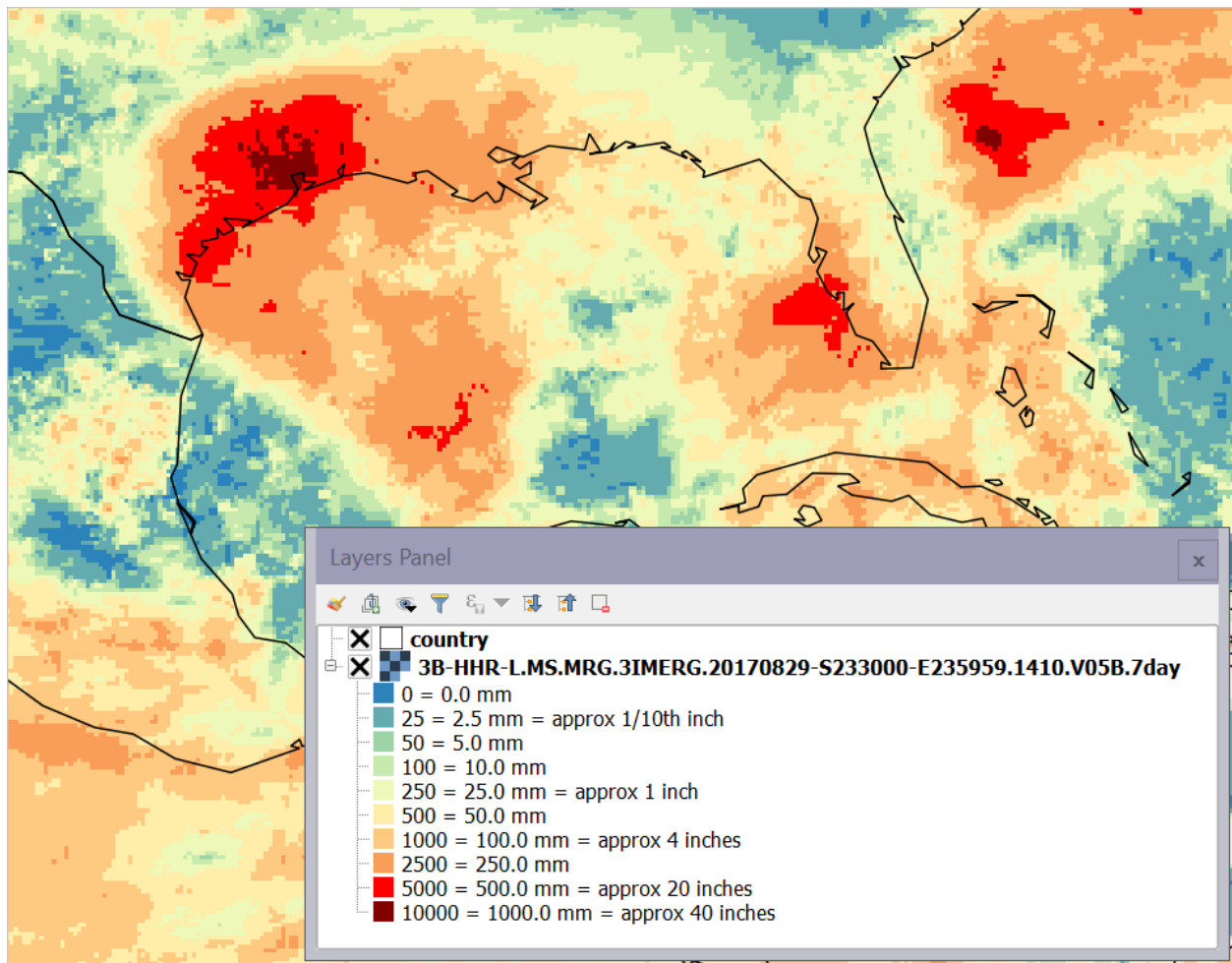# The IMERG multi-satellite precipitation estimates reformatted as 2-byte GeoTIFF files for display in a Geographic Information System (GIS)

The Precipitation Processing System (PPS) at NASA Goddard
https://pps.gsfc.nasa.gov/Documents/README.GIS.pdf

Updated on 22 April 2020
by Owen.Kelley@nasa.gov

Caption: The 7-day precipitation accumulation associated with Hurricane Harvey (2017) over Texas. This near-realtime file (Late IMERG GIS) covers the 7-day period ending at 2359 UTC on 29 August 2017. The data has 0.1° geographic resolution, 0.1 mm precision, and is displayed here using the QGIS application.

