NASA/NOAA's Global Land Data Assimilation System (GLDAS): Recent Results and Future Plans

Matthew Rodell

with contributions from

Paul Houser, Christa Peters-Lidard, Hiroko Kato, Sujay Kumar, Jon Gottschalck, Ken Mitchell, and Jesse Meng

Hydrological Sciences Branch
NASA Goddard Space Flight Center
Greenbelt, MD 20771, USA

Abstract

The goal of the Global Land Data Assimilation System (GLDAS) is to ingest satellite- and ground-based observational data products, using advanced land surface modeling and data assimilation techniques, in order to generate optimal fields of land surface states and fluxes. The software, which has been streamlined and parallelized by the Land Information System (LIS) project, drives multiple, offline (not coupled to the atmosphere) land surface models, integrates a huge quantity of observation based data, executes globally at high resolutions (2.5° to 1 km), and is capable of producing results in near-real time. A vegetation-based "tiling" approach is used to simulate sub-grid scale variability, with a 1 km global vegetation dataset as its basis. Soil and elevation parameters are based on high resolution global datasets. Observation-based precipitation and downward radiation products, as well as output fields from the best available global coupled atmospheric data assimilation systems, are employed to force the models. Intercomparison and validation of these products is being performed with the aim of identifying an optimal forcing scheme. Data assimilation techniques for incorporating satellite based hydrological products, including MODIS snow cover and leaf area index, are in various stages of development and implementation. The high-quality, global land surface fields provided by GLDAS support several current and proposed weather and climate prediction, water resources applications, and water cycle investigations. GLDAS is also a test bed for innovative modeling and assimilation capabilities. The project has resulted in a massive archive of modeled and observed, global, surface meteorological data, parameter maps, and output.

1. Introduction

Knowledge of land surface water, energy, and carbon conditions are of critical importance for improving weather and climate prediction and supporting real-world applications such as agricultural production, water resource management, and flood, weather and climate prediction. Land data assimilation systems use sophisticated land surface models (LSMs) to ingest satellite and ground-based observations, as parameters, forcing, and data for assimilation, in order to produce enhanced fields of land surface states and fluxes. Our knowledge of physical processes, which is embodied in the models, fills the observational gaps and merges data from disparate measurement systems. The multi-institution North American Land Data Assimilation System (NLDAS) project was the first to embrace this concept [Mitchell et al., 2004]. Its success led to the development of GLDAS [Rodell et al., 2004a], beginning in 2000, through the joint effort of scientists at NASA's Goddard Space Flight Center and NOAA's National Centers for Environmental Prediction (NCEP). The hypothesis that motivated GLDAS was that land surface simulation, observation, and analysis methods were sufficiently advanced to accurately determine global land surface energy and moisture stores for the initialization of prediction systems and to address land surface management issues. The high resolution fields produced by GLDAS support this hypothesis, and several current and proposed weather and climate prediction, water resources applications, and water cycle investigations rely on GLDAS output.

2. Methods

GLDAS enables multiple models to be driven using various combinations of parameters, forcing data, model resolutions, and advanced capabilities such as an elevation correction, all selected through a single, simple user interface. The software has been streamlined and parallelized by the Land Information System (LIS) sister project, which has achieved its goal of enabling global, 1 km resolution land surface model simulations to be executed on a Beowulf cluster (a single model day is typically completed in about six hours and requires on the order of a terabyte of storage capacity for output). Drivers have been installed for several LSMs including Mosaic [Koster and Suarez, 1996], the Community Land Model (CLM) [Dai et al., 2003], and Noah [Chen et al., 1996; Ek et al., 2003]. GLDAS runs globally and on user defined domains at resolutions ranging from 2.5° to 1 km. A "tiling" approach, whereby the model is run on multiple land cover specific "tiles" within each grid square, is used to simulate subgrid scale variability [Koster and Suarez, 1992], based on a 1 km global Advanced Very High Resolution Radiometer (AVHRR) derived vegetation dataset [Hansen et al., 2000]. Global 5' resolution soil property maps [Reynolds et al., 2000] also are used to parameterize the models.

