Global Precipitation Measurement (GPM) Mission Products and Services at the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC)

Z. Liu[a,b], D. Ostrenga[a,c], B. Vollmer[a], B. Deshong[a,c], K. MacRitchie[a,c], M. Greene[a,d], and S. Kempler[a]

[a]NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC)

[b]Center for Spatial Information Science and Systems (CSISS), George Mason University

[c]Adnet Systems, Inc.

[d]Wyle Information Systems, LLC

Submitted to:

Nowcast

Bulletin of the American Meteorological Society

Submitted on April 20, 2016

Corresponding author: Tel: +1 301 614 5764; fax: 301 614 5268

Email address: Zhong.Liu@nasa.gov

Abstract

This article describes NASA/JAXA Global Precipitation Measurement (GPM) mission products and services at the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC). Built on the success of the Tropical Rainfall Measuring Mission (TRMM), the next generation GPM mission consists of new precipitation measurement instruments and a constellation of international research and operational satellites to provide improved measurements of precipitation globally. To facilitate data access, research, applications, and scientific discovery, the GES DISC has developed a variety of data services for GPM. This article is intended to guide users in choosing GPM datasets and services at the GES DISC.

1. Introduction

Launched on 27 February 2014, the NASA/JAXA Global Precipitation Measurement (GPM) mission core satellite and a constellation of international satellites not only greatly extend the spatial coverage from its predecessor (the Tropical Rainfall Measuring Mission (TRMM)), but also provide improved measurements of precipitation globally. For example, a new Ka-band precipitation radar and additional high frequency channels in the microwave instrument have been added to the GPM core satellite for improving light rain and snowfall measurements. Furthermore, the Integrated Multi-satellitE Retrievals for GPM (IMERG) have been significantly improved over the TRMM Multi-satellite Precipitation Analysis (TMPA) in terms of spatiotemporal resolution, spatial coverage, and more.

GPM datasets are available for research and applications at the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC), home to the TRMM data archive as well. To new users, it can be a daunting task to locate a suitable GPM dataset. Even for experienced TRMM users, such activity can also be difficult due to many changes implemented because GPM datasets and services have been completely redesigned to accommodate changes in data structure, format, data volume, new technology, etc. Therefore, it is necessary to develop an overview document that guides users in locating datasets of interest and services that are suitable for their research and applications. Recognizing a very diverse user community consisting of users from different scientific disciplines, backgrounds, and countries with different levels of data downloading capabilities and Internet connectivity, the GES DISC has developed data services to facilitate GPM data access and exploration. This article is organized as follows: Section 2 describes GPM data products; Section 3 GPM data services; Section 4 GPM data exploration; Section 5 GPM data applications, followed by future plans in the final section.

2. GPM Data Products

GPM data products at the GES DISC are organized and archived based on three product levels defined by the NASA Earth Observing System Data and Information System (EOSDIS): Level-1, Level-2, and Level-3. In some satellite missions, Level-1 products are sub-divided into two categories: Level-1A and Level-1B. Level-1A is defined as, "Reconstructed, unprocessed instrument data at sensor's full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to Level-0 data." For Lelvel-1B, it is defined as, "Level 1A data that have been processed to sensor units (not all instruments have Level-1B source data)." For GPM, an additional Level-1category, Level-1C, has been added for common intercalibrated microwave brightness temperature (Tc) products from GPM constellation satellites, which is necessary to ensure no systematic differences for multi-sensor and multi-satellite precipitation retrieval algorithms such as GPM IMERG.

Table 1 lists GPM Level-1 datasets. Besides Level-1 datasets from the GPM Microwave Imager (GMI) and the Dual-frequency Precipitation Radar (DPR) onboard the GPM core satellite, there are Level-1 datasets from other satellites in the GPM constellation. There is only one Level-1A dataset containing GMI unpacked packet data or raw data. There are 3 Level-1B datasets (1 from GMI and 2 from DPR). The remaining datasets are Level-1C as described above. Figure 1a is a sample of GMI Level-1C common calibrated brightness temperatures at 37 GHZ showing Tropical Cyclone Nanauk over the Arabian Sea on 11 June 2014. As seen in Table 1, Level-1 GPM datasets consist of reconstructed and unprocessed instrument data at sensor's resolution and therefore are best suitable for algorithm development and other special activities.

