig. 2b, d) show significantly reduced O-Fs, especially in the northern mid-latitudes. Using only AMSU-A bias corrections (Fig. 2c) there is some reduction in the O-Fs, however the bias correction can tend toward the model bias (not shown). Results are best when both the MLS temperatures are included and the bias correction is on (Fig. 2d) since the cross-track bias is reduced somewhat. Figure 3 compares the globally averaged AMSU-A channel 14 bias correction as a function of time for the experiments that include (red curve) and exclude (black curve) MLS temperatures. Including MLS temperatures reduces the AMSU-A channel 14 bias correction by about a factor of 5 in April 2006. This implies that both AMSU-A channel 14 and MLS temperature profiles provide consistent measurements that work to support each other.

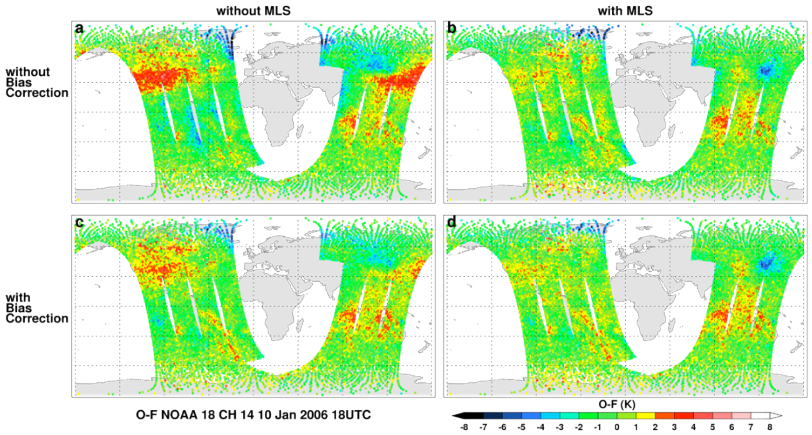


Figure 2. Observation minus forecast (O-F) error of brightness temperature K) from GEOS-5 for NOAA-18 AMSU-A channel 14 on 10 January 2006, 18 UTC. Results are shown for four assimilation experiment: without (a, c) and with (b, d) MLS temperatures and without (a, b) and with (c, d) variational bias correction applied to AMSU-A channel 14 radiances.

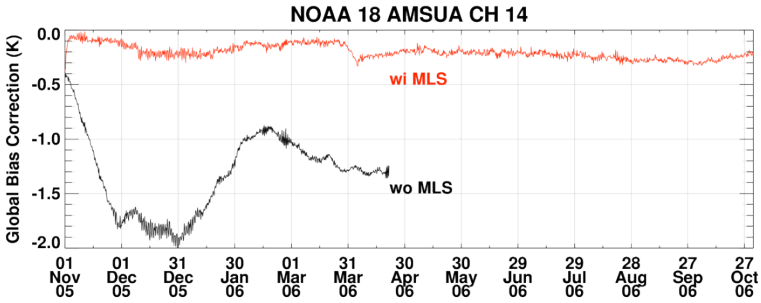


Figure 3. Global averaged bias correction for NOAA-18 AMSU-A channel 14 radiances, with (red curve) and without (black curve) MLS temperature assimilation.

In conclusion, MLS temperatures can provide an important constraint on analyses, allowing for meaningful bias correction of the high-peaking AMSU-A channels. The improved stratopause temperatures can in turn guide future model improvements. This work motivates the use of MLS temperatures in future reanalyses with GEOS-5.

References:

Manney, G.L., and coauthors, 2008: The evolution of the stratopause during the 2006 major warming: Satellite data and assimilated meteorological analyses. *J. Geophys. Res.*, **113**, D11115, doi:10.1029/2007JD009097.

Rienecker, M.M., and coauthors, 2011: MERRA. *J. Climate,* **24**, doi:10.1175/JCLI-D-11-00015.1.

Tropospheric Ozone Derived by Assimilating EOS Aura Data

Krzysztof Wargan and Steven Pawson (GMAO)
Mark Olsen, Jacquie Witte, Anne Douglas, and Susan Strahan (Atmospheric Chemistry & Dynamics Lab)

Accurate representation of the global distributions and variability of ozone in the troposphere is critical for air quality forecasting and for our understanding of tropospheric chemistry and radiative forcing. The sources of ozone include reactions that involve the odd nitrogen family (NO_x) associated with industrial pollution as well as natural processes such as lightning. Stratosphere–Troposphere Exchange also contributes to the tropospheric ozone budget. At present, these contributions are not well quantified.

A seven-year (2005-2011) analysis of ozone data from EOS Aura was performed using a modified version of the GEOS-5.7.2 Atmospheric Data Assimilation System. The assimilation used the near-total-column data from the Ozone Monitoring Instrument (OMI) and stratospheric profiles from the Microwave Limb Sounder instrument (MLS). The limited sensitivity of the OMI measurements to near-surface ozone was explicitly taken into account using the weighting functions provided with the retrieved data. This approach differs from previous methods of deriving daily tropospheric ozone information from OMI and MLS (Schoeberl et al., 2007) in that it explicitly takes into account the global ozone transport from assimilated meteorological fields. Doughty et al. (2011) showed that tropospheric ozone from assimilated EOS-Aura data exhibits a good agreement with ozone sonde observations over North America during the INTEX-B campaign in 2006.

Figure 1. Comparison of ozone analysis and ozonesondes for winter (2005-2011), 30°N – 60°N. (a) Surface-to-tropopause ozone columns; (b) 500-hPa-to-tropopause columns (c) The probability distribution of sonde-minus-analysis differences (step line) and a Gaussian fit (smooth line) from (a). (d) The same using data for 500 hPa to the tropopause (from (b)).

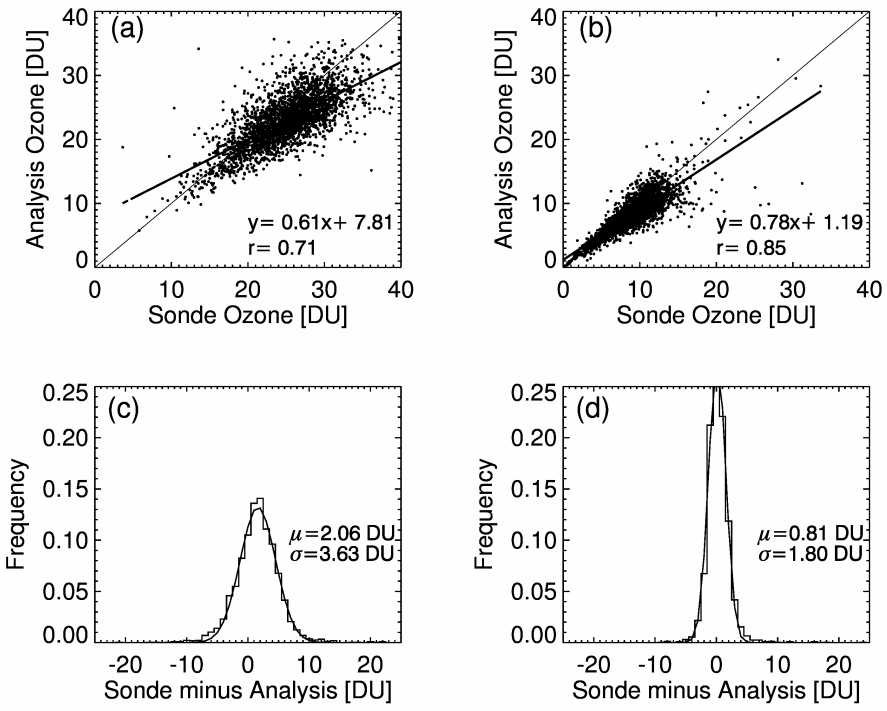


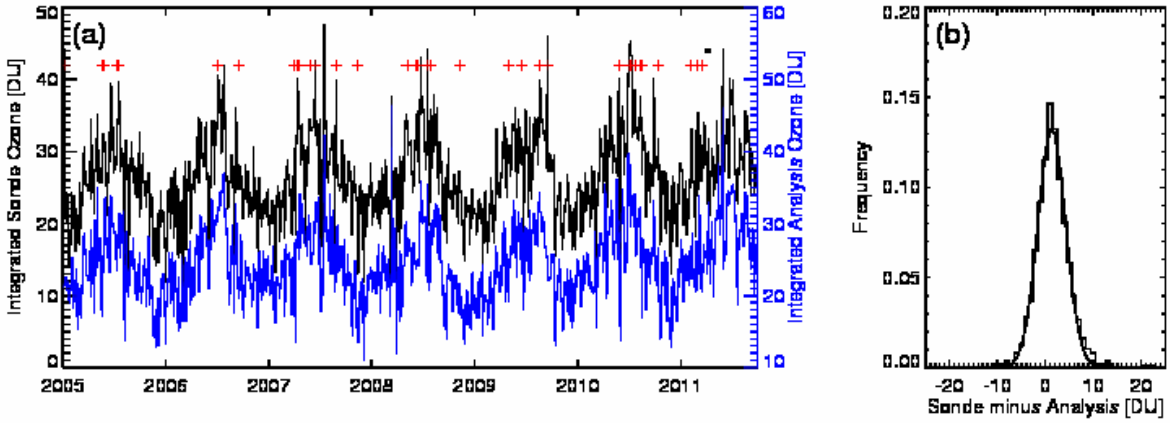
Figure 1 compares the assimilated tropospheric wintertime (December – February, 2005-2011) ozone and electrochemical concentration cell (ECC) ozone sonde data from the World Ozone and Ultraviolet Radiation Data Centre in the northern middle latitudes (30°N – 60°N). Results in panels (a) and (c) show ozone integrated between the surface and the tropopause. The correlation between the assimilated

product and sonde data is 0.71 and the slope of a linear fit is 0.61, indicating a better agreement with the sondes than for the product of Schoeberl et al. (2007), which did not assimilate the ozone data. There is a low bias of 2.06 Dobson Units, which we attribute to a low sensitivity of OMI observations to the lower troposphere, particularly below clouds, and the lack of a representation of NO_x–ozone chemistry in the model. This is supported by a much better agreement between assimilation and sonde data seen in the upper troposphere as seen in panels (b) and (d). Away from the surface sources, concentrations of ozone above 500 hPa are closer to their background values and are well observed by OMI. In addition, a high bias of MLS ozone in the lower stratosphere may contribute to the systematic error below.

The ability of this assimilation-based approach to reproduce temporal variability of ozone is illustrated in Figure 2, which shows the time series of ozone integrated between the surface and the tropopause from a series of balloon sondes at Hohenpeissenberg, Germany and from assimilation. The analysis reproduces the annual cycle as well as high frequency variations seen in the sonde data. However, some rapid excursions are not present in the assimilated product. This may be due, in part, to a relatively coarse resolution used for this analysis (2°×2.5°).

In summary, this study demonstrates the ability of data assimilation to reproduce the distribution and variability of integrated ozone fields in the troposphere.

Figure 2. (a) Time series of tropospheric ozone column from Hohenpeissenberg, Germany (47°48'N, 11°0'E) sondes (black) and analysis (blue), offset by 10 DU. Red plus signs show when the sonde minus analysis differences exceed 7 DU. (b) The distribution of sonde minus analysis differences (step line) and a Gaussian fit to the data (smooth line).



References

Doughty, D. C., and coauthors, 2011: An intercomparison of tropospheric ozone retrievals derived from two Aura instruments and measurements in western North America in 2006. *J. Geophys. Res.*, **116**, D06303, doi:10.1029/2010JD014703.

Schoeberl, M.R., and coauthors, 2007: A trajectory-based estimate of the tropospheric ozone column using the residual method. *J. Geophys. Res.,* **112**, D24S49, doi:10.1029/2007JD008773.

Publications

Wargan K., and coauthors, 2013: Tropospheric Ozone from Assimilation of the Ozone Monitoring Instrument and the Microwave Limb Sounder Data into GEOS-5. The Northern Extratropics (*in preparation*).

Gridded Innovations and Observations (GIO) from MERRA

Michael Bosilovich and Arlindo da Silva

Reanalyses assimilate a wide variety of observing systems into a single data product. The output from the data assimilation system includes the observations used to create the analysis, as well as the observation minus forecast (O-F) and observation minus analysis (O-A) residuals at the time, level, latitude and longitude of each observation, i.e., in observation space. To the extent that the observation and forecast errors are un-correlated, the O-F variance can be understood as an upper bound for the forecast (and analysis) error variance. Observation file formats – conventional observations (in-situ stations, radiosonde, dropsonde and retrieved satellite data) and satellite radiance observations (for each instrument and channel) – can be diverse, complicating the evaluation of several decades of data. While the observation-space data provides precise calculations for various assimilation statistics, we have simplified this data to a gridded version of MERRA's innovations. Since a common grid is used for all observations and is consistent with MERRA's native model grid, the resulting output files are easily accessible for the multiple decades of the reanalysis. Collectively, this processing of the reanalysis observations is called Gridded Innovations and Observations (GIO). We expect that the systematic examination of GIO can help improve the use of data in future reanalyses. For example, Bosilovich et al. (2011) used the GIO for radiosondes to examine low precipitation biases over Africa.

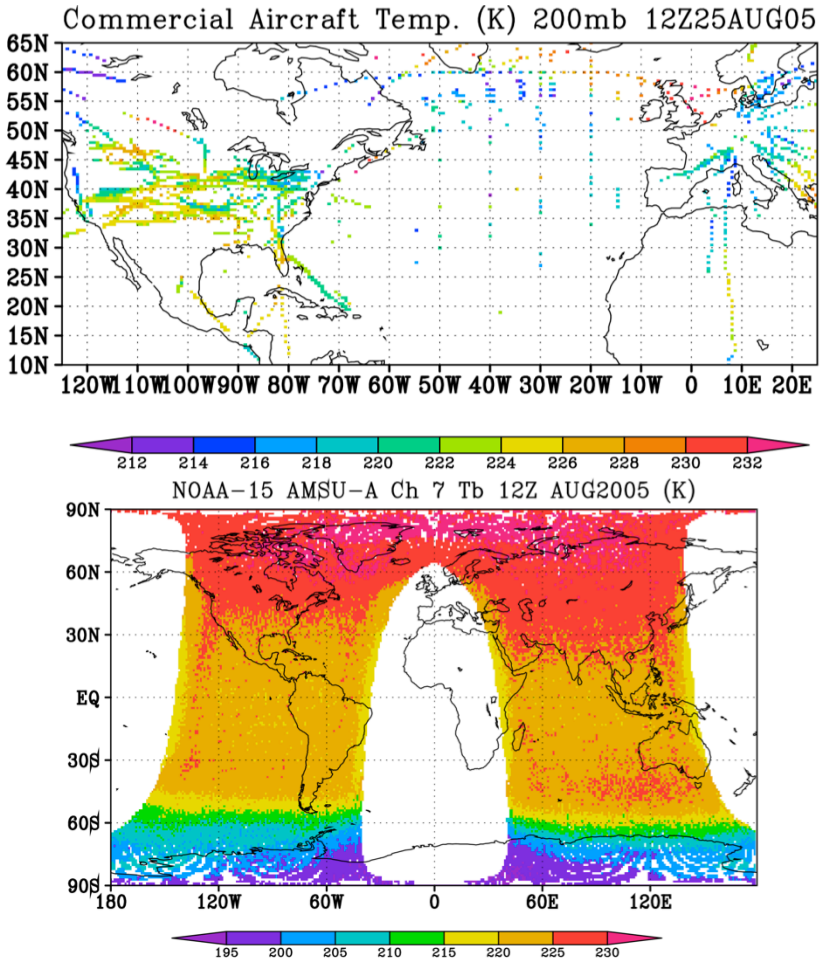


Figure 1. Examples of observations available through GIO, including instantaneous commercial aircraft temperature (200 hPa) at 12 UTC 25 August 2005 (top) and the August 2005 monthly average of 12 UTC brightness temperature (Ch 7, peaking near 200 hPa, regridded to 1 degree resolution) from NOAA-15 AMSU-A (bottom).

Figure 1 provides two examples of observations as presented by GIO. Commercial aircraft observations include three dimensional temperature and wind. The number of aircraft observations changes in time, with the volume increasing after 1991 when automated reporting became widely available. Even so, there are many regions not covered by the aircraft measurements. Upper levels are more densely observed, due to cruise level measurements, while mid and lower levels are sampled on ascent and descent flight paths. Aircraft routinely adjust their flight paths to optimize flight times and fuel usage, or safety considerations. Hence, evaluating decades of commercial aircraft observations or innovations in observation space is challenging. The gridded data simplifies large-scale and long time period computations.

MERRA and other modern reanalyses assimilate level 2 satellite radiances rather than retrieved atmospheric profiles. Satellite instruments can provide global cover, but there are many instruments and channels, each with their own limited duration throughout the historical period. The GIO product provides the radiances for each assimilated instrument and channel binned to the MERRA grid. The gridding simplification may not be appropriate for all uses, but it provides easy access to assimilated radiances for general evaluation. Figure 1 shows the monthly-averaged radiances from Channel 7 on AMSU-A for 12 UTC in August 2005.

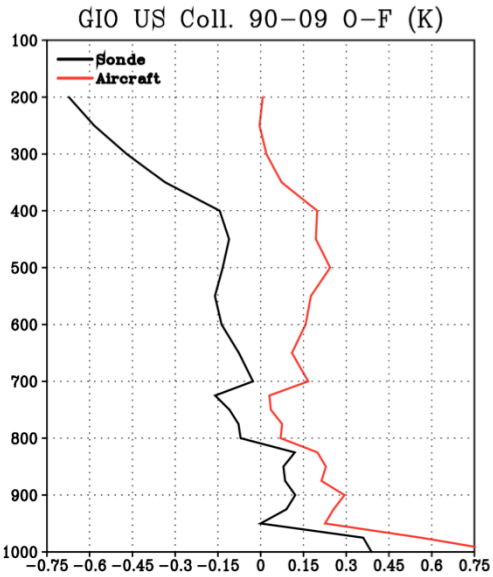


Figure 2. Vertical profile of observation minus forecast residuals (O-F) for collocated radiosonde (black curve) and commercial aircraft observations (red curve) assimilated in MERRA from 1990-2009 over the continental United States.

Figure 2 provides an example of the strengths of this simplification of the reanalysis innovations. Using the GIO, commercial aircraft data were filtered to coincident RAOB measurements over the United States. This collocation was performed for every six-hour analysis cycle for the period of 1990-2009 when there were high numbers of aircraft observations over the U.S. The collocations are limited in that the range of a match includes the six-hour analysis window and the area of a MERRA grid box. However, the data can be evaluated over decades, whereas most collocation comparisons in the literature consider only limited space and time domains. Figure 2 shows the temperature O-F for this collocation averaged for the 20-year period. In the upper levels, the cruise level aircraft observations dominate the analysis and so have small forecast error. Previous research identified warm biases in the aircraft data, so that the impact on radiosonde forecast error shows substantial increases in the upper levels. Subsequent versions of the GIO processing will include the bias correction information used in the data assimilation system. The processing will also be automated for new reanalyses and the GMAO near real-time system.

References

Bosilovich, M.G., F.R. Robertson and J. Chen, 2011: Global Energy and Water Budgets in MERRA. J. Climate 24, 5721–5739. doi: 10.1175/2011JCLI4175.1

Bosilovich, M.G. and A. M. da Silva, 2013: Evaluation of the Relative Contribution of Observing Systems in Reanalyses: Aircraft Temperature Bias and Analysis Innovations. *17th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS),* Austin TX, 7 January 2013.

Publication

da Silva, A.M. and M.G. Bosilovich, 2013: Evaluation of the Relative Contribution of Observing Systems in Reanalyses (*in preparation*).

Including Linearized Moist Physics in GMAO's Data Assimilation Tools

Daniel Holdway, Ronald Errico, Ronald Gelaro, and Ricardo Todling

An essential component in variational data assimilation is the linear version of the forecast model. Including moist physics in the linear atmospheric model is highly beneficial in terms of assimilating observations of atmospheric moisture. Further, it will improve the capability of monitoring and research tools, such as adjoint-based observation impacts and adjoint sensitivity studies that have been an important part of GMAO's research. Linearized versions of the moist physics schemes have been developed and tested in GMAO's data assimilation tools (Holdaway et. al., 2013).

Convection in the nonlinear system is modeled using the Relaxed Arakawa-Schubert (RAS) scheme. Using Jacobian sensitivities, an extensive study of linearity and stability in the RAS convection scheme has been performed (Holdaway and Errico, 2013). The scheme is found to be close to linear, especially for mid to deep convection, and very close to stable. Based on these findings an exact linearization of the RAS scheme has been developed, with some minor modifications. Due to complexity in the large-scale precipitation scheme, a highly simplified scheme that converts super-saturation to precipitation is linearized instead of the full scheme.

For a small number of profiles, an unrealistically large growth in the perturbation trajectory is encountered. Efficient filtering of these profiles is achieved though the diagnosis of steep gradients in the operator of the tangent linear model. With filtering turned on, these profiles are controlled and inclusion of linearized moist physics increases the correlation between the nonlinear perturbation trajectory and the linear approximation of the perturbation trajectory. Figure 1 compares the difference between the nonlinear perturbation trajectory and the tangent linear model perturbation trajectory at 500 hPa for the cases when dry physics only is used in the linear model and when moist physics is included. The large reduction in magnitudes in the right-hand plot shows that for a six-hour window (applicable to variational data assimilation) the moist physics improves the linear representation of the specific humidity perturbation trajectory.

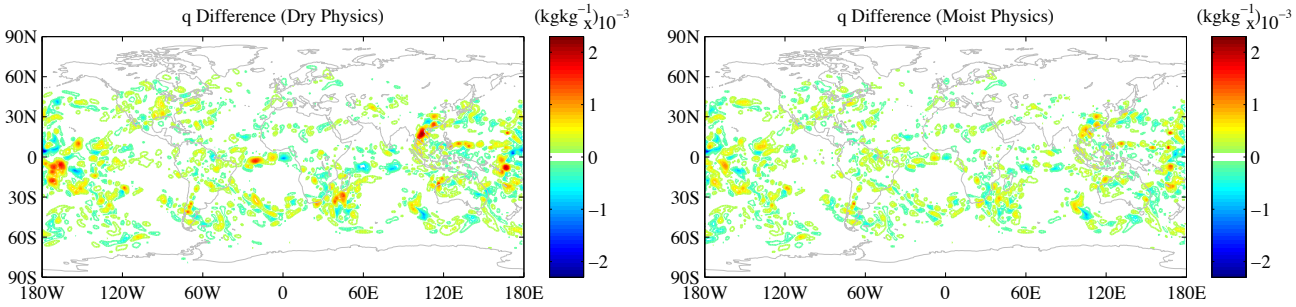


Figure 1. The difference between the nonlinear and linear perturbation trajectories for specific humidity when using only dry physics (left) and when including moist physics (right).

A month long observation impact experiment has been performed to compare the impacts obtained when dry versus moist physics is used in the adjoint model. Figure 2 shows the impact per analysis (left) and the fraction of beneficial observations (right). Impacts are for the period March 17th to April 17th, 2012. When moist physics are included in the adjoint the total reported impact increases for most instruments. The largest positive differences are seen for moist sensitive instruments such as AIRS, MHS, HIRS and wind observations based on feature tracking. For some instruments the number of beneficial observations is reported as decreasing while the total impact increases. Experiments were repeated for a moist norm (not shown). Using this metric to determine impacts further increases the reported impact from moist sensitive instruments. Overall, the ability of the linear model to capture the nonlinear impacts increases around 5% when moist physics are included in the adjoint model.

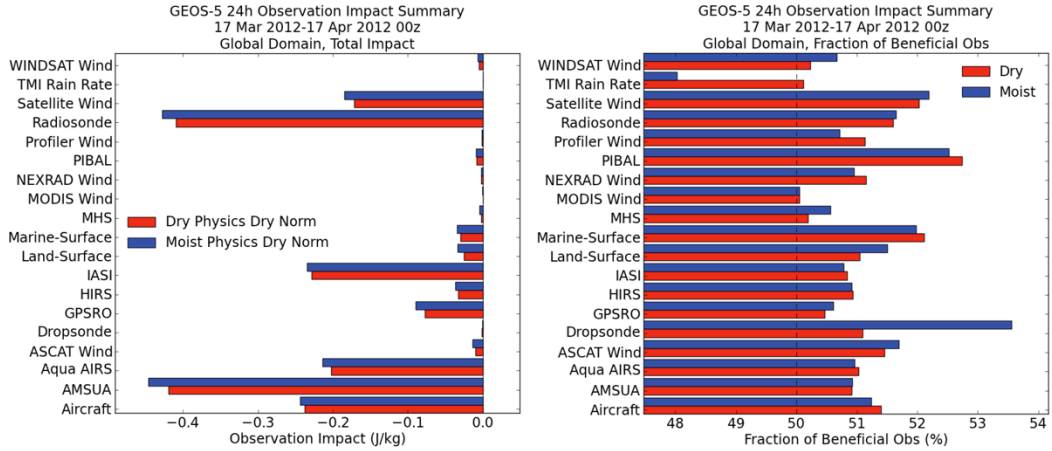


Figure 2. Observation impacts for March 17th to April 17th, 2012; total impact (left) and fraction of beneficial observations (right). Red bars show the current dry physics impacts; blue bars show the impacts when moist physics is included in the adjoint. The forecast metric uses a dry energy norm.

A case study examining the adjoint sensitivity for an intensifying North Atlantic storm is considered. Beginning on March 25th, 2012 at 00 UTC the adjoint, initialized with a 4° by 4° box around the storm's center, is integrated 24 hours. Figure 3 shows the adjoint sensitivity field for specific humidity after 24 hours (March 24th, 00 UTC) in red-yellow shading. Also shown is sea level pressure and convective precipitation. When only dry physics are included in the adjoint (left), sensitivity with respect to specific humidity is small and negligible compared with sensitivities to other fields. When moist physics is included in the adjoint (right), the sensitivity with respect to specific humidity increases three orders of magnitude and is comparable to the sensitivity with respect to the other fields. This shows both the importance of including linearized moist physics when performing adjoint sensitivities and also the importance of accurate initial moisture fields when forecasting the total energy of a storm of this nature.

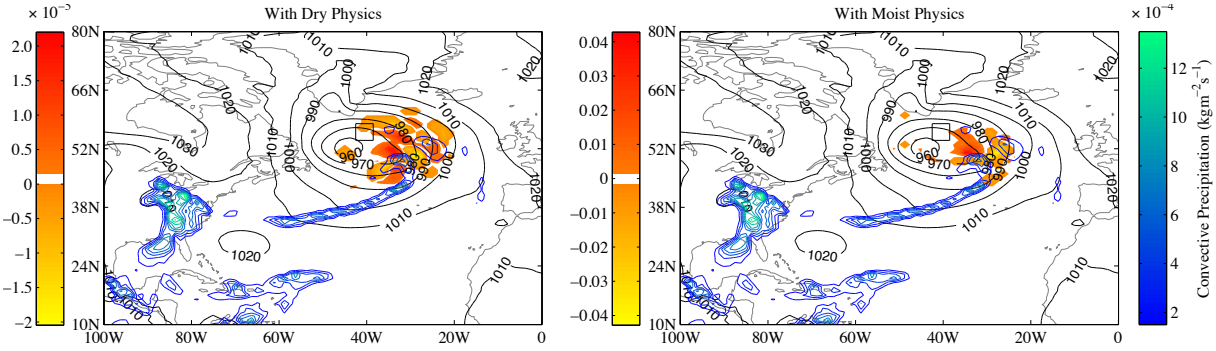


Figure 3. Specific humidity adjoint sensitivity for an Atlantic storm. Dry physics in the adjoint (left) and moist physics in the adjoint (right). Also shown is convective precipitation and sea level pressure.

Publications

Holdaway, D. and R. Errico, 2013: Using the Jacobian to assess a linearisation of the Relaxed Arakawa-Shubert convection scheme. *Quart. J. Roy. Meteorol. Soc.* (*in press*).

Holdaway, D., R. Errico, R. Gelaro and R. Todling, 2013: Inclusion of linearized moist physics in NASA's Goddard Earth Observing System Data Assimilation Tools. *Mon. Wea. Rev. (submitted*).

Assimilation of Atmospheric Motion Vectors in GEOS-5

Ronald Gelaro, Dagmar Merkova, Will McCarty, and King-Sheng Tai (GMAO)
Patricia Pauley and Nancy Baker (NRL)

The GMAO and NRL-Monterey are conducting a study to help understand the overall larger beneficial impact of atmospheric motion vectors from geostationary satellites (GeoWinds) in the U.S. Navy's global NWP system as compared with other operational forecast systems and, if possible, improve the use of these data in both the NASA and NOAA systems. Possible reasons for the different impact include (1) the greater number of GeoWinds assimilated in the Navy system, (2) the Navy's data selection and quality control procedures, including the use of carefully averaged "super-ob" wind vectors, and (3) the influence of other data types in the assimilation system, especially the greater number of satellite radiances assimilated in the NASA and NOAA systems. Cycling data assimilation experiments, including forecasts and adjoint-based observation impact calculations, have been conducted with GEOS-5 for a two-month period during the 2010-2011 boreal winter season. Results from a control experiment that includes all GeoWinds and other data types assimilated operationally in GEOS-5 were compared with those in which the GEOS-5 GeoWinds were replaced by ones used by NRL. In contrast with the control GeoWinds, the NRL GeoWinds are assimilated as super-obs and include GeoWinds produced by the University of Wisconsin's Cooperative Institute for Meteorological Satellite Studies (CIMSS) and the Air Force Weather Agency (AFWA), in addition to those produced by the operational data providers (NESDIS, EUMETSAT and JMA). On average, there are roughly three times as many GeoWinds in the NRL set as in the control.