Atmospheric data assimilation products from the operational weather forecast systems of NASA/GSFC's Global Modeling and Assimilation Office (GMAO), NCEP, and the European Centre for Medium-Range Weather Forecasts (ECMWF) provide the baseline forcing data. Observation-based precipitation (including near-real time satellite based products and a spatially and temporally disaggregated adaptation of the NOAA Climate Prediction Center's (CPC) Merged Analysis of Precipitation (CMAP), which integrates gauge and satellite observations [Xie and Arkin, 1997]) and solar radiation (derived using algorithms and data from the U.S. Air Force Weather Agency's (AFWA) Agricultural Meteorological modeling system (AGRMET), based mainly on satellite observations) optionally replace the baseline forcing fields. An elevation adjustment is applied to standardize the forcing data. GLDAS also utilizes satellite derived leaf area index (LAI) from both the AVHRR and MODIS sensors in order to further improve specification of surface characteristics. Various other modeling options also can be selected by the user, including the model and output timesteps, initialization specifications, and subgrid tiling constraints.

The GLDAS based LIS software has following characteristics: (1) High-resolution simulation capability through several independent community land surface models and utilizing distributed data management; (2) Web-based accessibility of tools for data mining, numerical modeling, and visualization; (3) Software portability to both small and large systems; and (4) adherence to Earth System Modeling Framework (ESMF) standards to facilitate coupling. Furthermore, LIS is thoroughly documented, including instructions for software development.

3. Recent Results

Intercomparison and validation of the forcing fields used by GLDAS has been a major thrust for the project. Precipitation is arguably the most important meteorological forcing variable in land surface modeling as it is the sole input to the water budget. We have performed a comparison of six global and two U.S precipitation products for the period from March 2002 – February 2003 [Gottschalck et al., 2004a]. Evaluation of total accumulated precipitation for the continental United States illustrated that CMAP had the closest agreement with a CPC rain gauge dataset for all seasons except winter. ECMWF performed the best of the atmospheric analyses. The satellite-only products suffered from a few deficiencies – most notably an overestimation of summertime precipitation in the central United States (Figure 1). The CMAP precipitation, which was spatially and temporally disaggregated using NCEP's GDAS precipitation product, was the most closely correlated with daily rain gauge data in spring, fall, and winter, while the satellite only estimates performed best in summer. We also analyzed GLDAS/Mosaic simulations 14 months in duration which were forced with the various products. The sensitivity of land surface states was shown to be substantial. For volumetric soil water content, the span of differences between the runs ranged from 30-45% and 20-30% of the typical total range of volumetric soil water content (0-0.50 m^3/m^3) for the spring/summer and fall/winter months respectively. The soil temperature spread between GLDAS runs was considerable and varied by up to +/- 3.0°C among the simulations.

