# Analysis of Multiple Precipitation Products and Preliminary Assessment of Their Impact on Global Land Data Assimilation System Land Surface States

JON GOTTSCHALCK AND JESSE MENG

*University of Maryland, Baltimore County, Goddard Earth Sciences and Technology Center (GEST), Baltimore, and Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland*

MATT RODELL AND PAUL HOUSER

*Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland*

(Manuscript received 21 June 2004, in final form 31 January 2005)

ABSTRACT

Precipitation is arguably the most important meteorological forcing variable in land surface modeling. Many types of precipitation datasets exist (with various pros and cons) and include those from atmospheric data assimilation systems, satellites, rain gauges, ground radar, and merged products. These datasets are being evaluated in order to choose the most suitable precipitation forcing for real-time and retrospective simulations of the Global Land Data Assimilation System (GLDAS). This paper first presents results of a comparison for the period from March 2002 to February 2003. Later, GLDAS simulations 14 months in duration are analyzed to diagnose impacts on GLDAS land surface states when using the Mosaic land surface model (LSM).

A comparison of seasonal total precipitation for the continental United States (CONUS) illustrates that the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) has the closest agreement with a CPC rain gauge dataset for all seasons except winter. The European Centre for Medium-Range Weather Forecasts (ECMWF) model performs the best of the modeling systems. The satellite-only products [the Tropical Rainfall Measuring Mission (TRMM) Real-time Multi-satellite Precipitation Analysis and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)] suffer from a few deficiencies—most notably an overestimation of summertime precipitation in the central United States (200–400 mm). CMAP is the most closely correlated with daily rain gauge data for the spring, fall, and winter seasons, while the satellite-only estimates perform best in summer. GLDAS land surface states are sensitive to different precipitation forcing where percent differences in volumetric soil water content (SWC) between simulations ranged from −75% to +100%. The percent differences in SWC are generally 25%–75% less than the percent precipitation differences, indicating that GLDAS and specifically the Mosaic LSM act to generally "damp" precipitation differences. Areas where the percent changes are equivalent to the percent precipitation changes, however, are evident. Soil temperature spread between GLDAS runs was considerable and ranged up to ±3.0 K with the largest impact in the western United States.

## 1. Introduction

Over the past three decades, it has become clear that the land surface exerts a significant impact on the atmospheric boundary layer via fluxes of momentum, energy, and water and therefore impacts weather and climate (Charney 1975; Charney et al. 1977; Shukla and Mintz 1982; Sud and Smith 1985; Meehl and Washington 1988). Current weather and climate forecast models can benefit from more accurate land surface states via initialization of land surface boundary conditions (Koster and Suarez 2003). The evolution of key land surface states such as skin temperature, soil moisture, and snow that determine the critical fluxes above are largely dictated by precipitation. The amount of water and/or ice in the soil impacts the energy cycle by modulating the partitioning of energy at the land surface between sensible and latent heat and also can affect the carbon cycle through control of transpiration.

*Corresponding author address:* Jon Gottschalck, Climate Prediction Center, NOAA/National Centers for Environmental Prediction, 5200 Auth Road, Camp Springs, MD 20746.
E-mail: Jon.Gottschalck@noaa.gov

There have been many efforts to produce accurate precipitation estimates for use as part of land surface process studies and for weather and climate model predictions. These include rain-gauge-only estimates, mainly model-based estimates as part of atmospheric data assimilation systems, ground-based radar estimates, satellite-only-based estimates using data from both infrared and microwave retrievals, and merged rain gauge and satellite estimates.

All of these estimates have pros and cons (Barrett and Martin 1981; Arkin and Ardanuy 1989). Gauge estimates are typically the most accurate available and are generally taken to be the "truth" as they provide a direct measure of precipitation at the surface and do not rely on model parameterizations or remote sensing methods. Rain gauges, however, can feature underestimates due to wind and evaporation as well as instrument and human errors (Neff 1977; Sevruk 1982; Legates and Willmott 1990). Model-based precipitation estimates represent synoptic-scale precipitation processes well but often err under convective regimes at the mesoscale (Sperber and Palmer 1995). In addition, many of the model-based products provide estimates at a coarser resolution than observation-based estimates as a result of computational constraints. Ground-based radar estimates provide excellent spatial and temporal resolution but suffer from error associated with elevation angle, ground clutter, virga (precipitation evaporating before reaching the ground), and anomalous propagation (Fulton et al. 1998). Satellite estimates based on infrared retrievals provide excellent temporal coverage but are dependent on cloud-top temperatures and can very often mistake high-level (cold) cloud tops for precipitation (Griffith et al. 1981; Wylie 1979; Arkin and Meisner 1987; Arkin and Xie 1994; Arkin et al. 1994). Microwave precipitation estimates are more physically based and can provide more accurate instantaneous measures of precipitation but are hampered by poor temporal sampling (mounted on polar-orbiting satellites) and complexities in the retrieval due to cloud microphysics and land surface characteristics (Wilheit et al. 1991; Spencer 1993).

Xie and Arkin (1995) compared infrared and microwave satellite estimates with gauge observations and found good spatial agreement during the warm season over the tropical Pacific Ocean but poor continental results during the cold season, especially those based on IR retrievals. Kondragunta and Gruber (1997) showed that merged satellite/gauge products were more realistic in the depiction of the annual and interannual cycles as compared to model-based assimilation systems. Janowiak et al. (1998) compared the Global Precipitation Climatology Project (GPCP; Huffman et al. 1997) merged gauge and satellite precipitation estimates with the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis precipitation product and found comparable large-scale features but substantial regional differences. Moreover, Gruber et al. (2000) compared the GPCP and the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997; Xie et al. 2003) merged gauge and satellite precipitation datasets and found that although the spatial and temporal correlations were high, significant differences did exist as a result of slightly different input data. The Third Precipitation Intercomparison Project (PIP-3; Adler et al. 2001) analyzed many products and found that the model-based precipitation estimates are poor in the Tropics but comparable over midlatitude continental areas.

The Global Land Data Assimilation System (GLDAS) is a global, high-resolution, offline (uncoupled to the atmosphere) terrestrial modeling system that incorporates satellite and ground-based observations in order to produce optimal fields of land surface states and fluxes in near–real time (Rodell et al. 2004). More details about the GLDAS system are provided in section 2. Because precipitation is one, if not the most critical, atmospheric forcing variable, a comprehensive comparison is being conducted to determine the best precipitation forcing for GLDAS real-time simulations.

Many of the precipitation-related studies cited above have evaluated datasets at coarse spatial and temporal resolution and within coupled global modeling systems. Since GLDAS runs in near–real time and at high global resolution (1/4°), it is important to analyze products of higher temporal and spatial resolution. This paper is unique in that it aims to evaluate multiple higher-resolution precipitation datasets (in both time and space) and conduct land surface model simulations offline in order to evaluate the impact on land surface states. The land modeling section of this paper focuses on the impacts and range of sensitivity for GLDAS land surface states when using different precipitation forcing as opposed to validation of the states themselves with ground truth measurements.

This paper presents the current state of work in the analysis of precipitation datasets available to GLDAS. Section 2 describes the precipitation datasets, outlines the GLDAS framework including descriptions of the Mosaic land surface model (LSM), and identifies the types of experiments that are used to diagnose the impacts on GLDAS land surface states. Section 3 provides results in two forms: 1) an offline comparison of precipitation datasets for the continental United States (CONUS) and 2) analysis of the sensitivity of land surface states to different precipitation datasets. Section 4 discusses the results and implications from this study and outlines future plans.