Figure 1 compares the relative impacts of selected observing systems assimilated in GEOS-5 for the control experiment and experiments using various configurations of the NRL GeoWinds in terms of their fractional contributions to the reduction of a global measure of 24-h forecast error. The measure combines errors in wind, temperature and surface pressure with respect to the verifying GEOS-5 analysis. In all cases, radiosondes and AMSU-A radiances have the largest beneficial impact, with each providing 20-25% of the total error reduction due to the assimilation of all observations. GeoWinds rank fifth and provide 8% of the total error reduction in the control experiment, but rank third and have roughly double the contribution to the total error reduction when the entire set of NRL GeoWinds is assimilated (NRLAMV). Note that the contributions from other leading observing systems are reduced by 10-15% in the NRLAMV experiment, highlighting the compensating effects of changing the number and mix of observations in the assimilation system. Reducing the number of satellite radiances assimilated in GEOS-5 by roughly 40% increases the impact of the NRL GeoWinds slightly (NLoRAD), but not enough to off-set the detrimental effects of the reduced number of radiances on 5-day forecast skill, especially in the Southern Hemisphere (not shown). Interestingly, assimilating only a limited subset of the NRL GeoWinds corresponding to those types in the control set (NGMAO) also increases the GeoWind impact relative to the control experiment, suggesting that the super-ob procedure used by NRL may be beneficial.

The overall greater volume of NRL GeoWinds appears to be the primary reason for their large impact. The effects of the use of super-obs, as well as fewer radiance data, are secondary and vary by season and location. Experiments are ongoing to examine these and other aspects of the results in more detail.

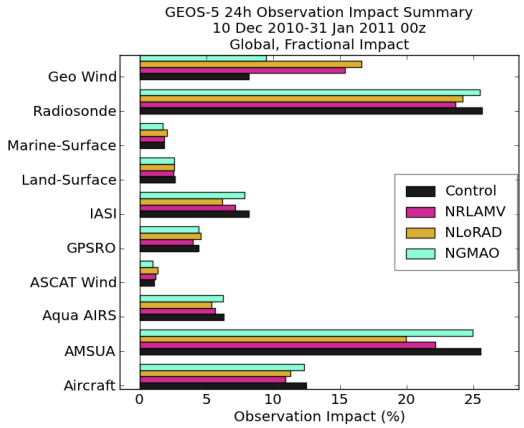


Figure 1. Impacts of various observing systems on the GEOS-5 24-hour forecasts from 00 UTC during the period 10 Dec 2010 – 31 Jan 2011. Results are shown for experiments using different configurations of the control and NRL GeoWinds (see text), and are expressed in terms of the fraction of the total error reduction due to assimilation of all observations based on a global error measure. The observation impacts are computed using the adjoint of the GEOS-5 atmospheric data assimilation system.

A Comparison of Clear-Sky Land Skin Temperature from GEOS-5 and from Geostationary Satellite Retrievals

Rolf Reichle and Qing Liu (GMAO)
Benjamin Scarino, Patrick Minnis, and Rabindra Palikonda (NASA LaRC)

Atmospheric models rely on accurate initial radiometric and surface conditions for better short-term meteorological forecasts and improved evaluation of global climate models. Remote sensing of the Earth's energy budget, particularly with instruments flown on geostationary satellites, allows for near-real-time evaluation of cloud and surface radiation properties. The persistence and coverage of geostationary remote sensing instruments grant the frequent retrieval of quasi-global skin temperature. Among other cloud and clear-sky retrieval parameters, NASA Langley Research Center (LaRC) provides a non-polar, high-resolution, near-real-time land surface skin temperature dataset from the Geostationary Operational Environmental Satellite (GOES) series. This study characterizes the diurnal and seasonal biases and root-mean-square differences (RMSDs) between the satellite retrievals and modeled land surface temperature (LST) values from GEOS-5.7.2 Atmospheric Data Assimilation System (ADAS) at 0.25° × 0.3125° resolution in preparation for assimilation of the retrievals into GEOS-5.

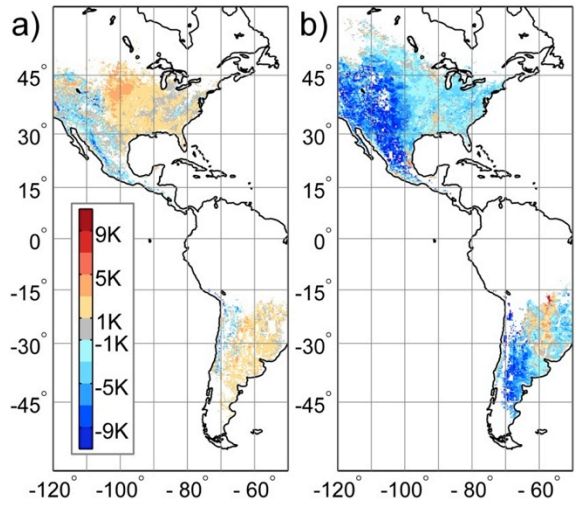


Figure 1. Mean difference (K) between GEOS-5 land surface temperature and GOES-13 HRTP for clear-sky conditions at (a) 06:00 UTC (nighttime) and (b) 18:00 UTC (daytime), averaged over 1 August 2011 to 31 July 2012. Areas in white indicate that metrics were not computed due to insufficient data.

Figure 1 shows the nighttime (06:00 UTC) and daytime (18:00 UTC) mean differences between the model and satellite estimates (GEOS-5 minus GOES-13) for the 12-month period from 1 August 2011 to 31 July 2012. The mean differences during the night are typically within ±3 K (Fig. 1a). During the day, the mean differences are considerably larger (Fig. 1b). In the western US, western Mexico, and western Argentina GEOS-5 LSTs are cooler than the GOES-13 retrievals by up to 10 K. Such large daytime differences, however, can be expected. While the satellite retrievals describe a true "skin" temperature, the GEOS-5 LST estimates are for a prognostic model surface layer with a small but non-zero heat capacity. The peak daytime model temperatures are therefore considerably lower than the satellite-retrieved skin temperatures. The situation is different in east-central Argentina, where GEOS-5 LSTs exceed the GOES-13 retrievals by as much as 9 K. An analysis of GEOS-5 cloud optical depth estimates suggests that these counter-intuitive differences are most likely caused by errors in the GEOS-5 moist physics and cloud parameterizations, which lead to strongly biased model temperatures that adversely impact model skin temperature estimates even during clear days (not shown).

Figure 2 shows daytime (18:00 UTC) seasonal RMSD values (excluding seasonal mean differences) between GEOS-5 LST and GOES-13 retrievals. The RMSD values are usually between 1 and 4 K, except during summer. The RMSD values exceed 4 K during JJA in much of the north-central US and during DJF in portions of Argentina and Uruguay. The lowest RMSD values of 1-2 K are found during cooler conditions (e.g., in the eastern and central US during DJF, in the southern US during SON, and in eastern Argentina and Uruguay during JJA). As can be expected, the nighttime (06:00 UTC) RMSD values are considerably lower and are typically 1-3 K across the domain and throughout the year (not shown).

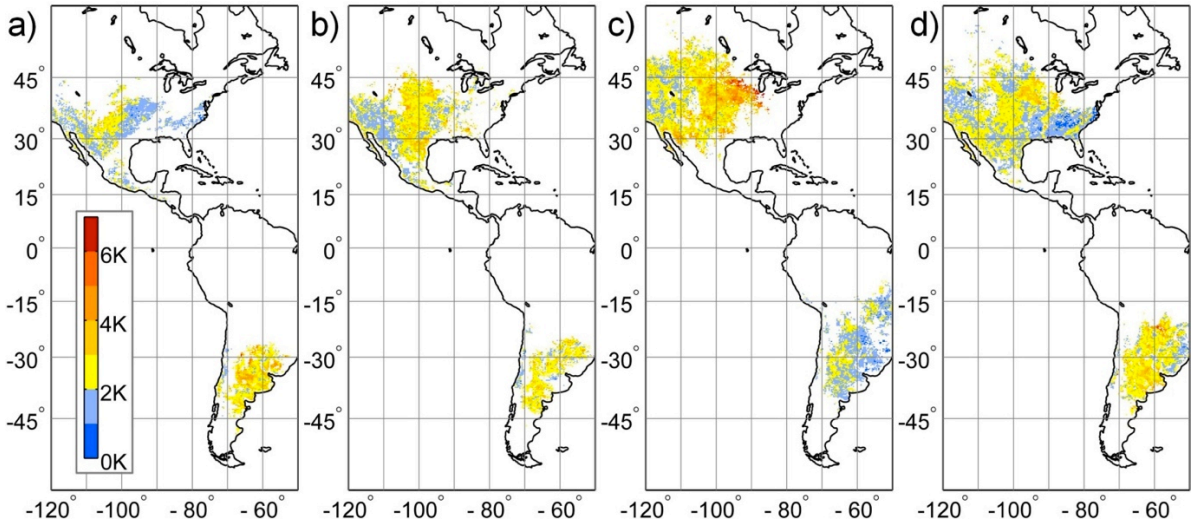


Figure 2. Daytime (18:00 UTC) seasonal RMSD (excluding seasonal mean difference) in degrees Kelvin between GEOS-5 land surface temperature and GOES-13 HRTP for clear-sky conditions for (a) DJF (Dec 2011-Feb 2012), (b) MAM (Mar-May 2012), (c) JJA (Aug 2011, Jun-Jul 2012), and (d) SON (Sep-Nov 2011).

The area-average diurnal cycles of the annual bias and RMSDs (excluding annual mean differences) are shown in Figure 3. The domain-averaged bias is typically below 1 K during the night (00:00 UTC to 09:00 UTC) and is considerably larger during the day (up to -4 K at 18:00 UTC). Similarly, the RMSD values are typically below 2 K during nighttime, including early morning and evening (00:00 UTC to 12:00 UTC). During the daytime (15:00 UTC to 21:00 UTC), however, the RMSD values exceed 2 K, with a maximum of 3 K at 18:00 UTC.

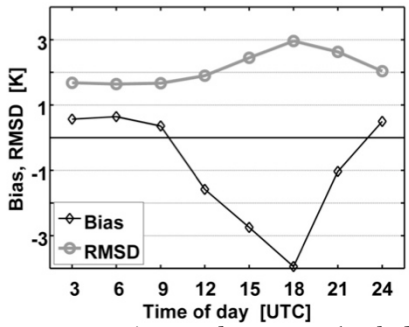


Figure 3. Bias and RMSD (excluding annual mean differences) in degrees Kelvin between GEOS-5 land surface temperature and GOES-13 skin temperature for clear-sky conditions, averaged over 1 August 2011 to 31 July 2012 and the areas in North, Central, and South America shown in Figures 1 and 2.

Looking forward, the comparison of the GOES-13 skin temperature retrievals with LST estimates from the GEOS-5 atmospheric analysis system is a necessary step towards the assimilation of the geostationary retrievals into the GEOS-5 system. The spatial and temporal variations of the biases shown in Figures 1 and 3 must be addressed as part of the assimilation system, which can be accomplished using a bias estimation and correction approach. Furthermore, the seasonal variations of the RMSD values (Fig. 2) should be reflected in the model and observation error covariances that are required for the assimilation of the satellite observations into GEOS-5. Our analysis also reveals deficiencies in the modeling of clouds in the GEOS-5 atmospheric model during summer over portions of Argentina. These issues manifest themselves in large and counter-intuitive mean differences between the GEOS-5 land surface temperatures and the GOES-13 satellite retrievals, which can only be fully resolved through improvements in the GEOS-5 atmospheric model.

Publication

Scarino, B., P. Minnis, R. Palikonda, R.H. Reichle, D. Morstad, C. Yost, B. Shan, and Q. Liu, 2013: Retrieving clear-sky surface skin temperature for numerical weather prediction applications from geostationary satellite data. *Remote Sensing,* **5**, 342-366, doi:10.3390/rs5010342.

Global Calibration of the GEOS-5 L-band Microwave Radiative Transfer Model over Non-Frozen Land Using SMOS Observations

Gabriëlle De Lannoy and Rolf Reichle (GMAO)
Valentijn Pauwels (Monash University)

Assimilating low-frequency (1-10 GHz) passive microwave observations into land surface models should improve estimates of land surface conditions and hence weather and climate predictions. Global observations of brightness temperatures (Tb) are available from the L-band Soil Moisture Ocean Salinity (SMOS), and similarly low frequency Tb observations are expected from the upcoming NASA Soil Moisture Active Passive (SMAP) mission. In this study, a microwave radiative transfer model (RTM) is coupled to the Catchment land surface model in GEOS-5 in preparation for the assimilation of global Tb observations from SMOS and SMAP as part of a radiance-based soil moisture analysis.

Figure 1 shows differences between climatological mean SMOS Tb observations and Tb forecasts using three different literature-based sets of RTM-parameters:

- "Lit1" refers to RTM parameters that are proposed for the future SMAP L2/3 radiometer retrieval products,
- "Lit2" refers to RTM parameters collected from published studies using the L-band Microwave Emission of the Biosphere Model (L-MEB), the Land Surface Microwave Emission Model (LSMEM), or the Community Microwave Emission Modelling Platform (CMEM), and
- "Lit3" parameters are identical to the Lit2 parameters except for the microwave roughness *h*, which is set to values used for SMOS monitoring at ECMWF.

All three sets of literature-based RTM parameters lead to substantial biases when used in the GEOS-5 system, with Lit1 being too cold by up to 50 K and Lit3 being too warm all over the globe. Even though Lit2 estimates are nearly unbiased in the global average, there are still significant regional biases in the simulated Tbs. Such biases would hamper the assimilation of Tb from SMOS or SMAP. Therefore, we calibrated the relevant RTM parameters to obtain climatologically unbiased Tb estimates. After calibration, the Tb simulations show a global absolute bias of 2.7 K, as shown in Figure 1d. It should be emphasized that an RMSE of

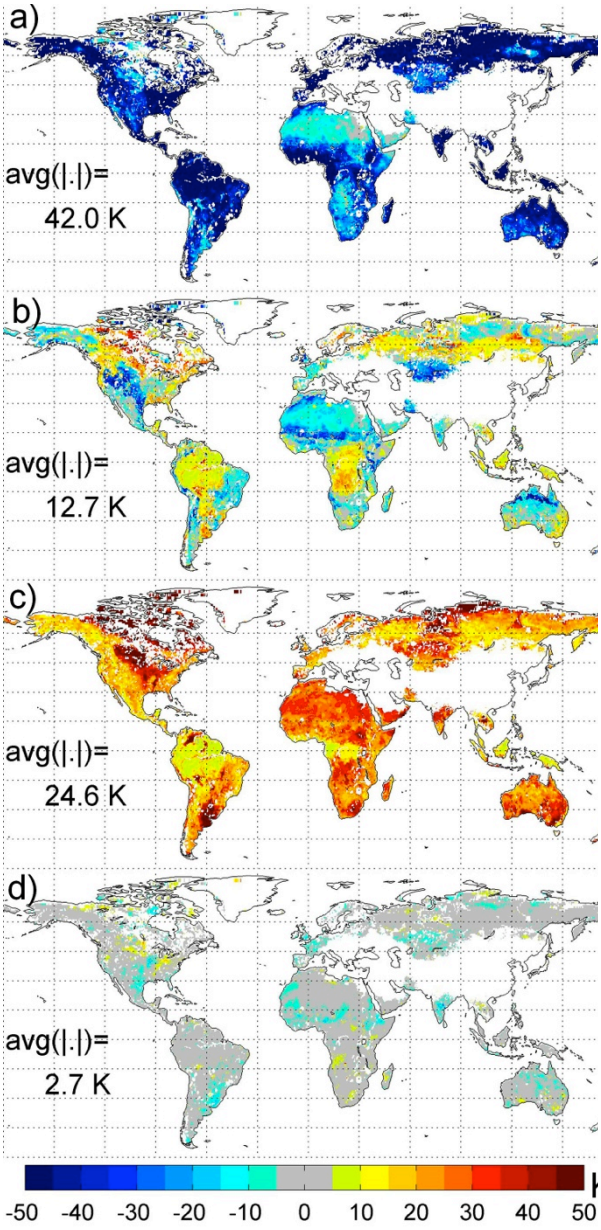


Figure 1. Difference in degrees Kelvin between one-year (1 July 2010 – 1 July 2011) mean values of horizontally polarized Tb at 42.5° incidence angle from GEOS-5 and SMOS observations for (a) Lit1, (b) Lit2, (c) Lit3, and (d) calibrated parameters. Within each subplot, "avg(|.|)" indicates the average absolute difference across the globe (excluding regions impacted by open water or radio-frequency interference that are shown in white).

approximately 10 K remains, due to seasonal biases and short-term errors, which will be addressed in the Tb data assimilation system.

The calibration minimized the difference between modelled and observed climatological means and standard deviations, as well as the deviation of the optimized parameters from prior guesses (Lit1, Lit2, and Lit3), at each individual location. Through a number of experiments, we determined that it is best to simultaneously calibrate (i) the microwave roughness h, (ii) the vegetation opacity τ and (iii) the scattering albedo ω.

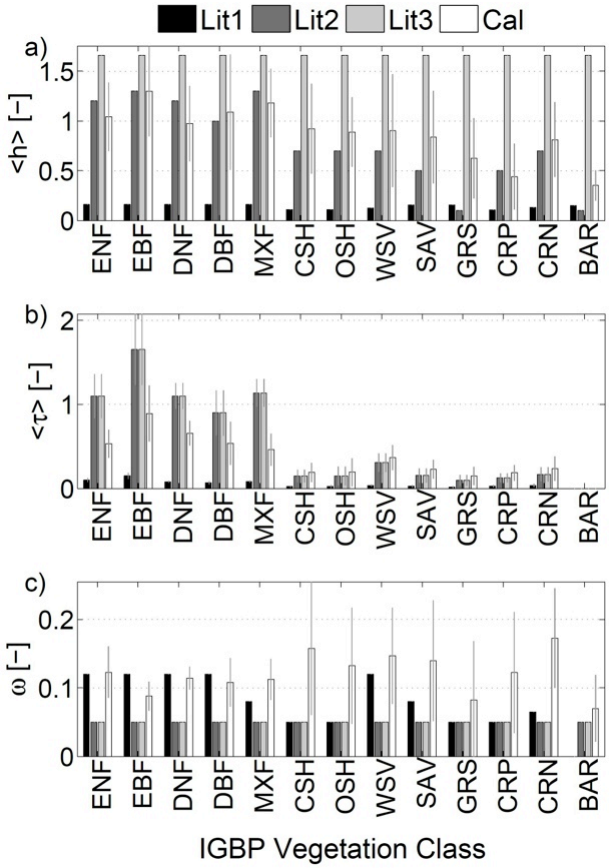


Figure 2. (a) Time-mean <h> (1 July 2010 – 1 July 2011), (b) time-mean <τ>, and (c) time-invariant ω; (Lit1, Lit2 and Lit3) before calibration, and (Cal) after calibration, spatially averaged by vegetation class. International Geosphere-Biosphere Programme (IGBP) vegetation classes are (ENF) Evergreen Needleleaf Forest, (EBF) Evergreen Broadleaf Forest, (DNF) Deciduous Needleleaf Forest, (DBF) Deciduous Broadleaf Forest, (MXF) Mixed Forest, (CSH) Closed Shrublands, (OSH) Open Shrublands, (WSV) Woody Savannas, (SAV) Savannas, (GRS) Grasslands, (CRP) Croplands, (CRN) Cropland and Natural Vegetation, and (BAR) Barren or Sparsely Vegetated. Thin gray lines for Cal indicate the spatial standard deviation within each vegetation class.

Figure 2 shows the prior guesses (Lit1, Lit2, and Lit3) for the RTM parameters along with the optimized parameter values after calibration, averaged by vegetation class. Both the microwave soil roughness h and vegetation opacity τ are time dependent (because h depends on soil moisture and τ depends on the leaf area index) and are therefore presented as temporal averages (denoted with <.>). The figure suggests that the h is too low in Lit1 and too high in Lit3. After calibration, h values are closest to Lit2. The τ estimates distinguish clearly between high and low vegetation. Finally, ω is increased over low vegetation to reduce the vegetation effect in the simulated Tb.

Publication

De Lannoy, G.J.M, R.H. Reichle, and V.R.N. Pauwels, 2013: Global calibration of the GEOS-5 L-band microwave radiative transfer model over non-frozen land using SMOS observations. *J. Hydrometeor.*, **14**, 765-785, doi:10.1175/JHM-D-12-092.1.

Estimating Root Mean Square Errors in Remotely Sensed Soil Moisture over Continental Scale Domains

Clara Draper and Rolf Reichle (GMAO)
Richard de Jeu and Robert Parinussa (VUA, Netherlands)
Wolfgang Wagner and Vahid Naeimi (TU Wien, Austria)

Making the best use of recently developed remotely sensed soil moisture data sets requires confident understanding of their errors around the globe. However, current evaluation strategies are limited to a handful of intensely observed locations. Hence, this study compares two recently developed methods to provide global distributed estimates of the Root Mean Square Error (RMSE) in remotely sensed surface soil moisture: i) triple colocation and ii) error propagation through the retrieval algorithms. Triple colocation estimates errors using three independent data sets. Simulations from the GEOS-5 Catchment land surface are used as the third data set. The RMSEs in surface soil moisture derived from the Advanced Scatterometer (ASCAT), and the Advanced Microwave Scanning Radiometer (AMSR-E) are examined over a continental scale domain in North America, from January 2007 to October 2010 (4.75 years).

In the absence of a consensus about the climatology of soil moisture over large spatial scales, presenting a RMSE in soil moisture units requires that it be specified relative to an arbitrarily selected reference data set. In this study the RMSE is presented as a fraction of the time series standard deviation, fRMSE. The fRMSE metric is demonstrated to have several advantages over the traditional approach of presenting the soil moisture RMSE relative to a reference.

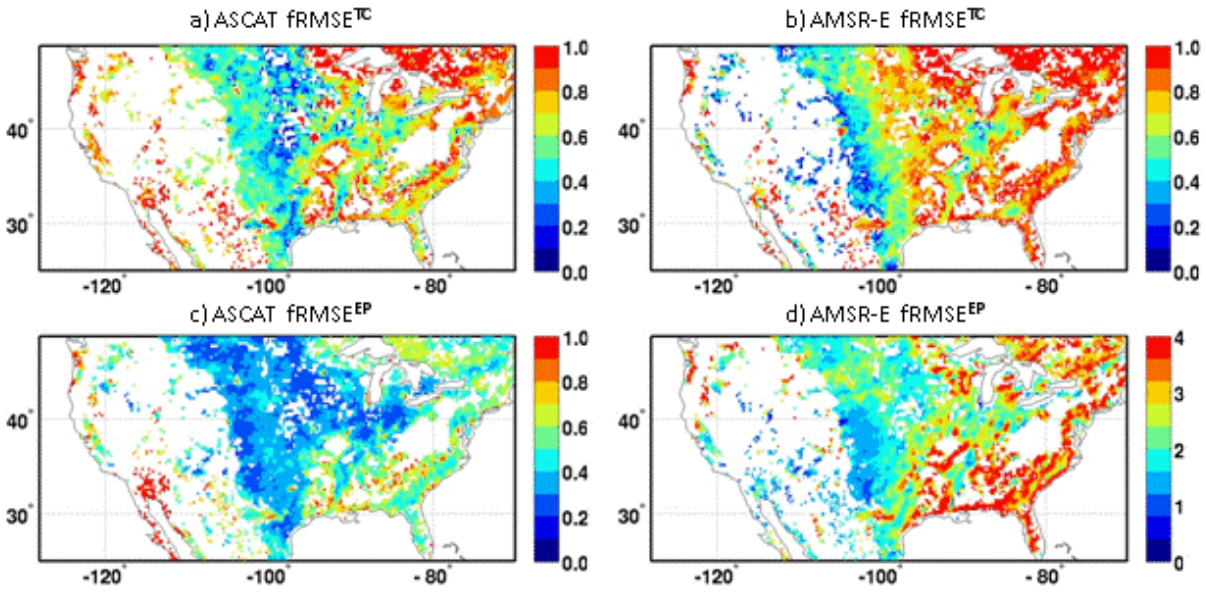


Figure 1. fRMSE of (left) ASCAT and (right) AMSR-E surface soil moisture retrievals, estimated using (upper) triple colocation and (lower) error propagation and plotted only where triple colocation results are available. Note the different color scale for the AMSR-E fRMSE^{EP} in d.

Figure 1 shows maps of the ASCAT and AMSR-E fRMSE calculated using triple colocation (fRMSE^{TC}) and error propagation (fRMSE^{EP}). Error propagation is designed to determine the spatial and temporal variability in the uncertainty of a given data set. The magnitude of the error propagation output depends on the uncertainties specified for the retrieval model parameters, and these uncertainties are not well understood at scales relevant to remote sensing. Hence, the unrealistically large fRMSE^{EP} for AMSR-E are not surprising (values above one indicate RMSE greater than the time series standard deviation).

In terms of the spatial variability, both methods accurately detect instances of relatively low/high fRMSE in surface soil moisture retrievals. For each sensor, the fRMSE^{TC} and fRMSE^{EP} show similar patterns of relatively high/low errors in Figure 1, and the mean fRMSE for each land cover class in Figure 2 is

consistent with expectations. For example, Figure 2 shows that there is a general increase in the fRMSE with increasing vegetation leaf area index (LAI). Also, the ASCAT fRMSE is very high for open shrubs (Fig. 2a), which are located in the arid southwest of the domain, consistent with known deficiencies in the ASCAT change-detection retrieval model in arid regions. While they generally agree, there are some discrepancies between the $fRMSE^{TC}$ and $fRMSE^{EP}$ in Figure 2, from which shortcomings in the retrieval algorithms that lead to errors in the $fRMSE^{EP}$ can be identified.

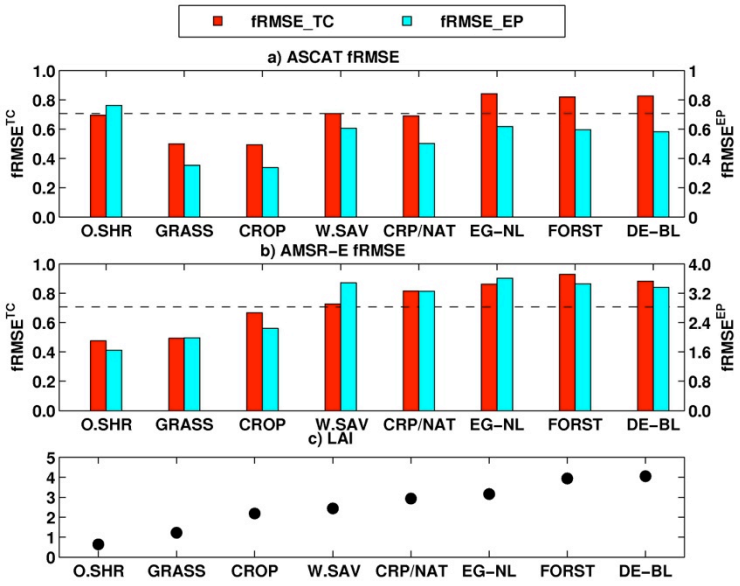


Figure 2. Mean of $fRMSE^{TC}$ and $fRMSE^{EP}$ for a) ASCAT and b) AMSR-E, and for c) LAI, as a function of land cover type.

Triple colocation accurately estimates the RMSE in remotely sensed soil moisture anomaly time series, and it is shown to be surprisingly robust to representativity differences between the soil moisture data sets used. In contrast to error propagation, triple colocation can be used to compare the errors in different data sets. Figure 2 shows that both sensors have similar accuracy across the land cover classes, with the exception of the high ASCAT fRMSE in arid regions noted above, and the relatively low ASCAT fRMSE over croplands – although neither of these differences is significant at 5%. As expected, both data sets perform well under sparse to moderate vegetation (LAI<3), which covers 63% of the domain plotted in Figure 1.

In summary, the substantial spatial variability in the fRMSE in Figure 1 highlights that an evaluation of soil moisture based on a limited number of locations is not necessarily representative of a larger domain. The RMSE defined by triple colocation (unbiased RMSE of anomalies from the seasonal cycle) and error propagation (errors associated with model input and parameters only) differs from the definition commonly used to specify target accuracies for remote sensing data products. However, both methods can indicate regions where the accuracy from intensely observed locations can be confidently extrapolated, and regions where otherwise unforeseen problems may occur. For most applications, triple colocation will be more useful, since in addition to predicting the spatial variability in the errors, it also produces RMSE estimates with a reliable magnitude.

Publication

Draper, C., R.H. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner, 2013: Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. *Remote Sens. Environ.* (*submitted*).

Estimating Passive Microwave Brightness Temperature over Snow-covered Land in North America Using the GEOS-5 Catchment Land Surface Model and an Artificial Neural Network

Bart Forman (University of Maryland), Rolf Reichle (GMAO), and Chris Derksen (Environment Canada)

It is well established for atmospheric data assimilation systems that the assimilation of satellite radiance observations is preferable to the assimilation of geophysical retrievals. The former approach incorporates the radiative transfer model (RTM) into the assimilation system and thereby avoids inconsistencies in the use of ancillary data between the assimilation system and the (pre-processed) geophysical retrievals.