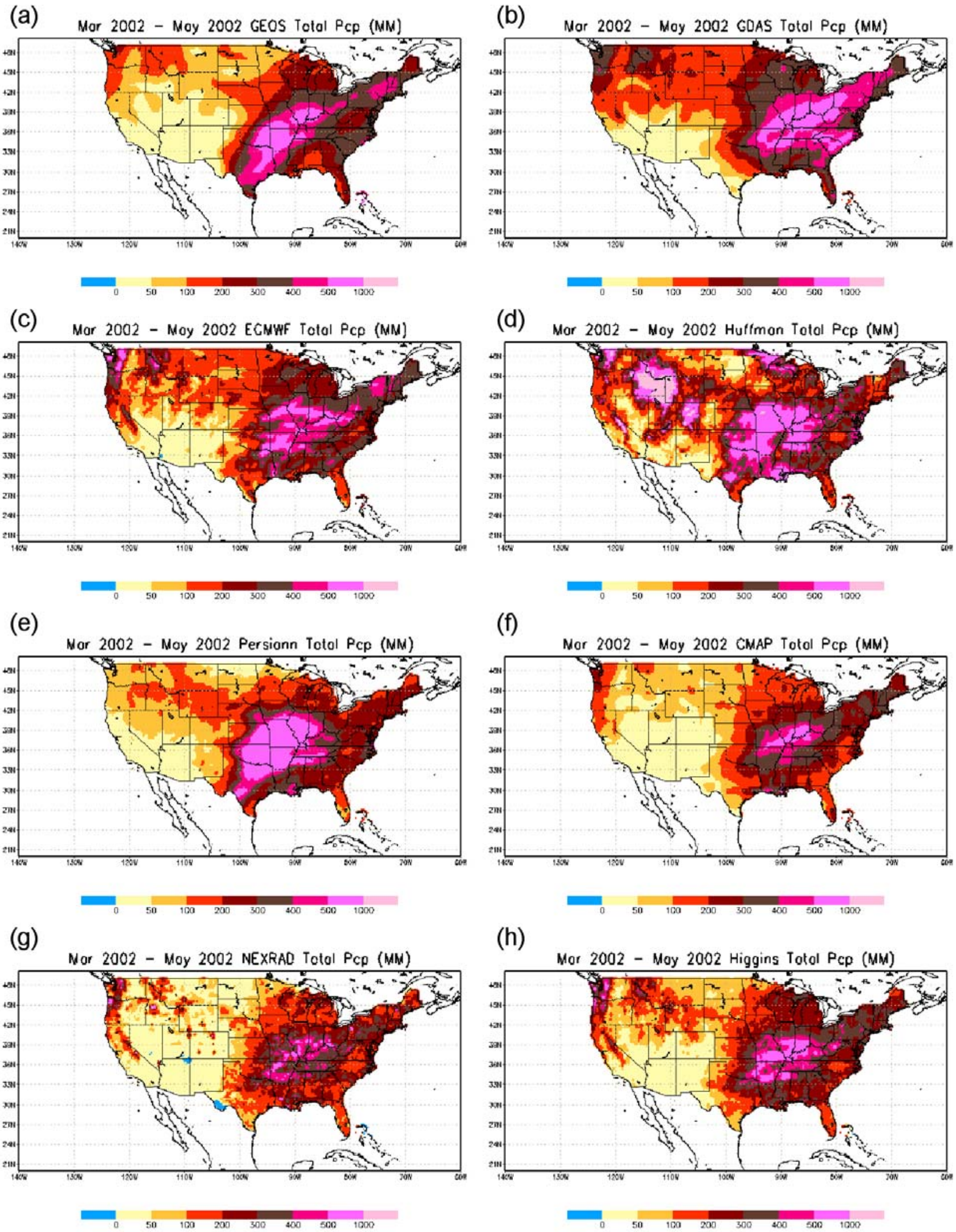

Figure 1 Total precipitation [mm] for March, April, and May 2002 over CONUS for (a) GEOS, (b) GDAS, (c) ECMWF, (d) Huffman, (e) Persiann, (f) CMAP, (g) NEXRAD, and (h) Higgins gauge.

We have likewise evaluated the available downward radiation products [Meng et al., 2004]. In terms of zonal-average, downward shortwave fluxes over the land surface, AGRMET and GEOS exhibited very similar patterns, except in the tropics. The differences in mid- and high-latitudes were within 10%. On the other hand, GDAS was about 10% lower than AGRMET in most of the winter hemisphere, and was about 20-25% lower in most of the summer hemisphere. Monthly mean surface downward shortwave fluxes from the models were compared with the ground observations from six U.S. Surface Radiation (SURFRAD) sites (Figure 2). In general, AGRMET had an approximately 10% high bias in both summer and winter. GDAS generally performed well, though GEOS and GDAS appeared to overestimate winter surface shortwave fluxes, suggesting that these two models underestimate cloud cover. ECMWF tended to overestimate summer incoming shortwave radiation in the morning and underestimate it in midday and afternoon.

