Integrating Gridded NASA Hydrological Data into CUAHSI HIS

Hualan Rui[1,3], William Teng[1,4], Bruce Vollmer[1], David M. Mocko[2,6], Hiroko K. Beaudoing[2,5], Tim Whiteaker[7], David Valentine[8], David Maidment[7], Richard Hooper[9]

[1]Goddard Earth Sciences Data and Information Services Center, Goddard Space Flight Center (GSFC), NASA, Greenbelt, MD 20771, USA
[2]Hydrological Sciences Branch, Goddard Space Flight Center, NASA, Greenbelt, MD 20771, USA
[3]ADNET Systems, Inc., 164 Rollins Avenue, Rockville, MD 20852, USA
[4]Wyle Information Systems, Inc., 1651 Old Meadow Road, McLean, VA 22102, USA
[5]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20742, USA
[6]SAIC, 4600 Powder Mill Road, Suite 400, Beltsville, MD, 20705, USA
[7]University of Texas, Austin, TX, 78712, USA
[8]Universityof California San Diego, CA, 92093, USA
[9]CUAHSI, 196 Boston Avenue, Suite 2100, Medford, MA 02155, USA

CUAHSI June 22 - 24, 2011

http://disc.gsfc.nasa.gov/hydrology
Email: Hualan.Rui@nasa.gov

Hydrology Data and Information Services Center (HDISC)
NASA Goddard Earth Sciences (GES)
Data and Information Services Center (DISC)
Help Desk: help-disc@listserv.gsfc.nasa.gov

Introduction

The amount of hydrological data available from NASA remote sensing and modeling systems is vast and ever-increasing; but, one challenge persists: increasing the usefulness of these data for, and thus their use by, end user communities. The Hydrology Data and Information Services Center (HDISC), part of the Goddard Earth Sciences DISC, has continually worked to better understand the hydrological data needs of different end users, to thus better able to bridge the gap between NASA data and end user communities. One effective strategy is integrating the data into end user community tools and environments. There is an ongoing collaborative effort between NASA HDISC, NASA Hydrological Sciences Branch, and CUAHSI to integrate NASA gridded hydrology data into the CUAHSI Hydrologic Information System (HIS).

Hydrological Data at HDISC NASA

- The goal of a land data assimilation system (LDAS): generate optimal fields of land surface states and fluxes and, thereby, facilitate hydrology and climate modeling, research, and forecast.
- NLDAS: North American Land Data Assimilation System (Mitchell et al., 2004)
- GLDAS: Global Land Data Assimilation System (Rodell et al., 2004)

Table 1. Basic characteristics of the NLDAS and GLDAS products

	NLDAS	GLDAS
Content	Water and energy budget data, forcing data	
Spatial coverage	Conterminous US, parts of southern Canada and northern Mexico	All land north of 60° South
Spatial resolution	0.125°	0.25° and 1.0°
Temporal coverage	Phase-1: Aug.1, 1996 - Dec. 31, 2007 Phase-2: Jan. 1, 1979 – present	Version-1 1.0°: Jan. 1, 1979 - present 0.25°: Feb. 24, 2000 - present Version-2: Jan. 1, 1948 - present
Temporal resolution	Hourly and monthly	3-hourly and monthly
Forcing	Multiple data sets derived from satellite measurements, radar estimation, precipitation gauges, and atmospheric analyses	Multiple data sets derived from satellite measurements and atmospheric analyses
Land surface models	Mosaic, NOAH SAC, VIC	CLM, Mosaic, NOAH VIC
Output format	GRIdded Binary (GRIB)	
Elevation definition	GTOPO 30	
Vegetation definition	University of Maryland, 1 km	

Both NLDAS and GLDAS data sets have recently been improved.

