ENHANCING DISCOVERY, SEARCH, AND ACCESS OF NASA HYDROLOGICAL DATA BY LEVERAGING GEOSS

William Teng
ADNET Systems, Inc.
NASA Goddard Earth Sciences Data and Information Services Center
Code 610.2
Greenbelt, MD 20771

ABSTRACT

An ongoing NASA-funded project has removed a longstanding barrier to accessing NASA data (i.e., accessing archived time-step array data as point-time series) for selected variables of the North American and Global Land Data Assimilation Systems (NLDAS and GLDAS, respectively) and other EOSDIS (Earth Observing System Data Information System) data sets (e.g., precipitation, soil moisture). These time series ("data rods") are pre-generated. Data rods Web services are accessible through the CUAHSI Hydrologic Information System (HIS) and the Goddard Earth Sciences Data and Information Services Center (GES DISC) but are not easily discoverable by users of other non-NASA data systems.

The Global Earth Observation System of Systems (GEOSS) is a logical mechanism for providing access to the data rods. An ongoing "GEOSS Water Services" project aims to develop a distributed, global registry of water data, map, and modeling services cataloged using the standards and procedures of the Open Geospatial Consortium and the World Meteorological Organization. The ongoing "data rods" project has demonstrated the feasibility of leveraging the GEOSS infrastructure to help provide access to time series of model grid information or grids of information over a geographical domain for a particular time interval. A recently-begun, related NASA-funded "ACCESS-GEOSS" project expands on these prior efforts. Current work is focused on both improving the performance of the generation of on-the-fly (OTF) data rods and the Web interfaces from which users can easily discover, search, and access NASA data.

MOTIVATION: REMOVING BARRIERS TO ACCESSING NASA DATA

A longstanding and significant "Digital Divide" in data representation exists between hydrology and climatology and meteorology. Typically, in hydrology, earth surface features are expressed as discrete spatial objects such as watersheds, river reaches, and point observation sites; and time varying data are contained in time series associated with these spatial objects. Long time histories of data may be associated with a single point or feature in space. In meteorology and climatology, remotely sensed observations and weather and climate model information are expressed as continuous spatial fields, with data sequenced in time from one data file to the next. Hydrology tends to be narrow in space and deep in time, while meteorology and climatology are broad in space and narrow in time (Maidment et al., 2010).

An ongoing NASA-funded "data rods" (time series) project has demonstrated the bridging of this Digital Divide (i.e., accessing archived time-step array data as point-time series) for selected variables of the North American and Global Land Data Assimilation Systems (NLDAS and GLDAS, respectively) and other NASA EOSDIS (Earth Observing System Data Information System) data sets (Rui et al., 2012; Teng et al., 2012; Rui et al., 2013). Figure 1 summarizes the main components of the "data rods" project, with the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Hydrologic Information System (HIS) as the main end user.