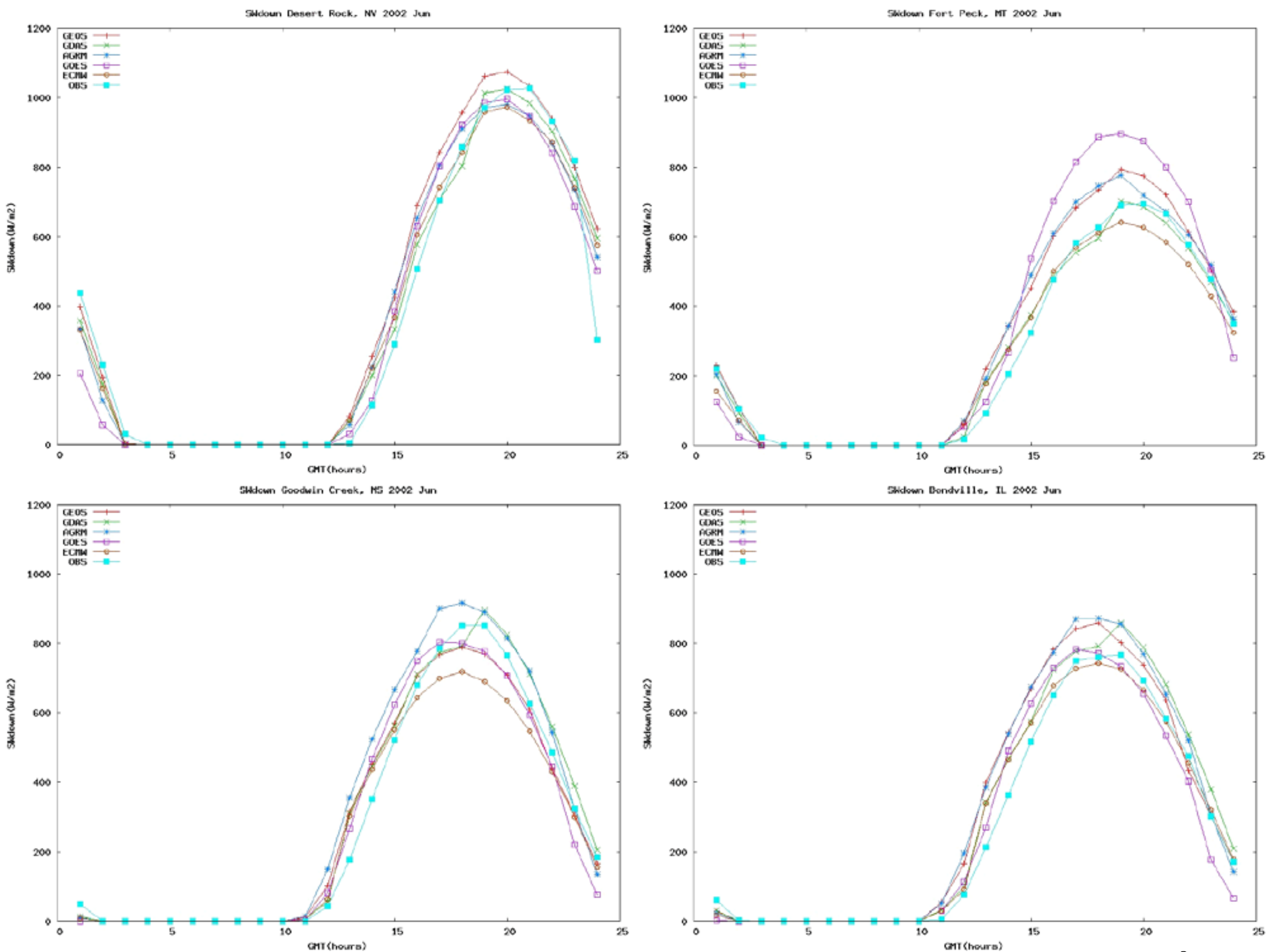

Figure 2 Comparison of the June 2002 mean diurnal cycle of downward shortwave radiation [W/m^2] from three atmospheric analyses, two observation based products and surface radiation network measurements at four sites in the U.S.A.

A scheme for incorporating leaf area index (LAI) observations from the Advanced Very High Resolution Spectroradiometer (AVHRR) was designed and tested in GLDAS [Gottschalck et al., 2004b]. Improvement in simulated land surface temperature was demonstrated, where geostationary infrared derived surface temperature retrievals over the CONUS were used for verification (Figure 3). However, it was determined that an algorithm that scales the observed LAI to the modeled LAI using the spatially and temporally varying ranges of both would be preferable.