The GEOS-5 Catchment land surface model supports the use of a physically-based microwave RTM for warm-season processes (De Lannoy et al., 2013). However, the snow model components in the Catchment model, like those in other global land surface models, are too simplistic to support physically-based RTM modeling in the presence of snow. Specifically, global snow models lack reliable estimates of snow microphysical properties (such as grain size, ice layers, and depth hoar) that are needed for physically-based forward modeling of the microwave brightness temperatures. We therefore constructed an empirical forward RTM for snow-covered land surfaces based on an Artificial Neural Network (ANN) and the GEOS-5 Catchment model.

The Catchment model state variables used as input to the ANN include the density and temperature of the snowpack at multiple depths, the temperature of the underlying soil, the overlying air, and the vegetative canopy, and the total amount of water equivalent within the snowpack. In addition, a cumulative temperature gradient index (TGI) is used as a proxy for snow grain size evolution in the presence of a vapor pressure gradient. Using the above inputs, the ANN is trained and (independently) validated using 10.7, 18.7, and 36.5 GHz microwave brightness temperatures at H- and V-polarization from AMSR-E. The independent validation is accomplished as follows: From the 9-year AMSR-E data record, each single year is withheld in turn from the ANN training, and skill metrics for the resulting ANN predictions are computed only against the AMSR-E data that have been withheld from the ANN training.

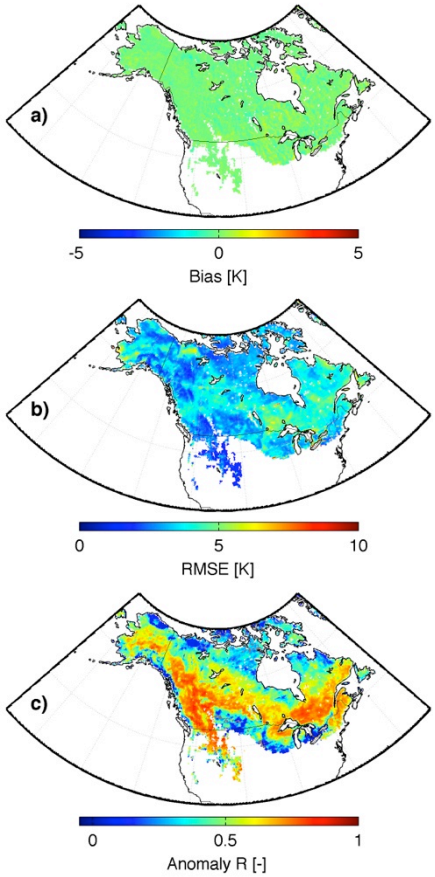


Figure 1. (a) Bias, (b) RMSE, and (c) anomaly R for ANN simulated 10 GHz V-polarized Tb from 1 September 2002 to 1 September 2011 vs. AMSR-E observations not used in training. Anomaly R values not statistically different from zero at the 95% significance level based on a Fisher Z transform are shown in gray. Such non-significant R values occur in only a few very small regions.

Figure 1 demonstrates the performance of the ANN predictions relative to AMSR-E measurements that were not used during training. The figure illustrates the overall ability of the ANN to predict Tbs for the 10 GHz V-polarized channel. The ANN predictions are essentially unbiased (relative to the AMSR-E measurements) across the 9-year period (Fig. 1a). The RMSE is typically less than 5 K (Fig. 1b). In addition, the ANN demonstrates skill in predicting interannual variability, with anomaly R values well above 0.5 over large parts of North America (Fig. 1c). Relatively low skill can be seen in areas along the southern periphery, where the snowpack is relatively thin and ephemeral, as well as in areas north of the boreal forest, where sub-grid scale lake ice (which is not modeled in the land surface model) is common. In short, Figure 1 suggests considerable skill by the ANN at predicting interannual variability in 10 GHz V-polarized Tbs across North America with negligible bias and a reasonable RMSE. The RMSE is

somewhat higher but still reasonable (less than 10 K) for the higher frequencies and for H-polarization Tb (see Figures 4-6 of Forman et al., 2013).

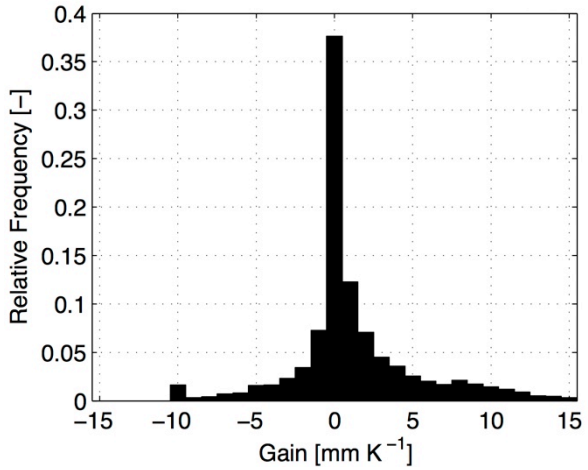


Figure 2. Histogram of the Kalman gain on 6 February 2003 for SWE versus ΔTb=(TbV(18 GHz)-TbV(36 GHz)).

We also assessed the potential for using the ANN as a forward observation operator in radiance-based snow assimilation. For this demonstration, the observations are considered to be in the form of spectral differences in V-polarization brightness temperatures, $\Delta Tb \equiv Tb_V(18\ GHz) - Tb_V(36\ GHz)$. Since ΔTb typically increases with increasing snow water equivalent (SWE), this spectral difference is commonly used to estimate SWE in retrieval algorithms. For the demonstration of the radiance-based assimilation considered here, observations of ΔTb imply that the resulting Kalman gain is proportional to error correlations between modeled SWE and ANN predictions of ΔTb. To obtain analysis increments, the Kalman gain would be multiplied by innovations in ΔTb (that is, the difference between actual AMSR-E observations of ΔTb and ANN predictions of ΔTb).

The Kalman gain computed for 6 February 2003 ranges from -10 mm K^{-1} to 15 mm K^{-1} as illustrated in Figure 2. A gain of 1 mm K^{-1} equates to an increase of 1 mm in the posterior (updated) modeled SWE for a 1 K innovation (that is, for a difference of 1 K between AMSR-E ΔTb measurements and ANN ΔTb predictions). Similarly, a negative Kalman gain in the presence of a positive-valued innovation would equate to a reduction in modeled SWE. Most importantly, the results suggest that there is a non-zero error correlation between the model SWE forecasts and the simulated ΔTb measurements across much of the North American domain. Overall, the results suggest that the ANN could serve as a computationally efficient observation operator for radiance-based snow data assimilation at the continental scale.

Reference

De Lannoy, G.J.M, R.H. Reichle, and V.R.N. Pauwels, 2013: Global calibration of the GEOS-5 L-band microwave radiative transfer model over non-frozen land using SMOS observations. *J. Hydrometeor.*, **14**, 765-785, doi:10.1175/JHM-D-12-092.1.

Publication

Forman, B.A., R.H. Reichle, and C. Derksen, 2013: Estimating passive microwave brightness temperature over snow-covered land in North America using a land surface model and an artificial neural network. *IEEE Trans. Geosci. Rem. Sens.* (*in press*), 10.1109/TGRS.2013.2237913.

Soil Moisture Assimilation and Precipitation Error Modeling in the Ensemble-based GEOS-5 Land Data Assimilation System

Rolf Reichle (GMAO)
Viviana Maggioni and Emmanouil Anagnostou (University of Connecticut)

A key issue in soil moisture data assimilation is that observational and modeling uncertainties are poorly known, and incorrect assumptions about these errors may compromise the efficiency of the land data assimilation system (LDAS). It is thus crucial to investigate the impact of the error characterization on the assimilation of soil moisture observations, in particular because LDASs often use very simplistic error models. As rainfall is the dominant meteorological forcing input to the land surface model for soil moisture estimation, a more comprehensive characterization of rainfall uncertainty may improve soil moisture estimates. Since soil moisture temporally integrates antecedent precipitation and is subject to lower and upper limits, the variability of errors in soil moisture is typically smaller than that of errors in precipitation. This error variance relationship is not linear and depends on the error properties of the rainfall fields (Maggioni et al., 2012b). Maggioni et al. (2011) showed that the use of a complex error model to characterize the spatial variability of rainfall errors could better capture soil moisture error properties. Furthermore, in a synthetic numerical assimilation experiment, Maggioni et al. (2012a) demonstrated that using the more elaborate rainfall error model may slightly improve surface and root zone soil moisture estimates obtained from assimilating soil moisture retrievals.

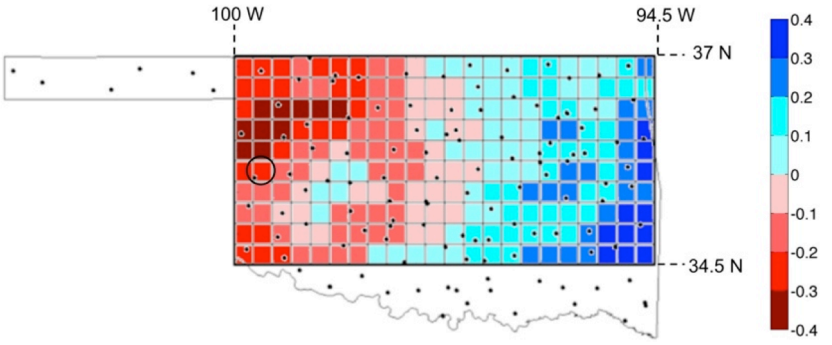


Figure 1. Experiment domain in Oklahoma, USA. Colors indicate the rainfall climatology parameter (dimensionless) used to bin the results in Figure 2. Blue (red) tones indicate "wet" ("dry") areas.

This study expands the synthetic experiment of Maggioni et al. (2012) by evaluating the assimilation of near-surface soil moisture retrievals from AMSR-E. Data were assimilated for a three-year period (2004-2006) over a domain in Oklahoma (Fig. 1) into the Catchment land surface model with the ensemble-based GEOS-5 LDAS.

Two different rainfall error models were considered: a complex, multi-dimensional model (SREM2D) and the simpler model (CTRL) used in the GEOS-5 LDAS. Specifically, five experiments were conducted: a "benchmark" simulation forced with Stage IV radar rainfall, and four experiments obtained by perturbing satellite rainfall fields with the two rainfall error models of different complexity, with and without the assimilation of AMSR-E soil moisture retrievals. Satellite rainfall was from the NOAA Climate Prediction Center morphing (CMORPH) product. Surface and root zone soil moisture outputs from each experiment were compared against Oklahoma Mesonet in-situ measurements.

Figure 2 shows that the assimilation of satellite soil moisture retrievals provides a significant improvement of surface and root zone soil moisture estimates, indicating the ability of the model update to propagate to deeper soil levels. The improvement due to assimilation is apparent also in comparison to the AMSR-E retrievals themselves, i.e., starting from two poorer estimates of soil moisture, data assimilation provides a superior estimate. We also note that soil moisture estimates from data assimilation exhibited correlations higher than, or at least as high as, those estimated from the model

forced with the most accurate rainfall (i.e., Stage IV). Furthermore, we observe that all simulation and assimilation results (but not the AMSR-E retrievals) perform better for the wetter conditions. Finally, the use of the more complex SREM2D rainfall error model leads to only marginally better soil moisture analyses, which suggests that the simpler CTRL rainfall error model of the GEOS-5 LDAS may be adequate for soil moisture data assimilation. Nevertheless, the use of a more sophisticated error model, such as SREM2D, is suggested in future land data assimilation studies in order to provide a more realistic representation of the sources and nature of errors in precipitation retrievals.

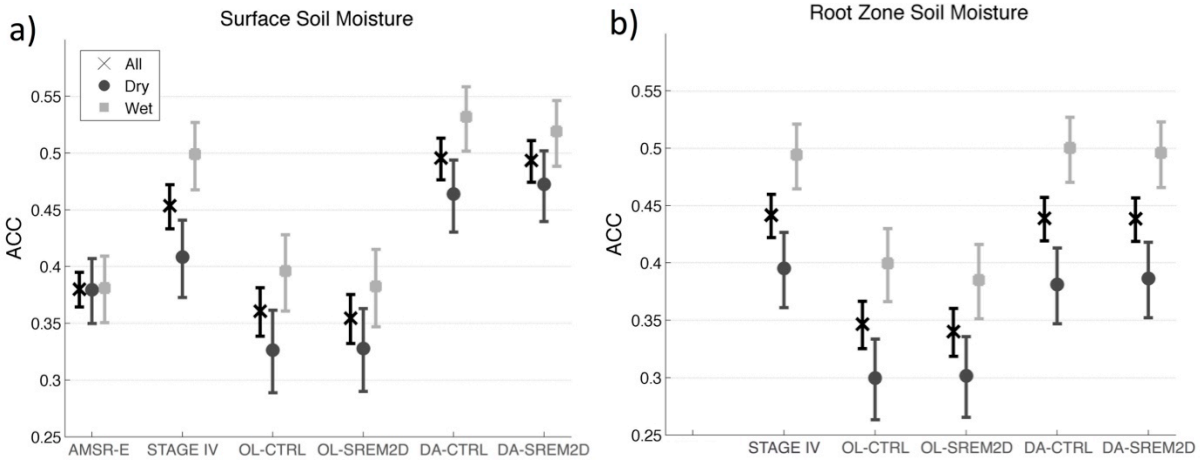


Figure 2. Domain-average of daily anomaly correlation coefficients (ACCs; dimensionless) calculated against MESONET station observations for (a) surface and (b) root zone soil moisture with 95% confidence intervals for all grid cells and binned by dry and wet conditions. Simulation experiments without assimilation of AMSR-E retrievals are labeled "OL" (for open loop). Experiments with assimilation of AMSR-E retrievals are labeled "DA" (for data assimilation). The rainfall error model (CTRL or SREMS2D) used in the simulation or assimilation experiments is also indicated. "Stage IV" is a simulation experiment that uses high-quality radar precipitation forcing.

The results are encouraging towards the use of satellite retrievals (of both soil moisture and precipitation) in LDAS, especially because of the increasing availability of satellite soil moisture and rainfall observations from the Soil Moisture and Ocean Salinity mission, the Soil Moisture Active Passive mission, and the Global Precipitation Measurement mission.

References

Maggioni, V., R.H. Reichle, and E.N. Anagnostou, 2011: The effect of satellite-rainfall error modeling on soil moisture prediction uncertainty. *J. Hydrometeor.,* **12**, 413-428, doi:10.1175/2011JHM1355.1.

Maggioni, V., R.H. Reichle, and E.N. Anagnostou, 2012a: The impact of rainfall error characterization on the estimation of soil moisture fields in a land data assimilation system. *J. Hydrometeor.,* **13**, 1107-1118, doi:10.1175/JHM-D-11-0115.1.

Maggioni, V., E.N. Anagnostou, and R.H. Reichle, 2012b: The impact of model and rainfall forcing errors on characterizing soil moisture uncertainty in land surface modeling. *Hydrology and Earth System Sciences,* **16**, 3499-3515, doi:10.5194/hess-16-3499-2012.

Publication

Maggioni, V., R.H. Reichle, and E.N. Anagnostou, 2013: The efficiency of assimilating satellite soil moisture retrievals in a land data assimilation system using different rainfall error models. *J. Hydrometeor.,* **14**, 368-374.doi:10.1175/JHM-D-12-0105.1.

Impact of the Assimilation of Aquarius Sea Surface Salinity in the GEOS-5 System

Guillaume Vernieres, Robin Kovach, Christian Keppenne, and Anna Borovikov

GMAO's ocean sea-ice data assimilation system aims to make the best possible use of satellite observations relevant to the ocean. While sea surface temperature (SST) has been measured remotely since 1981, until recently, estimates of sea surface salinity (SSS) have relied on very sparse in-situ observations. The advent of relatively high-resolution, global space-based measurements of SSS from the Aquarius/SAC-D mission has the potential to improve our ocean state estimate and forecasting ability.

Our most recent work consists of including Level-2 retrieved SSS from Aquarius into our integrated ocean data assimilation system (iODAS; Vernieres et al., 2012). Figure 1 illustrates the global probability density functions (PDFs) of the retrieved SSS and collocated in-situ observations of salinity in the upper 10 meters of the ocean. The large differences between the two PDFs are due to:

- Large biases and errors in the retrieval due to instrumental calibration problems.
- Possible differences between the in-situ observations, which represent the average salinity in the upper 10 m of the ocean, and SSS, especially in strongly stratified regions.

The bottom left panel of Figure 1 shows the regression plot of Aquarius SSS versus in-situ salinity. While the correlation is relatively high (0.8), implying that the instrument is capturing the proper variability, the root mean square error is too high for the data to be usable in a state estimation framework.

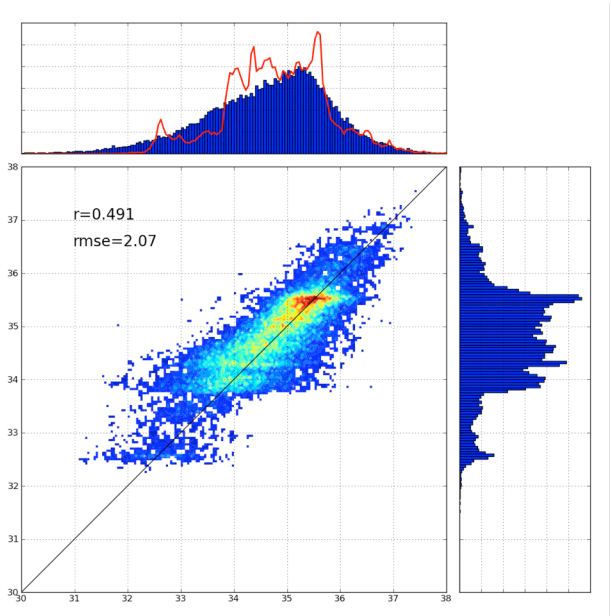


Figure 1. (Bottom right) PDF of upper ocean in-situ salinity measurements collocated with Aquarius SSS. (Top) PDF of Aquarius SSS at the in-situ location in blue. The red line corresponds to the PDF of upper ocean in-situ salinity measurement collocated with Aquarius SSS. (Bottom left) Regression plot of Aquarius SSS versus upper ocean in-situ observations, r is the correlation coefficient and rmse the root mean square error.

To compensate for these discrepancies, we developed a mapping from the retrieved SSS onto an estimate of upper ocean salinity that can directly be ingested by our assimilation system. The map is based on a feed forward neural network (FANN). The network topology is shown in Figure 2. As input, it takes the location of the retrieval, the day of the year, SST, and Aquarius SSS; it outputs a corrected bulk salinity.

In order to test our methodology and quantify the impact of assimilating Aquarius salinity retrievals, five simulations have been carried out, starting on November 3, 2011 from initial conditions identical to those used for the GMAO seasonal forecast and using the same external forcing from MERRA. Two simulations serve as references; one does not include any assimilation and is referred to as BASENODA. The other assimilates upper ocean in-situ salinity observations (from the surface down to 100 m) and will be referred to as BASEDA. The third simulation (RAW) assimilates the Aquarius retrieval without any pre-processing and is meant to illustrate the difficulties associated with the assimilation of the original data set. The fourth simulation assimilates the Aquarius retrieval, pre-processed with the neural network and

is referred to as ANN. Finally, our last experiment assimilates the upper ocean in-situ salinity observations from BASEDA and the same Aquarius salinity as ANN and is referred to as ALL.

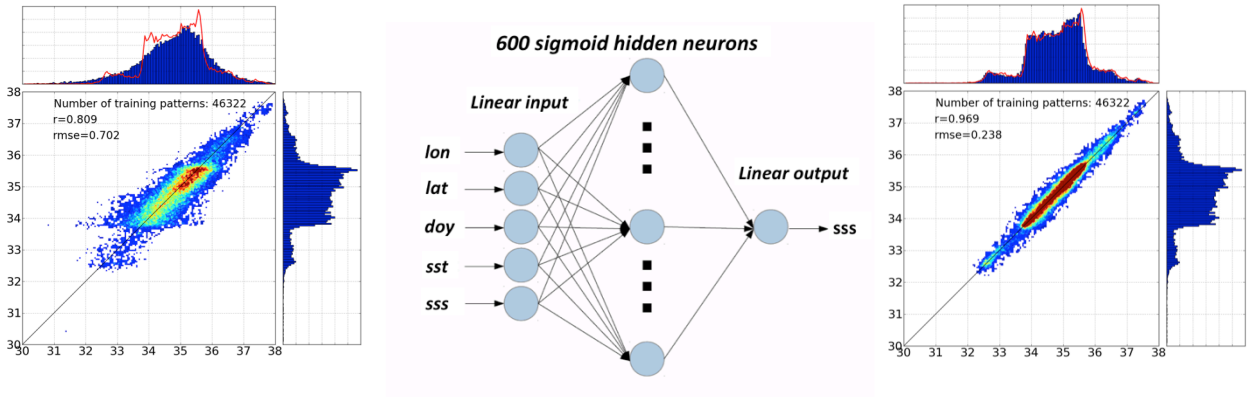


Figure 2. Input Aquarius SSS retrieval (left), FANN topology (middle) and FANN output after training (right).

Figure 3 shows the global RMS and mean of salinity observation minus forecast (OMF) differences using Argo observations. The mean OMFs from RAW (green line) are the worst of all experiments, with a maximum of 0.12 in October 2012, although this experiment is slightly better than the BASENODA in terms of RMS OMFs. The results from the other experiments are much as one would expect from a high quality data set, in which more observations result in a better estimate of the state and in a better forecast. The smallest mean and RMS OMFs is observed in the ALL experiment, which assimilates both Argo upper ocean salinity and the pre-processed Aquarius retrievals.

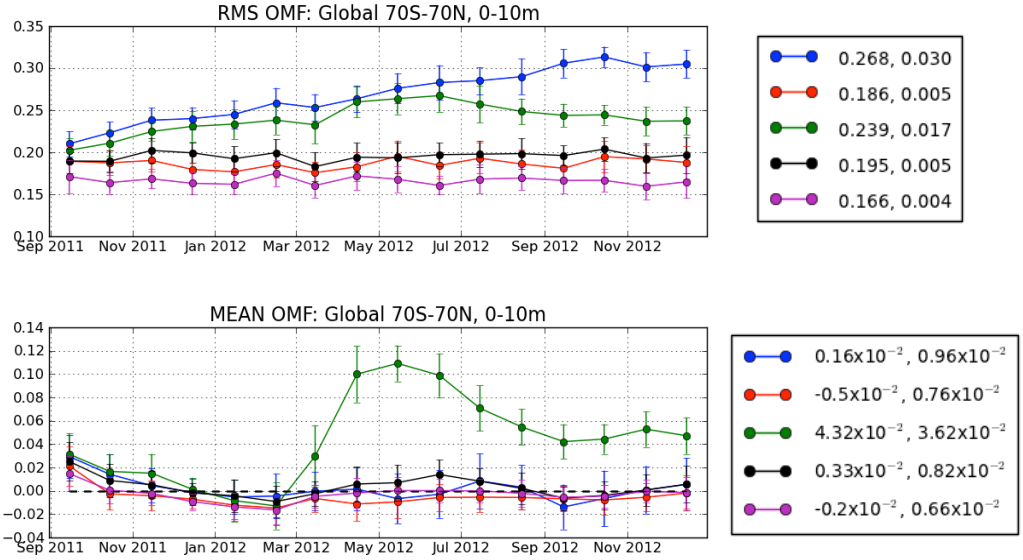


Figure 3. Global statistics of SSS OMFs at Argo locations.

Publications

Vernieres, G., M. Rienecker, R. Kovach, and C. Keppenne, 2012: The GEOS-iODAS: Description and Evaluation. *NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM—2012-104606*, Vol. 30, 61 pp.

Vernieres, G., M.M. Rienecker, R. Kovach, S. Akella, C. Keppenne, and A. Borovikov, 2013: The Impact of assimilation of Aquarius sea surface salinity data in the GEOS system. *(Draft ms)*.

Assimilating Global Ocean Chlorophyll from Suomi-NPP/VIIRS: Prospects for Extending the Ocean Color Time Series

Watson Gregg, Nancy Casey, and Cecile Rousseaux

The Suomi-NPP mission is intended to continue the time series of global ocean color data. It follows the succession of OCTS (1996-1997), SeaWiFS (1997-2010) and MODIS-Aqua (2002-present). As with all new ocean color sensors, VIIRS was expected to exhibit discrepancies with previous and concurrent missions, including SeaWiFS and MODIS-Aqua.

In an evaluation of the first 10 months of global ocean data from VIIRS, we find that the VIIRS global chlorophyll data compares favorably with that from MODIS-Aqua. The mean difference between the two is 3.6%, with VIIRS being higher. VIIRS modestly overestimates chlorophyll in the tropics. The differences are much larger in the northern high latitudes: VIIRs has a large negative bias along the western sides of both the North Pacific and North Atlantic and a positive bias in many other locations.

However, global mean chlorophyll representations are distorted by data gaps. Ocean color missions typically observe only about 15% of the ocean per day, due to inter-orbit gaps, insufficient light for detection at high latitudes, sun glint, clouds, and aerosols. Thus, biases caused by sampling are considerable. Data assimilation should remove such biases by producing complete global fields each day. Unfortunately, differences between VIIRS and MODIS-Aqua data assimilation products were even larger, increasing to nearly 16%. Differences between the satellite data in terms of location and extent of data gaps, as well as chlorophyll concentrations in the vicinity of the gaps, led to differences in daily assimilated fields that then aggregated to produce monthly mean relative biases.

Using a variation of a bias-correction method called the Empirical Satellite-In situ Data algorithm, we modified Aqua and VIIRS chlorophyll to reduce disparities. Instead of using in-situ data for bias correction, we use MODIS chlorophyll and tune the VIIRS radiances for a best fit (now called ESRIDS). At 9 km resolution, the results indicate strong agreement between the two data sets, with a satellite-weighted bias of 0.2%, uncertainty of 14.0%, and regression coefficient of 0.964 (significant at the 95% level).

The ESRIDS bias correction improves the consistency substantially (by nearly 40%), but discrepancies are still apparent, mainly in the tropics (Fig. 1). Minor modifications to the bias-correction procedure might reduce the differences even more. Accounting for the changes in temporal calibration would be helpful. The results hold promise for a consistent ocean color time series across missions.

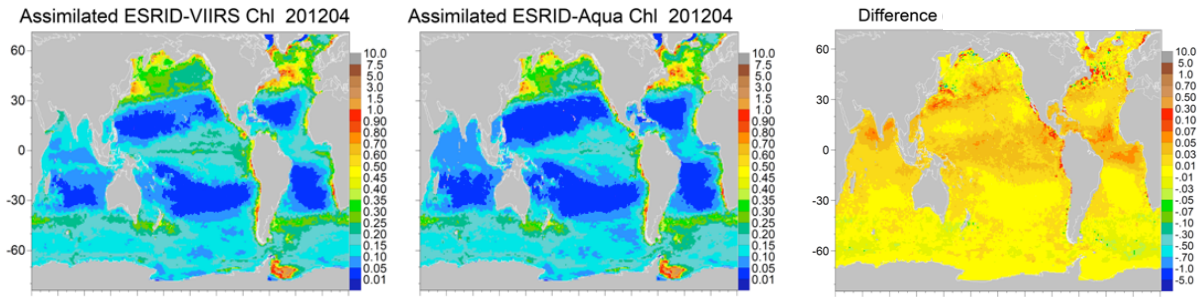


Figure 1. Assimilated ESRIDS VIIRS and MODIS-Aqua chlorophyll for April 2012 show good spatial agreement, but there are biases in the tropics. These biases are on the order of 10-20%, with VIIRS higher than Aqua. There are several possible explanations, such as improper bi-directional reflectance and polarization corrections, but the most likely is lack of cloud shadow masking in the data.

Simulation of MODIS Cloud Retrievals from GEOS-5

Galina Wind, Arlindo da Silva, Peter Norris (GMAO) and Steven Platnick (Code 610)

A project is underway to simulate full MODIS (MODerate resolution Imaging Spectroradiometer) cloud and aerosol retrievals from GEOS-5 model output. The method first simulates pixel scale cloud optical properties using GEOS-5 output; it then generates equivalent MODIS radiances from this simulated atmosphere. Finally, these radiances are substituted into the operational MODIS Adaptive Processing System (MODAPS) chain to produce a variety of remote sensing products that would normally be produced from the actual sensor output.

The project brings remote sensing methods together with model-generated fields. MODIS geo-location is used to sample GEOS-5 fields as if the MODIS instrument were flying over the model fields instead of Earth's surface. For cloud related fields, the sampling is not straightforward because cloud properties typically vary on small scales not adequately resolved by the ¼-degree resolution currently used for the GEOS-5 near-real-time assimilation and forecast system. To sample cloud fields, MODIS pixels for each GEOS-5 grid column are collected and the same number of sub-columns is generated for each grid column using a statistical model of sub-grid column variability. The general approach of Norris et al. (2008) is followed, namely using parameterized probability density functions (PDFs) of total water content for each model layer and a Gaussian copula to correlate layers in the vertical. In this application, we use skewed triangle PDFs, and for the Gaussian copula correlation matrix we use a fixed vertical correlation scale (100 hPa), modified by a Riishojgaard (1998) flow-dependent correlation in total water. The sub-columns thus generated are independent, but are subsequently "clumped," or given horizontal spatial coherence (of scale roughly 5 km), by using a horizontal Gaussian copula applied to condensed water path. This acts to give the generated clouds in each model grid column a reasonable horizontal structure, which is necessary because the MODIS cloud optical and microphysical properties retrieval algorithm has some spatial variance tests for potentially partly-cloudy pixels (removing cloud edges by so-called "clear-sky restoral") (Platnick et al., 2003).