Level-2 datasets are defined as, “Derived geophysical variables at the same resolution and location as Level 1 source data.” Table 1 lists GPM Level-2 datasets distributed at the GES DISC. It is seen that GPM Level-2 datasets include those from GPROF (the Goddard Profiling Algorithm) from the GPM constellation satellites, GPM DPR, and their combined datasets as well as latent heating products from DPR. Figure 1b is a sample of GMI Level-2 GPROF surface precipitation, showing Hurricane Arthur near the South Carolina and Georgia coasts on 3 July 2014. Since Level-2 GPM datasets contain geophysical variables at sensor's resolution, their usage is typically wider than Level-1 datasets, for example, Level-2 precipitation can be used in case studies, ground validation, model verification, etc.

Level-3 datasets are “Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.” Table 1 lists Level-3 half-hourly, daily and monthly gridded datasets. Half-hourly datasets consist of IMERG products only. Daily datasets include daily gridded orbital mosaic (or ascending/descending for DPR) datasets from microwave sensors in the GPM constellation satellites and DPR as well as daily GMI and DPR combined datasets. Monthly datasets include all GPROF datasets from the constellation and one from IMERG. The most popular datasets are the multi-satellite, multi-sensor, and multi-algorithm GPM IMERG products that include Early, Late, and Final Run. The Early and Late Run of IMERG consist of near-real-time monitoring products with climatological gauge calibration. For the Final Run, the Global Precipitation Climatology Centre (GPCC) monthly monitoring gauge dataset is used for bias correction. The latencies from observation to public distribution are, 6 hours (Early Run), 18 hours (Late Run) and 4 months (Final Run), respectively. Both spatial (0.1 degree) and temporal (half-hourly) resolutions of IMERG have been significantly improved compared to 0.25 degree and 3-hourly resolutions in TMPA. These improvements are important

for hydrometeorological research and applications as well as other applications. Details about the IMERG datasets can be found in their technical documents. Figure 1c is an example from the half-hourly IMERG Final Run, showing heavy precipitation at 03Z 15 June 2014 in the Midwestern United States. Other monthly datasets derived from different satellites in the GPM constellation are useful for the understanding of uncertainties in global precipitation measurements.

3. GPM Data Services

GPM data services are crucial to facilitate data evaluation and access in order to maximize the use of datasets in research and applications. Precipitation dataset users are very diverse, consisting of college professors, researchers, operational forecasters, citizen scientists, high school students, etc. Some of them are first-time users of remote sensing products and human-readable data formats such as ASCII are needed. The HDF5 data format is used in all GPM standard products. Special software and knowledge are required to handle such complex data structures. Format conversion is often needed for many users from different backgrounds. In addition, not all users need a global coverage and a subsetting capability is necessary to minimize data transfer and storage, which is particularly important for users from developing countries where Internet bandwidth can be very limited. On the other hand, hydrologic applications are closely associated with watersheds and some applications use political boundaries such as states or counties. Having a GIS shapefile capability is necessary to allow users downloading data only in an irregular shape area.

Mirador is a Google-based data search interface that allows searching, browsing, and retrieving of Earth science datasets at the GES DISC. Mirador will soon be replaced by a more

124 powerful data service system called the Unified User Interface (UUI) to unify several existing
125 user services and provide data, services, and information in one unified user interface. Without
126 the UUI, users will have to visit different websites or portals for data subsetting, visualization,
127 document information, data recipes, etc. In short, the UUI will save users time and expedite data
128 access.

129 The Simple Subset Wizard or SSW (Fig. 2a) provides a simple and easy way to subset
130 Level-3 and limited Level-2 datasets not only from the GES DISC but also from other NASA
131 data centers such as the NASA Global Hydrology Resource Center, the NASA Langley
132 Atmospheric Science Data center, etc. SSW contains a text input area for keyword search, a
133 calendar for selecting beginning and ending times and a spatial bounding box for choosing an
134 area of interest (Fig. 2a). SSW allows parameter subsetting and format conversion (Fig. 2b). For
135 example, SSW can convert the original HDF5 format in the IMERG Final Run dataset to either
136 NetCDF or ASCII (Fig. 2b). For those who are not familiar with HDF5 or NetCDF, ASCII is a
137 user-friendly and human-readable format. After all these, SSW generates a list of URLs which
138 can be used for batch download with popular off-the-shelf software packages such as wget.
139 Currently, SSW provides data subsetting and format conversion services for all GPM Level-3
140 products listed in Table 1, including the popular IMERG datasets, except the latent heating
141 datasets.

142 The Open Source Project for a Network Data Access Protocol (OPeNDAP) provides
143 interoperability and remote access to individual variables within datasets in a form usable by
144 many tools including IDV, McIDAS-V, Panoply, Ferret and GrADS. Format conversion can be
145 achieved through OPeNDAP and available formats are ASCII, NetCDF 3, NetCDF 4, and

146 binary. In addition to interoperability, OPeNDAP is very useful for supporting operational
147 activities because users can write a script to automatically pull data from OPeNDAP on a fixed
148 schedule. All datasets listed in Table 1 can be accessed through OPeNDAP.