- With the motivation of creating more climatologically consistent data sets, GLDAS-2 data have been generated by using the Princeton meteorological dataset (Sheffield et al, 2006) and upgraded versions of Land Surface Models (LSMs).
- The NLDAS Phase 1 data (1996 – 2007) were added to the GES DISC archives and released to the public, to allow easier comparisons between the two phases of NLDAS.

Table 2. LSM model versions for GLDAS-1 and GLDAS-2.

Model	Resolution	GLDAS-1	GLDAS-2	Remarks
NOAH	1.0°	Version 2.7	Version 2.7.1	Updated model parameters that specify the initial soil temperature
CLM	1.0°	Version 2.0	Version 3.5	Used MODIS based parameter data sets, stand alone
VIC	1.0°	Version 4.4 Water balance mode	Energy balance mode	Includes all variables
Catchment	1.0°	Mosaic model	Catchment	Model switch
NOAH	0.25°	Version 2.7, Snow DA (data assimilation): direct insertion	Version 2.7.1, Snow DA: forward-looking	Updated model parameters that specify the initial soil temperature

More information about GLDAS and NLDAS and model data validation can be found at Land Data Assimilation Systems Web site at http://ldas.gsfc.nasa.gov/nldas/.

GLDAS and NLDAS Data Access

All NLDAS and GLDAS data are accessible from the HDISC NASA http://disc.sci.gsfc.nasa.gov/hydrology, with several convenient data access methods.

- **Mirador** searching and downloading (Lynnes et al., 2009) - Includes keyword searching, hierarchical navigation based on projects and on Science Areas. Provides spatial and parameter subsetting and data format conversion. http://mirador.gsfc.nasa.gov/
- **GrADS Data Server (GDS)** access - Provides parameter and spatial subsetting. Outputs data in binary, ASCII, or image. Performs any operation that can be expressed in a single GrADS expression. http://hydro1.sci.gsfc.nasa.gov/dods/
- **Anonymous ftp** downloading - Navigation based on model year, and date, simple and fast direct data downloading. ftp://hydro1.sci.gsfc.nasa.gov/data/s4pa/
- **Giovanni** visualization and analysis - a Web-based application developed by the GES DISC NASA.

Giovanni GLDAS Portal: http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=GLDAS10_M

Table 3. GLDAS and NLDAS data access at http://disc.sci.gsfc.nasa.gov/hydrology/data-holdings.

Data Type (Short Name)	Description	FTP	GDS	Mirador Navigation	Mirador Search	Giovanni * (visualization)
NLDAS-1, 0.125 degree, North America						
NLDAS_FOA0125_H.001	Hourly forcing	✔	✔	✔	✔	
NLDAS-2, 0.125 degree, North America						
NLDAS_FORA0125_H.002	Hourly primary forcing	✔	✔	✔	✔	
NLDAS_FORB0125_H.002	Hourly secondary forcing	✔	✔	✔	✔	
NLDAS_MOS0125_H.002	Hourly Mosaic	✔	✔	✔	✔	
GLDAS-2, 1.0 degree. Global						
GLDAS_NOAH10_3H_E1.002	3 hourly Noah experiment 1	✔	✔	✔	✔	
GLDAS_NOAH10_M_E1.002	Monthly Noah experiment 1	✔	✔	✔	✔	Coming soon
GLDAS-1, 0.25 degree, Global						
GLDAS_NOAH025SUBP_3H	3 hourly Noah	✔	✔	✔	✔	
GLDAS_NOAH025_M	Monthly Noah	✔	✔	✔	✔	
GLDAS-1, 1.0 degree, Global						
GLDAS_CLM10SUBP_3H	3 hourly CLM	✔	✔	✔	✔	
GLDAS_CLM10_M	Monthly CLM	✔	✔	✔	✔	✔
GLDAS_MOS10SUBP_3H	3 hourly Mosaic	✔	✔	✔	✔	
GLDAS_MOS10_M	Monthly Mosaic	✔	✔	✔	✔	✔
GLDAS_NOAH10SUBP_3H	3 hourly Noah	✔	✔	✔	✔	
GLDAS_NOAH10_M	Monthly Noah	✔	✔	✔	✔	✔
GLDAS_VIC10_3H	3 hourly VIC	✔	✔	✔	✔	
GLDAS_VIC10_M	Monthly VIC	✔	✔	✔	✔	✔