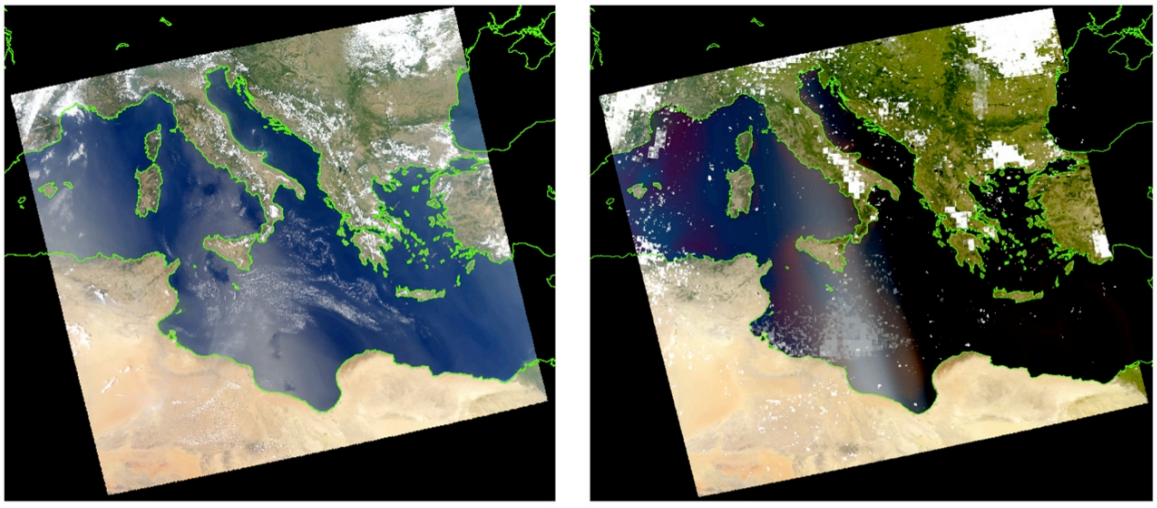


Figure 1. (Left) actual RGB composite image for Aqua MODIS granule for August 15, 2012 at 12:00 UTC and (right) equivalent aerosol-free image simulated from GEOS-5. GEOS-5 does not assimilate cloudy radiances and so there should be little expectation of a granule-level feature match. However, in this case, the model does remarkably well with cloud placement.

The equivalent sensor radiance data for the MODIS geometry is generated using the Discrete Ordinate Radiative Transfer (DISORT) code (Stamnes et al., 1988). We then replace the contents of the MODIS Level-1B radiance files with these simulated radiances and insert the resulting alternate data stream at the

head of the MODAPS operational algorithm processing chain; a typical RGB image generated from these radiances is shown in Figure 1. The data stream is fully transparent to the system so that Level-2 retrievals can be aggregated all the way to Level-3 global 1 x 1 degree gridded products just as regular MODIS-produced data would be. There are many potential uses for the resulting Level-2 and Level-3 data. Level-3 data can be compared to MODIS Level-3 and GEOS-5 source data fields directly as a means of model validation. Level-2 data can be used to study retrieval algorithm behavior and sensitivities since all retrievals are performed with known truth.

This process can also be generalized to sensors other than MODIS. Different sensors or different retrieval algorithms for the same sensor could be compared and analyzed in a controlled environment, leading to improvements in remote sensing algorithms. The output can be used for a wide variety of purposes such as model parameter verification, model parameterization inter-comparisons, remote sensing algorithm validation, theoretical future sensor studies, and validation of "fast simulators," such as in the CFMIP Observation Simulator Package (COSP).

References

Norris P.M., L. Oreopoulos, A.Y. Hou, W.-K. Tao, and X. Zeng, 2008: Representation of 3D heterogeneous cloud fields using copulas: Theory for water clouds. *Q. J. Roy. Meteorol. Soc.*, **134**, 1843–1864. doi:10.1002/qj.321.

Platnick, S., M.D. King, S.A. Ackerman, W.P. Menzel, B.A. Baum, J.C. Riedi, and R.A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Trans. Geosci. Remote Sens.,* **41**, 459–473.

Riishojgaard, L.P., 1998: A direct way of specifying flow- dependent background error correlations for meteorological analysis systems. *Tellus,* **50A**, 42–57.

Stamnes, K., S.C. Tsay, W. Wiscombe, and K. Jayaweera, 1988: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.,* **27**, 2502–2509.

Development, Validation, and Investigations of the GMAO Observing System Simulation Experiment Framework

Ronald M. Errico and Nikki C. Privé

An extensive validation of an early version of the observing system simulation experiment (OSSE) framework developed at the GMAO was reported in a pair of papers (Errico et al., 2013; Privé et al., 2013a). This is the most extensive validation of any OSSE framework reported to date. While there are shortcomings, notably in slightly over-optimistic medium-range forecast scores compared with assimilation of real data, most corresponding statistics compare quite well as illustrated in Figure 1. This OSSE framework therefore appears suitable to investigate many outstanding and otherwise elusive questions in data assimilation.

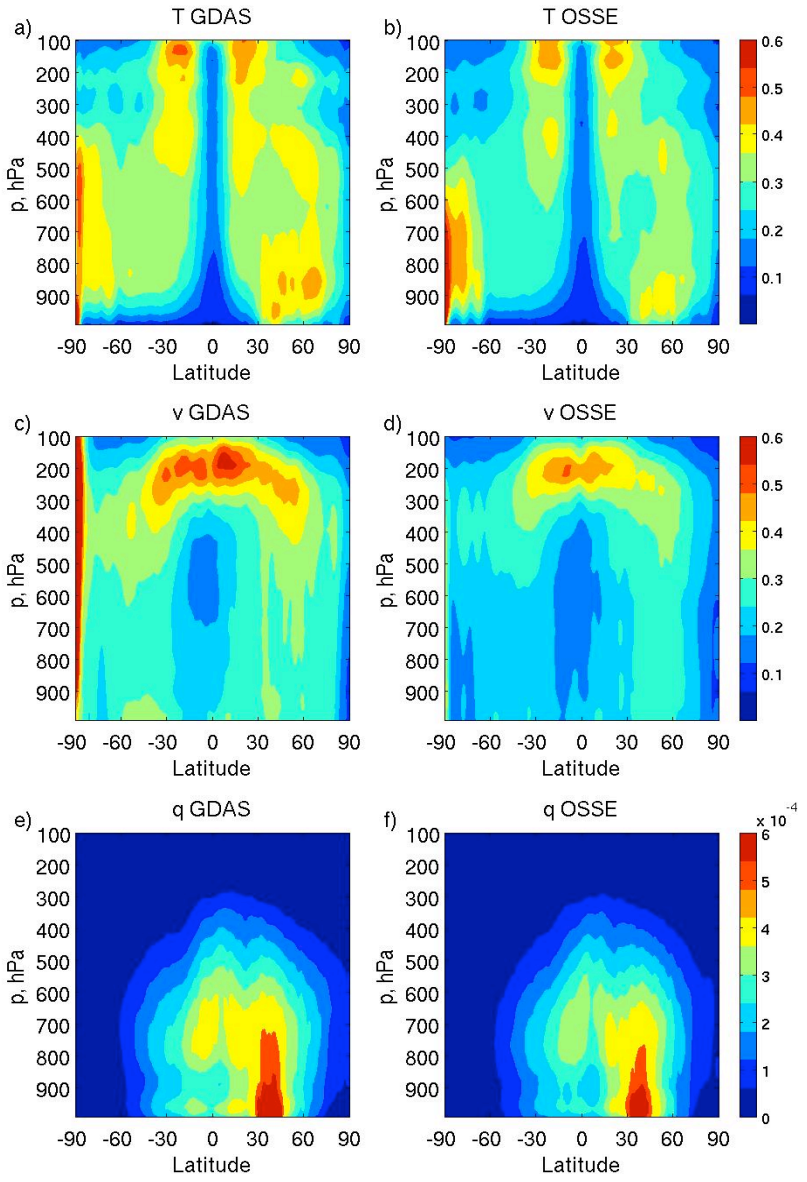


Figure 1. Square roots of zonal means of temporal variances of analysis increments of temperature, meridional wind, and specific humidity for the OSSE (right) compared with real data (GDAS, left). Units are K for T, m/s for v and g/kg for q.

Upgrades and improvements to the GMAO OSSE have now been undertaken, including the addition of IASI, MHS, GPS radio-occultation, and ASCAT data to the types of observations assimilated. Cloud and water vapor tracked wind data locations are now consistent with locations of those features in the nature run. Radiosonde observations are also now computed with wind drift and significant-level reports consistent with the nature run. Continuing upgrades to the OSSE are anticipated in order to keep abreast of changes to the global observing network and data assimilation system.

The GMAO OSSE was applied to several problems. One was determination of characteristics of analysis error (Errico and Privé, 2013). Since an analysis produced by a modern data assimilation system is generally the best description of the global atmosphere at a given time, it is very difficult to validate such an analysis against any better rendition of the state. Realizations of analysis error can be explicitly calculated in the OSSE context by simple differencing with the known true state that is the nature run. Results reveal significant longitudinal as well as latitudinal variation in many of these statistics, with large percentage errors in the divergent wind analysis at almost all horizontal scales.

A preliminary investigation of the fidelity of the "NMC method" for prescribing background error covariances was undertaken within the OSSE context. The NMC method uses scaled statistics generated from 48-hour minus 24-hour forecasts validating at the same time as proxies for statistics of 6-hour forecasts errors that define background errors. The NMC method was originally invented because no more direct method was available to estimate background error statistics globally, but in the OSSE the latter calculation is simple and direct. The preliminary results indicate large differences between results using the two methods regarding estimated correlation distances, revealing an inherent weakness of the NMC method. More conclusive studies will be performed after the GMAO OSSE framework is updated to the current GEOS-5 operational atmospheric data assimilation system.

A series of experiments was also performed in order to examine the relative impacts of observation error, model error, and initial condition error on the evolution of forecast error and on the effectiveness of the data assimilation system (Privé et al., 2013b). The magnitude of observation errors was varied from low to high, and the global observing network was tested both in a standard realistic configuration and in an idealized global network of high-quality sounding observations. An 'identical twin' case employing a perfect model scenario was also developed to explore the case with no model error. Model error appears to play a larger role in degrading the forecast skill compared to initial condition error in the GMAO OSSE. Further experiments are underway to determine if the very rapid initial forecast error growth is primarily due to systematic model error or due to numerical artifacts of the forward integration process.

Publications

Errico, R.M., R. Yang, N. Privé, K.-S. Tai, R. Todling, M. Sienkiewicz, and J. Guo. 2013: Validation of version one of the Observing System Simulation Experiments at the Global Modeling and Assimilation Office. *Quart. J. Roy. Meteor. Soc.,* doi: 10.1002/qj.2027 (*in press*).

Errico, R.M. and N.C. Privé. 2013: An estimate of some analysis error statistics using the GMAO observing system simulation framework. *Quart. J. Roy. Meteor. Soc.* (*in press*).

Privé N.C., R.M. Errico, and K.-S. Tai. 2013a: Validation and forecast skill of the Global Modeling and Assimilation Office observing system simulation experiment. *Quart. J. Roy. Meteorol. Soc.,* doi: 10.1002/qj.2029 (*in press*).

Privé, N.C., R.M. Errico, and K.-S. Tai, 2013b: The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *J. Geophys. Res. - Atmos.* doi: 10.1002/jgrd.50452 (*in press*).

Aerosols in GEOS-5: Simulations of the UV Aerosol Index and Comparisons with OMI Retrievals

Virginie Buchard-Marchant and Arlindo da Silva

The UV Aerosol Index (AI) at 354-380 nm is useful for detecting the presence of absorbing aerosols in the atmosphere. We have developed a radiative transfer interface to simulate the AI from GEOS-5 aerosol assimilated fields. Each 3D GEOS-5 aerosol concentration profile is interpolated to the Ozone Monitoring Instrument (OMI) observation location, and a Mie-type calculation is performed to produce profiles of aerosol extinction, single scattering albedo, and the components of the scattering phase matrix. These parameters are then input to the radiative transfer code VLIDORT (Vector LInearized Discrete Ordinate Radiative Transfer) to produce top-of-atmosphere radiances at OMI wavelengths and calculating the AI at 354 nm. This calculation considers the appropriate viewing geometry, with surface albedo taken from the OMI files. We have compared model produced AI with the corresponding OMI measurements, identifying regions where the model representation of absorbing aerosols were deficient. An example of the comparison between the AI simulated by GEOS-5 and retrieved by OMI for July 2007 appears in Figure 1. The GEOS-5 simulated AI captures major features of the global absorbing aerosol observations, although the model tends to underestimate the AI over most of North America and in the biomass-burning region in Southwest Africa.

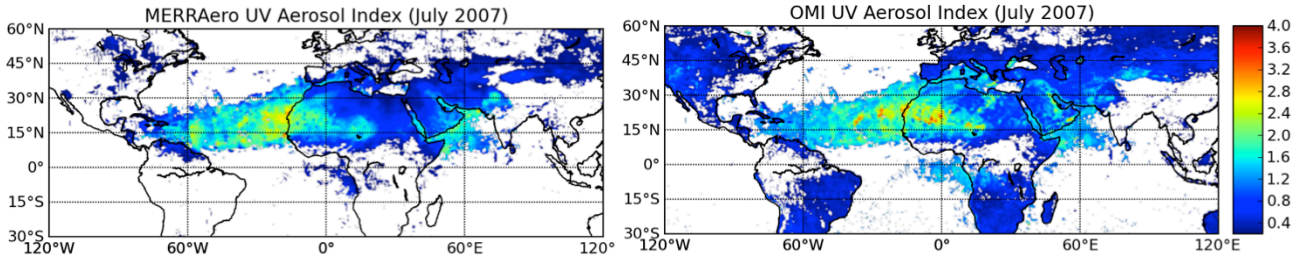


Figure 1. Comparison of the monthly mean simulated MERRA Aerosol Reanalysis AI to OMI observations.

Because the AI is dependent on both aerosol concentration and altitude of the aerosol layer, it is important to use complementary observations to fully diagnose the model. Making use of CALIPSO measurements we can investigate the impact of the altitude of the aerosol layer on OMI derived AI, allowing us to ascertain the misplacement of plume height by the model. Figure 2 shows the comparison of MERRAero attenuated backscatter profiles to CALIOP measurements for a case of South African biomass burning. For this comparison, we have sampled the MERRAero along the CALIPSO orbit track for a three-month period, June, July and August 2011. It is apparent that the simulated attenuated backscatter signal by GEOS-5 is a little bit higher in magnitude than the CALIOP estimates. Additionally we notice that the GEOS-5 biomass-burning plume does not remain elevated in the simulation as it does in the observations.

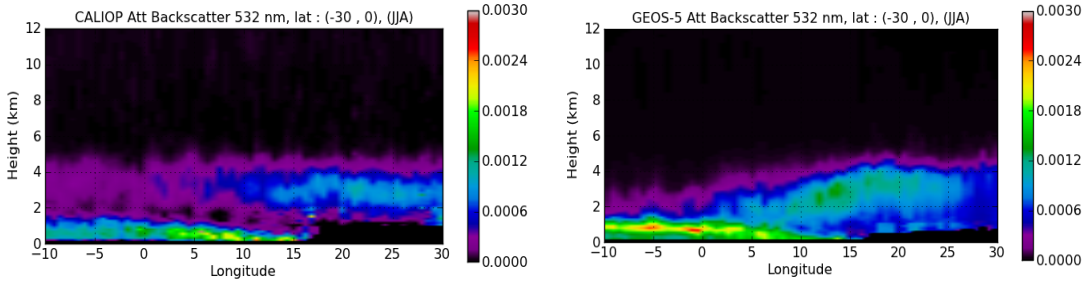


Figure 2. Left: Three-month mean (June, July, and August 2011) of CALIOP attenuated backscatter coefficient for a case of South African biomass burning. Right: MERRAero attenuated backscatter coefficient sampled on the CALIPSO track for the same period.

Characterization of Representativeness Errors Associated with Satellite CO_2 Observations Using High Resolution Global Model Simulations

Lesley Ott, Steven Pawson, Bill Putman, Watson Gregg, Jim Collatz (Code 618), Arlindo da Silva, and Anton Darmenov

Column CO_2 observations from current and future remote sensing missions promise a major advancement in our understanding of the carbon cycle and are expected to help constrain source and sink distributions. However, the use of these observations in combination with data assimilation and inversion methods is challenged by the difference in scale of models and observations. For example, OCO-2 footprints represent an area of several square kilometers while NASA's future ASCENDS lidar mission is likely to have an even smaller footprint. In contrast, the resolution of models used in global inversions are typically hundreds of kilometers wide and often cover areas that include combinations of land, ocean and coastal areas and areas of significant topographic, land cover, and population density variations. Mismatches in the resolution of models and observations can lead to "representativeness errors" that have the potential to bias source and sink estimates calculated by inversion studies. To date, understanding of such errors and their implications is poor.

In this study, representativeness errors are estimated using 12-km global CO_2 distributions calculated by the GEOS-5 atmospheric general circulation model. This resolution, typical of mesoscale atmospheric models, represents an order of magnitude increase in resolution over typical global simulations of atmospheric composition allowing new insight into small-scale CO_2 variations across a wide range of surface flux and meteorological conditions. These simulations include estimates of fossil fuel CO_2 emissions from the Emission Database for Global Atmospheric Research (EDGAR), provided at approximately 10 km resolution globally; estimates of fire emissions from NASA's Quick Fire Emission Dataset (QFED) computed using MODIS fire products are also provided at 10 resolution. Biosphere flux datasets provided by NASA's Carnegie-Ames-Stanford-Approach – Global Fire Emissions Database version 3 (CASA-GFED3) at 1/2° resolution have been downscaled to 12 km using MODIS EVI data as an indicator of small scale net ecosystem exchange variability. Ocean-atmosphere CO_2 exchange is calculated within GEOS-5 simulations using ocean pCO_2 distributions calculated by the NASA Ocean Biogeochemical Model and GEOS-5 atmospheric CO_2 and wind fields. Combining realistic, high-resolution flux estimates informed by a suite of remote sensing products and the high resolution atmospheric modeling capabilities of GEOS-5 allows insight into the magnitude of representativeness errors.

Column CO_2 distributions were saved hourly and means were calculated over areas typical of global model resolution (1°, 2°, 4°, and 10°). To estimate potential representativeness errors, differences between the 12-km grid cell column CO_2 mixing ratio and the coarse area mean were calculated every hour for each coarse model grid cell. Figure 1 shows standard deviations of these distributions for different coarse grid cell sizes calculated during July 2005. During the northern hemisphere summer months, column CO_2 distributions are characterized by low mixing ratios in northern hemisphere mid- and high latitudes because of drawdown by the terrestrial biosphere. In these regions, we see the largest potential representativeness errors in all coarse resolution cases examined. Not surprisingly, the magnitude of potential errors depends strongly on the disparity between the native resolution (12 km) and the resolution assumed for the coarse grid cell. For a 1° global model, the standard deviation of differences due to resolution disparity are typically less than 0.5 ppmv over northern hemisphere growing regions with smaller variations over oceans. At 10°, standard deviations are significantly larger over continental growing regions (typically greater than 1 ppmv) and greater than 0.5 ppmv over oceans. During January (not shown), standard deviations at all coarse grid cell resolutions are less than their July values in the northern hemisphere while standard deviations in the southern hemisphere are increased during the growing season.

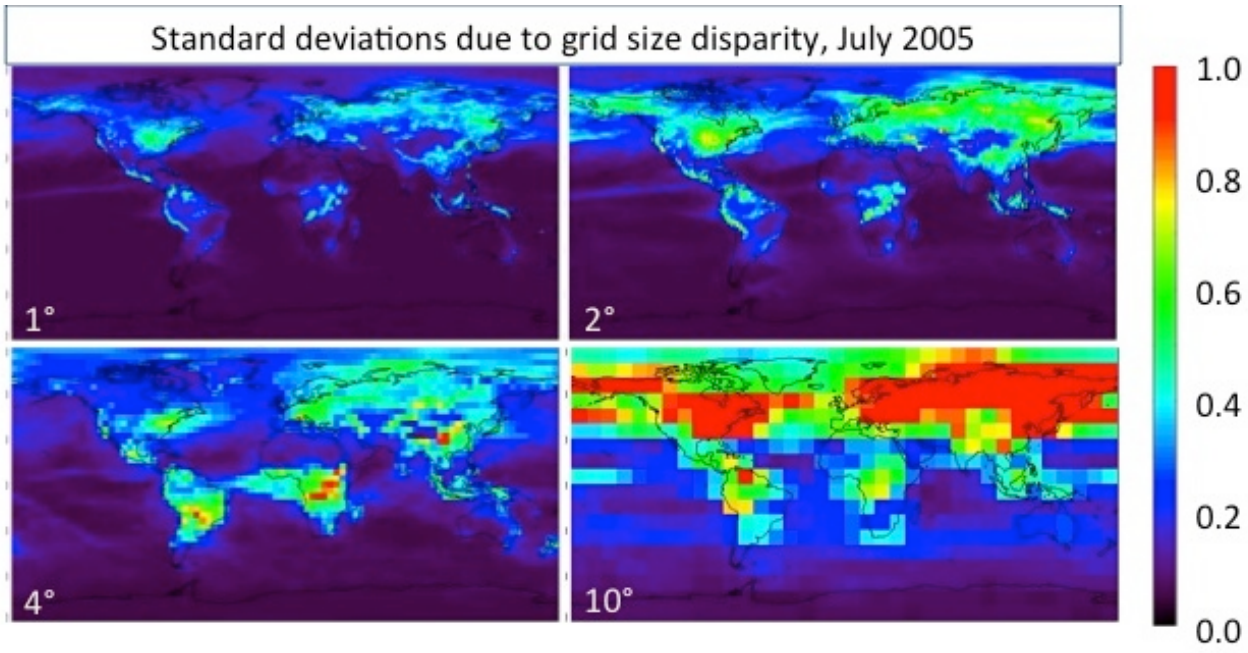


Figure 1. Standard deviations of errors due to spatial aggregation of 12-km simulated column CO_2 mixing ratio. Differences between native resolution fields and mean CO_2 aggregated over an area representative of a typical global model are calculated hourly and standard deviations of these differences calculated over the course of the one month simulation for July 2005. Resulting errors for 1° (upper left), 2° (upper right), 4° (lower left), and 10° (lower right) resolutions are shown.

It is important to note that these estimates represent a simplified view of satellite representativeness error due only to the mismatch in resolution between an observation and a global model typical of assimilation and inversion studies. Further, complicating factors include the ability of satellites to sample only under clear sky conditions, which can introduce additional biases when satellite data are used in conjunction with coarse resolution models, and errors in satellite observations that may be related to surface reflectivity and the presence of atmospheric aerosols. Still, these results provide some of the only global estimates to date of potential representativeness errors and their variations as a function of coarse model resolution and season. Future work will include the generation of realistic pseudo-datasets using longer term high resolution global CO_2 simulations and process-based instrument error estimates to further improve the characterization of representativeness errors. These error estimates will be used in the GEOS-5 carbon data assimilation and in inversion studies to maximize the impact of satellite CO_2 observations and to better constrain global source and sink distributions.

Implementation of the Modal Aerosol Model (MAM) in GEOS-5

Anton Darmenov and Arlindo da Silva

The aerosol module currently used in GEOS-5 is the online version of the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model that belongs to the family of sectional aerosol models that trace the mass of aerosol particles within non-overlapping size ranges (bins). However, the discretization of the size distribution imposes limits on the accuracy of this approach when using a small number of bins. Moreover, deriving the number of aerosol particles, which is important for modeling of cloud condensation nuclei, ice nuclei and aerosol optical parameters, requires prescribing sub-bin size distributions. Another class of aerosol models uses smooth analytical functions to represent the aerosol size distribution. A common approach is to use several log-normal modes with fixed widths, but with geometric mean size, and hence aerosol size distribution, that vary in space and time. For the computation of the size distribution parameters, the total number of particles and mass of aerosol species are required. An example of a modal scheme is the 7-mode Modal Aerosol Model (MAM7) (Liu et al., 2012). The MAM7 scheme provides an extensive treatment of primary and secondary aerosol components while still offering competitive performance when compared to sectional schemes such as GOCART. This served as a motivation for incorporating the MAM7 scheme into GEOS-5.

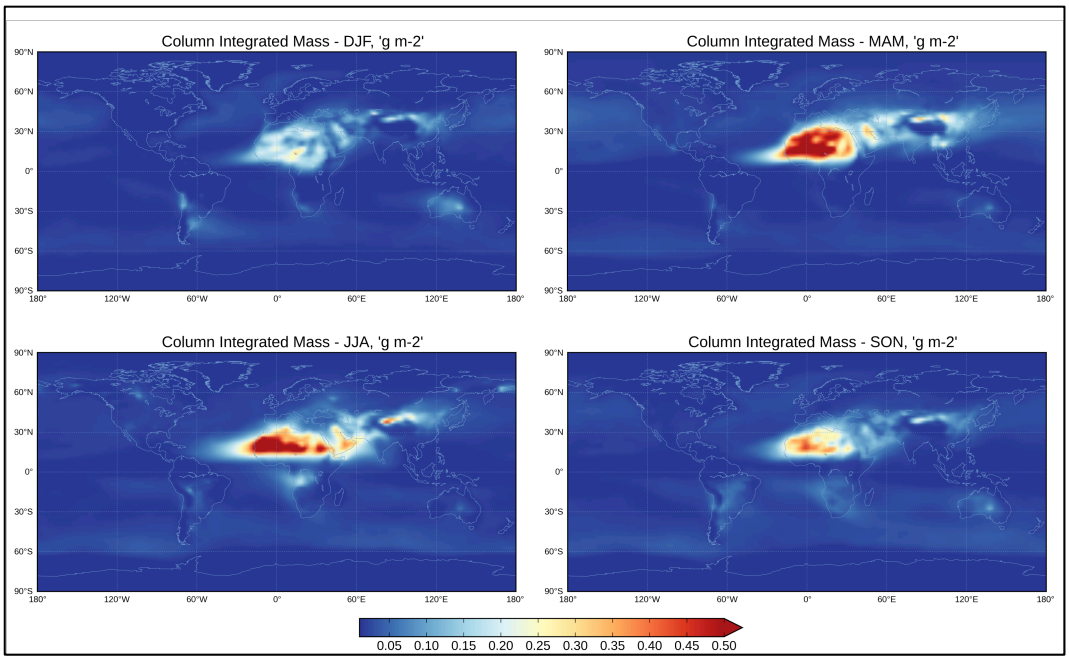


Figure 1. Seasonal variation of column-integrated aerosol mass from a year-long intermittent replay run with MAM integrated into GEOS-5 on a cubed-sphere grid at approximately 2 ×2° horizontal resolution.

Two major development goals were achieved: wrapping the aerosol microphysics core from the Community Atmospheric Model (CAM5) MAM7 scheme into an ESMF gridded component and creating an atmospheric chemistry component tailored to the needs of the new aerosol module. The new aerosol chemistry component utilizes the Kinetic PreProcessor (KPP; Sandu and Sander, 2006) to perform the numerical integration of the sulfur and ammonia gas-phase chemical reaction mechanism. The use of an automatic code generator such as KPP allows extension of the chemical mechanism. The next stage will focus on coupling MAM to the radiation and cloud microphysics. Extensive testing, tuning, and validation of the system will be performed to ensure stable and well-understood modeling system.

References

Liu, X., and coauthors, 2012: Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5. *Geosci. Model Dev.*, **5**, 709-739.

Sandu, A., and R. Sander, 2006: Technical note: Simulating chemical systems in Fortran90 and Matlab with the Kinetic PreProcessor KPP-2.1. *Atmos. Chem. Phys.*, **6**, 187-195.

Continued Development of Cloud Microphysics in GEOS-5

Donifan Barahona, Andrea Molod, and Max Suarez (GMAO)
Julio Bacmeister, Andrew Gettelman, and Hugh Morrison (NCAR)
Athanasios Nenes (Georgia Institute of Technology)

The description of the formation cloud and precipitation within GEOS-5 has been recently updated to account for the effect of aerosol emissions on cloud properties. The new two-moment microphysics scheme predicts particle number concentration and mass mixing ratio for ice and liquid cloud species. The initial implementation of the microphysics showed a significant improvement in the representation of cloud fields over the operational version of GEOS-5. However it depended on several unconstrained empirical parameters regarding the subgrid scale dynamics, the role of convective detrainment in the development of clouds, and the presence of supercooled water. Several new parameterizations were developed to address these issues. First, a new representation of subgrid scale motion was formulated and implemented. This parameterization links large-scale dynamical fields to the distribution of vertical velocities of individual cloudy parcels. It accounts for both turbulence and gravity wave motion and therefore is applicable to different cloud regimes, including cumulus convection and cirrus. Two-moment microphysics for convective clouds was also formulated. In the initial implementation the number concentration of particles detrained from convective sources was prescribed following empirical correlations without a link to aerosol sources. This has been updated and the current microphysics implementation accounts for the effects of aerosol emissions on convective development. Finally, the description of the partitioning between liquid and ice in mixed-phase clouds was also improved. Explicit parameterizations for immersion and contact freezing were formulated and implemented. The fraction of supercooled liquid existing in mixed-phase clouds, initially assumed as a linear function of temperature, is now linked to aerosol sources and dynamical conditions. Ongoing testing of the improved cloud microphysics shows more realistic cloud fields when compared to observations.