149 **4. GPM Data Exploration**

150 Giovanni (an acronym for the Geospatial Interactive Online Visualization ANd aNalysis
151 Infrastructure) is an online tool, developed at the GES DISC, to facilitate access, evaluation, and
152 exploration of Earth science datasets. All IMERG datasets can be easily visualized and analyzed
153 online with Giovanni without the need to download data and software. For novices, using
154 satellite remote sensing datasets can be a daunting task and numerous issues can be encountered
155 in data processing such as data format, data structure, data volume, Internet connectivity or
156 bandwidth, etc. Moving a large amount of remote sensing data over the Internet can be time
157 consuming and problematic for countries with low bandwidth and unreliable Internet
158 connections. Sending a graphic result or time series in ASCII instead, other than the entire
159 dataset, can make a significant difference to users in those countries. Nonetheless, online tools
160 like Giovanni can provide a convenient way to bridge GPM data and users.

161 Recently, Giovanni has been completely redesigned due to an increasing demand for
162 integrated analysis and visualization of a large collection of Earth science datasets at the GES
163 DISC and other NASA data centers. Meanwhile, Giovanni evolves with modern software
164 technologies and development to make it more user-friendly and increase its performance for
165 data exploration. Giovanni contains only one landing page (Fig. 3a). Keyword and facet search
166 capabilities (Fig. 3a) make searching a large amount of datasets a simple process. Due to a large
167 amount of variables (over ~1400 as of this writing) in Giovanni, the list of search results

168 sometimes can be very long and difficult for users to locate a variable of interest. Faceting makes
169 picking a dataset easy. For example, if one looks for calibrated precipitation in the monthly
170 IMERG Final Run dataset, a search for "precipitation" in Giovanni returns a list of 102 variables
171 from TRMM, GPM, MERRA (Modern Era Retrospective-Analysis for Research and
172 Applications), NLDAS (Global Land Data Assimilation System), etc. By choosing GPM from
173 the facet list, the list is shortened to 17 variables and after clicking on "monthly" in Temporal
174 Resolutions, only 4 variables are available and they all belong to the IMERG monthly product.
175 Of course, one can simply search "IMERG Monthly" without doing any filtering work. IMERG
176 Early, Late, and Final Run are available in Giovanni. For example, a trio of typhoons in the
177 Western Pacific is shown in the rainfall intensity from the IMERG Early Run at 02Z 7 July 2015
178 (Fig. 4a). For the time being, facets in Giovanni contain disciplines, measurements,
179 platform/instrument, spatial resolutions, temporal resolutions, wavelengths, depths, special
180 features, and portals. For users who are familiar with the TRMM Online Visualization and
181 Analysis System (TOVAS), simply type in "TOVAS" in the search box and it retrieves all
182 TRMM and GPM related variables in Giovanni.

183 Table 2 lists Giovanni plot types and formats available for data downloads. New
184 functions are still being added to Giovanni. In the map group (Table 2), the "Accumulated"
185 function allows users to generate an accumulated precipitation map either from a rectangular box
186 or a shape (countries, major watersheds, and states in the United States). A sample is presented
187 in Fig. 3b. The "User-Defined-Climatology" allows defining custom climatology for a user-
188 defined time period so one can compare climatologies with different time lengths. In the
189 comparison group, the interactive scatter map allows picking a point in a scatter plot to show its
190 geographic location, which can be useful for investigating unusual points in a scatter plot. In the

time series group, the “Seasonal” allows users to choose one or multiple seasons and plot the time series. To compare with a gauge-based time series, users can input the latitude and longitude of a gauge location in the landing page and use the time series function to obtain the plot and the ASCII data in CSV (Comma Separated Values) for comparison. To create maps and time series for an irregular shape such as countries, states, and watersheds, one can click on the "Show Shapes" button and select a shape. Figure 3b is a sample rainfall map, showing heavy rainfall (in mm) in the capital region of Tokyo due to the passages of Super Typhoons Phanfone and Vongfong in October 2014.

Giovanni output can be downloaded as well. Users can download images in the GeoTIFF, KMZ, and PNG formats. Digital map data can be downloaded as NetCDF, which is a very common format in many scientific communities and easy to be imported into GIS software packages such as ArcGIS. Non-gridded 2-D data from time series, zonal mean, etc. can be downloaded as ASCII CSV which can be imported into Microsoft Excel for further analysis.