Contents

1. Introduction

NASA has worked for decades to improve global, high-resolution, satellite-based estimates of surface precipitation. It is difficult to quantify how much precipitation falls around the Earth. This is especially true if one specifies a short time interval, a high spatial resolution, or rapid dissemination (i.e., timeliness). The difficulty with obtaining such precipitation estimates arises because precipitation is often patchy and short-lived and because observations from multiple kinds of instruments must be combined to produce the best global estimates.

NASA's most advanced algorithm for this purpose is IMERG, the Integrated Multi-satEllite Retrieval for Global Precipitation Measurement (GPM), which was first released in 2014 (Huffman et al. 2019a, 2019b). The IMERG science algorithm writes its output in the Hierarchical Data Format Version 5 (HDF5) file format (http://www.hdfgroup.org/HDF5/). The HDF5 format widely used by the scientific community but is poorly supported by Geographic Information Systems (GIS) such as such as ESRI's ArcMap and the open-source QGIS application. Furthermore, GIS users often require periods of accumulation other than half-hourly or monthly, the only two periods covered by IMERG HDF5 files.

For these reasons, NASA produces a GIS translation of IMERG for various accumulation periods that maintains the 0.1-degree resolution and global coverage of the original HDF5 data product. In April 2019, Version 6 of IMERG was released in both the HDF5 and the GIS formats. Version 6 is the first version to cover the approximately 20-year period from June 2000 to the present. This long-duration archive will be a boon to scientific research and will provide an expanded training set for realtime applications such as disaster monitoring.

The purpose of this document is to describe the Geographic Information System (GIS) translation of the IMERG data product, hereafter referred to as "IMERG GIS." IMERG GIS is essentially a data file containing two-byte integer data stored as a Georeferenced Tagged Image File Format (GeoTIFF) file. A GeoTIFF file contain geographic metadata, but PPS also stores the same the geographic metadata in a separate Environmental Systems Research Institute (ESRI) WorldFile, a human readable text file. The WorldFile is for the use of GIS applications that are unable to read the geographic metadata that is stored inside of a GeoTIFF file.

IMERG GIS files are somewhat different than typical TIFF images, which often contain a 8-bit or 24-bit digital image with pixels indexed to a 256-element color table or assigned red-green-blue values. In contrast, IMERG GIS files are 16-bit TIFF files that contain grids of unsigned-two-byte-integer data with values from 0 to 65535. Some kinds of IMERG GIS files contain data that is scaled by 10 or 1000 and then rounded to the nearest integer. This scaling allows the TIFF file to preserve an additional digit or digits of accuracy.

Within a few hours of the satellite observations being collected, the Precipitation Processing System (PPS) creates "realtime" IMERG, which is supplemented by the higher-

quality "research" IMERG that PPS creates several months after the observations are collected. Realtime IMERG is intended for applications such as disaster monitoring and flood and landslide risk assessment. Research IMERG is intended for meteorological and climatological research. In this document, the terms "realtime" and "research" are used in an informal sense. The official terms are "Early IMERG" and "Late IMERG," which are created in near-realtime, and "Final IMERG," which is intended for use by researchers, as described in Section 1.1.

While the purpose of this document is to describe the GIS translation of IMERG, some information about the original IMERG HDF5 files is provided as well. The science behind the IMERG algorithm is described in NASA's Algorithm Theoretical Basis Document (ATBD) that is available on the PPS website (http://pps.gsfc.nasa.gov/). IMERG HDF5 files are the official multi-satellite data product of the U.S. GPM team. PPS creates IMERG HDF5 files for periods of either 30 minutes or 1 month. PPS translates the research IMERG into 30-minute, daily, or monthly files in a GIS-friendly format. In near-realtime, PPS creates GIS precipitation accumulations for these same period and also for several intermediate periods (3 hours to 7 days) that are intended to assist with disaster-risk assessment (i.e., floods and landslides). The time periods of Early, Late, and Final IMERG GIS are depicted in Figure 1. Both the HDF5 and GIS versions of IMERG contain rectangular latitude-longitude grids that have 3600 columns for longitude and 1800 rows for latitude. This grid covers the globe (90°S to 90°N and 180°W to 180°E) at 0.1° resolution.