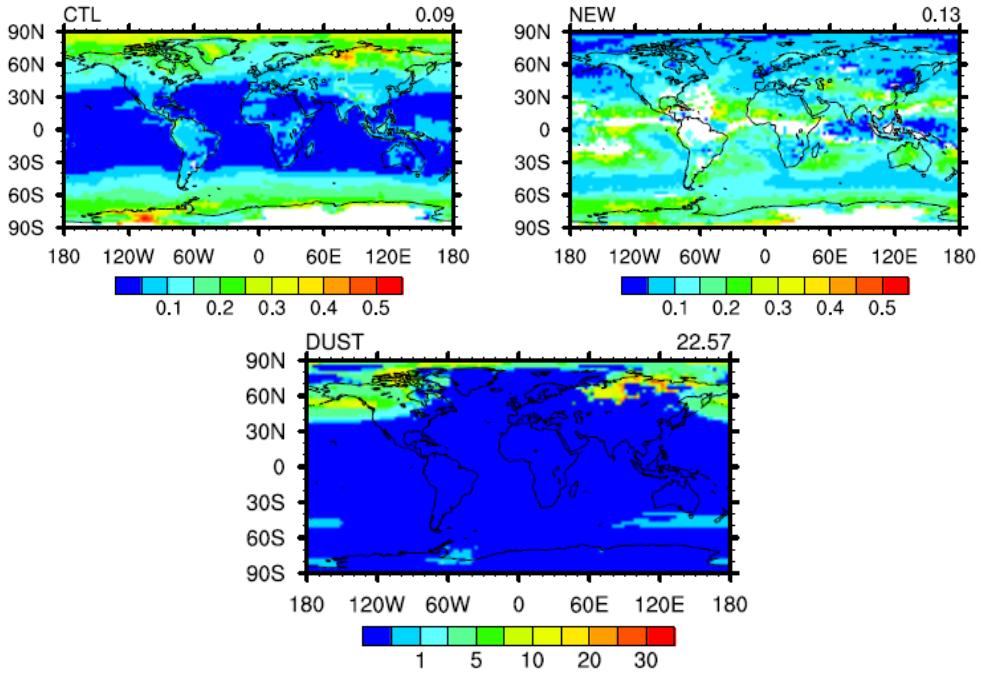


Figure 1. Annual mean fraction of cloud condensate existing as liquid at -20 °C from five-year simulations of the operational GEOS-5 (CTL, upper left) and using the new microphysics (NEW, upper right). Also shown (lower) is the annual mean dust concentration (10^3 cm^{-3}). Due to the freezing properties of dust, regions of high dust concentration coincide with regions low supercooled water content. This behavior is better captured with the new microphysics.

Representation of Polar Ice Sheets in GEOS-5

Richard Cullather, Bin Zhao, and Max Suarez

The ability of models to predict future sea level change is limited by crude representations of polar ice sheets and their complex dynamical ice flow characteristics. These limitations were highlighted in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC/AR4). To address the problem, a collaborative project with the GSFC Cryospheric Sciences Lab seeks to couple a dynamical ice sheet model (ISM) with the GEOS-5 General Circulation Model (GCM). The prospective ISM is the Ice Sheet System Model of NASA JPL and the University of California at Irvine, which utilizes an adaptive mesh to adequately simulate small-scale ice stream drainage features.

Two input fields of critical importance to an ISM are surface temperature and surface mass balance (SMB). Most current GCMs use a simplistic surface representation over glaciated land surfaces. The existing GEOS-5 parameterization used a fixed subsurface temperature, and snow hydrology was not represented. This produces an erroneous annual mean net cooling of the atmosphere by the ice sheets. Additionally, the parameterization does not allow for melt runoff, and therefore no surface mass balance field is produced. Competing surface redesigns have been tested with the objective of having a unified snow model in GEOS-5. These configurations differ in the representation of the snow-ice interface. In the first configuration, snow depth is a fixed value over glaciers, and snow is allowed to compact to the approximate density of ice. In the second configuration, snow and ice are separately represented, and snow is allowed to accumulate to a maximum depth or melt to a bare ice surface (Fig. 1).

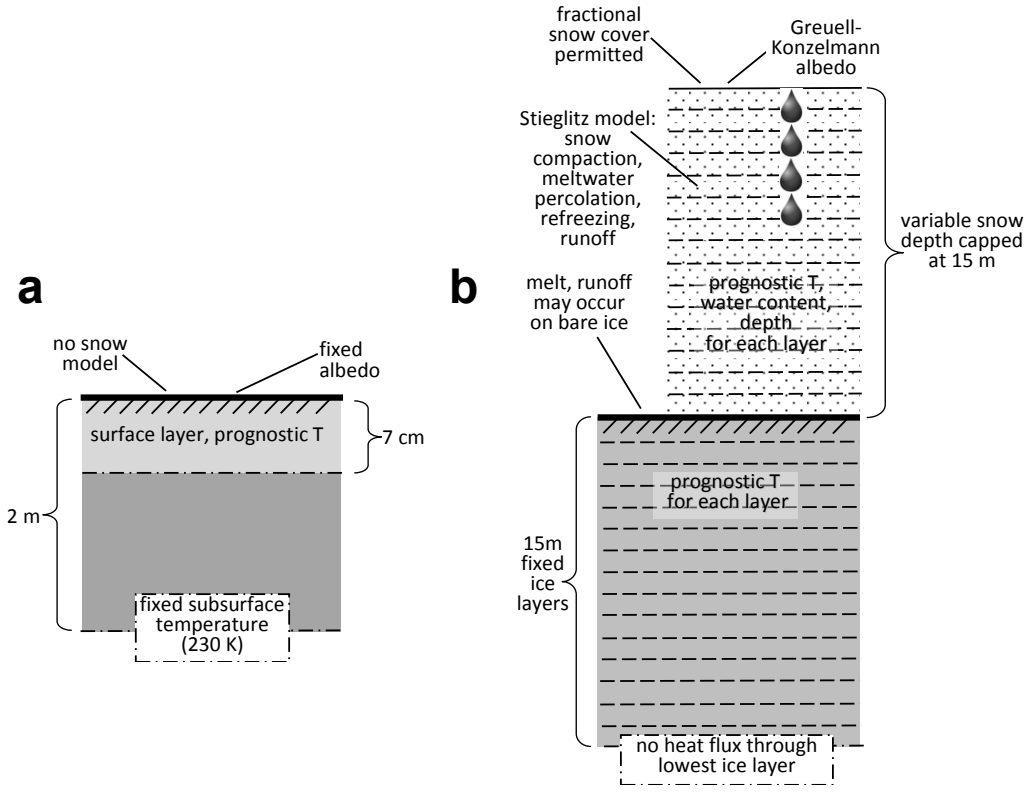


Figure 1. Schematic indicating the representation glaciated ice surfaces in the GEOS-5 model: (a) in original configuration as used in MERRA, and (b) the applied second test candidate, with application of snow model and ice representation.

These configurations have been evaluated over the Greenland Ice Sheet (GIS) against the output of an accepted high-resolution regional climate model and available in-situ observations using an ensemble of 30-yr AMIP simulations. The principal characteristics in the regional climate model SMB include large values in the southeast and along the higher elevations of the west coast, and negative values along

peripheral margins, particularly in the west and north, associated with surface melt in summer. At 2° resolution, GEOS-5 reproduces these basic features. Both configurations indicate the ablation zone along the western periphery. Differences are apparent in southern Greenland, where both configurations indicate negative SMB while the regional climate model is positive, and in northeast Greenland, where the second configuration is clearly superior. A comparison of the second configuration at 2° and ½° resolutions with the regional climate model (Fig. 2) indicate comparable performance.

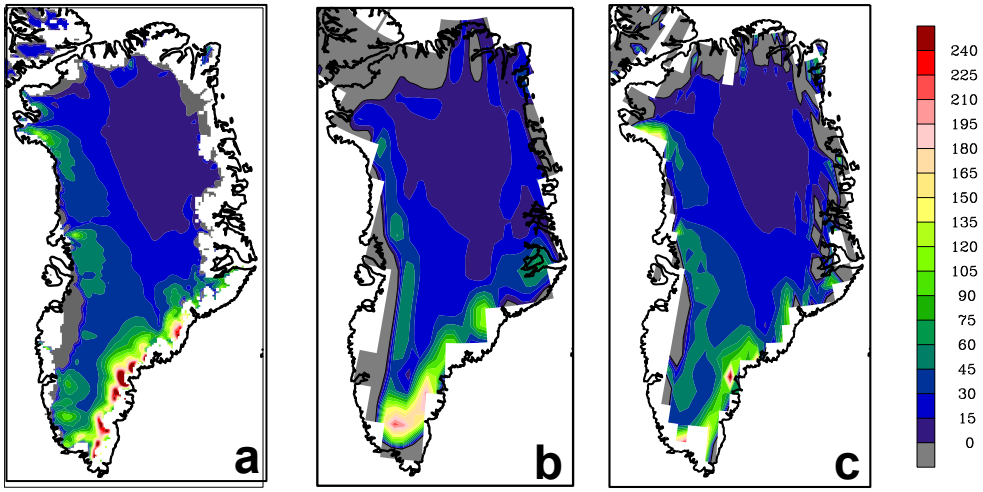


Figure 2. Comparison of GIS surface mass balance (in cm yr⁻¹) from (a) RACMO2 regional climate model, (b) GEOS-5 2° simulation, and (c) GEOS-5 ½° simulation.

The elimination of the annual mean heat flux results in improved surface temperatures (not shown). This is particularly true along the ice sheet margins which are 3-4°C warmer than the control simulation. These local heat flux differences are found to impact general circulation in the GEOS-5 model.

Catastrophe Concept-based Cumulus Parameterization: Correction of Systematic Errors in the Diurnal Cycle of Precipitation over Land in GEOS-5

Winston C. Chao

The onset of cumulus convection in a grid column is a catastrophe, also known as a subcritical instability. Accordingly, in designing a cumulus parameterization scheme the onset of cumulus convection requires that a parameter crosses a critical value and the termination requires that the same (or perhaps different) parameter crosses a different critical value. Once started, cumulus convection stays active, regardless of whether the onset criterion is still met or not, until the termination criterion is met. In addition, the intensity of cumulus precipitation is related to how far the state is from the termination, not the onset, criterion. In contrast, the cumulus parameterization schemes currently in use treat the onset of cumulus convection as a supercritical instability. Namely, convection is active when a parameter exceeds a critical value and is inactive when the same parameter falls below the same critical value. In addition, the intensity of cumulus precipitation is related to how much this critical value has been exceeded.

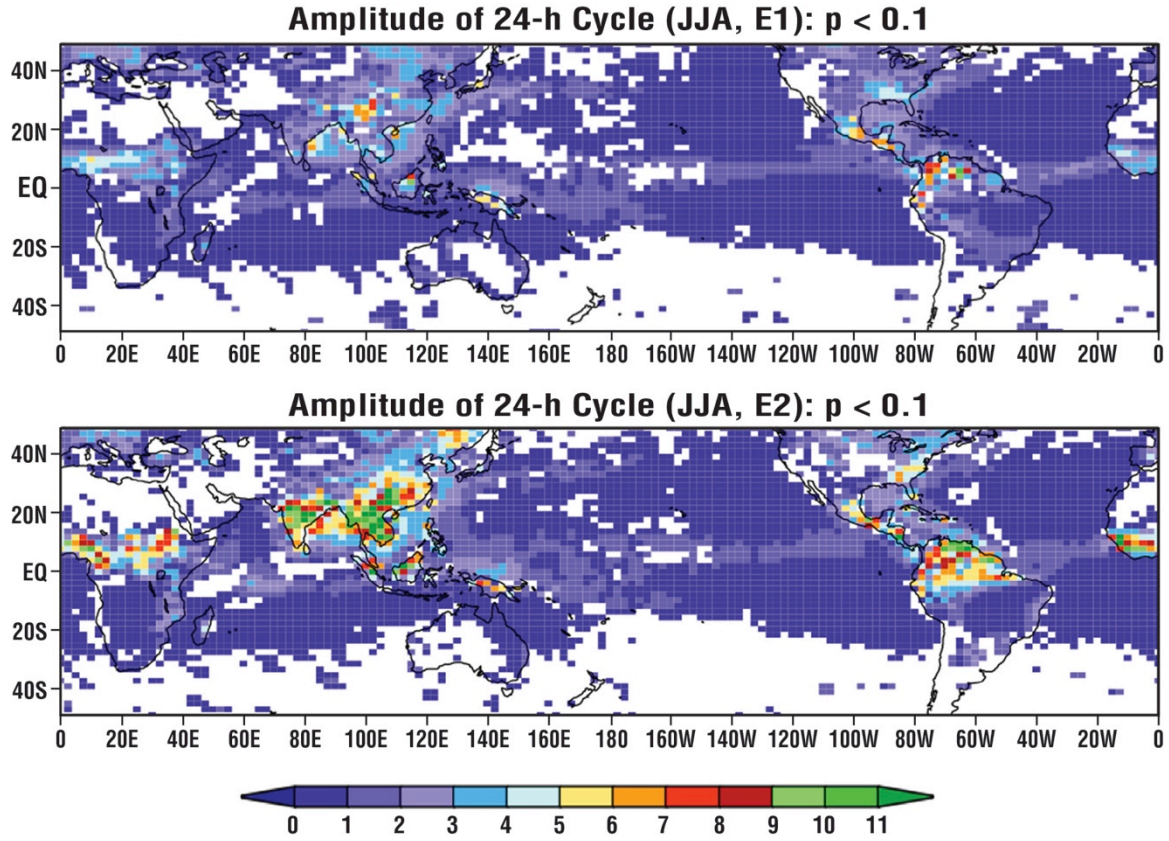


Figure 1. The amplitude of the precipitation diurnal cycle (mm/day) in one Jun-Jul-Aug season before (upper panel) and after (lower panel) the introduction of the catastrophe concept in the cumulus parameterization.

Among the adverse consequences of the supercritical-instability-concept-based cumulus parameterization schemes are that over relatively flat land the precipitation peak occurs around noon – 4~6 hours too soon – and that the amplitude of the diurnal cycle of precipitation is too weak.

Based on the above concept, a catastrophe concept-based cumulus parameterization (C-CUP) scheme has been designed by taking advantage of the existing infrastructure in the relaxed Arakawa-Schubert scheme

(RAS), but replacing RAS's guiding principle with the catastrophe concept. Test results using GEOS-5 (Figs. 1 and 2) show dramatic improvement in the amplitude and phase of the diurnal cycle of precipitation over relatively flat terrain.

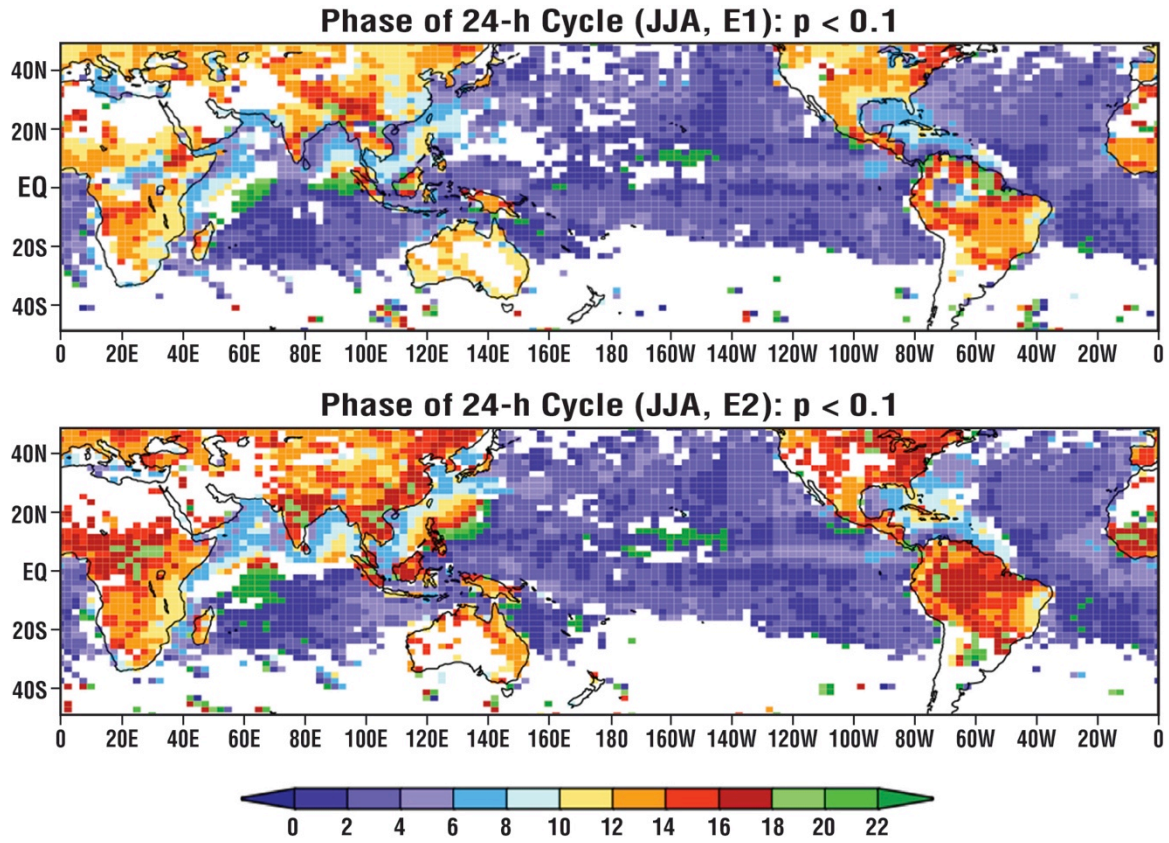


Figure 2. The phase of the precipitation diurnal cycle (the local solar time of the peak of the precipitation diurnal cycle) in one Jun-Jul-Aug season before (upper panel) and after (lower panel) the introduction of the catastrophe concept in the cumulus parameterization.

Publication

Chao, W., 2013: Catastrophe concept-based cumulus parameterization: Correction of systematic errors in the precipitation diurnal cycle over land in a GCM. *J. Atmos. Sci.* (*submitted*).

PBL Depth Diagnostics in GEOS-5

Erica McGrath-Spangler and Andrea Molod

The planetary boundary layer (PBL) is crucial to surface-atmosphere exchanges of energy, moisture, aerosols, carbon, and other chemical tracers. The ability to accurately model the PBL character and height in global atmospheric general circulation models (AGCMs) has significant implications for climate and weather predictions, and is made even more difficult due to the lack of adequate observations. In addition, there are several PBL depth definitions and these can give different results (Seidel et al., 2010) necessitating a comparison between observations and models using the same definition. The depth of the PBL is of particular importance for air quality studies since pollutants, aerosols, and carbon dioxide are mixed by turbulent motions throughout the layer resulting in lower concentrations for deep PBLs and higher concentrations for shallow PBLs for a given surface flux. Carbon inversion studies, which seek to estimate surface fluxes given surface concentrations, are particularly sensitive to PBL depth errors in atmospheric models (Gurney et al., 2002).

This study incorporates 6 different PBL depth definitions into the GEOS-5 AGCM based on the vertical profiles of the scalar diffusivity, the bulk Richardson (Ri) number, and the horizontal turbulent kinetic energy (TKE). A single climate simulation at approximately 0.5° horizontal resolution is analyzed to compare PBL heights based on the different definitions. The AGCM PBL heights are aggregated onto the Köppen-Geiger climate classes (Peel et al., 2007) in order to provide insight into implications on the regional and global climate scale.

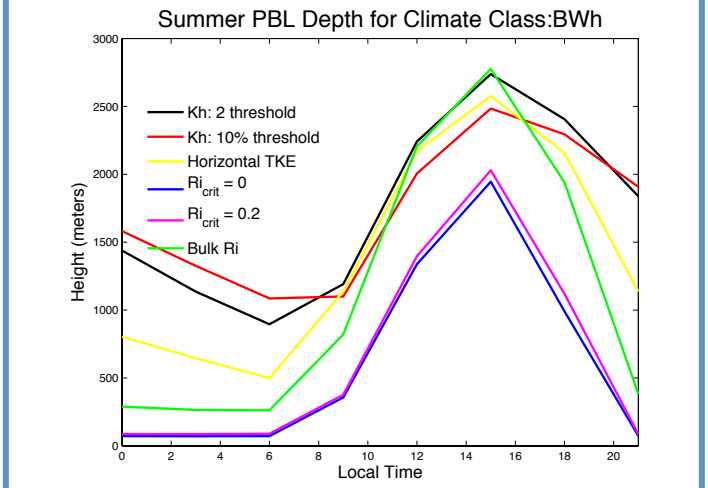


Figure 1. Seasonal mean diurnal cycle of PBL depth in summer conditions over hot, arid deserts.

In general, the definitions based on the scalar diffusivity (Kh) using a threshold of 2 $m^2 s^{-1}$ (the default in GEOS-5), Kh using a threshold of 10% of the column maximum, the horizontal TKE, and bulk Richardson number produce similar midday values while definitions based on the local Richardson number definitions (with Ri_{crit} equal to 0 or 0.2) are considerably shallower. The bulk Richardson number and local Richardson number definitions collapse properly at night while the other definitions remain high. An example of this is shown in Figure 1, based on PBL heights during summer desert conditions. This region is especially important because of implications of PBL depth and turbulent mixing on the transport of desert dust.

In the temperate and cold extratropics, the definition using the horizontal TKE produces deeper midday PBL depths during winter (not shown) than the other definitions. This difference is likely attributed to the fact that the horizontal TKE is sensitive to the wind profile and is thus sensitive to the passing of mid-latitude cyclones and their associated fronts. A deeper PBL in the wintertime is most likely related to enhanced storm activity in this season producing enhanced wind shear and increasing the vertical extent of turbulent mixing and consequently the depth of the PBL.

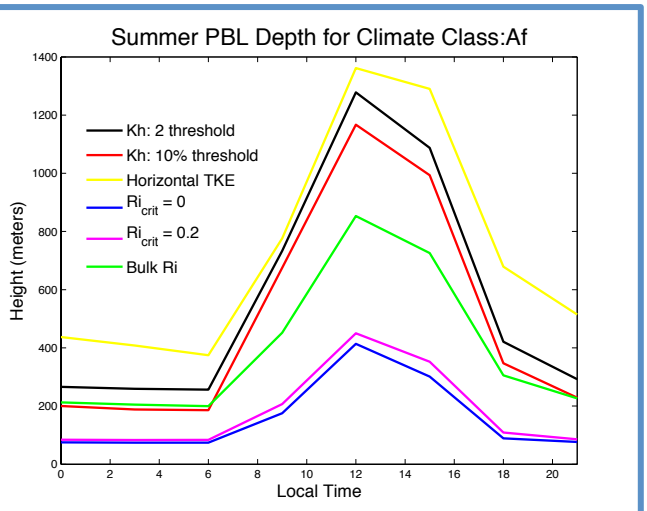


Figure 2. Seasonal mean diurnal cycle of PBL depth in summer conditions over the tropical rainforest.

Under warm, moist conditions, such as is found in the tropical rainforest, the midday bulk Richardson number definition is considerably shallower than the scalar diffusivity definitions (Fig. 2). A similar result is found, for instance in the equatorial Pacific near the dateline, but not at other tropical locations characterized by stratocumulus and stratus conditions. It is currently presumed that this is related to the increased sensitivity of the Lock (2000) turbulence scheme (which is in current use in the GEOS-5 AGCM) to buoyant sources of turbulence, which would result in a scalar diffusivity profile with a deep well-mixed layer. The bulk Richardson number, which is sensitive to both buoyant and shear sources of turbulence, produces a shallower PBL in regions of high buoyancy. This could have implications for cumulus convection and its associated fluxes.

References

Gurney, K.R., et al., 2002: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models. *Nature*, **415** (6872), 626-630, doi:10.1038/415626a.

Lock, A.P., A.R. Brown, M.R. Bush, G.M. Martin, and R.N.B. Smith, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. *Mon. Weather Rev.*, **128**, 3187-3199, doi:10.1175/1520-0493(2000)128<3187:anblms>2.0.co;2.

Peel, M.C., B.L. Finlayson, and T.A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sc.*, **11**(5), 1633-1644.

Seidel, D.J., C.O. Ao, and K. Li, 2010: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis. *J. Geophys. Res.-Atmos.*, **115**, doi:10.1029/2009jd013680.

Hydroclimatic Controls on Vegetation Phenology

Randal Koster, Gregory Walker, G. James Collatz (Code 618)
Peter Thornton (DOE/ORNL)

The character of vegetation varies significantly across the globe, with some areas (e.g., forests) featuring lush vegetation, with many leaves per unit area (i.e., a high leaf area index, or LAI), and other regions (e.g., shrublands) featuring much sparser vegetation. Phenological variables such as LAI are useful for describing the variation of vegetation, not only in space but also across the seasons at a given location – the variables can characterize the timing of bud-burst, subsequent canopy growth, and autumnal senescence. Because the dynamics of vegetation phenology underlie much of the dynamics of the land surface component of the global carbon cycle, and because the GMAO has a strong interest in addressing carbon cycle questions, the accurate modeling of phenological variations is a key objective of the land modeling group. In this study, we examine the hydroclimatic controls on observed vegetation phenology and evaluate whether our dynamic vegetation model adequately captures these controls.

The vegetation observations used for the study are derived from the latest version of the Global Inventory and Mapping Studies, or GIMMS, dataset (Tucker et al., 2005). We focus in particular on the FPAR ("fraction of absorbed photosynthetically active radiation") product. This variable ranges from 0 to 1, with higher values indicating that the vegetation is able to utilize the incoming radiation more completely; higher values thus imply a lusher and greener vegetation state. The FPAR data used below span the years 1997-2009 (though will soon be expanded to cover 1982-present); it has been aggregated to monthly values at 1/4°×1/4° resolution. The precipitation data we used consist of monthly precipitation totals for 1979-present at 2.5°×2.5° degree resolution, as produced by the Global Precipitation Climatology Project as part of their Version 2 Satellite-Gauge dataset (Adler et al. 2003; see also ftp://precip.gsfc.nasa.gov/pub/gpcp-v2/doc/V2_doc).

Figure 1a shows how the observed FPAR at each land point (for the peak month in the local mean seasonal cycle of FPAR) varies with the mean precipitation at that point (x-axis) and the standard deviation of annual precipitation, σ_P, at the point (y-axis). For clarity, Figure 1b presents a binned version of this figure. Two features are clearly seen in the observations: (i) as expected, vegetation "lushness", as characterized by FPAR, tends to increase with increasing mean precipitation; and (ii) interestingly, vegetation lushness tends to *decrease* with increasing interannual variability of precipitation. Apparently, vegetation requires some stability in annual precipitation (low σ_P) to flourish.

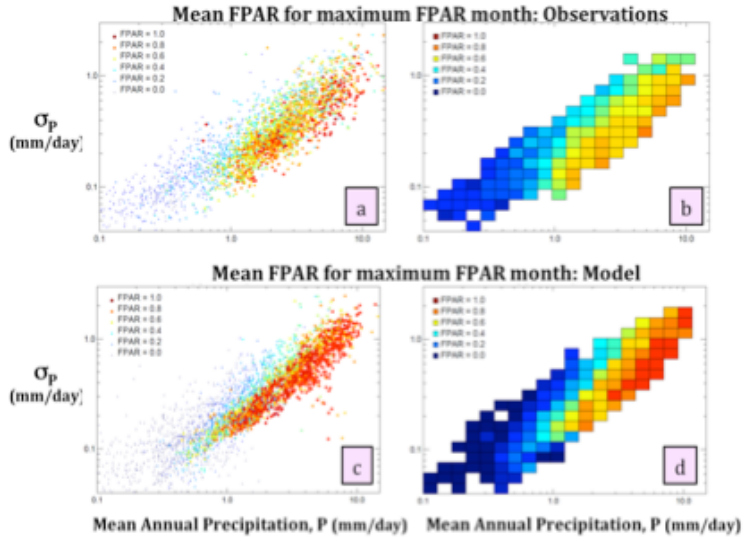


Figure 1. (a) Average GIMMS FPAR in peak FPAR month as a function of the mean precipitation (x-coordinate) and the standard deviation of annual precipitation (y-coordinate). Each dot represents a single land grid cell. (b) Same as (a), but with the individual values in the scatter plot averaged over bins. At least 5 dots must lie within a bin for the binned value to be plotted. (c) and (d): Same as (a) and (b) , but for model-simulated FPAR data.

Figures 1c and 1d show the corresponding plots from a multi-decadal simulation with the GMAO vegetation model (see the GMAO 2011-2012 annual report). The comparison of the observations and model results shows that the model does overestimate FPAR in wetter climate regimes. However, the

model captures, reasonably accurately, the sensitivities of long-term average vegetation lushness to variations in precipitation means and variability. These are indeed the sensitivities we need to capture if we are to use GMAO systems to tie together the global energy, water, and carbon cycles.