Giovanni allows users to explore other precipitation datasets such as those from TRMM, MERRA, NLDAS, etc. For precipitation, it is well known that different units are used in different disciplines, for example, the units in IMERG (mm/hr) and MERRA (kg/m^2/s) monthly precipitation products are different. In addition, their grid structures are different. Unit conversion and regridding algorithms are available in Giovanni, making the comparison of these monthly products possible. Figure 4b shows a difference map between the TMPA and IMERG Final Run monthly datasets in July 2014. It is seen that the IMERG precipitation is in general higher than that of TMPA over land and lower over oceans for July 2014 (Fig. 4b). Figure 4c is a scatter plot between the two variables, showing a close relationship.

5. GPM Data Applications

Societal impacts are an important component of the GPM mission. Since the TRMM era, the GES DISC has developed data services to support domestic and international users in their precipitation related applications. Based on user's reports, these applications range from flood/drought monitoring activities, crop monitoring, disease studies/monitoring, hurricane watch, insurance industries, etc. In addition to the data services that have been mentioned above, the GES DISC is working closely with U.S. federal agencies such as the United States Department of Agriculture (USDA) Foreign Agriculture Service (FAS) to develop data services and support their worldwide operation. For example, the near-real-time TMPA 10-day product and anomaly have been in operation in the USDA Crop Explorer since TRMM. As soon as the retrospective processing of the IMERG data in the TRMM era is finished, we will work with USDA FAS to replace the TMPA product with the higher spatial resolution (0.1 degree) IMERG near-real-time product.

6. Future Plans

Future plans consist of two areas: value-added products and services. Value-added products are being developed to facilitate data access and scientific investigation activities. For example, not all users need half-hourly IMERG products and daily products are sufficient to meet their requirements. Such daily products are available now. We will work closely with users and algorithm developers to develop additional value-added precipitation products. As for data services, more can be added to the existing services. Subsetting Level-1 and Level-2 datasets is needed to avoid downloading unwanted data outside an area of interest. Although OPeNDAP can perform such task, it is not as straightforward as SSW where a web interface is available for collecting user's input and generating a list of URLs for batch data download. Visualization and

analysis of GPM Level-1 and Level-2 datasets at sensor's resolution in Giovanni are helpful for case studies, dataset evaluation, and algorithm development. For Level-3 products, custom datasets are necessary for those who use different grid structures, spatial and temporal resolutions, and projections in their activities. The GES DISC is also home to many NASA satellite missions or projects. Capabilities to integrate, analyze, and visualize datasets from other satellite missions or projects such as CloudSat, TRMM, etc. are also necessary for data exploration and scientific discovery. Event-based subsetting services can save time because users do not need to use different subsetters (if available) for obtaining data subsets, which is particularly useful for case studies.

Acknowledgments: The authors would like to thank GPM science team members, in particular, Dr. George Huffman and many users for providing comments and suggestions during the dataset and service development at the GES DISC. Thanks extend to Andrey Savtchenko for the contribution and three anonymous reviewers for their comments and suggestions that have significantly improved and strengthened the manuscript. GPM datasets are processed and provided by the Precipitation Processing System (PPS) that also distributes GPM data and provides services.

FOR FURTHER READING

GPM and TRMM data access through Mirador: http://mirador.gsfc.nasa.gov/

GPM data access through UUI: http://disc.sci.gsfc.nasa.gov/uui/datasets?keywords=GPM

GPM documents: http://pps.gsfc.nasa.gov/GPMprelimdocs.html

GPM IMERG data in Giovanni:
http://giovanni.sci.gsfc.nasa.gov/giovanni/#service=TmAvMp&starttime=&endtime=&bbox=-180,-90,180,90&dataKeyword=imerg

GPM overview: Hou, Arthur Y., Ramesh K. Kakar, Steven Neeck, Ardeshir A. Azarbarzin, Christian D. Kummerow, Masahiro Kojima, Riko Oki, Kenji Nakamura, and Toshio Iguchi, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722. doi: http://dx.doi.org/10.1175/BAMS-D-13-00164.1

NASA Data Processing Levels: http://science.nasa.gov/earth-science/earth-science-data/data-processing-levels-for-eosdis-data-products/

NASA Precipitation Measurement Missions web portal: http://pmm.nasa.gov/

OPeNDAP access: http://gpm1.gesdisc.eosdis.nasa.gov/opendap/

Simple Subset Wizard: http://disc.sci.gsfc.nasa.gov/SSW/#keywords=GPM

TRMM overview: Garstang, Michael, Kummerow, Christian D. 2000: The Joanne Simpson Special Issue on the Tropical Rainfall Measuring Mission (TRMM). Journal of Applied Meteorology: Vol. 39, No. 12, pp. 1961-1961.