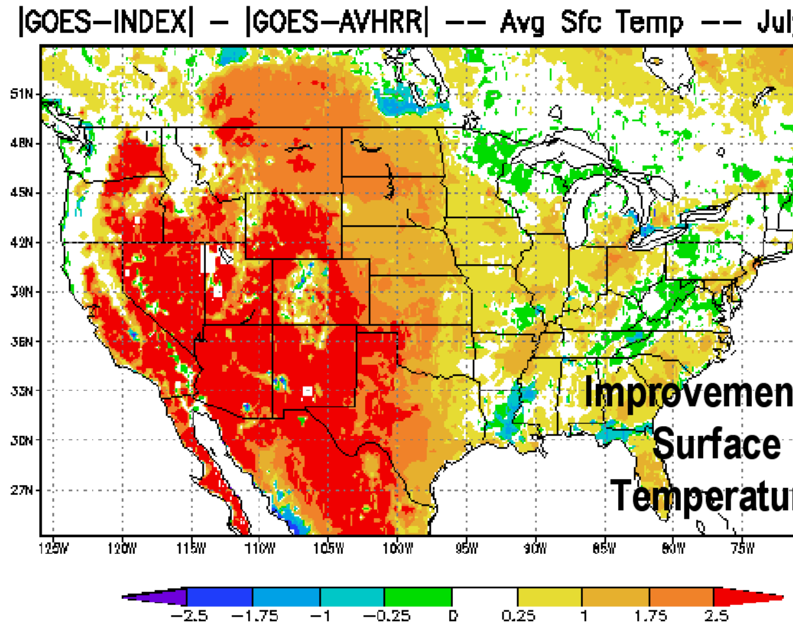

Figure 3 Improvement, compared with GOES observation, in surface temperature (° C) when satellite derived LAI is used, July 2001

An updating routine which makes use of Moderate Resolution Imaging Spectroradiometer (MODIS) derived snow cover fields has been installed and tested in GLDAS [Rodell et al., 2004b]. Simulated snow water equivalent is adjusted when and where the model and MODIS observation differ, following an internal accounting of the observation quality, by either removing the simulated snow or adding a thin layer. The scheme was tested in a global simulation of GLDAS/Mosaic which spanned January to April 2003. Output from this simulation was compared to that from a control (not updated) simulation, and both were assessed using an operational snow cover product and data from ground based observation networks over the continental U.S. In general, output from the updated simulation displayed more accurate snow coverage and compared more favorably with in situ snow water equivalent time series than output from the control. Both simulations had serious deficiencies on occasion and in certain areas when and where the precipitation and/or surface air temperature forcing inputs were unrealistic, particularly in mountainous regions. Additional algorithms have been developed for the assimilation of soil moisture [Zhan et al., 2002] and surface temperature [Radakovich et al., 2001].

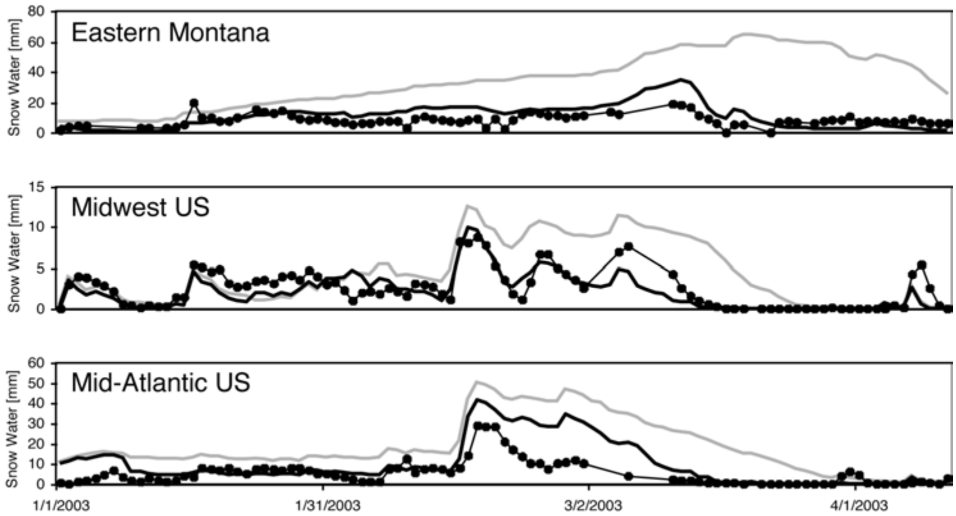

Figure 4 Regional average time series of sbow water equivalent [mm] from CO-OP ground based observations (black dots) and GLDAS/Mosaic control (gray line) and assimilation run (black line) outputs).