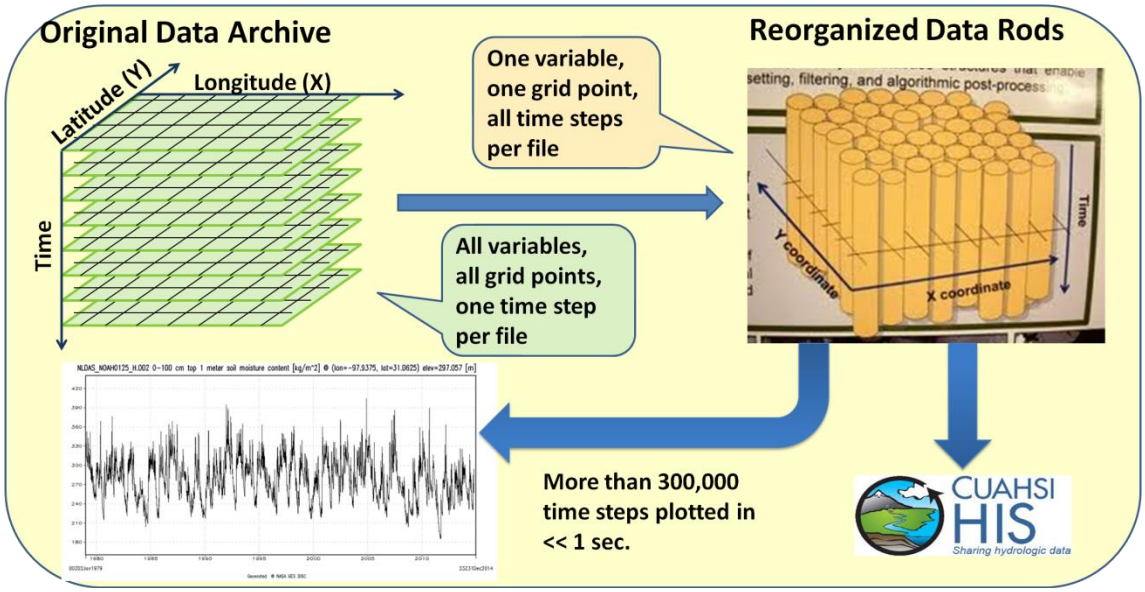

Figure 1. Schematic of data reorganization for optimal time series access.

In order to make data rods more widely accessible, beyond the CUAHSI/HIS user community, we have also made the data rods available via a Web interface, providing a probability description at each grid cell and for each day. Current values can be seen in the context of a probability distribution of past values, for that location and time (Fig. 2). Surface soil moisture for the U.S. was the initial variable made available this way; variables to be added include precipitation, runoff, evapotranspiration, and surface temperature. In addition to the values of the displayed variable map, display options also include anomaly and percentile. Clicking anywhere on the displayed map will show the spatio-temporal statistics of that geographic location, in a pop-up window, along with a plot of the cumulative distribution function (CDF) and current value and a plot of the values of the previous 30 days (respectively, top right and bottom right of Fig. 2).

The "data rods" project has greatly enhanced the accessibility of certain NASA hydrological data for the large community of CUAHSI HIS users, as well as users of the NASA GES DISC.