TRMM product and service overview: Liu, Z. D. Ostrenga, W. Teng and S, Kempler, 2012, Tropical Rainfall Measuring Mission (TRMM) Precipitation Data Services for Research and Applications, Bulletin of the American Meteorological Society, doi: http://dx.doi.org/10.1175/BAMS-D-11-00152.1

USDA Crop Explorer: http://www.pecad.fas.usda.gov/cropexplorer/

Figure Captions:

Figure 1. Samples of GPM datasets at different levels: a) Level-1C GMI common calibrated brightness temperatures at 37 GHZ showing Tropical Cyclone Nanauk over the Arabian Sea on 11 June 2014; b) Level-2 GMI GPROF surface precipitation showing Hurricane Arthur near the South Carolina and Georgia coasts on 3 July 2014; and c) Level-3 half-hourly IMERG Final precipitation showing heavy precipitation at 03Z 15 June 2014 in the Midwestern United States.

Figure 2. The Simple Subset Wizard (SSW) provides a simple way to subset Level-3 and Level-2 (limited) GPM datasets not only from the GES DISC, but also other NASA data centers. a): The landing page of SSW. b): Sample output showing different options for subsetting data.

Figure 3.The GES DISC Giovanni allows visualization and easy access to IMERG data: a): A screenshot of the Giovanni landing page. Features such as keyword and facets make dataset search simple. b): A sample rainfall map from IMERG Final Run showing heavy rainfall (in mm) in the capital region of Tokyo due to the passages of Super Typhoons Phanfone and Vongfong in October 2014.

Figure 4. Samples of IMERG in Giovanni: a) Rainfall intensity (mm/hr) from the IMERG Early Run at 02Z 7 July 2015 showing a trio of typhoons in the Western Pacific; b) A monthly precipitation difference map (mm/hr) between 3B43 and IMERG Final Run for July 2014; and c) their scatter plot.

Table 1. GPM datasets at the GES DISC.

Level-1 Dataset (Reconstructed and unprocessed observations at sensor's resolution)	Level-2 Dataset (Derived Geophysical variables at sensor's resolution)	Level-3 Dataset (Variables on uniform space-time grid)
GMI unpacked packet data	GPM DPR environment	GPM DPR Daily (Ascending, Descending)
GMI Brightness Temperatures	GPM DPR Ku Precipitation	GPM DPR Daily and Monthly Precipitation Profiles
GMI Common Calibrated Brightness Temperatures Collocated	GPM DPR Ka Precipitation	GPM DPR, GMI Combined Daily and Monthly Precipitation
Common Calibrated Brightness Temperatures from the constellation of satellites (SSMI F16, SSMI F17, SSMI F18, AMSR2 GCOMW1, GMI, METOPA MHS, METOPB MHS, MT1 SAPHIR, NOAA-18, NOAA-19, TRMM TMI, ATMS SUOMI-NPP)	Radiometer Profiling from the constellation of satellites (SSMI F16, SSMI F17, SSMI F18, AMSR2 GCOMW1, GMI, METOPA MHS, METOPB MHS, MT1 SAPHIR, NOAA-18, NOAA-19, TRMM TMI, ATMS SUOMI-NPP)	Daily and Monthly GPROF Profiling from the constellation of satellites (SSMI F16, SSMI F17, SSMI F18, AMSR2 GCOMW1, GMI, METOPA MHS, METOPB MHS, MT1 SAPHIR, NOAA-18, NOAA-19, TRMM TMI, ATMS SUOMI-NPP)
GPM DPR Level-1B Ku-band Received Power	GPM DPR and GMI Combined Precipitation	IMERG Half Hourly and Daily (Early Run, Late Run, and Final Run)
GPM DPR Level-1B Ka-band Received Power	GPM DPR Convective Stratiform Heating	IMERG Monthly (Final Run)
	GPM DPR Spectral Latent Heating	GPM DPR Daily and Monthly Convective Stratiform Heating and Spectral Latent Heating

Table 2. Giovanni plot types. File formats for downloads are PNG, GeoTIFF, KMZ, NetCDF, and ASCII (non-gridded 2-D data only)

Maps	Comparisons	Time Series	Vertical Plots	Miscellaneous
Time-averaged map	Correlation map	Area-averaged differences	Cross-section map, latitude-pressure	Zonal mean
Animation	Scatter plot, area-averaged (static)	Area-averaged	Cross-section map, longitude-pressure	Histogram
Difference of time-averaged maps	Scatter plot (interactive)	Seasonal (inter-annual)	Cross-section map, time-pressure	
Accumulated map	Scatter plot (static)	Hovmöller, longitude-averaged	Vertical profile	
User-defined climatology	Scatter plot, time-averaged (interactive)	Hovmöller, latitude-averaged		