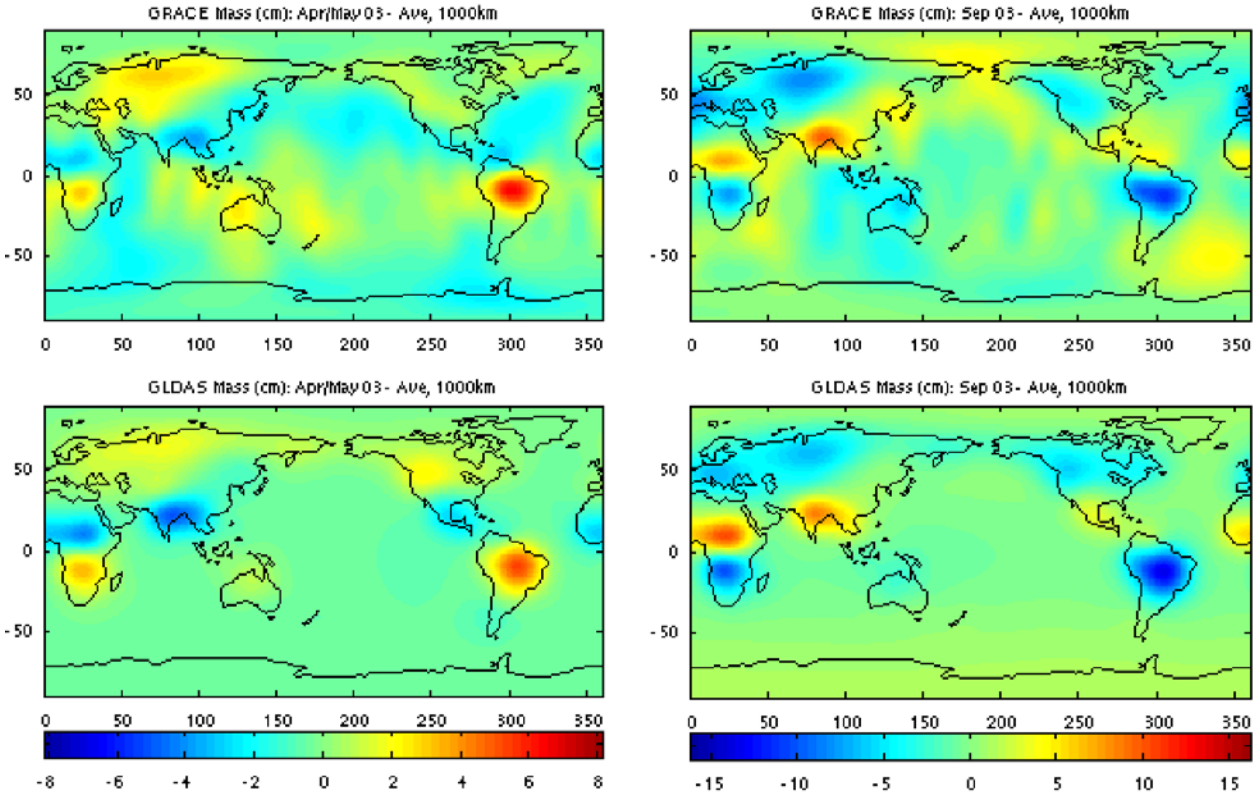

Figure 5 Comparison of GRACE derived terrestrial water storage anomalies [cm] (top) with GLDAS soil moisture plus snow anomalies [cm] (bottom} for April/May (left) and September (right) 2003.

Validation of the GLDAS output products is ongoing. We are obtaining data for validation from U.S. and global in situ observation networks. The latter are being provided through the GEWEX sponsored Coordinated Enhanced Observing Period (CEOP) initiative. Further, we have been comparing simulated soil and snow water storage with terrestrial water storage anomaly maps and river basin time series now being derived from Gravity Recover and Climate Experiment (GRACE) satellite gravity observations [Rodell et al., 2004c]. Figure 5 shows a comparison of GRACE terrestrial water storage anomalies with GLDAS/Noah soil moisture plus snow water storage anomalies for April/May and September 2003. The GLDAS fields have been smoothed using a 1000 km radius averaging function to match the GRACE retrieval. The agreement is generally good, and differences likely reflect the lack of a groundwater component in GLDAS as well as uncertainty in the model forcing and physics.

4. Future Plans

GLDAS output surface state fields are being tested in weather and climate model initialization studies at NOAA/NCEP and NASA/GMAO. GLDAS soil moisture fields already have been shown to improve the predictability of precipitation when used to initialize the GMAO seasonal prediction model [Koster et al., 2004]. The LIS software is currently being installed at NCEP for further initialization testing, and we hope to continue to transfer technology to that and other operational centers as the system evolves. Additionally, GLDAS soil moisture and snow fields are being tested as input to the water management decision support system at the U.S. Bureau of Reclamation. Joint projects with several other U.S. agencies are now being proposed to utilize the LIS software and GLDAS data archive for improving water resources and natural hazards forecasts.