The IMERG HDF5 files contain variables that describe the surface precipitation rate and the phase of the precipitation reaching the Earth's surface. In the HDF5 files, precipitation rate is expressed in millimeters per hour. In the 30-minute HDF5 files, the likelihood that the precipitation reaching the Earth's surface is in either liquid phase or mixed phase, not ice phase, is expressed as a percentage (0% to 100%). See section 3.3 of this document for details about how precipitation phase is represented in IMERG GIS files.

The IMERG HDF5 files are generated from a constellation of polar-orbiting satellites carrying passive microwave sensors, from geosynchronous satellites carrying infrared sensors, and from calibration data derived from either satellite observations or rain gauges. At NASA Goddard, PPS generates the IMERG HDF5 files and also a legacy data product called TMPA that was first created in 1998. TMPA stands for Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis. Section 5 of this document describes the TMPA GIS data product, while the rest of the document describes the IMERG GIS data product.

IMERG is the most popular data product among more than a dozen data products associated with the GPM mission and generated by PPS at NASA Goddard. Between April 2015 and 2019, about 10% of realtime IMERG files that were downloaded were in the GIS format and the other 90% were in the HDF5 format (a total of over 100 million Early and Late IMERG file downloads). During the same period, approximately 1% of the ~100-million Final IMERG file downloads were in the GIS format, with the other 99% in the HDF5 format.

A team of scientists works to improve the IMERG algorithm. PPS began running Version 6 of the IMERG algorithm in the spring of 2019. Version 6 IMERG is the first version to be

available for both the TRMM era and the GPM era, i.e., June 2000 to the present. Versions 5, 4, and 3 IMERG were released in late 2017, early 2017, and late 2016, respectively, and they were generated only for the GPM era (April 2014 onward). Research and realtime applications should use the latest version of IMERG because it is believed to be the most accurate. Publications should always state the version of IMERG being used, to avoid ambiguity.

In IMERG HDF5 files, missing-data values (-9999.9) fill the portion of the grid where precipitation estimates cannot be calculated. Version 6 is the first version of IMERG to morph microwave-based precipitation estimates that occur poleward of 60° latitude, rather than only report microwave-based precipitation estimates at the time of observation at these high latitudes. Still, Version 6 uses missing-data values over any snow-covered continent and ice-covered ocean that lies north of 60°N or south of 60°S.