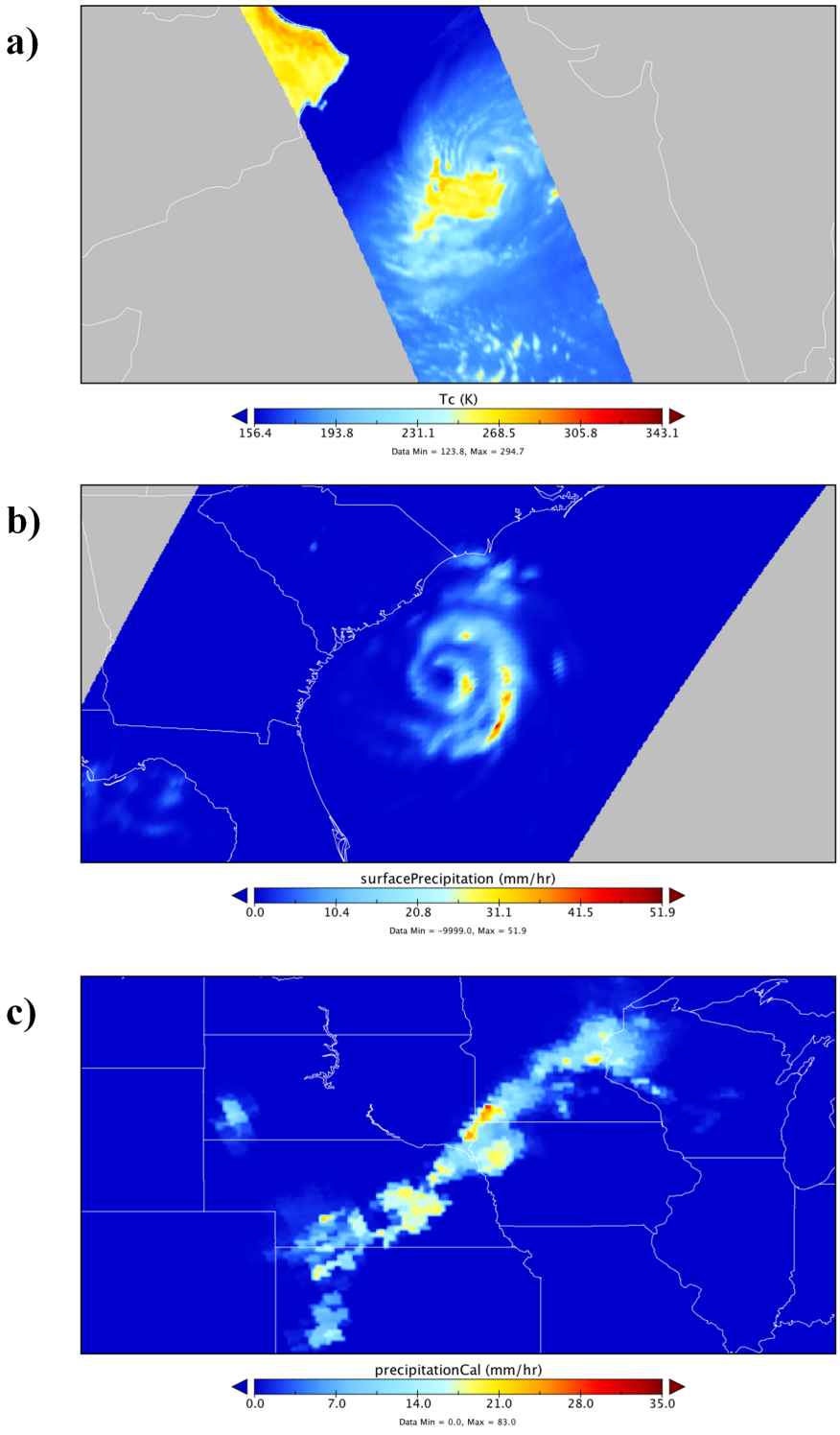
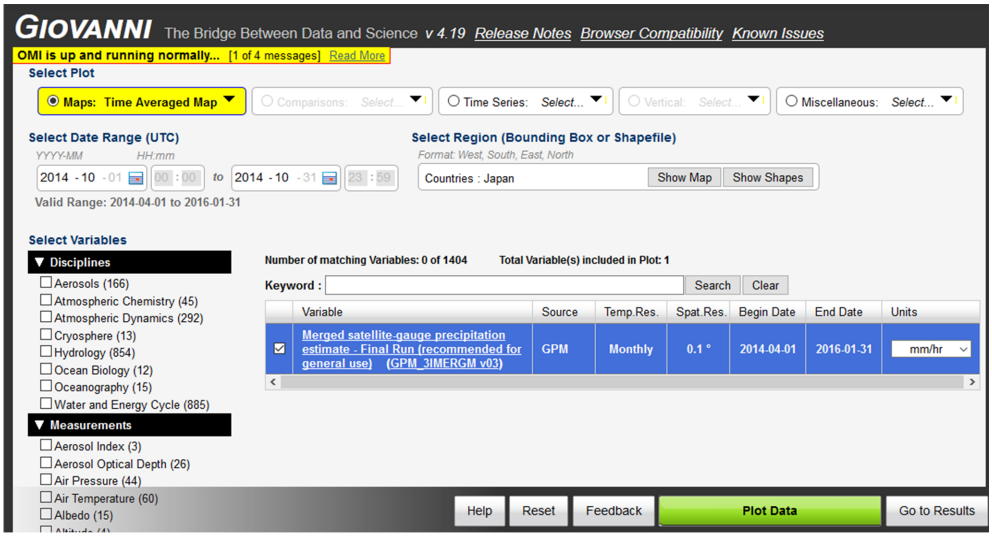