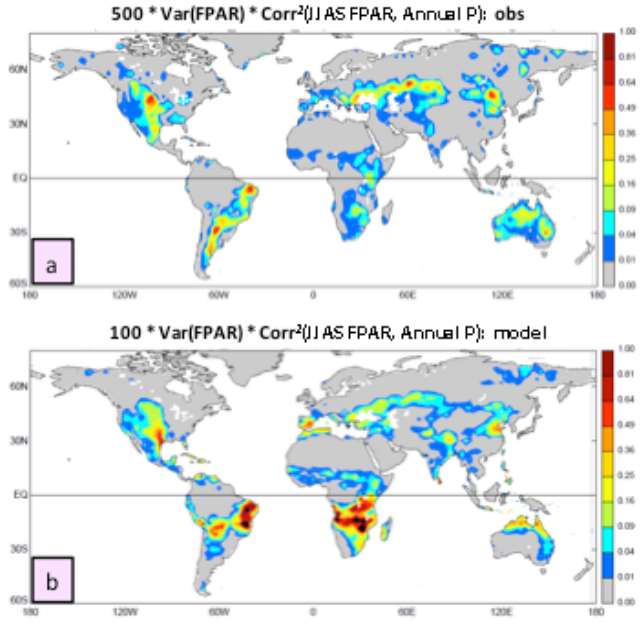


Figure 2. (a) Product of the interannual variance of GIMMS FPAR data averaged over the maximum FPAR season and the square of the correlation between FPAR and annual precipitation, multiplied by 1000. (b) Same, but for model-generated FPAR data, and with the scaling factor changed to 500.

Whereas Figure 1 focuses on long-term phenological means, Figure 2 focuses on interannual phenological variability – in particular, the product of the interannual variance of June-September FPAR and the square of the correlation between this FPAR and annual precipitation. This quantity is, in effect, the portion of the interannual variance of FPAR directly associated with ("explained by") interannual variations in water supply. Aside from some large biases in the Southern Hemisphere (for which the period chosen no longer represents peak foliage season), the model and observational results agree – for both datasets, water-related FPAR variability is highest in grassland regimes. The model captures, to first order, observed hydroclimatic controls on year-to-year variations in vegetation phenology.

In considering Figures 1 and 2, one might wonder if the highlighted FPAR controls are determined in part by the world's distribution of vegetation type. Given, for example, that we prescribe the locations of grassland in the vegetation model (though not its phenological variables), one might hypothesize that the model distributions shown in Figure 2b reflect the imposition of grassland at these locations rather than the model's correct capture of hydroclimatic controls over phenology. We tested this with an additional model simulation, one for which all global land areas were forced to be covered with grassland. The results (not shown) are similar to those shown in Figures 1 and 2, implying that the hydroclimatic controls are indeed properly captured. The GMAO vegetation model is proving itself valid for studies of the interactions of the global energy, water, and carbon cycles.

References

Adler, R.F., and co-authors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeor.*, **4**, 1147– 1167.

Tucker, C.J., J.E. Pinzon, M.E., Brown, D. Slayback, E.W. Pak, R. Mahoney, E. Vermote, and N. Saleous, 2005: An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Rem. Sens.*, **26**(20), 4485–4498. DOI: 10.1080/01431160500168686.

Publication

Koster, R.D., G.K. Walker, P. Thornton, and G.J. Collatz, 2013: Hydroclimatic controls on the means and variability of vegetation phenology. *J. Climate* (*in preparation*).

Using the GMAO Land Modeling System to Address Unresolved Issues in Hydrological Prediction

Randal Koster, Greg Walker, Sarith Mahanama, and Rolf Reichle

Accurate streamflow forecasts on seasonal time scales have obvious societal benefits, particularly in areas of limited water supply. A forecast technique that is growing in popularity involves the separation of a continental region into a high-resolution array of land elements and the subsequent use of a land surface model at each element to translate predicted precipitation forcing (say, from a global seasonal forecast system) into local streamflow. Much of the skill obtained with such a system depends on the accuracy of the initial conditions applied at each land element – in particular, on the accuracy of the assumed initial status of the land's soil moisture and snowpack. Thus, NASA soil moisture and snow mapping missions have the potential to contribute significantly to streamflow forecasts.

Tapping into this potential requires investigations into the still largely unexplored science of streamflow prediction. In this study, we used the GMAO's offline land modeling system to address two of the science's unanswered questions. Both exercises involved driving the GMAO's Catchment LSM with multiple decades of observations-based forcing across the continental United States and then analyzing its simulated streamflows in depth.

In the first exercise, we quantified the degree to which errors in soil moisture initialization degrade a streamflow forecast. We first performed a series of "control" forecasts, utilizing our best estimates of soil moisture initial conditions; we quantified the skill of these forecasts relative to observed streamflows. We then performed several series of "experiment" forecasts that differed from the control only in the initialization of soil moisture – in these experiments, we added well-quantified, artificial error to the soil moisture initial conditions. The addition of a given level of soil moisture error leads to a nearly proportional decrease in streamflow forecast skill, as shown in Figure 1. Through extrapolation, this linearity allows us to estimate the *increase* in forecast skill attainable with improved soil moisture measurement, as expected from the upcoming NASA SMAP satellite mission.

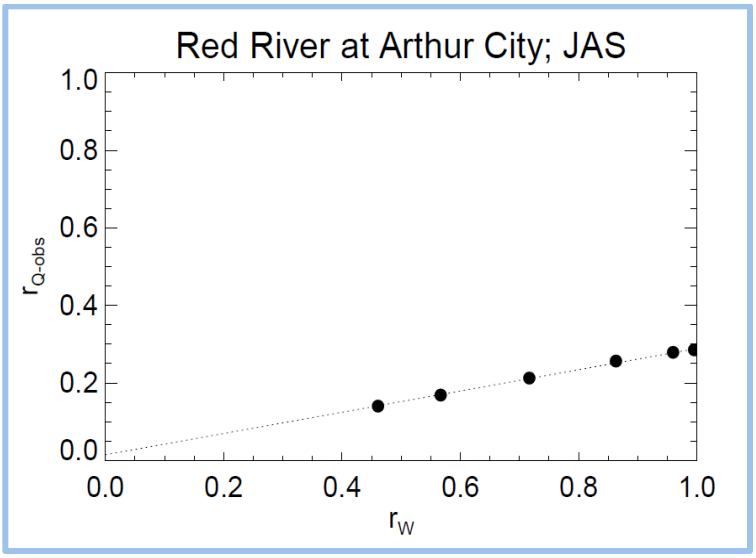


Figure 1. Plot showing how imposed soil moisture error acts to reduce streamflow forecast skill. Each dot corresponds to a different forecast experiment. A given dot's abscissa is the average correlation (across the Red River Basin) between the time series of initial soil moisture contents used by the forecast experiment and those used by the control forecast experiment (the one with no imposed soil moisture error); thus, a smaller abscissa implies a larger imposed error. The dot's ordinate is the correlation between the forecasted streamflow for the basin (i.e., the time series of July-September totals) and the observed streamflow. The dot at r_W=1 thus shows the maximum skill attainable with the system for this basin and season. The linearity shown is found for almost all basins and seasons.

In the second exercise, we examined how the information content of high-resolution precipitation data translates to streamflow forecast skill. We performed two series of forecasts across the continental United States: one with high-resolution precipitation data and the other with the same data, but first coarsened to a lower resolution. The impact of the coarsening on large-area streamflow forecasts was found to be negligible in the eastern U.S., as shown in Figure 2. Subsequent analysis isolated the reason: in the eastern U.S., evaporation doesn't vary much from year to year. (The eastern U.S. lies in an energy-limited evaporation regime.) Because streamflow, over long time scales, is essentially equal to precipitation minus evaporation, the use of the same large-scale precipitation and the production of the same evaporation rates in the two series of forecasts led to the same large-scale runoff, even if the high resolution details of the runoff differed. The added information content in the high-resolution data was not able to affect the large-scale, spatially-averaged fields.

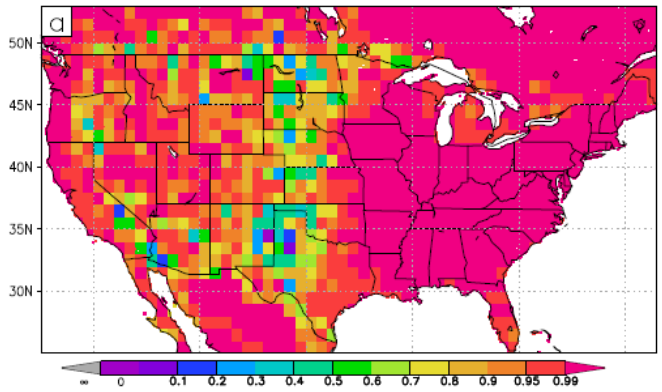


Figure 2. Correlation between the time series of large-scale (1° × 1°) annual runoffs generated in the experiment with coarse resolution (1°×1°) precipitation forcing and those generated in the simulation with high resolution (1/8°×1/8°) precipitation forcing. Given that the resolution of the forcing was the only difference in the simulations, the high (essentially unity) correlations in the eastern half of the continental U.S. is an indication that the added information in the high resolution forcing there did not translate into a change in the runoff when the latter is aggregated up to coarse (1°×1°) scales.

Taken together, the two exercises demonstrate the usefulness of offline land modeling systems as powerful and highly efficient tools for addressing important yet still unresolved issues in seasonal streamflow forecasting. We expect that such systems have great, untapped potential as testbeds for basic hydrological research.

Publication

Koster, R.D., G.K. Walker, S.P.P. Mahanama, and R. H. Reichle, 2013: Soil moisture initialization error and subgrid variability of precipitation in seasonal streamflow forecasting. *J. Hydrometeor.* (*submitted*).

ENSO Connection to U.S. Summertime Precipitation in Reanalyses

Michael Bosilovich

While reanalyses tend to represent the large-scale atmospheric variability in dynamical quantities accurately, physics parameterizations provide a number of diagnostic quantities that are important to society and decision making. These diagnostics are also connected, such that features apparent in one field are also reflected in other related fields. By evaluating the physical quantities with independent observations, we may be able to better understand the accuracy of the parameterized physical processes. This is especially true for northern summer, when convective instability drives precipitation variability.

Over the United States, the network of gauge precipitation stations allows for a very complete representation of precipitation and a reasonable source of independent validation data in recent history. Reanalyses integrate satellite and in-situ observations to constrain the day to day weather in an atmospheric numerical model. That model connects all the observations, but also produces more data than is assimilated, for example, radiation, clouds and precipitation. At large time and space scales, we would expect that the reanalysis is able to replicate the variability of precipitation well.

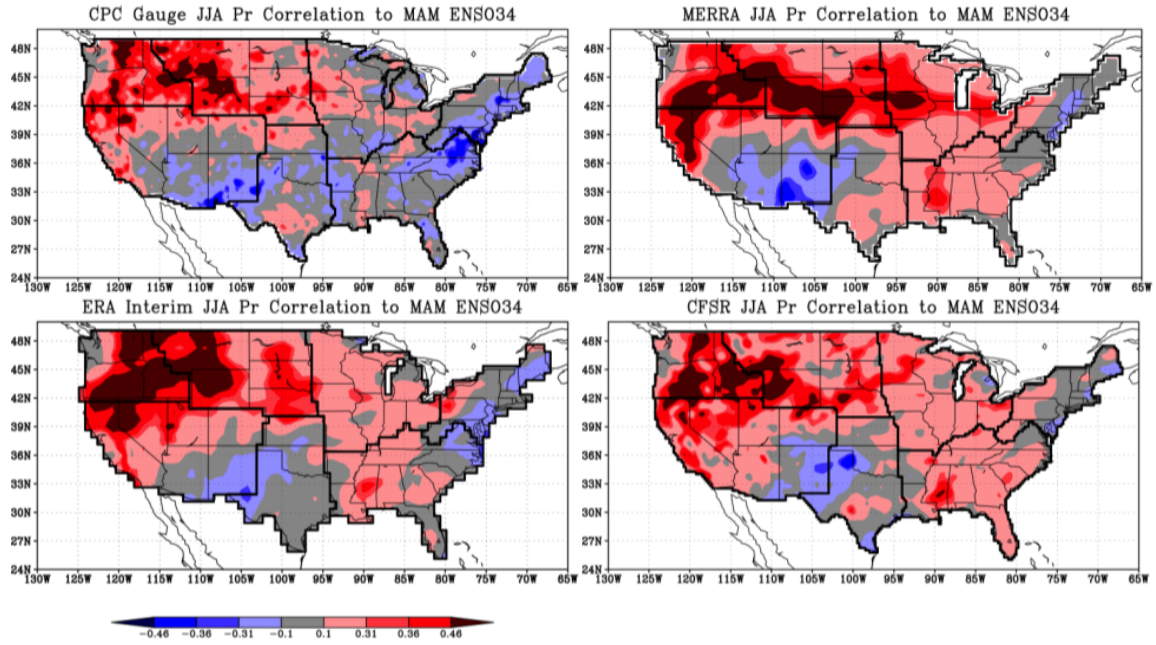


Figure 1. Spatial distribution of the correlation between mean summertime precipitation and spring Niño 3.4 index. The colors indicate level of significance of the correlation (0.31~90%, 0.36~95% and 0.46~99%). While the lowest color contours have less statistical significance, the United States summer precipitation in all reanalyses is more closely related to ENSO than observations suggest. However, in the northwestern United States, this seasonal variability is comparable to gauge observations.

Figure 1 shows summertime correlation of various precipitation data sets to springtime Niño3.4 (an index representing the intensity of the El Niño Southern Oscillation). Observations indicate that the Northwest US has high correlation of the summer precipitation to Niño3.4, and lower or no correlation elsewhere. MERRA shows a similar pattern, as do other reanalyses. The significant correlation in the northwestern US is likely related to the large-scale circulation and its replication in reanalyses comes from the assimilated satellite observations over the Pacific Ocean.

In other regions, notably the Midwest and Southeast US, there is little or no correlation to Niño3.4 in the summertime observations, yet the reanalyses, especially MERRA, show positive correlation, sometimes with statistical significance. In these regions, we expect that the atmospheric boundary layer and land-atmosphere interactions have a dominant influence over precipitation occurrences. The overly correlated

reanalyses may have too strong large-scale circulations, or limited local circulations related to coarse spatial resolution, relative to mesoscale convective complexes that occur during summer.

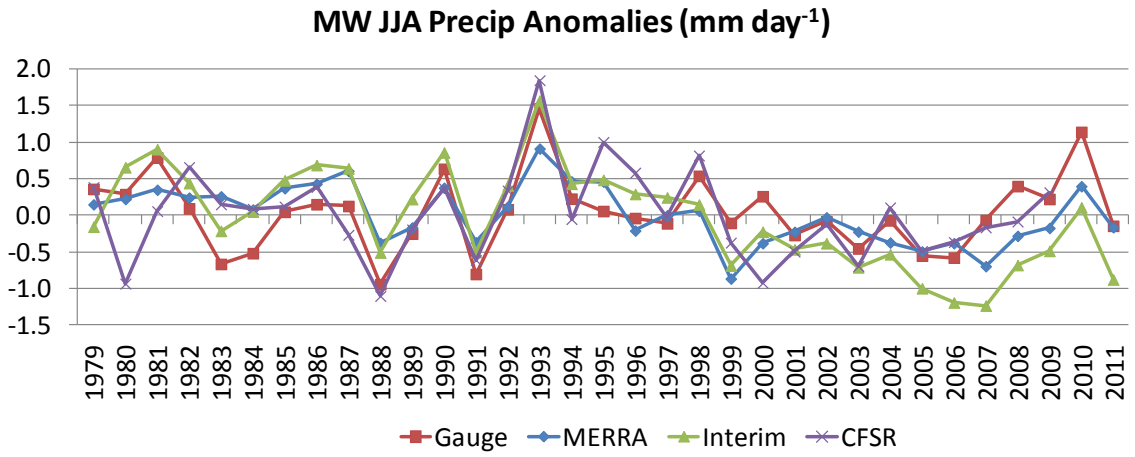


Figure 2. Time series of Midwest US seasonal precipitation anomalies for CPC Gauge observations, MERRA, ERA-Interim and CFSR. Each reanalysis has some weakness in this region.

Figure 2 shows the time series of precipitation anomalies from the observation data and reanalyses for the Midwest US. While the Northwest US shows nearly indistinguishable time series among these data (not shown), the Midwest time series of all the reanalyses have weaknesses in their interannual variability. ERA-Interim has a persistent declining trend, far more significant than that which is observed. NCEP's CFSR exhibits extreme seasons in years when none were observed. MERRA is not able to replicate the extent of the wings of the distribution, specifically, anomalously dry seasons are not as dry and wet seasons are not as wet (note the summers of 1988 and 1993, and others). While the reanalyses show some similarity in the interannual variability mentioned above, closer inspection also shows substantial differences in how the interannual variability is reproduced in each reanalysis.

Given that resource management and decision making are done on smaller scales than these, reanalyses data will need improvements through system development and likely spatial resolution to extend their usefulness beyond that of the large-scale environment.

Publication

Bosilovich, M.G., 2013: Regional climate and variability of NASA MERRA and recent reanalyses: US summertime precipitation and temperature. *J. Appl. Meteor. Climatol.* (*accepted*).

Decadal Prediction Skill in the GEOS-5 Forecast System

Yoo-Geun Ham, Michele Rienecker, Max Suarez, Yury Vikhliaev, Bin Zhao, Jelena Marshak, Guillaume Vernieres, and Siegfried Schubert

As a contribution to the IPCC's Fifth Assessment Report (AR5), we conducted a suite of decadal predictions with the GEOS-5 Atmosphere-Ocean General Circulation Model (AOGCM). Such predictions, which examine the impact of initialization of the AOGCM compared to the uninitialized simulations and projections of climate variations and trends, are a natural extension of the GMAO's Seasonal-to Interannual (SI) prediction efforts.

The hindcasts were initialized every December 1st from 1959 to 2010, following the CMIP5 experimental protocol for decadal predictions. The initial conditions were obtained from a multi-variate ensemble optimal interpolation ocean and sea-ice reanalysis conducted in the coupled system with the atmosphere constrained by MERRA. The configuration of the atmospheric component used a 2.5° longitude x 2° latitude grid, with 72 vertical levels extending to 0.01hPa; the ocean component used 1° longitude and latitude, with a meridional equatorial refinement to 1/3°, and 50 vertical levels. The atmospheric component includes a river runoff routing scheme and an aerosol model based on the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model using emissions prescribed by the CMIP5 protocol. Only those volcanic aerosols from continually outgassing volcanoes have been included, i.e., the explosive volcanic eruptions are not included.

Here, we examine the forecast skill of a three-member-ensemble mean compared to that of an experiment without initialization but also forced with observed greenhouse gases. Figure 1 shows the anomaly correlation coefficient (ACC) of the 4-year-averaged SST from the 20th century (C20C) simulations and the difference of the ACC for the initialized decadal forecasts from that for the C20C simulations. The calculation of the anomaly corrects the bias by removing the mean drift away from the observations (Ham et al., 2013). The HadISST1 is used for validation since it is a gridded synthesis of available observations and provides the appropriate metric for the C20C simulations.

Based on the ACC, there is a predictable signal in the C20C simulation over the Indian Ocean, western and south-central Pacific, and equatorial Atlantic. In addition, the ACC exceeds 0.6 over the mid-latitude western Atlantic. This skill is entirely due to the external forcing because the C20C simulation is not initialized. Therefore, the comparison of the skill in the C20C simulation to that in the decadal forecasts provides some information on added skill or on skill degradation due to initialization. According to the anomaly correlation measure, the initialized prediction skill of 2-to-5-year lead forecasts over the subtropical, mid- and eastern North Atlantic is greater than that for the C20C simulation. This is significant with 90% confidence level using a bootstrap approach following Smith et al. (2010). This implies that the initialization has been successful in enhancing the prediction skill of Atlantic decadal variability. The Mean Squared Skill Score (MSSS; Goddard et al., 2013) shows more robust improvement with initialization compared with the anomaly correlation metric. The MSSS values are positive over most regions of the globe. However, exceptions are evident in localized areas over the mid-North Pacific and parts of the subpolar North Atlantic.

For the 6-to-9-year lead forecasts, the improvement in ACC decreased over the subpolar North Atlantic compared with years 2-5, while that over the mid-latitude Atlantic and Pacific is even higher than in the C20C simulations. This indicates that the GEOS-5 AOGCM has some success in predicting the Atlantic climate up to about 10-year lead times except for the subpolar regions. The MSSS results are similar for both leads. Thus, by these measures, initialization gives a systematic increase of prediction skill for SST over many areas of the globe, with the notable exception of the subpolar North Atlantic.

Not surprisingly, areas of significant skill for the ocean heat content in the upper 500 m (HC500) were found to be larger than for SST (not shown). Improvements in ACC were found in the central basins and high latitudes as well as the western Indian Ocean. On the other hand, the ACC over the eastern Indian Ocean and eastern equatorial Pacific as well as over the western and subpolar North Atlantic degraded in the initialized forecasts. While the 6-9-year forecast skill measured by the MSSS showed 50% improvement for HC500 over the subtropical and mid-latitude Atlantic, prediction skill was relatively low in the subpolar gyre. This low skill is due in part to the fact that the spatial pattern of the dominant simulated decadal mode in upper ocean heat content over this region appears to be unrealistic. An

analysis of the large-scale temperature budget shows that this is the result of a model bias, implying that realistic simulation of the climatological fields is crucial for skillful decadal forecasts.

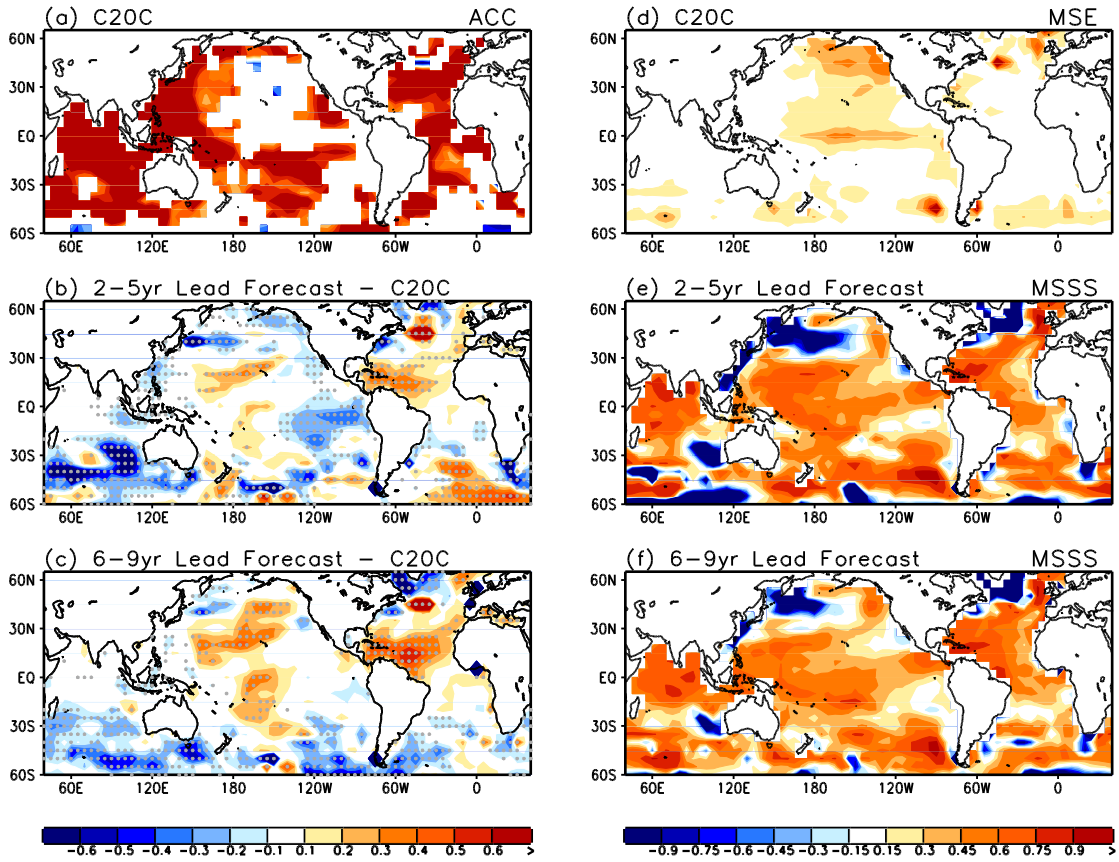


Figure 1. (a) The anomaly correlation of 4-year moving averaged SST from the HadISST1 with the C20C simulation. The difference of anomaly correlation in the (b) 2-5-year lead forecast and (c) 6-9-year lead forecast from that in C20C. The right panel shows the (d) Mean Square Error (MSE) of the C20C simulation, Mean Squared Skill Score (MSSS) for the (e) 2-5-year lead forecast and (f) 6-9-year lead forecast. The grey dots in panel (b), and (c) denote that the difference in correlation is significant with 90% confidence level. The procedure for the significant test is the same as in Smith et al. (2010).

References

Goddard, L. and co-authors, 2013; A verification framework for interannual-to-decadal predictions experiments. *Clim. Dyn.*, **40**, 245-272, DOI: 10.1007/s00382-012-1481-2.

Smith, D.M., R. Eade, N.J. Dunstone, D. Fereday, J.M. Murphy, H. Pohlmann, and A.A. Scaife, 2010: Skilful multi-year predictions of Atlantic hurricane frequency. *Nature Geosci.*, **3**, 846-949, DOI:10.1038/NGEO1004 supplementary material.

Publication

Ham, Y.-G., M.M. Rienecker, M.J. Suarez, Y. Vikhliaev, B. Zhao, J. Marshak, G. Vernieres, and S. Schubert, 2013: Decadal prediction skill in the GEOS-5 forecast system. *Clim. Dyn.* (*accepted*).

An Assessment of the Skill of GEOS-5 Seasonal forecasts

Yoo-Geun Ham, Siegfried Schubert, and Michele Rienecker

The seasonal forecast skill of the GEOS-5 AOGCM is evaluated based on an ensemble of 9-month lead forecasts for the period 1993 to 2010. Comparisons are made with an earlier version (V1) of the system in which the AGCM (the NSIPP model) was coupled to the Poseidon Ocean Model. The current version (V2) is a major upgrade of that system consisting of the GEOS-5 AGCM coupled to the MOM4 ocean model.

The correlation skill of the Sea Surface Temperature (SST) forecasts is generally better in V2, especially over the sub-tropical and tropical central and eastern Pacific. V2 also shows improved correlation skill and Mean Square Skill Score (MSSS) over the Atlantic and Indian Ocean for all forecast lead months. The skill of Niño3.4 is greater than 0.7 for up to 7-month forecast lead times – a substantial improvement compared with V1. Figure 1 shows the correlation skill and MSSS score of the SST with respect to the initial (x-axis) and target month (y-axis). Improvements in V2 come mainly from the better forecasts of the developing phase of ENSO from boreal spring to fall. The skill of forecasts initiated during the peak phase of ENSO are however slightly worse in V2.

An analysis of the SST tendencies shows that the magnitude of ENSO in V1 is excessive during its developing phase due to a model bias characterized by a too strong climatological zonal temperature gradient (Ham et al., 2013). This is due to a cold bias located over the eastern Pacific. In contrast, V2 has a more zonally uniform cold bias extending from the central to eastern Pacific that, to a large extent, preserves the observed climatological zonal temperature gradient in that region. In addition, an excessive climatological meridional current in V1 appears to be responsible for the development of ENSOs that are stronger than observed.

V2 forecasts that are initiated during the ENSO mature phase exhibit a phase transition of ENSO-related geostrophic zonal current that is too slow, due to the weak discharge of equatorial Warm Water Volume (WWV). In addition, the discharge process in V2 is systematically slower than that in V1. This slower discharge process is caused by a positive meridional zonal wind stress gradient over the central-eastern Pacific Intertropical Convergence Zone (ITCZ) that leads convergent meridional flow within the thermocline layer. Ultimately this leads to an ENSO in V2 that grows excessively compared with both the observed and that in V1 when the forecast is initialized in the boreal winter season (i.e. ENSO peak phase).

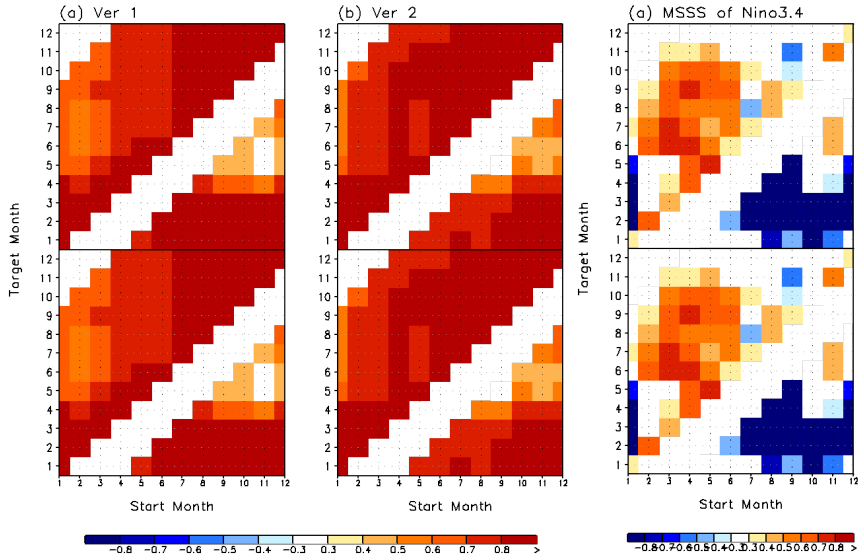


Figure 1. The correlation skill of SST in V1 (left) and V2 (middle), and MSSS scores of V2 compared to the V1 (right panel) with respect to the initial (x-axis) and target month (y-axis).