Gap between Data and End User Communities

The hydrologic science community commonly requires data to be at specific geo-locations, often as time series. In order to retrieve a single point complete time series for one parameter, e.g., a user has to go through the entire data archive, often of volumes in the Terabytes. This is the "Digital Divide" that exists between the world of discrete spatial objects in geographical information systems (GIS) and associated time series and the world of continuous space-time arrays as used in weather and climate science (Maidment et al., 2010).

- **Data Format:** NLDAS/GLDAS in GRIB (GRIdded Binary) and many other data sets in HDF (Hierarchical Data Format)

  GRIB is a mathematically concise data format commonly used in meteorology to store historical and forecast weather data. It is standardized by the World Meteorological Organization (WMO) . Although it has been popularly used for archiving remote sensing data, along with related software packages and tools for handling GRIB files, some users from communities other than meteorology still have difficulty handling data in GRIB format.

  HDF is a self-describing file format designed to store and organize large amounts of numerical data and for transfer of various types of data between different machines. Although it has been popularly used for archiving remote sensing data, along with software packages and tools, it is still a complicated data format to end user communities.
- **Data Organization:** Many variables one time step per file, inefficient for time series retrieving.
- **Data Volume:** Vast and increasing. NLDAS/GLDAS estimated total around 20 TB

Bridging the Gap ("Digital Divide")

- **HDISC has continued efforts to better bridge the gap between NASA data and end user communities.**
  - Giovanni online visualization and analysis system that provides Time Series plot and ASCII output without users needing to download the entire data.
  - Mirador subsetting service that provides parameter and spatial subsetted files.
  - GrADS Data Server that provides parameter and spatial subsetting service and outputs data in binary and ASCII (good for short time range).
- **One effective strategy is integrating the data into end user community tools and environments.**
  - End user community tools are specifically designed and implemented for their communities.
  - End user community tools have more complete data collections for their communities.
  - End users are more familiar with their tools.
- **HDISC ongoing efforts:**
  - Integrate NASA Hydrological data into CUAHSI HIS. (This poster focuses on this effort.)
  - Integrate NASA NLDAS precipitation data into U.S. Environmental Protection Agency (EPA), Better Assessment Science Integrating Point & Nonpoint Sources (BASINS) (Rui et al, 2011).

In collaboration with the Hydrologic Sciences Branch (HSB) at NASA's GSFC and CUAHSI HIS, NASA HDISC has integrated NLDAS data into CUAHSI HIS.