Figure 1. Samples of GPM datasets at different levels: a) Level-1C GMI common calibrated brightness temperatures at 37 GHZ showing Tropical Cyclone Nanauk over the Arabian Sea on 11 June 2014; b) Level-2 GMI GPROF surface precipitation showing Hurricane Arthur near the South Carolina and Georgia coasts on 3 July 2014; and c) Level-3 half-hourly IMERG Final precipitation showing heavy precipitation at 03Z 15 June 2014 in the Midwestern United States.

Figure 2. The Simple Subset Wizard (SSW) provides a simple way to subset Level-3 and Level-2 (limited) datasets not only from the GES DISC, but also other NASA data centers. a): The landing page of SSW. b): Sample output showing different options for subsetting data.

a)



b)

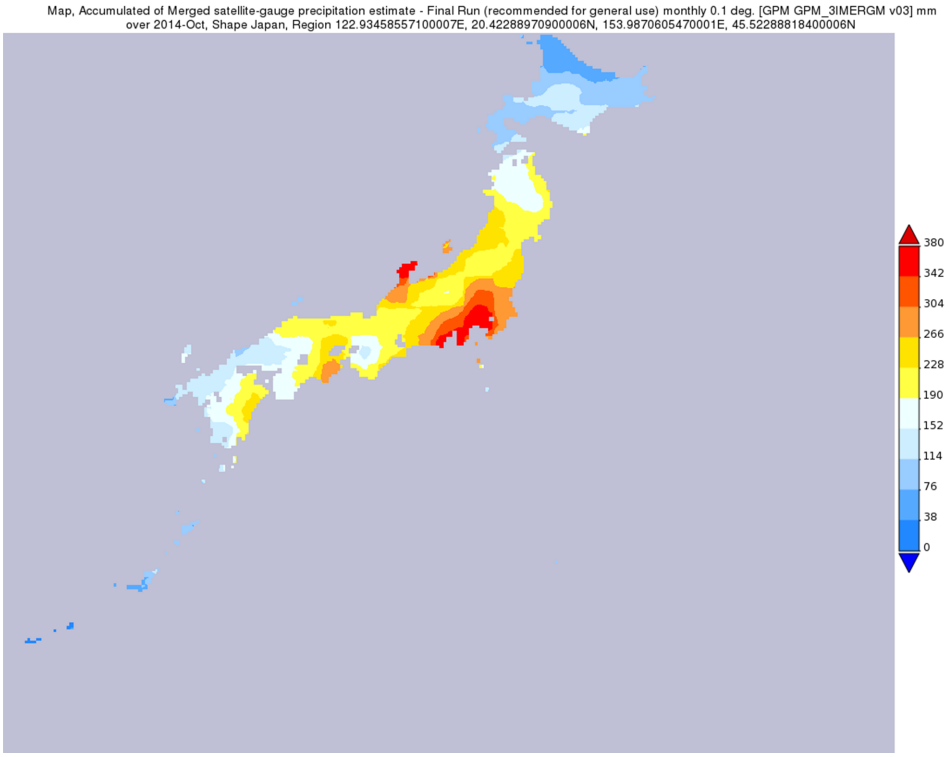


Figure 3. The GES DISC Giovanni allows visualization and easy access to IMERG data: a): A screenshot of the Giovanni landing page. Features such as keyword and facets make dataset search simple. b): A sample rainfall map from IMERG Final Run showing heavy rainfall (in mm) in the capital region of Tokyo due to the passages of Super Typhoons Phanfone and Vongfong in October 2014.