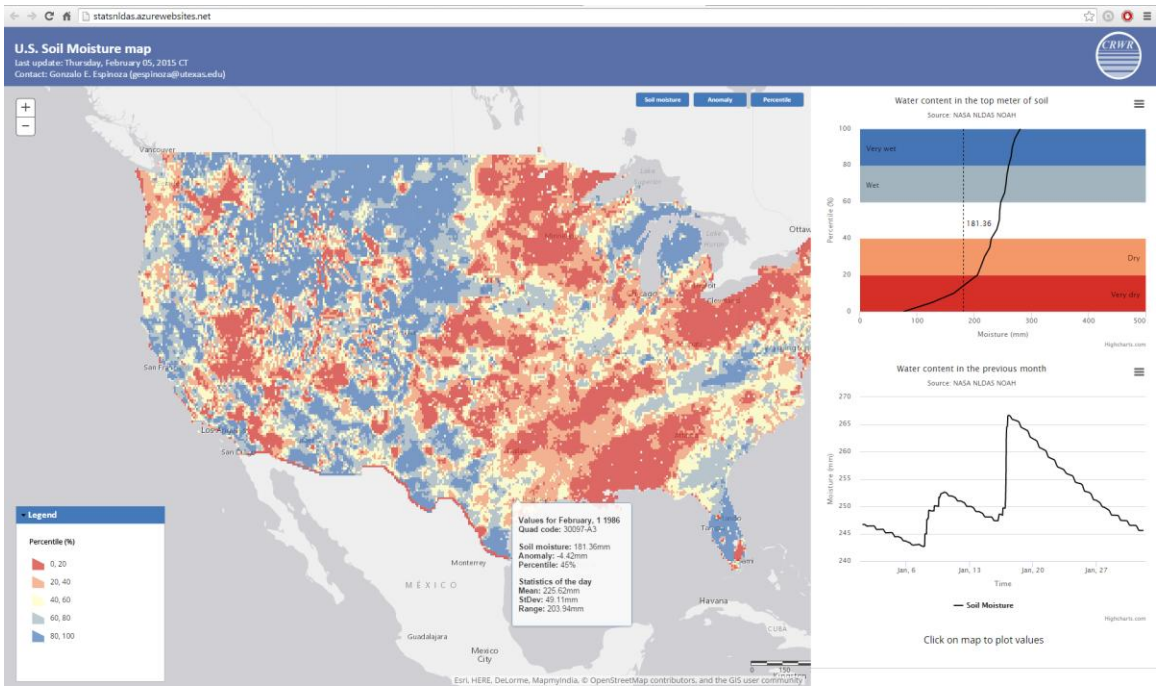

Figure 2. Probability presentation of data rods.

NASA HYDROLOGICAL DATA VIA GEOSS

To further enhance the discovery, search, and access of NASA data by users of other non-NASA data systems, another NASA-funded "ACCESS-GEOSS" project was recently begun, leveraging the Global Earth Observation System of Systems (GEOSS). There is an ongoing series of multi-organizational GEOSS Architecture Implementation Pilots, now in Phase-8 (AIP-8) and with a strong water sub-theme, that is aimed at the GEOSS Water Strategic Target, which is "to produce comprehensive sets of data and information products to support decision-making for efficient management of the world's water resources, based on coordinated, sustained observations of the water cycle on multiple scales." As part of GEOSS AIP-x, the "GEOSS Water Services" project aims to develop a distributed, global registry of water data, map, and modeling services cataloged using the standards and procedures of the Open Geospatial Consortium and the World Meteorological Organization. The AIP-x has demonstrated that individual services registered with the GEOSS Common Infrastructure (GCI) can be discovered. Each feature in a map service describes one observed time series accessible at one location and, if available, contains a link to that time series service. Also demonstrated was that GCI can be adapted to provide time series access to model grid information (e.g., NLDAS, GLDAS) or grids of information over a geographic domain for a particular time interval.

The "ACCESS-GEOSS" project expands on these previous efforts, including the generation of on-the-fly (OTF) data rods, leveraging the NASA Simple Subset Wizard (SSW), a gateway to NASA data centers (Teng et al., 2014). The advantage of OTF processing of data rods is that it potentially avails users many more variables than are currently available as pre-generated data rods, from both the GES DISC and, via SSW, the other participating (in SSW) data centers. The tradeoff is a shorter allowable requested time period. Current benchmark for OTF-processing performance is 90 seconds for 10,000 time steps, which, for hourly data, such as NLDAS, is about a one-year time period. Figure 3 shows the flow of pre-generated and OTF data rods to Web maps for time series (service endpoints), either directly or via the GEOSS Portal. Registration of the data rods leverages existing Directory Interchange Format (DIF) documents of the Global Change Master Directory (GCMD). Figure 4 shows the access to data rods Web services, as WaterML 1.0, ASCII, or plot. Current work is focused on both improving the performance of the

generation of OTF data rods and the Web interfaces from which users can easily discover, search, and access NASA data.