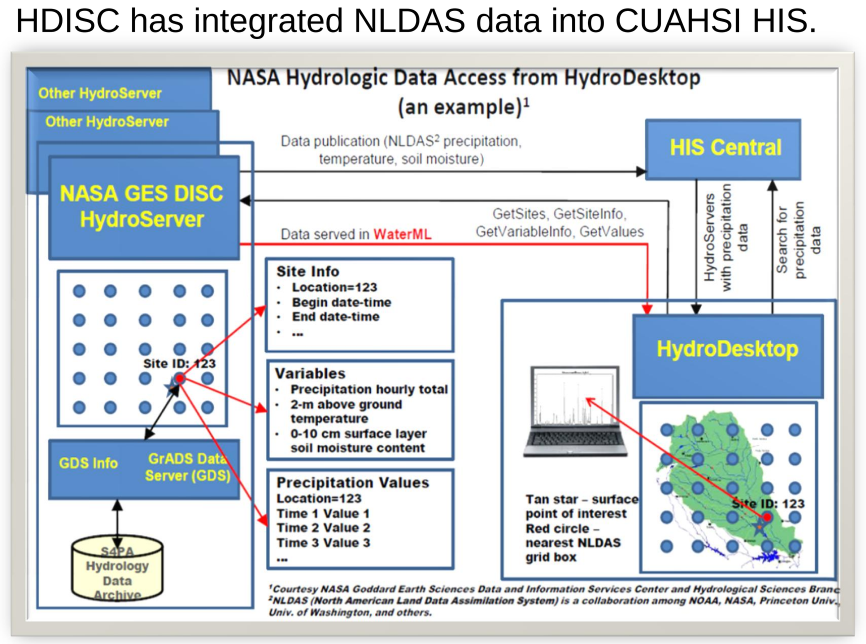


Figure 1 . Schematic of NASA hydrological data access from CUAHSI HIS' client, HydroDesktop.

http://hiscentral.cuahsi.org/pub_network.aspx?n=180

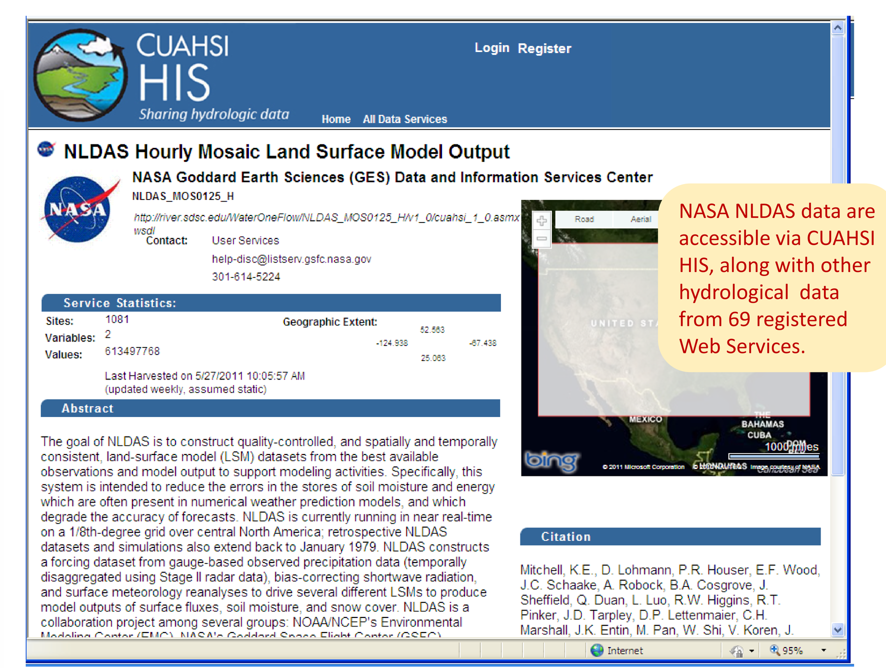


Figure 2 . NASA Web Service for NLDAS data is a registered public data service at CUAHSI HIS.

Integrate NASA Hydrological Data into CUAHSI HIS

- The CUAHSI Hydrologic Information System (HIS), http://his.cuahsi.org/, is an internet-based system for sharing hydrologic data. It comprises databases and servers, connected through web services to client applications, allowing for the publication, discovery, and access of data.
- Most NASA hydrological data are gridded, with high temporal resolution, of huge data volumes (TB's), archived in Linux systems in GRIB or HDF format (i.e., not in databases), and accessible via direct ftp, http, or GDS (OPeNDAP).
- Therefore, NASA data do not directly fit in HIS' ODM (Observations Data Model), a relational schema for storing point hydrologic observations in a relational database management system used by HydroServer.

Major work for integrating NASA data into CUAHSI HIS

- Develop a Web service that serves NASA hydrological data, as time series, and corresponding metadata in WaterML. The Web service provides four methods/functions.
  - **GetSites**: Given a data product, this method returns the site (grid point) metadata for each grid point.
  - **GetSitesInfo**: Given a site number (grid point), this method returns the site's metadata.
  - **GetVariableInfo**: Given a variable code, this method returns the variable's metadata.
  - **GetValues**: Given a site number (grid point), a variable, a start date, and an end date, this method returns a time series.
- Develop a SOAP Proxy for registering the NASA Web service at CUAHSI HIS .