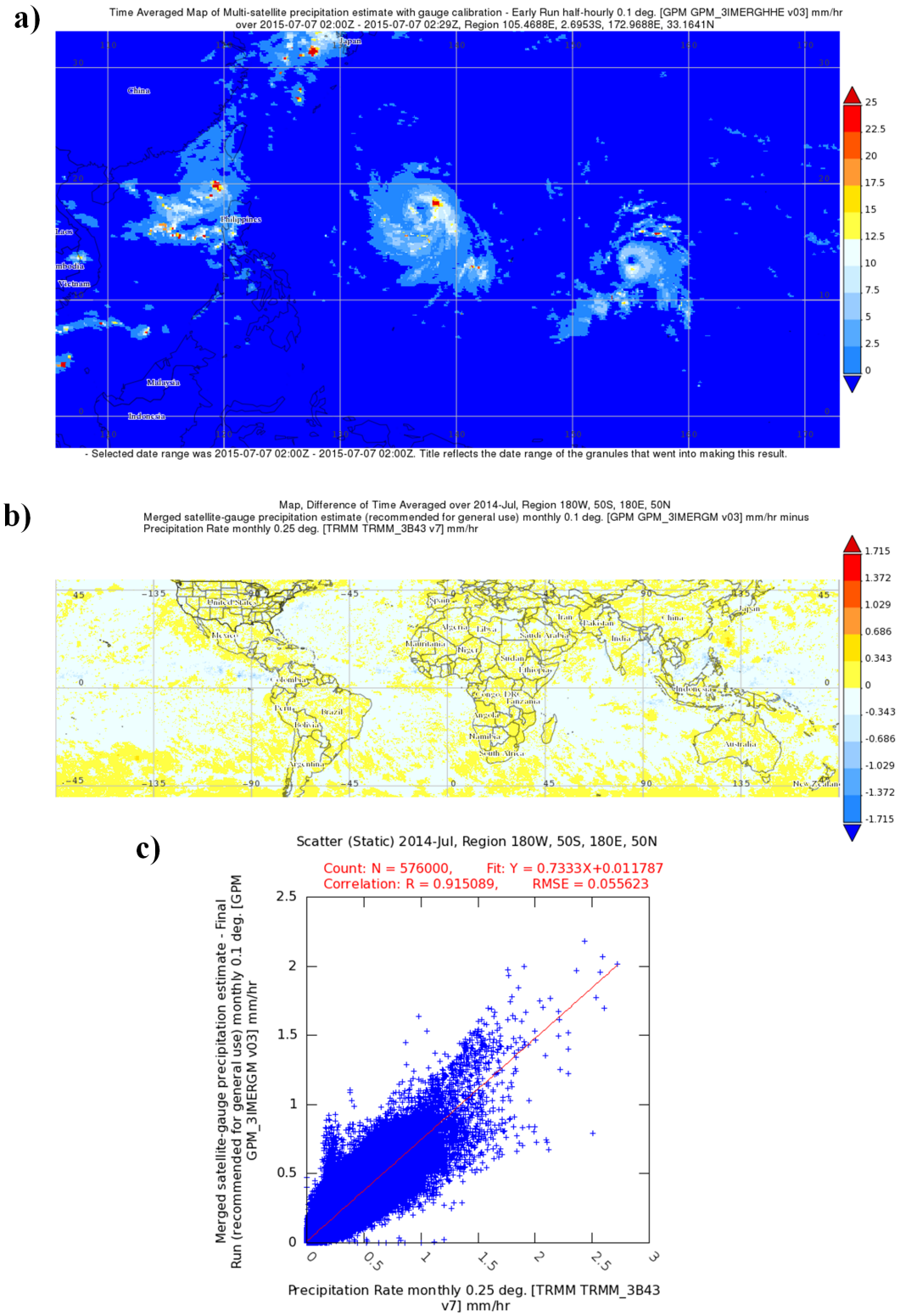


Figure 4. Samples of IMERG in Giovanni: a) Rainfall intensity (mm/hr) from the IMERG Early Run at 02Z 7 July 2015 showing a trio of typhoons in the Western Pacific; b) A monthly precipitation difference map (mm/hr) between 3B43 and IMERG Final Run for July 2014; and c) their scatter plot.