Questions related to the science of the IMERG algorithm should be sent to Dr. George Huffman (george.j.huffman@nasa.gov) at NASA Goddard. Questions about the GIS reformatting of IMERG can be sent to the PPS Helpdesk (helpdesk@mail.pps.eosdis.nasa.gov). For more information about GPM, please visit the Precipitation Measurement Missions (PMM) science-team website (http://pmm.nasa.gov).
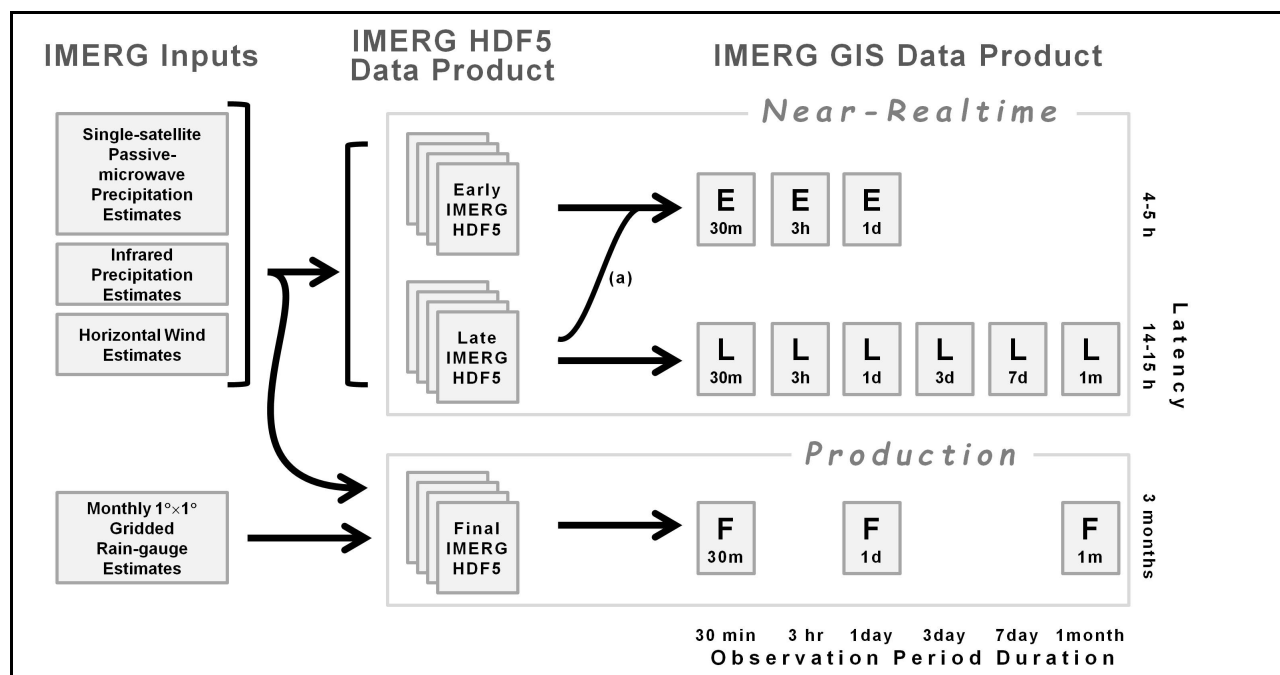
Figure 1. Schematic diagram of the inputs that go into the IMERG HDF5 and IMERG GIS data products. Line (a) indicates that Early IMERG GIS product uses the Early IMERG HDF5 product only for any 30-minute time period for which the higher-quality Late IMERG HDF5 product is not yet available.

1.1. The Early, Late, and Final versions of IMERG: latency and retention

Latency is the amount of time that passes between an observation being collected by the satellite and a data product being produced. There are three flavors of IMERG, each with a different latency. These three flavors are called Early, Late, and Final and their associated latencies are approximately 6 hours, 18 hours, and 4 months, respectively. In the Early IMERG data product, some regions that lack high-quality microwave estimates of precipitation are filled in with regions of microwave-estimated precipitation morphed in from previously observed high-quality observations. Where morphed-microwave information is not available, IMERG grid cells are filled with estimates that are based on geosynchronous infrared sensors. The longer latency of Late IMERG allows high-quality microwave-based precipitation estimates to be morphed both forward and backward in time to reach the time of interest. The much longer latency of Final IMERG allows the precipitation estimates to be calibrated by ancillary data sets (such as rain gauges) that become available months after the satellite and gauge observations were collected.

PPS takes several actions to assist developers using near-realtime IMERG. Specifically, PPS does not delete realtime IMERG HDF5 or GIS products once the higher-quality Final IMERG product is available. Realtime IMERG HDF5 and GIS products (Early and Late) are retained until a new version of the Early and Late IMERG algorithm replaces them. This retention period gives application developers a long-duration data set for calibrating their realtime applications. Researchers, however, should use the Final IMERG product whenever possible instead of Early or Late IMERG because of the higher accuracy of Final IMERG.

Since Version 5 IMERG, PPS creates Early and Late IMERG GIS files back to the beginning of the GPM mission whenever a new version of the IMERG algorithm is delivered, rather than just start Early and Late IMERG GIS processing with current observations. This long-term data set is called the "retrospective" run of Early and Late IMERG GIS. The retrospective run is a service to developers, to assist with calibrating their realtime applications. The retrospective run of Early and Late IMERG GIS may be downloaded from the same online archive as the most recently generated Early and Late IMERG GIS files.

1.2. IMERG HDF5 variables used to create IMERG GIS files

The IMERG GIS algorithm reads two variables from IMERG HDF5 files in order to create IMERG GIS files. These two HDF5 variables are called "/Grid/precipitationCal" and "/Grid/probabilityLiquidPrecipitation." The precipitationCal variable is an estimate of surface precipitation rate based infrared satellite observations, microwave satellite observations, and any calibration data sets available at the time that the IMERG science algorithm is run. The precipitationCal variable contains the best available estimate of surface precipitation. For a particular half-hour period, the probabilityLiquidPrecipitation variable is the probability that all of the precipitation fell in liquid phase or mixed phase (not ice phase) within a given grid box.