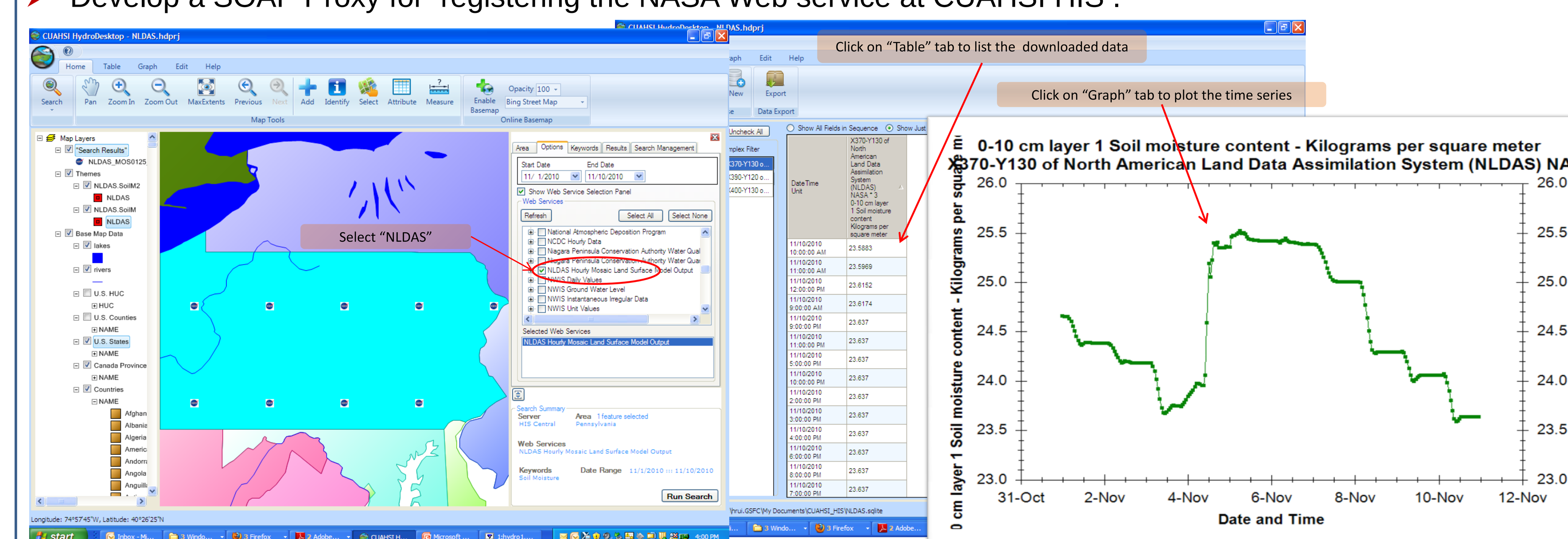
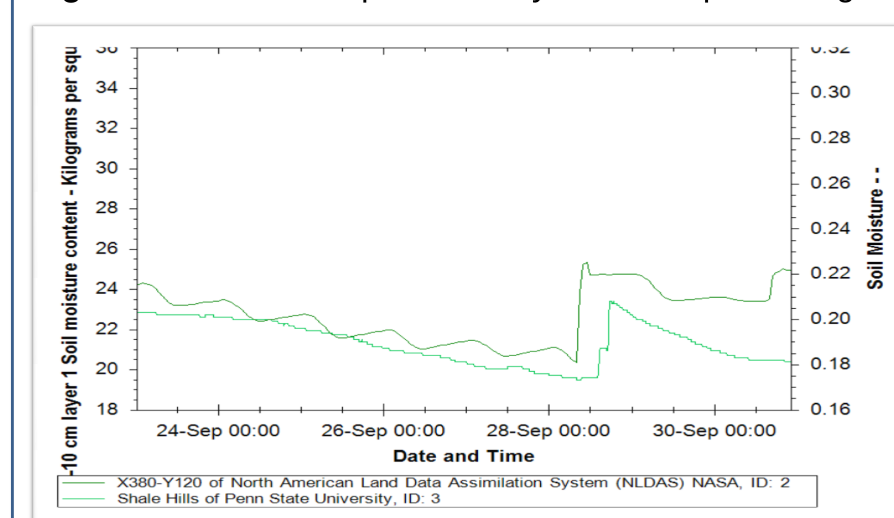