Near term development will focus on installing and testing the interactions and effects of mature assimilation capabilities. Sensitivity analyses and error assessments will be performed with the objective of characterizing the value of multivariate data assimilation and the uncertainty in the fields that result. The output fields and associated errors will contribute to improved understanding and prediction of the energy and water cycles. The enhanced LIS software will remain in the public domain, making it an invaluable resource for the Earth Science community. Specific objectives are:

- Implementing mature assimilation capabilities within LIS, including AMSR-E soil moisture assimilation, MODIS snow cover assimilation, and multiple sensor surface temperature assimilation,
- Incorporating new data products, including root zone depth maps and others as they are developed,
- Utilizing a runoff routing scheme and coupled boundary layer model (being explored under separate projects),
- Incorporating subgrid precipitation variability (developed under a separate project),
- Assessment of the impacts of each new dataset and capability,
- Production of long term GLDAS analyses and associated uncertainty assessments, using optimal system configurations with each of the linked LSMs, and
- Data distribution through a public, web-based interface.

5. References

Chen, F., and Coauthors, Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res*., **101 (D3),** 7251-7268, 1996.

Dai ,Y., and 12 Coauthors, The Common Land Model (CLM), *Bull. Amer. Meteor. Soc.*, **84**, 1013-1023, 2003.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, *J. Geophys. Res*., **108(D22),** 8851, doi:10.1029/2002JD003296, 2003.

Gottschalck, J., J. Meng, M. Rodell, P. Houser, U. Jambor, and K. Arsenault, Analysis of multiple precipitation products and preliminary assessment of their impact on Global Land Data Assimilation System (GLDAS) land surface states, submitted to *J. Hydromet*., 2004a.

Gottschalck, J., P. Houser, M. Rodell, U. Jambor, J. Meng, K. Arsenault, and B. Cosgrove, Impact of remotely sensed leaf area index as part of the global land data assimilation system, in preparation, 2004b.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg, Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, **21**, 1331-1364, 2000.

Koster, R. D., and M. J. Suarez, Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *J. Geophys. Res*., **97**, 2697-2715, 1992.

Koster, R. D., and M. J. Suarez, Energy and Water Balance Calculations in the Mosaic LSM. *NASA Technical Memorandum 104606*, 9, 76 pp, 1996.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, M. Kistler, A. Berg, R. Reichle, M. Rodell, and J. Famiglietti, Realistic initialization of land surface states: impacts on subseasonal forecast skill, *J. Hydromet*., in press, 2004.

Meng, C. J., K. Mitchell, P. R. Houser, M. Rodell, G. Gayno, U. Jambor, J. Gottschalck, and K. Arsenault, Assessment of global land surface solar radiation budget within the Global Land Data Assimilation System (GLDAS), in preparation, 2004.

Mitchell, K., and 21 coauthors, The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res.*, **109**, D07S90, doi: 10.1029/2003JD003823, 2004.

Radakovich, J. D., P. R. Houser, A. da Silva, and M. G. Bosilovich, Results from global land-surface data assimilation methods, *AMS 5th Symposium on Integrated Observing Systems*, Albuquerque, NM, 14-19 January, 132-134, 2001.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls, Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions, *Wat. Resour. Res*., **36**, 3653-3662, 2000.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, *Bull. Amer. Meteor. Soc*., **85** (3), 381–394, 2004a.

Rodell, M., and P. R. Houser, Updating a land surface model with MODIS derived snow cover, J. *Hydromet*., in press, 2004b.

Rodell, M., J. S. Famiglietti, J. Chen, S. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson, Basin scale estimates of evapotranspiration using GRACE and other observations, *Geophys. Res. Lett*., **31**, L20504, doi:10.1029/2004GL020873, 2004.

Xie, P., and P. Arkin, Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates and Numerical Model Outputs, *Bull. Amer. Meteor. Soc*., **78**, 2539-2558, 1997.

Zhan, X., J. K. Entin, P. R. Houser, J. P. Walker, and R. H. Reichle, Application of Kalman filtering for soil moisture data validation with NASA's Land Data Assimilation System, American Geophysical Union, Spring Meeting, *EOS Trans. AGU*, **83**, 2002.