Publication

Ham, Y.-G., S. Schubert, and M.M. Rienecker, 2013: An assessment of the skill of GEOS-5 seasonal forecasts. (*draft ms*).

Sea Surface Temperature in the North Tropical Atlantic as a Trigger for El Niño/Southern Oscillation Events

Yoo-Geun Ham, Jong-Seong Kug (KIOST), Jong-Yeon Park (KOIST), and Fei-Fei Jin (IPRC)

Although it's well-known that El Niño/Southern Oscillation (ENSO) events can influence the weather and sea surface temperatures (SSTs) in ocean basins outside of the tropical Pacific, only recently have the influences of the tropical Atlantic on the ENSO cycle been explored. In a recent study (Ham et al., 2013), data from MERRA and other sources were analyzed in order to identify how anomalously warm SSTs in the northern tropical Atlantic (NTA, here referred to as the region 90° W - 20° E, 0° - 15° N) influence ENSO variability by teleconnections.

To assess the impact that anomalously warm SSTs in the NTA have on ENSO, we analyzed monthly mean wind, geopotential height, and precipitation data from MERRA for the thirty years from 1980 to 2010, along with SST, subsurface temperature, ocean currents, and vertical velocity data from NCEP ocean reanalyses. Since NTA SSTs are influenced, with a few months' lag, by ENSO variability, ENSO's effect had to be removed from the analysis. For that, a linear regression was performed with respect to Niño3.4 SSTs (the Niño3.4 region is 170° - 120° W, 5° S - 5° N) for the December, January, and February (DJF) season of the previous year and subtracted from a lagged regression, with respect to NTA SSTs, for the February, March, and April (FMA) season. The results, summarized in Figure 1, show that warm SSTs in the NTA in boreal spring can remotely influence ENSO variability, leading to a La Niña event the following winter.

Anomalously warm SSTs in the NTA in March, April, and May (MAM) were found to result in enhanced convective activity, and therefore enhanced diabatic heating, especially over the equatorial Atlantic. In response, a low level cyclonic circulation develops in the subtropical eastern Pacific, thereby creating a northerly flow at the west flank of the circulation that results in surface cooling from both the enhanced winds and from the advection of colder/drier air from higher latitudes. This cold, dry air acts to suppress convection and precipitation and creates a clear area in part of the subtropical Pacific. During the boreal summer months of June, July, and August (JJA), the area of suppressed convection moves north with the typical northward migration of the Pacific Intertropical Convergence Zone (ITCZ), and generates a low level anticyclonic circulation that enhances the northerly flow already present on its eastern flank, thereby reinforcing the subsidence.

As the summer progresses, the negative precipitation anomaly and associated anticyclonic flow that are created by the subsidence induced by enhanced northerly flow expands westward and induces an easterly flow over the western Pacific at the equator, generating an upwelling Kelvin wave that cools SSTs in the equatorial Pacific, and possibly making conditions favorable for a La Niña event. Westerly winds over the eastern Pacific suppress upwelling. Hence, cold SSTs are mainly confined over the central Pacific.

In the eastern Pacific the local and remote response of SST through both mechanisms are of opposite sign and the response to the NTA is small. Under the influence of both mechanisms, the SST response is mainly confined to the central Pacific.

This investigation of reanalysis data has been verified by a modeling study in which ocean-atmosphere coupled global circulation models (CGCMs) were run with prescribed SST fields, where the only difference between runs was the SST anomaly over the NTA. In runs where the NTA SSTs were warm in the spring season, a La Niña event was more likely to take place the following winter than in the control, and the La Niña event had a high probability of being strong. In instances where an El Niño event was underway when NTA SSTs warmed, that El Niño event was more likely to be terminated early than in the control. In simulations where SSTs in the NTA were anomalously cool, an El Niño event was more likely to develop, but unlike a canonical El Niño event, the warm SST anomaly was more likely to be located in the central Pacific.

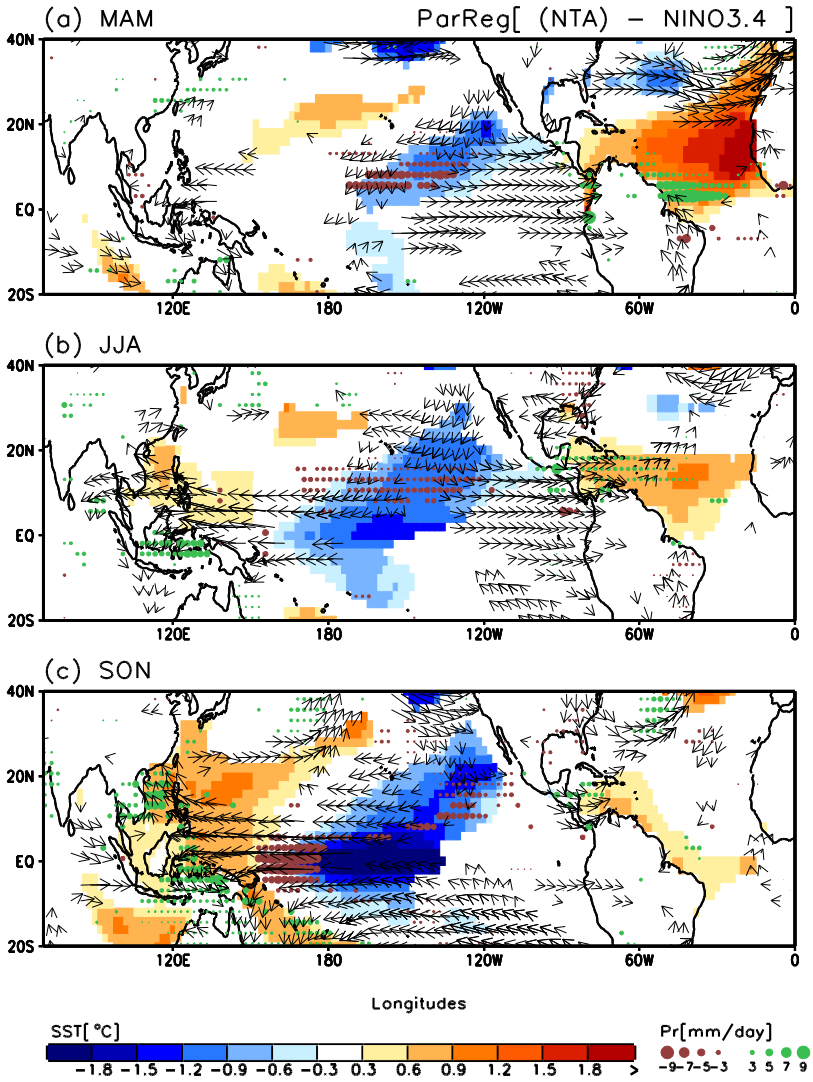


Figure 1. Lagged regressions between NTA SST (90°W-20°E, 0-15°N) averaged over February-to-April (FMA season) and SST, wind vector at 850 hPa (vector), and precipitation during (a) MAM, (b) JJA, and (c) SON seasons, after excluding the impact of Niño3.4 SST (170-120°W, 5°S-5°N) during the previous DJF season. Only the values at the 95% confidence level or higher are shown.

Publication

Ham Y.-G., J.-S. Kug, J.-Y. Park, and F.-F. Jin, 2013 : Sea surface temperature in the north tropical Atlantic as a trigger for El Niño/Southern Oscillation events. *Nature-Geoscience*, **6**, 112-116, doi:10.1038/ngeo1686.

A Modeling Study of the Spring 2011 Extreme U.S. Weather Activity

Siegfried Schubert, Yehui Chang, Hailan Wang, and Max Suarez

During April 2011, the United Sates experienced record-breaking precipitation in the Ohio Valley accompanied by record high tornadic activity over much of the southeast as well as flooding on the Mississippi. These climate anomalies were juxtaposed with continued record drought over Texas. The large-scale circulation during that month was characterized by enhanced moisture flux from the Gulf of Mexico and from the Atlantic along the eastern seaboard. Other aspects of the circulation include a Pacific jet with an anomalous extension across the northern tier of states and Canada, and an extremely large amplitude positive North Atlantic Oscillation (NAO). The SST anomalies showed continuing but weakening La Niña conditions.

This study examines the causes of the extreme US weather activity by running a number of simulations using the GEOS-5 AGCM run at ¼° latitude/longitude resolution. Specific questions that we address are: 1) what is the role of the SST in the development of the April anomalies; 2) are the late winter/early spring soil moisture and snow conditions important for the subsequent development; and 3) what are the important large-scale atmospheric controls on the weather extremes?

The AGCM simulations consistent of a series of hindcasts initialized on March 1 and forced with observed SST. The atmosphere and land were initialized from MERRA. The perturbations (in the atmosphere only) were computed as the scaled differences between two November states taken 1 day apart. For each year (1980-2008), four-member ensembles were generated to provide a forecast climatology. For 2011, an additional 18 ensemble members were generated for a total of 22. In addition, an auxiliary set of 22 hindcasts was generated for 2011 in which the land states were randomly swapped out with the land states from the previous 10 years, allowing us to assess the impact of the initial soil moisture conditions.

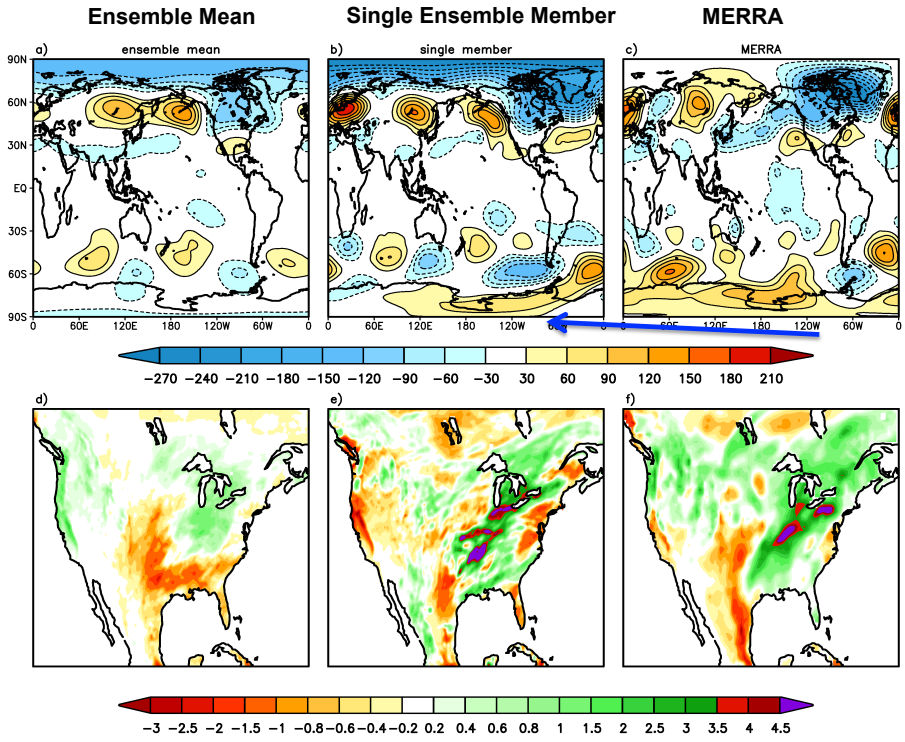


Figure 1. Top panels from left to right: The ensemble mean, a single ensemble member, and the observed 250mb height anomalies (meters) during April 2011. The bottom panels: Same as top panels but for precipitation (mm/day).

The modeling results show that the extreme weather activity and record-breaking precipitation over the US during April 2011 was in part a forced response to the prevailing SST (La Niña conditions), and in part the result of the development of an unusually large amplitude positive NAO-like (free) mode. Figure 1 highlights the forced response (upper tropospheric height and precipitation) and compares that with a single ensemble member chosen to most closely resemble the observations in the height field. Figure 2 shows that the intra-ensemble noise includes the NAO as an important component, allowing us to reconstruct the observations as a linear combination of the forced response and some of the leading noise patterns. The NAO has limited predictability at 2-month lead-time, and appears to play a major role in limiting the predictability of the precipitation over the US at these time scales (signal to noise ratios less than 0.5). We find some evidence that the SST-forced (La Niña) response in weather variability caused a predilection for a positive NAO, though this still needs to be better quantified. Experiments with modified initial soil moisture suggest it contributed to the wet conditions in the upper Midwest and dry conditions in the Southeast. The La Niña related SST anomalies also contributed to the continuing drought over Texas.

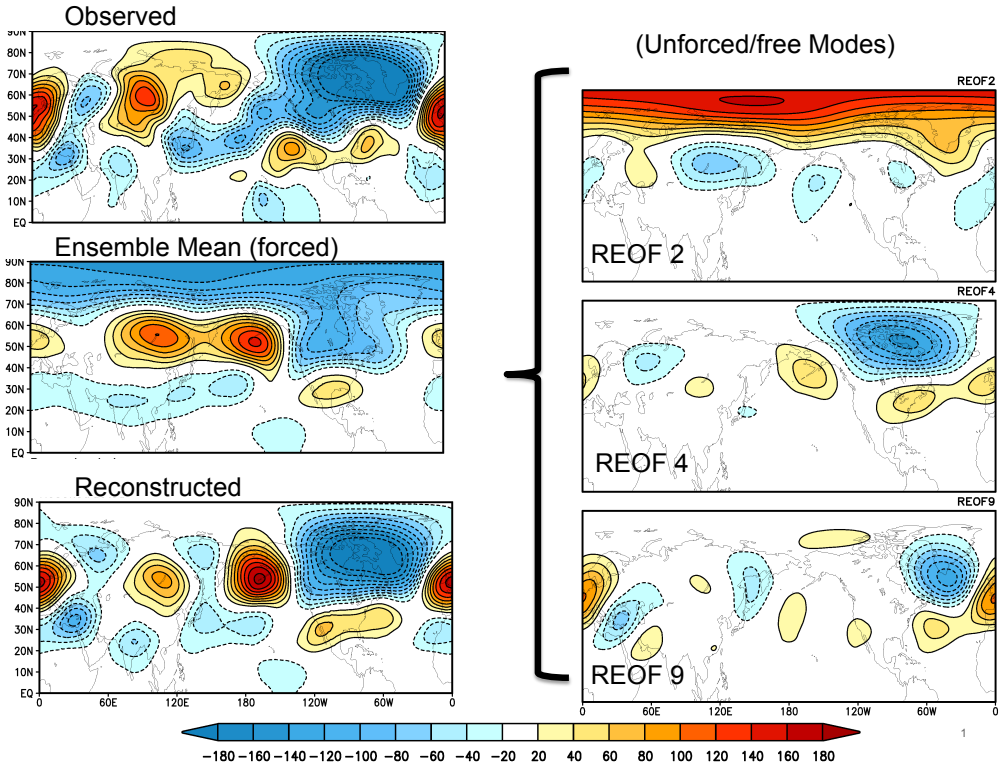


Figure 2. The April 2011 250mb height anomalies. The top left is the observed anomaly. The middle left is the ensemble mean of the model simulations forced with observed SST. The bottom left is the height field that has been reconstructed in order to most closely match the observed (via regression), as a sum of the forced response and three appropriately weighted REOFS of the intra-ensemble variability (or noise).

SST Forcing and Soil Moisture Feedbacks during the 2012 U.S. Drought

Hailan Wang, Siegfried Schubert, Randal Koster, and Max Suarez

The drought of 2012 was characterized by record-breaking precipitation deficits throughout the central and northern Great Plains. It has been referred to as a "flash drought", developing rather suddenly during May and June, and peaking in July. The drought was accompanied by extreme heat that covered much of the same region. Here we investigate the causes and predictability of that extreme event, focusing on the temperature anomalies and the role of SST forcing in the various ocean basins as well as the role of land-atmosphere feedbacks. A series of twenty 1-year simulations were carried out with the GEOS-5 AGCM initialized 15 November 2011. The various runs were forced with either global or regional (tropical Pacific, north Pacific, tropical Atlantic, north Atlantic) observed SST. In a second set of twenty runs forced with global SST, the land coupling was disabled (by specifying climatological soil moisture). We also examined predictions of this event from our seasonal forecasts carried out with the fully coupled GEOS-5 model.

The drought was accompanied by SST anomalies that are generally conducive to drought conditions over the U.S., though the previous La Nina conditions were largely gone by the summer of 2012, with the remaining anomalies characterized by a cold PDO and warm AMO (Fig. 1 left panels). Figure 1 also shows the development (May through July) of the surface temperature anomalies over the US from MERRA, from the ensemble mean of the AGCM simulations, and for a single ensemble member chosen to most resemble the observations. MERRA shows generally warm conditions throughout the central and eastern US during May: these developed into a more concentrated region of warming over the central plains during June and July. The model results show that the SST produce a rather weak (compared to the MERRA) pattern of warming largely confined to the west and southern Plains, and along the east coast during May and June. By July, the anomalies had moved northward extending into the central and western Plains. The single ensemble member shows strong warming in the central Plains throughout the summer, with magnitudes comparable to those from MERRA.

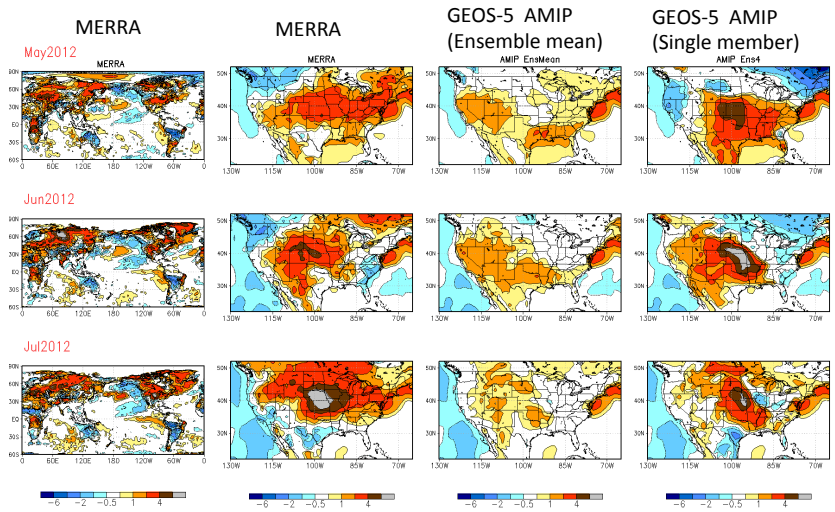


Figure 1. Left panels: The global T2m anomalies from MERRA for May, June and July of 2012. Second from left: The T2m anomalies over the US from MERRA. Second from right: The ensemble mean T2m anomalies from the GEOS-5 model simulations forced by the observed SST. Right panels: The T2m anomalies from a single ensemble member. Units: °C.

Figure 2 shows the response of the model to the SST anomalies in the different ocean basins, with the tropical Pacific SST providing the main forcing of the warm anomalies over the southern US during May and June, while the tropical Atlantic SST force the warm anomalies over the central Plains during July. The North Atlantic SSTs appear to be the main forcing of the warmth on the eastern seaboard. These results indicate that the "flash drought" that developed in the central plains during June and July was in part driven by the tropical Atlantic SST, though a large component of that event appears to be unforced by SST. In fact, further analysis shows that the anomalous upper tropospheric high that developed over the central and northern U.S. associated with the drought is primarily maintained by transient vorticity flux convergence over the western U.S. which appears to be linked to atmospheric internal variability.

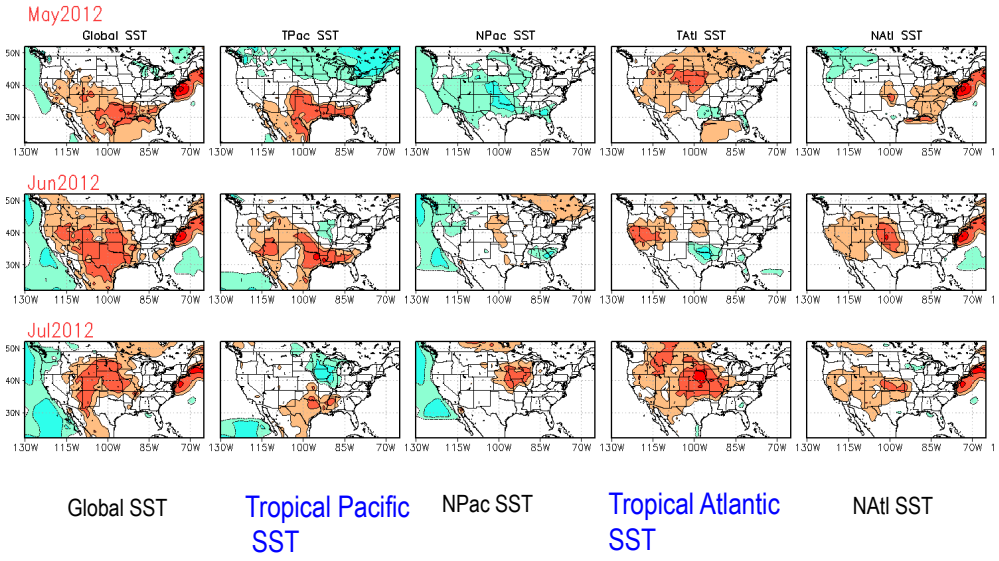


Figure 2. The ensemble mean of the T2m anomalies over the US from simulations with the GEOS-5 AGCM for May, June and July 2012. Left panels: runs forced with global SST. Second from left: runs forced by tropical Pacific SST. Third from left: runs forced by north Pacific SST. Fourth from left: runs forced by tropical Atlantic SST. Right panels: runs forced by north Atlantic SST.

The simulations with specified climatological soil moisture (not shown) indicate that soil moisture feedback is critical for the development of the surface temperature anomalies over the US. Finally, Figure 3 shows that land initial conditions appear to play an important role in providing skillful July predictions for forecasts initialized at the beginning of June and July, helping to overcome the limited predictability afforded by the SST on these time scales. Further information about this drought and a similar analysis carried out for the 2011 US drought can be found in Wang et al. (2013).

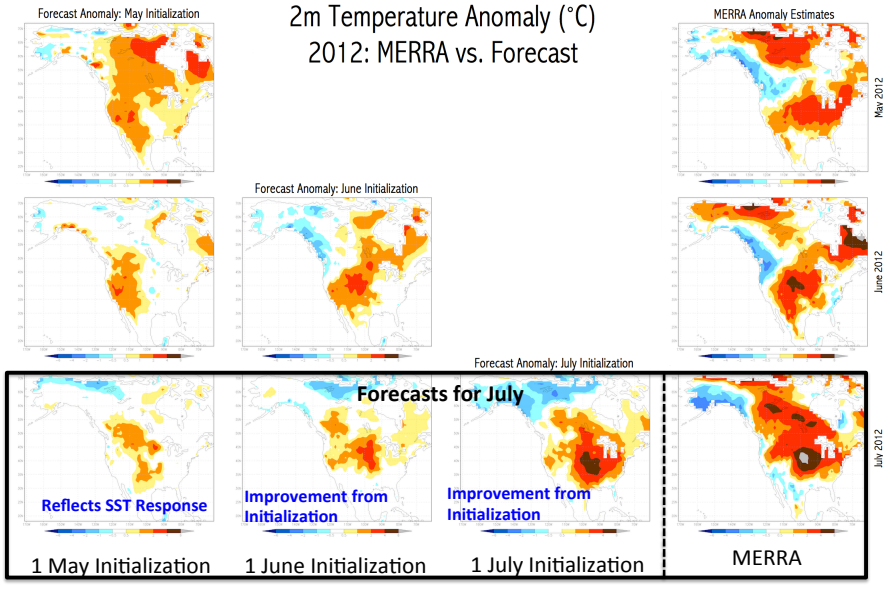


Figure 3. First column: Surface temperature anomalies (ensemble means) produced in GEOS-5 coupled model forecasts initialized in early May, for May (top), June (middle), and July of 2012. Second column: Same, but for forecasts initialized in early June. Third column: Same, but for forecasts initialized in early July. Fourth column: MERRA estimates of the surface temperature anomalies.

Publication

Wang, H., S. Schubert, R. Koster, Y.-G. Ham, and M. Suarez, 2013: On the role of SST forcing in the 2011 and 2012 extreme U.S. heat and drought: A study in contrasts. Submitted to the special collection of the *J. Hydrometeor.* on "Advancing Drought Monitoring and Prediction".

Satellite Views of Global Phytoplankton Community Distributions using an Empirical Algorithm and a Numerical Model

Cecile S. Rousseaux, Taka Hirata (Hokkaido University), and Watson W. Gregg

Phytoplankton composition plays a major role in biogeochemical cycles in the ocean. The intensity of carbon fixation and export is strongly dependent on the phytoplankton community composition. The approaches to characterize the phytoplankton community composition at a global scale can be roughly classified in two categories: modeling approaches and satellite-derived approaches. In the modeling approach, data assimilation techniques can also be used to constrain the model to track observations time series and to optimize certain variables.

Models and satellite-derived approaches have been used to assess global changes in phytoplankton biomass and community composition at various time scales. Temporal oscillations of phytoplankton biomass are often very variable. The number, timing and magnitude of annual blooms may differ remarkably among locations. In this study we compared the functional response of a numerical model (NASA Ocean Biogeochemical Model, NOBM; Gregg et al., 2003) versus an empirical algorithm (Hirata et al., 2011) in describing the spatial distribution and seasonal variation of four phytoplankton groups (diatoms, cyanobacteria, coccolithophores and chlorophytes) globally and in 12 major oceanographic basins. Global mean differences of all groups were within ~15% of an independent observational data base for both approaches except for satellite-derived chlorophytes. Diatoms and cyanobacteria concentrations were significantly ($p < 0.05$) correlated with the independent observation data base for both methods. Coccolithophore concentrations were only correlated with the in-situ data for the model approach and the chlorophyte concentration was only significantly correlated to the in-situ data for the satellite-derived approach.

This comprehensive global comparison of a data assimilating biogeochemical model and a satellite-derived approach with extensive in-situ archives indicate an overall good agreement of the spatial distribution and seasonal variation in phytoplankton community composition. Using the model, high diatom concentrations were spatially more widespread than using the satellite-derived approach (Fig. 1). For example, the North Pacific was the region where the model identified the highest diatom concentration (0.33 mg m^{-3}) whereas the satellite-derived average concentration was 0.15 mg m^{-3}. Although not as productive (in terms of chlorophyll concentration) as the northernmost latitudes, Antarctic was also a region of abundant diatoms and where the two approaches differed. Both approaches agreed on the overall distribution of cyanobacteria, except for high latitude and upwelling regions. For example, the satellite-derived approach identified the North Atlantic and Pacific as the regions with highest cyanobacteria concentrations (~0.05 mg m^{-3}) whereas the model detected low cyanobacteria concentrations (<0.01 mg m^{-3}) in these two regions. This was also the case for the Antarctic where satellite-derived cyanobacteria concentrations were of ~0.04 mg m^{-3}, while the model did not identify the presence of any cyanobacteria in this region.

Using monthly means from 2002-2007, the seasonal variation from the satellite-derived approach and model were significantly correlated in 11 regions for diatoms and in 9 for coccolithophores but only in 3 and 2 regions for cyanobacteria and chlorophytes. Most disagreement on the seasonal variation of phytoplankton composition occurred in the North Pacific and Antarctic where, except for diatoms, no significant correlation could be found between the monthly mean concentrations derived from both approaches. Chlorophytes were the group for which the two approaches differed most and were furthest from the in-situ data. These results highlight the strengths and weaknesses of both approaches and allow us to make some suggestions to improve our approaches to understanding phytoplankton dynamics and distribution.

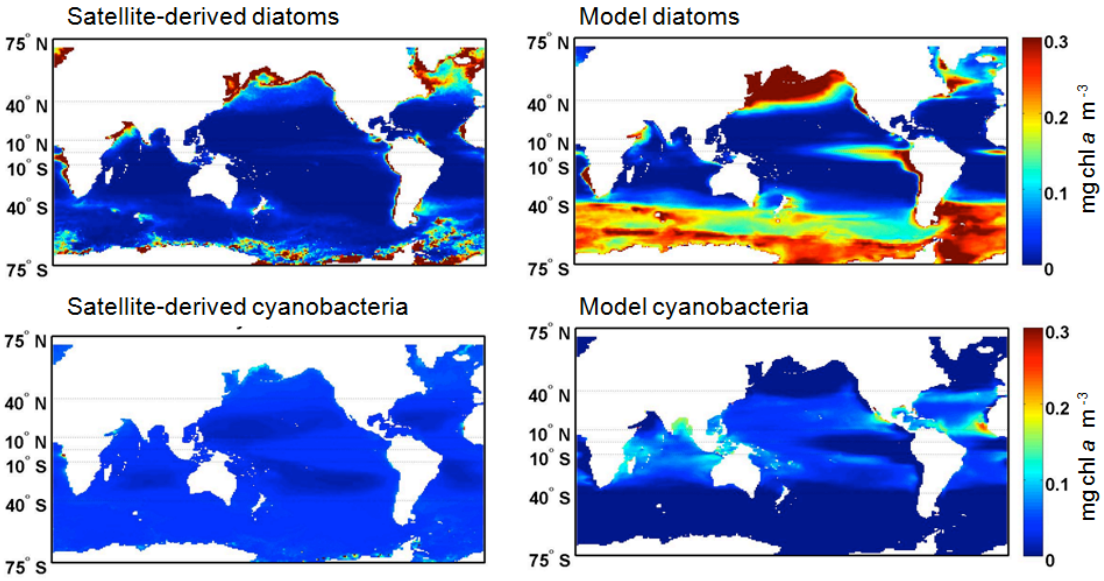


Figure 1. Climatology (1998-2007) of the spatial distribution of diatoms and cyanobacteria using the satellite-derived approach and the NASA Ocean Biogeochemical Model. Note that the satellite-derived prokaryotes are compared to cyanobacteria from the model.