Figure 3. Screen snapshots of HydroDesktop showing selected NLDAS data and time series graph.



With the HDISC Web service registered in CUAHSI HIS, NLDAS, GLDAS, and other NASA hydrological data become searchable, retrievable, and analyzable, along with hydrologic data from other sources available via HIS. This enhanced data access will facilitate, for the broad CUAHSI HIS user community, the use of NASA hydrological data. Figures 3 and 4 illustrate how NLDAS data can now be easily retrieved, analyzed, and inter-compared with other hydrological data in CUAHSI HIS.

Figure 4 (left) . Time Series of NLDAS-2 Mosaic 0-10 cm soil moisture and SRBHOS soil moisture at Shale Hills of Penn State University.

Future Improvement

- The prototype NASA Web service is built on GDS for parameter and spatial subsetting and time stitching. Due to the limited web session time and performance of GDS for time series retrieval, the maximum time range is limited to 240 time steps, for performance reason. This limitation should be lifted.
  - The data need to be reprocessed and archived for optimal time series retrieval to remove the limit on the maximum number of retrievable time steps.
  - Data service should be improved for serving time series more effectively.
  - Before the limitation is lifted, a message should be provided for informing the users about the limitation.
- A better method for presenting and handling site information for gridded data should be developed.
  - Listing grid points as sites, with more informative site names.
  - Masking out grid points over oceans/waters, so only land grids are listed.
  - Listing all grid points over land (current prototype lists only one tenth of grid points).
- Error handling method should be enhanced, for providing more comprehensive error messages.
- More NASA hydrological data will be integrated into the CUAHSI HIS.

References

Mitchell, K.E., D. Lohmann, P.R. Houser, E.F. Wood, J.C. Schaake, A. Robock, B.A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R.W. Higgins, R.T. Pinker, J.D. Tarpley, D.P. Lettenmaier, C.H. Marshall, J.K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A.A. Bailey, 2004. 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.

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, 2004. The Global Land Data Assimilation System, Bull. Amer. Meteor. Soc., 85(3), 381-394.

Maidment, D.R., F. Salas, B. Domenico, and S. Nativi, 2010. Crossing the Digital Divide: Connecting GIS, time series and space-time arrays, Abst. IN13A-1095, 2010 AGU Fall Meeting, San Francisco.

Rui, H. W. Teng, B. Vollmer, D.M. Mocko, H.K. Beaudoing, J. Nigro, M. Gray, D. Maidment, and R. Hooper, 2011. Bridging the gap between NASA hydrological data and the geospatial community, *in* Proc. ASPRS 2011 Annual Conference, Milwaukee.

Conclusions

- The gap, or "Digital Divide," between NASA hydrological data and the geospatial community is a longstanding one and still to be bridged. The key to bridging this gap is a better understanding of the hydrological data needs of the geospatial end users, which is a central focus of the NASA HDISC.
- To bridge the gap, one effective strategy is integrating the data into end user community tools and environments.
- In collaboration with CUAHSI HIS, NASA HDISC has integrated NLDAS data into CUAHSI HIS, which has already demonstrated the potential of customized Web services for enhanced access to and use of NASA data.