2. How to Download IMERG GIS Files from PPS

2.1. Downloading IMERG GIS files

Before downloading data files from the Precipitation Processing System (PPS), register your email address with PPS by visiting http://registration.pps.eosdis.nasa.gov/ . After registering an email address, one can download Final IMERG GIS files using the PPS data-ordering system called STORM: https://storm.pps.eosdis.nasa.gov/ . Other options for downloading Early, Late, and Final IMERG GIS files are given below. Please note that as of June 1, 2020, realtime IMERG GIS files (Early IMERG GIS and Late IMERG GIS) will no longer be available for download via FTP. PPS has made other download options available as of March 2020 because NASA has mandated that NASA data centers phase out the use of FTP servers.

Through the summer of 2020, FTP will remain the preferred method for downloading Final IMERG GIS, even after Early and Late IMERG GIS are only available through non-FTP mechanisms starting on 1 June 2020. To download Final IMERG GIS, visit ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/ using a web browser or connect to arthurhou.pps.eosdis.nasa.gov using some other FTP client, such as the Linux "ftp" command. Then, descend into the year, month, and day of interest. Last, go into the "gis" directory. For example, for 1 June 2014, one would go to the "gpmdata/2014/06/01/gis" directory.

From 2014 until the end of May 2020, a way to download realtime IMERG GIS (Early IMERG GIS and Late IMERG GIS) is to visit ftp://jsimpson.pps.eosdis.nasa.gov/data/imerg/gis/ using a web browser. Alternatively, connect to jsimpson.pps.eosdis.nasa.gov using some other FTP client such as the Linux "ftp" command and then go to the "data/imerg/gis" directory.

Starting in March 2020, you have three new options for downloading realtime IMERG GIS. If one wishes to use a web browser interactively, you can use the first of the three new options. Go to https://jsimpsonhttps.pps.eosdis.nasa.gov/ and give your PPS-registered email address when prompted for username and password. You will be presented with a web page with clickable links that you can use to tunnel down into a directory that contains the files you are interested in. Within the data/imerg/gis directory, the most recent year of Late IMERG GIS files are placed in the "01" through "12" subdirectories which indicate January through December. Previous years of Late IMERG GIS are placed in the year-month (yyyy/mm) subdirectory tree of "data/imerg/gis". The Early IMERG GIS files are all placed in the "data/imerg/gis/early" directory. The most-recent month of Early IMERG GIS is placed directly in "data/imerg/gis/early". Previous months of Early IMERG GIS are placed in the year-month (yyyy/mm) subdirectory tree of "data/imerg/gis/early".

The second new option for downloading realtime IMERG GIS is best if you are writing a script to navigate the PPS archive and download files. One might use Python, BASH, IDL or other languages for this script. Have your script go to this URL https://jsimpsonhttps.pps.eosdis.nasa.gov/text/, which is the same as the previously mentioned

URL with the string "text/" added to its end. In this representation of the archive, PPS will send a text file with one item per line when you are looking at a directory in the archive. PPS will send the data file as a stream when one looks at a file in the archive. By successively tacking on directory names to the URL, you can descend to the level of the archive where the contents of a directory are data files rather than subdirectories.

The third new option for downloading realtime IMERG GIS is best if you can ask your system administrators to alter firewall rules to allow FTPS connections and if you system has an FTPS client installed such as the "curl" command on linux and Mac OS X systems. The syntax for FTPS access to the PPS Realtime archive varies from one operating system to another, and it will be up to you to figure out the exact syntax for your operating system. Here are an example that you might use as a starting point:

```
curl -O --ssl-reqd -u your.email@address.here:your.email@address.here
   ftp://jsimpsonftps.pps.eosdis.nasa.gov/data/imerg/gis/2016/06/3B-
   DAY-L.GIS.IMERG.20160602.V06B.zip
```

Where `-O` indicates that the file written to your hard disk is named the same as the file in the PPS archive, `--ssl-reqd` or `--ftp-ssl-reqd` indicates a secure connection will be used, -u indicates that the username:password sequence will follow. Insert the email address that you registered with PPS here as both the username and password, separating them with a colon character.