References

Gregg, W.W., P. Ginoux, P.S. Schopf, and N.W. Casey, 2003: Phytoplankton and iron: Validation of a global three-dimensional ocean biogeochemical model. *Deep Sea Res. Part II*, **50**, 3143-3169.

Hirata, T., and co-authors, 2011: Synoptic relationships between surface chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences*, **8**, 311-327.

Publication

Rousseaux C.S., T. Hirata, and W.W. Gregg, 2013: Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model. *Biogeosciences* (*discussion paper*).

GEOS-5 Product Generation

Gi-Kong Kim, Robert Lucchesi, and Jonathan Kelly

Successful GEOS-5 data production, delivery, and system upgrades highlight a busy year for production operations. One of the most significant upgrades came to fruition in March 2013 when we implemented a new near-real-time (NRT) forward processing data stream we developed in coordination with NASA instrument teams. This new stream, called GEOS-5 Forward Processing for Instrument Teams (FP-IT), is specifically tailored to better meet Instrument Team requirements for re-processing.

Concurrent with upgrading our instrument team support we continued to deliver GEOS-5.2 NRT data to NASA instrument teams and upgraded the core GEOS-5 model. Likewise, we delivered high quality MERRA data without interruption. It is a testament to the quality of and ease of access to the MERRA output that interest and use of MERRA products continues to grow worldwide.

GMAO's Products web page at http://gmao.gsfc.nasa.gov/products/ provides more detailed information on GEOS-5 products.

GEOS-5 Data Production

We generate GEOS-5 data products using High End Computing (HEC) resources at Goddard Space Flight Center's (GSFC) NASA Center for Climate Simulation (NCCS). One of our most significant upgrades this year was to split our NRT data production in two streams at the NCCS. One stream is named Forward Processing-Instrument Team (FP-IT) and provides ½° resolution data from a semi-frozen version of the GEOS-5 system. The other is called simply our Forward Processing (FP) stream and comes from our most recent GEOS-5 system. Currently, it generates ¼° data, including forecasts as well as assimilation products.

FP-IT System: Simultaneous with our preparation and introduction of the split streams, our legacy instrument team support continued without significant interruption. We delivered our NRT GEOS-5.2 data to the team via the Modeling and Assimilation Data and Information Services Center (MDISC) (http://disc.sci.gsfc.nasa.gov/mdisc/).

NRT FP-IT data delivery to the MDISC began in early March 2013. FP-IT data are produced using GEOS-5.9.1 at the same spatial resolution (½ degree) as the GEOS-5.2 data. The ½ degree processing is the only viable option that also allows both GEOS-5 forward processing and reprocessing. This is important because we are currently running three significant reprocessing streams for instrument teams dating back as far as 2000. All of current reprocessing efforts are expected to be complete by late summer 2013.

The Community Earth System Model (CESM) team and the MODIS L-3 AOD production team were added to the GEOS-5 FP-IT real time data user community this year. Additionally, the Ozone Mapping Profiler Suite – Limb Profiler (OMPS-LP) team switched to GEOS-5 data from NCEP data in March 2013 to take advantage of better quality of GEOS-5 temperature data in the stratosphere. The OMPS-LP team produces ozone profiles from cloud-top to 60 km.

FP System: GEOS-5 FP system continues to run twice daily with output available through the NCCS data portal. The FP system will soon be upgraded to V5.10.

Other upgrades: There were several other significant upgrades that affect both the FP and FP-IT systems. The GCM in both systems now uses a cube-sphere grid scheme, which not only avoids computational problems at the poles on the lat-lon grid but also helps increase processing speeds. Our upgrade to the GEOS-5.9.1 system incorporated improvements and new features in the General Circulation Model (GCM) and the Grid point Statistical Interpolation (GSI). GEOS-5.9.1 and the upcoming GEOS-5.10 upgrade are capable of assimilating new observation types that are not used in GEOS-5.2, including GPS radio occultation, Aura OMI ozone, and brightness temperature data from GOES-13, S-NPP ATMS, Metop-A, Metop-B, and NOAA-19. We added a module to assimilate MODIS aerosol observations from Aqua and Terra as well. Finally, the FP and FP-IT systems now use OSTIA Sea Surface Temperature (SST) data that has a higher horizontal and temporal resolution than that used with GEOS-5.2.

MERRA Data Downloads Still Trending Upward

The average volume of monthly MERRA products downloaded has continued to increase since MERRA data were released to the general user community in late 2008 (Fig. 1). As of February 2013, the cumulative total volume of data downloaded exceeded 1.6 PB. This is an increase of 0.6 PB in just the last year alone.

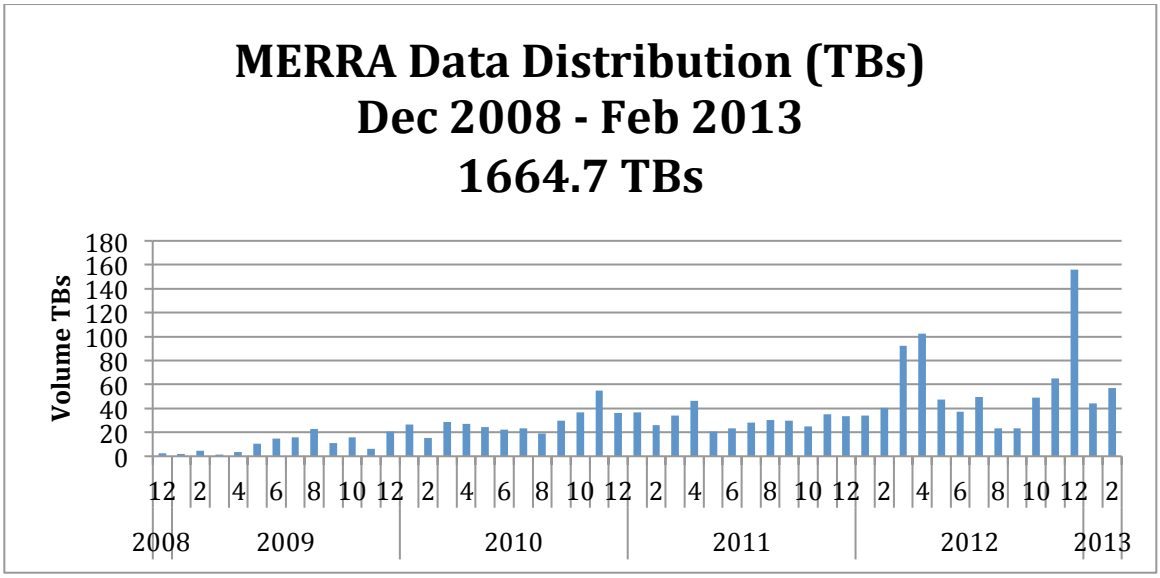


Figure 1. Monthly MERRA output volume accessed by users.

G5.4.1- Clouds and the Earth's Radiant Energy System (CERES)

GEOS-5 data delivery to the CERES IT continued without incident this year. Our GEOS-5.4.1 reprocessing effort for CERES, starting from February 2000, continues to make steady progress. We estimate this reprocessing will be completed in May 2013.

Field Campaign Support

GMAO real time support to field campaigns during 2012 and early 2013 included:

Campaign Name	Date
Two-Column Aerosol Project (TCAP)	June through July 2012
Hurricane and Severe Storm Sentinel (HS3)	August through October 2012
Airborne Tropical Tropopause Experiment (ATTREX)	January through March 2013
Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ)	January through February 2013
Polarimeter Definition Experiment (PODEX)	January through February 2013

Support generally consisted of 5-day forecasts customized specifically for the individual campaigns and made available on the NCCS data portal. GMAO web sites tailored for the field campaigns also provided tools for forecast visualization. For example, the online tool for PODEX and DISCOVER-AQ provided maps showing aerosol optical depth as well as meteorological forecasts over the region of interest.

Publications

Peer-reviewed publications

Albergel, C., W. Dorigo, **R. H. Reichle**, G. Balsamo, P. de Rosnay, J. Munoz-Sabater, L. Isaksen, R. de Jeu, and W. Wagner, 2013: Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. *J. Hydrometeorol. (submitted).*

Bosilovich, M. G., M. Rixen, P. van Oevelen, G. Asrar, G. Compo, K. Onogi, A. Simmons, K. Trenberth, D. Behringer, T. Hossain Bhuiyan, S. Capps, A. Chaudhuri, J. Chen, L. Chen, N. Colasacco-Thumm, M. G. Escobar, C. R. Ferguson, T. Ishibashi, M. L. R. Liberato, J. Meng, **A. Molod**, and co-authors, 2012: Report of the 4th World Climate Research Programme International Conference on Reanalyses. *4th World Climate Research Programme International Conference on Reanalyses, Silver Spring, Maryland,* 7-11 May 2012.

Bosilovich, M., A. Chauduri, and M. Rixen, 2013: Earth system reanalysis: Progress, challenges and opportunities. *B. Am. Meteorol. Soc. (submitted).*

Bosilovich, M., 2013: Regional climate and variability of NASA MERRA and recent reanalyses: U.S. summertime precipitation and temperature. *J. Appl. Meteorol. Clim. (submitted).*

Chao, W., 2012: Correction of excessive precipitation over steep and high mountains in a GCM. *J. Atmos. Sci.*, **69**, 1547-1561. DOI: 10.1175/JAS-D-11-0216.1.

Chao, W. C., 2013: Catastrophe concept-based cumulus parameterization: Correction of systematic errors in the precipitation diurnal cycle over land in a GCM. *J. Atmos. Sci. (submitted).*

Choi, S., Y. Wang, R. J. Salawitch, T. Canty, J. Joiner, T. Zeng, T. P. Kurosu, K. Chance, A. Richter, L. G. Huey, J. Liao, J. A. Neuman, J. B. Nowak, J. E. Dibb, A. J. Weinheimer, G. Diskin, T. B. Ryerson, **A. da Silva**, and co-authors, 2012: Analysis of satellite-derived Arctic tropospheric BrO columns in conjunction with aircraft measurements during ARCTAS and ARCPAC. *Atmos. Chem. Phys.*, **12**, 1255-1285. DOI: 10.5194/acp-12-1255-2012.

Cullather, **R. I**., and **M. Bosilovich**, 2012: The energy budget of the polar atmosphere in MERRA. *J. Climate*, **25**, 5-24. DOI: 10.1175/2011JCLI4138.1.

Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and **M. Bosilovich**, 2012: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Climate*, **25**, 1916-1944. DOI: 10.1175/JCLI-D-11-00004.1.

De Lannoy, G. J. M., R. H. Reichle, and V. R. N. Pauwels, 2013: Global calibration of the GEOS-5 L-band microwave radiative transfer model over land using SMOS observations. *J. Hydrometeorol.,* DOI: 10.1029/2011WR010588 *(in press).*

De Lannoy, **G. J. M., R. H. Reichle**, K. R. Arsenault, P. R. Houser, S. V. Kumar, N. E. C. Verhoest, and V. R. N. Pauwels, 2012: Multi-scale assimilation of AMSR-E snow water equivalent and MODIS snow cover fraction in Northern Colorado. *Water Resour. Res.*, **48**, W01522. DOI: 10.1029/2011WR010588.

Draper, **C. S**., **R. H. Reichle**, **G. J. M. De Lannoy**, and **Q. Liu**, 2012: Assimilation of passive and active microwave soil moisture retrievals. *Geophys. Res. Lett.*, **39**, L04401. DOI: 10.1029/2011GL050655.

Errico, R. M., R Yang**, N. Prive, K.-S. Tai, R. Todling, M. Sienkiewicz,** and **J. Guo,** 2013: Development and validation of observing system simulation experiments at the Global Modeling and Assimilation Office. *Q. J. Meteor. Soc.,* DOI: 10.1002/qj.2027 *(in press).*

Errico, R. M., and **N. Prive**, 2013: An estimate of some analysis error statistics using the GMAO observing system simulation framework. *Q. J. Meteor. Soc. (submitted).*

Feng, X., **M. Bosilovich**, P. Houser, and J.-D. Chern, 201e: Impact of Land Surface Conditions on 2004 North American Monsoon in GCM experiments. *J. Geophys. Res.,* DOI: 10.1029/2012JD018805 *(in press).*

Forman, B. A., **R. H. Reichle**, and C. Derksen, 2013: Estimating passive microwave brightness

temperature over snow-covered land in North America using a land surface model and an artificial neural network. *IEEE Trans. Geosci. Remote Sens. (in press).*

Forman, B. A., **R. H. Reichle**, and M. Rodell, 2012: Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resour. Res.*, **48**, W01507. DOI: 10.1029/2011WR011239.

Forman, B. A, and **R. H. Reichle**, 2013: The spatial scale of model errors and assimilated retrievals in a terrestrial water storage assimilation system. *Water Resour. Res. (submitted).*

Guan, B., D. E. Waliser, J.-L. F. Li, and **A. da Silva**, 2013: Evaluating the impact of orbital sampling on satellite climate model comparisons. *J. Geophys. Res.* DOI: 10.1029/2012JD018590 *(in press).*

Guo, Z., P. Dirmeyer, T. DelSole, and **R. Koster**, 2012: Rebound in atmospheric predictability and the role of the land surface. *J. Climate,* **25**, 4744-4749. DOI: 10.1175/JCLI-D-11-00651.1.

Gupta, P., M. Khan, **A. da Silva**, and F. Patadia, 2013: MODIS aerosol optical depth observations over urban areas in Pakistan: Quantity and quality of the data for air quality monitoring. *Atmos. Pol. Res.*, DOI: 10.5094/APR.2013.005 *(in press)*

Ham, Y.-G., I.-S. Kang, and J.-S. Kug, 2012: Coupled bred vectors in the tropical Pacific and the application to ENSO prediction. *Prog. Oceanogr.,* **105**, 90-101. DOI**:** 10.1016/j.pocean.2012.04.005.

Ham, Y.-G., H.-J. Song, and **R. Todling**, 2013: The non-stationary incremental analysis updates (NIAU) algorithm. *Tellus (submitted).*

Ham, Y.-G., and J.-S. Kug, 2012: How well do current climate models simulate two types of El Niño? *Clim. Dynam.,* **39**, 383-398. DOI: 10.1007/s00382-011-1157-3.

Ham, Y.-G., and **M. Rienecker**, 2012: Flow-dependent empirical singular vector with an ensemble Kalman filter data assimilation for El Niño prediction. *Clim. Dynam.*, **39**, 1727-1738. DOI: 10.1007/s00382-012-1302-7.

Ham, Y.-G., M. M. Rienecker, S. Schubert, J. Marshak, S.-W. Yeh, and S.-C. Yang, 2012: The decadal modulation of coupled bred vectors. *Geophys. Res. Lett.,* **39**, L20712. DOI: 10.1029/2012GL053719.

Ham, Y.-G., J.-S. Kug, and M.-J. Lim, 2012: Rectification feedback of high-frequency atmospheric variability into low-frequency zonal flows in the tropical Pacific. *J. Climate*, **25**, 5088-5101. DOI: 10.1175/JCLI-D-11-00303.1.

Ham, Y.-G., M. M. Rienecker, M. J. Suarez, Y. Vikhliaev, B. Zhao, J. Marshak, G. Vernieres, and **S. D. Schubert**, 2013: Decadal prediction skill in the GEOS-5 forecast system. *Clim. Dynam. (accepted).*

Ham, Y.-G., J.-S. Kug, D. Kim, Y.-H. Kim, and D.-H. Kim, 2012: What controls phase-locking of ENSO to boreal winter in coupled GCMs? *Clim. Dynam.,* DOI: 10.1007/s00382-012-1420-2.

Ham, Y.-G., J.-S. Kug, and J.-Y. Park, 2013: Sea surface temperature in the north tropical Atlantic as a trigger for El-Niño–Southern oscillation events. *Nat. Geosci. (in press).*

Holmes, T. R. H., T. Jackson, **R. Reichle**, and J. Basara, 2012: An assessment of surface soil temperature products from numerical weather prediction models using ground-based measurements. *Water Resour. Res.*, **48**, WO2531. DOI: 10.1029/2011WR010538.

Hong, S.-J., I.-S. Kang, and **Y.-G. Ham**, 2013: Climate response in the tropical pacific associated with Atlantic warming in recent decades. *Asia-Pac. J. Atmos. Sci. (submitted).*

Houborg, R., M. Rodell, B. Li, **R. H. Reichle**, and B. F. Zaitchik, 2012: Drought indicators based on model assimilated GRACE terrestrial water storage observations. *Water Resour. Res.*, **48**, W07525. DOI: 10.1029/2011WR011291.

Koster, R., and **S. P. Mahanama**, 2012: Land-surface controls on hydroclimatic means and variability. *J. Hydrometeorol.,* **13**, 1604-1620. DOI: 10.1175/JHM-D-12-050.1.

Kug, J.-S., **Y.-G. Ham**, J.-Y. Lee, and F.-F. Jin, 2012: Improved simulation of two types of El Niño in CMIP5 models. *Environ. Res. Lett.*, **7**, 034002. DOI: 10.1088/1748-9326/7/3/039502.

Kug, J.-S., and **Y.-G. Ham**, 2012: Indian Ocean feedback to the ENSO transition in a multi-model

ensemble. *J. Climate*, **25**, 6942-6957. DOI: 10.1175/JCLI-D-12-00078.1.

Kumar, S. V., **R. H. Reichle**, K. W. Harrison, C. D. Peters-Lidard, S. Yatheendradas, and J. A. Santanello, 2012: A comparison of methods for a priori bias correction in soil moisture data assimilation. *Water Resour. Res.*, **48**, W03531. DOI: 10.1029/2010WR010261

Langlois, T. J., B. T. Radford, K. P. Van Niel, J. J. Meeuwig, A. F. Pearce, **C. S. G. Rousseaux**, G. A. Kendrick, and E. S. Harvey, 2012: Consistent abundance distributions of marine fishes in an old, climatically buffered, infertile seascape. *Global Ecol. Biogeogr.,* **21**, 886-897. DOI: 10.1111/j.1466-8238.2011.00734.x.

Lahoz, W. A., and **G. J. M. De Lannoy**, 2013: Closing the gaps in our knowledge of the hydrological cycle of the Earth system: Conceptual problems. *Surv. Geophys. (submitted).*

Li, B., M. Rodell, B. F. Zaitchik, **R. H. Reichle**, **R. D. Koster**, and T. M. van Dam, 2012: Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in Western and Central Europe. *J. Hydrol.*, **446**, 103-115. DOI: 10.1016/j.jhydrol.2012.04.035.

Lu, J., and **B. Zhao**, 2012: The role of oceanic feedback in the climate response to doubling CO2. *J. Climate,* **25**, 7544-7563. DOI: 10.1175/JCLI-D-11-00712.1.

McGrath-Spangler, E. L., and A. S. Denning, 2013: Global seasonal variations of midday planetary boundary layer depth from CALIPSO space-borne LIDAR. *J. Geophys. Res. (in press).*

Maggioni, V., **R. H. Reichle**, and E. N. Anagnostou, 2012: The impact of rainfall error characterization on the estimation of soil moisture fields in a land data assimilation system. *J. Hydrometeorol.*, **13,** 1107-1118. DOI: 10.1175/JHM-D-11-0115.1.

Maggioni, V., E. N. Anagnostou, and **R. H. Reichle**, 2012: The impact of land model structural, parameter, and forcing errors on the characterization of soil moisture uncertainty. *Hydrol. Earth Syst. Sc.*, **9**, 2283-2319. DOI: 10.5194/hessd-9-2283-2012.

Maggioni, V., **R. H. Reichle**, and E. N. Anagnostou, 2013: The efficiency of assimilating satellite soil moisture retrievals in a land data assimilation system using different rainfall error models. *J. Hydrometeorol.,* DOI: 10.1175/JHM-D-12-0105.1 *(in press).*

Mahanama, S. P. P., B. Livneh, **R. D. Koster**, D. Lettenmaier, and **R. H. Reichle**, 2012: Soil moisture, snow, and seasonal streamflow forecasts in the United States. *J. Hydrometeorol.*, **13**, 189-203. DOI: 10.1175/JHM-D-11-046.1.

Molod, A., 2013: Constraints on the total water PDF in GCMs from AIRS and a high resolution model. *J. Climate,* DOI: 10.1175/JCLI-D-11-00412.1 *(in press).*

Molod, A. M., M. J. Suarez, and **G. Partyka,** 2013: The impact of limiting ocean roughness on GEOS-5 AGCM tropical cyclone forecasts. *Geophys. Res. Lett. (in press).*

Nearing, G. S., W. T. Crow, K. R. Thorp, M. S. Moran, **R. H. Reichle**, and H. V. Gupta, 2012: Assimilating remote sensing observations of leaf area index and soil moisture for wheat yield estimates: An observing system simulation experiment. *Water Resour. Res.*, **48**, W05525. DOI: 10.1029/2011WR011420.

Orth, R., **R. Koster**, and S. Seneviratne, 2013: Inferring soil moisture memory from runoff observations. *J. Hydrometeorol. (submitted).*

Palmer, P., M. Parrington, J. Lee, **S. Pawson**, and **A. da Silva**, 2013: Quantifying the impact of BOReal forest fires on tropospheric oxidants over the Atlantic using aircraft and satellites (BORTAS) experiment: Design, execution, and science overview. *Atmos. Chem. Phys. (submitted).*

Pauwels, V. R. N., **G. J. De Lannoy**, H.-J. Hendricks Franssen, and H. Vereecken, 2013: Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid kalman filter. *Adv. Water Resour. (submitted).*

Prive, N., **R. M. Errico,** and **K.-S. Tai**, 2013: Validation of forecast skill of the Global Modeling and Assimilation Office observing system simulation experiment. *Q. J. Roy. Meteor. Soc.,* DOI: 10.1002/qj.2029 *(in press).*

Privé, N. C., R. Errico, and K.-S. Tai, 2013: The impact of increased frequency of rawinsonde observations on forecast skill investigated with an observing system simulation experiment. *Mon. Weather Rev. (submitted).*

Privé, N. C., R. M. Errico, and K.-S. Tai, 2013: The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *J. Geophys. Res. (submitted).*

Rabier, F., S. Cohn, P. Cocquerez, A. Hertzog, L. Avallone, T. Deshler, J. Haase, T. Hock, A. Doerenbecher, J. Wang, V. Guidard, J.-N. Thepaut, R. Langland, A. Tangborn, G. Balsamo, E. Brun, D. Parsons, J. Bordereau, C. Cardinali, F. Danis, J.-P. Escarnot, N. Fourrie, **R. Gelaro**, and co-authors, 2013: The Concordiasi field experiment over Antarctica: First results from innovative atmospheric measurements. *Bull. Am. Meteorol. Soc.*, DOI: 10.1175/BAMS-D-12-00005.1 *(in press).*

Reichle, R. H., G. J. M. De Lannoy, B. A. Forman, **C. S. Draper, and Q. Liu**, 2013: Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surv. Geophys. (in press).*

Roberts, J. B., F. R. Robertson, C. A. Clayson, and **M. G. Bosilovich**, 2012: Characterization of turbulent latent and sensible heat flux exchange between the atmosphere and ocean in MERRA. *J. Climate*, **25**, 821-838. DOI: 10.1175/JCLI-D-11-00029.1.

Rousseaux, C. S., and **W. W. Gregg,** 2012: Climate variability and phytoplankton composition in the Pacific Ocean. *J. Geophys. Res.,* **117**, C10006. DOI: 10.1029/2012JC008083.

Rousseaux, C. S., T. Hirata, and **W. W. Gregg,** 2012**:** Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model. *Biogeosciences (submitted).*

Scarino, B., P. Minnis, R. Palikonda, **R. Reichle**, D. Morstad, C. Yost, B. Shan, and **Q. Liu**, 2013: Retrieving surface skin temperature for NWP applications from global geostationary satellite data. *Remote Sensing*, **5**, 342-366. DOI: 10.3390/rs5010342.

Seneviratne, S.-I., and **R. D. Koster**, 2012: A revised framework for analyzing soil moisture memory in climate data: Derivation and interpretation. *J. Hydrometeorol.*, **13**, 1525-755X. DOI: 10.1175/JHM-D-11-044.1.

Strode, S., **L. Ott, S. Pawson**, and T. W. Bowyer, 2012: Emission and transport of Cesium-137 from boreal biomass burning the summer of 2010. *J. Geophys. Res.,* **117**, D09302. DOI: 10.1029/2011JD017382.

Tangborn, A., L. L. Strow, B. Imbiriba, **L. Ott**, and **S. Pawson**, 2012: Evaluation of a new middle-lower tropospheric CO_2 product using data assimilation. *Atmos. Chem. Phys.*, **12**, 26685-26717. DOI: 10.5194/acpd-12-26685-2012.

Todling, R., 2012: Insights on observation impact measures. *Mon. Weather Rev. (submitted).*

Vinukollu, R. K., J. Sheffield, E. F. Wood, **M. Bosilovich,** and **D. Mocko**, 2012: Multimodel analysis of energy and water fluxes: Intercomparisons between operational analyses, a land surface model, and remote sensing. *J. Hydrometeorol.*, **13**, 3-26. DOI: 10.1175/2011JHM1372.1.

Wei, J., P. A. Dirmeyer, **M. Bosilovich**, and R. Wu, 2012: Water vapor sources for Yangtze River valley rainfall: Climatology, variability, and implications for rainfall forecasting. *J. Geophys. Res.*, **117**, DO5126. DOI: 10.1029/2011JD016902.

Wernberg, T., D. A Smale, F. Tuya, M. S. Thomsen, T. J. Langlois, T. de Bettignies, S. Bennett, and **C. S. Rousseaux**, 2013: An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change,* DOI: 10.1038/nclimate/627 *(in press).*

Wu, M.-L., O. Reale**,** and **S. Schubert**, 2013: A characterization of African Easterly Waves on 2.5-6 day and 6-9 day time scales. *J. Climate (submitted)*

Xue, Y., M. Balmaseda, T. Boyer, N. Ferry, S. Good, I. Ishikawa, A. Kumar**, M. Rienecker**, A. Rosati, and Y. Yin, 2012: A comparative analysis of upper ocean heat content variability from an ensemble of operational ocean reanalyses. *J. Climate,* **25**, 6905-6929. DOI: 10.1175/JCLI-D-11-00542.1.

Xue, Y., M. Balmaseda, T. Boyer, N. Ferry, S. Good, I. Ishikawa, **M. Rienecker,** T. Rosati, Y. Yin, and A. Kumar, 2012: A comparative analysis of upper ocean heat content variability from ensemble operations ocean analyses. *U.S CLIVAR Variations*, **9 (1),** 7-10.

Yeh, S.-W., **Y.-G. Ham**, and J.-Y. Lee, 2012: Changes in the tropical Pacific SST trend from CMIP3 to CMIP5 and its implication of ENSO. *J. Climate*, **25**, 7764-7771. DOI: 10.1175/JCLI-D-12-00304.1.

Yi, Y., J. S. Kimball, L. A. Jones, **R. H. Reichle**, R. R. Nemani, and H. A. Margolis, 2013: Recent climate and fire disturbance impacts on boreal and arctic ecosystem productivity estimated using a satellite-based terrestrial carbon flux model. *J. Geophys. Res. (submitted).*

Technical Memoranda

Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and **A. Eichmann**, 2012: The GEOS-5 Atmospheric general circulation model: Mean climate and development from MERRA to Fortuna. *Technical Report Series on Global Modeling and Data Assimilation*, Vol. **28**, NASA/TM-2012-104606/28.

Rienecker, M., D. Dee, J. Woollen, G. Compo, K. Onogi, **R. Gelaro, M. Bosilovich, A. da Silva, S. Pawson, S. Schubert, M. Suarez**, D. Barker, H. Kamahori, R. Kistler, and S. Saha, 2012: Atmospheric reanalyses – Recent progress and prospects for the future. A report from a technical workshop, April 2010. *Technical Report Series on Global Modeling and Data Assimilation*, Vol. **29**, NASA/TM-2012-104606/29.

Vernieres, G., **M. M. Rienecker**, **R. Kovach**, and **C. L. Keppenne**, 2012: The GEOS-iODAS: Description and evaluation. NASA Technical Report Series on Global Modeling and Data Assimilation, Vol. **30**. NASA/TM-2012-104606.

Book Chapters

Bosilovich, M., J. Kennedy, D. Dee, A. O'Neill, 2013: On the reprocessing and reanalysis of observations for climate. *IEEE Monographs (submitted).*

Schubert, S. D., and Y.-K. Lim, 2013: "Climate variability and weather extremes: model-simulated and historical data" in *Hydrologic Extremes in a changing climate – detection, analysis & uncertainty*. Ed. A. AghaKouchak, D. Easterling, K. Hsu, **S. Schubert**, and S. Sorooshian, *Springer*. ISBN: 978-9400744783.