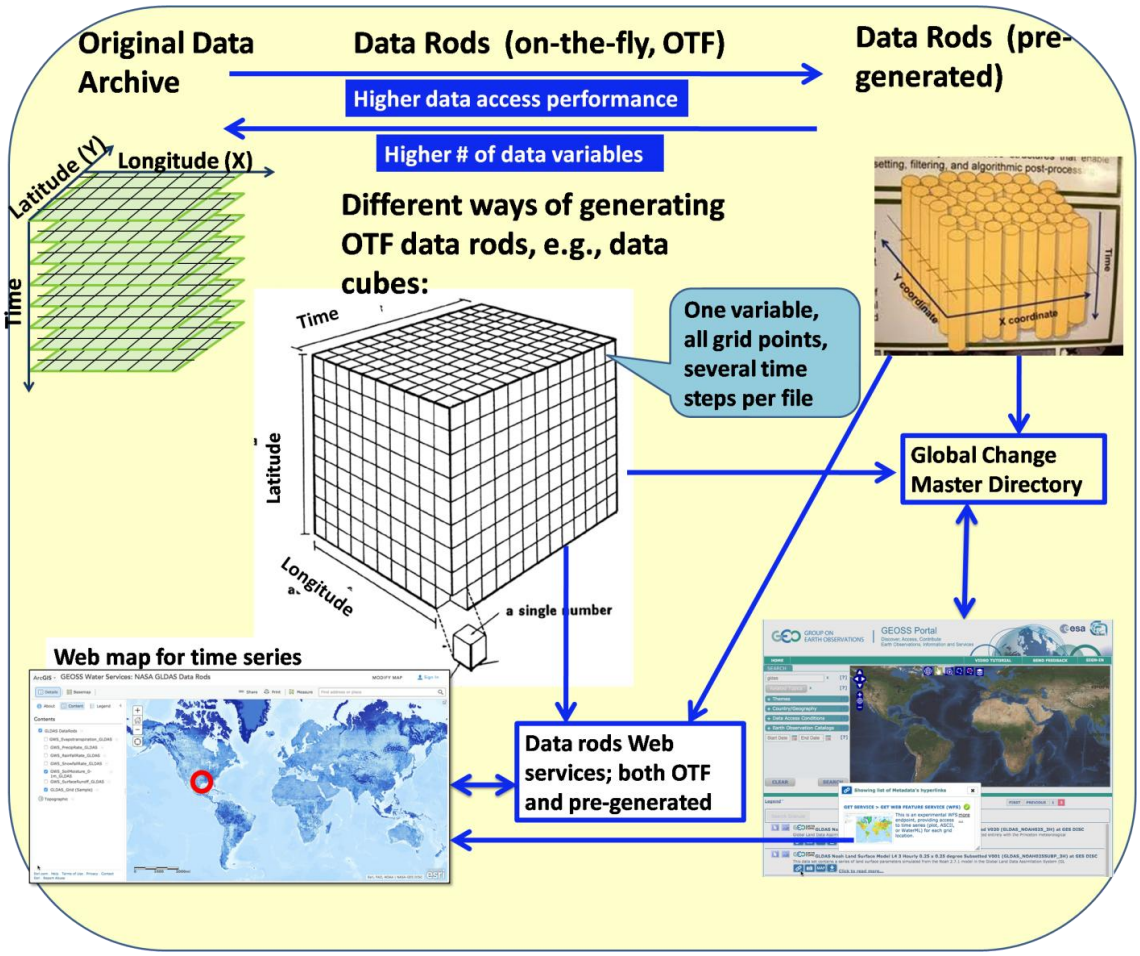

Figure 3. Schematic of pre-generated and on-the-fly (OTF) generated data rods accessible via Web maps for time series (service endpoints), directly or via the GEOSS Portal and Global Change Master Directory (GCMD).