2.2. GeoTIFF vs. WorldFiles

Each IMERG GIS grid is stored in a separate GeoTIFF file. A GeoTIFF file has a *.tif file extension. For GIS software to properly display the GeoTIFF file, you may need to download from PPS the *.tfw WorldFile metadata file whose basename matches the GeoTIFF file. The basename is the portion of the filename that appears before the *.tif or *.tfw filename extension. A WorldFile is a small ASCII text file that contains geographic metadata (Wikipedia 2015). If your GIS application can read the geographic metadata stored within the GeoTIFF file, then there is no need to download the WorldFile. The *.zip compressed files in the online archive contain pairs of GeoTIFF and WorldFile files.

3. IMERG GIS Filenames and File Format

3.1. Scaling of precipitation rate in Final IMERG GIS

IMERG GIS files store precipitation as one-byte or two-byte integers because some GIS applications historically have had difficulty reading TIFF files that contain floating-point values.

When necessary to preserve accuracy, the precipitation estimates are scaled by 10 or 1000 before being rounded off and written as integers to the TIFF file, as described below and in Table 1.

Thirty-minute and 1-day Final IMERG GIS files contain precipitation rates that are scaled by 10 before they are written to the two-byte TIFF file. A 1-month Final IMERG GIS file contains monthly-average precipitation rates that were scaled by 1000 before being written to the two-byte TIFF file. The reason for the larger scaling factor for monthly averages is that the monthly-average precipitation rate is typically much smaller (near ~1 mm/h) than is the average precipitation rate of an individual storm (10 to 100 mm/h can occur).

For example, a rain rate of 1.0 mm/h in a 30-minute or 1-day Final IMERG HDF5 file would be stored as a two-byte integer value of 10 in the 30-minute or 1-day Final IMERG GIS TIFF file. A rain rate of 1.0 mm/h in a monthly Final IMERG HDF5 file would be stored as a two-byte integer value of 1000 in the monthly Final IMERG GIS TIFF file. In other words, the units of a 30-minute or 1-day Final IMERG GIS file are 0.1 millimeters per hour. The units of a monthly Final IMERG GIS file are 0.001 millimeters per hour.

For consistency with Early and Late IMERG GIS, PPS also stores precipitation accumulation in millimeters in Final IMERG GIS files, not just average precipitation rate.

3.2. Scaling of precipitation accumulation in Early and Late IMERG GIS

Realtime IMERG (i.e., Early and Late IMERG) has precipitation accumulation stored in its GIS translation, not precipitation rate. For accumulation periods of 30 minutes to 7 days, the precipitation accumulation in millimeters is multiplied by 10 before being rounded off to make a two-byte integer that is written to the TIFF file. For example, an accumulation of 1 millimeter would be written as 10 in the two-byte TIFF file of the Early or Late IMERG GIS data product. In other words, the units of an Early or Late IMERG GIS file are 0.1 millimeters. For full-month Late IMERG accumulations, the data values are stored unscaled, i.e. a 10 mm accumulation would be stored as "10" in the full-month TIFF file.

Precipitation accumulation calculated for various time periods can be useful for assessing flood risk or landslide hazards in realtime (Hong et al. 2007), and this is the reason for converting rate to accumulation in Early and Late IMERG GIS files. Every time that PPS generates a 30-minute realtime IMERG HDF5 file, PPS immediately creates a pair of 30-minute GIS files. Every 3 hours, PPS also generates an updated accumulation for the 3-hour to 7-day accumulation periods listed for Early and Late IMERG in Table 1. In this way, realtime users have up-to-date accumulations for various durations. At the end of a month, a full-month accumulation is created for Late IMERG GIS.

Table 1. The precipitation variable and units of various IMERG GIS files.