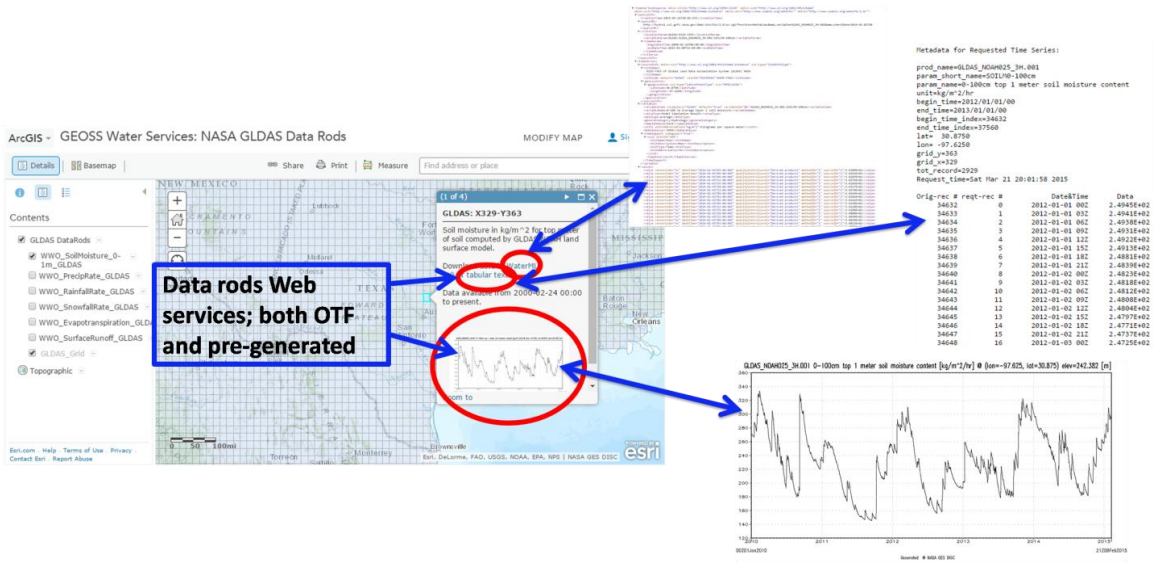

Figure 4. Enlargement of red circle in "Web map for time series" at lower left of Figure 3 (Texas and the Gulf coast), showing access to data rods Web services, as WaterML 1.0, ASCII, or plot.

FOR MORE INFORMATION

NASA GES DISC Hydrology Portal: http://disc.sci.gsfc.nasa.gov/hydrology
NASA GSFC Hydrological Sciences Laboratory: http://ldas.gsfc.nasa.gov/
NASA Giovanni Portal for NLDAS Hourly, GLDAS 3-hourly, TRMM precipitation, and other data products: http://giovanni.gsfc.nasa.gov/giovanni/

ACKNOWLEDGMENT

Both the "data rods" and "ACCESS-GEOSS" projects are truly team efforts, with contributions from NASA GES DISC, University of Texas at Austin, NASA GSFC Hydrological Sciences Laboratory, and Brigham Young University. The author would like to thank all the members from both project teams:

"Data rods": David Maidment, Bruce Vollmer, Christa Peters-Lidard, Hualan Rui, Richard Strub, Tim Whiteaker, David Mocko, and Dalia Kirschbaum

"ACCESS-GEOSS": David Maidment, Matthew Rodell, Richard Strub, David Arctur, Daniel Ames, Hualan Rui, Bruce Vollmer, and Edward Seiler

The work conducted for the two projects is supported by NASA ROSES NNH11ZDA001N-ACCESS and NNH13ZDA001N-ACCESS.

REFERENCES

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

Rui, H., B. Teng, R. Strub, and B. Vollmer, 2012, Data reorganization for optimal time series data access, analysis, and visualization, 2012 AGU Fall Meeting, San Francisco.

Rui, H., R. Strub, W. Teng, B. Vollmer, D. Mocko, D. Maidment, and T. Whiteaker, 2013, Enhancing access to and use of NASA hydrological data, 2013 AGU Fall Meeting, San Francisco.

Teng, W.L., D.R. Maidment, B. Vollmer, C.D. Peters-Lidard, H. Rui, R. Strub, T. Whiteaker, D.M. Mocko, and D.B. Kirschbaum, 2012, Bridging the Digital Divide between discrete and continuous space-time array data to enhance accessibility to and usability of NASA earth sciences data for the hydrological community, 2012 AGU Fall Meeting, San Francisco.

Teng, W., D. Maidment, M. Rodell, R. Strub, D. Arctur, D. Ames, H. Rui, B. Vollmer, E. Seiler, 2014, Enhancement of mutual discovery, search, and access of data for users of NASA and GEOSS-cataloged data systems, 2014 AGU Fall Meeting, San Francisco.