Official name	GIS file's precipitation variable and period	Precipitation scaling factor (physical units)	Integers that may be stored in the GIS file [a]	Corresponding precipitation values [b]

Early IMERG	Accumulation in
	  30 minutes
	  3 hours
	  1 day [c]
		x10 (0.1 mm)	0–29998	0.0–2999.8 mm

Late IMERG	Accumulation in
	  30 minutes
	  3 hours
	  1 day
	  3 days
	  7 days
		x10 (0.1 mm)	0–29998	0.0–2999.8 mm

	Accumulation in 1 month	x1 (1 mm)	0–29998	0.0–29998. mm

Final IMERG	Rate averaged over
	  30 minutes
	  1 day
	  1 month
		x10 (0.1 mm/h)
		x10 (0.1 mm/h)
		x1000 (0.001 mm/h)
			0–29998
			0–29998
			0–29998
				0.0–2999.8 mm/h
				0.0–2999.8 mm/h
				0.0–29.998 mm/h

	Accumulation in
	  30 minutes
	  1 day
	  1 month
		x10 (0.1 mm)
		x10 (0.1 mm)
		x1 (1 mm)
			0–29998
			0–29998
			0–29998
				0.0–2999.8 mm
				0.0–2999.8 mm
				0.0–29998. mm

[a] In the 2-byte unsigned TIFF files, a value of 29999 is used to indicate "missing" data.

[b] These very large ranges for precipitation rate and accumulation can be represented in the file, although the ranges that are likely to be encountered are smaller, as discussed in Section 4 of this document.

[c] The 1-day, 3-day, and 7-day references in this table refer to periods that are exactly 24, 3×24, or 7×24 hours long.

3.3. Precipitation-phase variables (liquid vs. solid precipitation)

The precipitation rates and accumulations discussed in this document are all liquid-equivalent rates and depths. For example, an IMERG file that reports 1 inch of frozen precipitation (i.e., 25.4 mm) means that whatever the depth that this precipitation had in its original frozen form, this precipitation would have a 1-inch depth were it melted first before being measured. According to section 3.10 of the IMERG Algorithm Theoretical Basis Document (ATBD), the current IMERG algorithm groups together mixed-phase and liquid-phase precipitation. Since hail is usually mixed with rain, IMERG categorizes all of the precipitation from hail events as liquid / mixed-phase.

As described below, the precipitation-phase variables in a Final IMERG HDF5 file should be interpreted differently depending on whether that HDF5 file covers a 30-minute period or a 1-month period (Section 3.10, IMERG ATBD). Similar considerations cause the precipitation-phase accumulated from multiple 30-minute Early or Late IMERG HDF5 files to be handled differently for accumulation periods that are longer than 1 day vs. accumulations periods that are up to 1 day long.

The IMERG GIS algorithm handles precipitation-phase information in a way that follows logically from the precise meaning of the precipitation-phase information that is stored in a single 30-minute IMERG HDF5 file. In a 30-minute IMERG HDF5 file, the precipitation-phase variable is the percentage likelihood ($PERCENT$ = 0% to 100%), at each 0.1° grid box, that all of the precipitation reaching the Earth's surface during that 30-minute period is in liquid phase or mixed phase. The complement of that percentage (100% − $PERCENT$) is the percentage likelihood that all of the precipitation that reached the Earth's surface in that 0.1° grid box was frozen during that 30-minute period. It follows from this definition of precipitation phase that one should use a 50% threshold (or a similar percentage) to decide if the precipitation in short-duration accumulations (30 minutes to 1 day) is 100% liquid/mixed phase or 100% frozen. This method of determining precipitation phase is subsequently called the "50%-threshold method."

In a monthly IMERG HDF5 file, the precipitation-phase variable ($PERCENT$) is computed as the percentage of total monthly accumulation that fell in liquid/mixed phase. If a monthly IMERG HDF5 file reported a precipitation rate p (liquid-equivalent depth), then one should assume that a rate of p $PERCENT$ of precipitation reached the Earth's surface in liquid/mixed phase and that p (1 − $PERCENT$) of precipitation reached the Earth's surface frozen. This method of calculating precipitation phase is subsequently called the "product method" because it takes the product of the probability and the total precipitation rate that are both stored in the IMERG HDF5 file.

If one is interested in the phase of the precipitation, one should download the compressed file with the *.zip extension. The *.zip file contains GeoTIFF and WorldFiles (*.tif and *.tfw files) for total precipitation (TP), liquid/mixed precipitation (LP), liquid-equivalent ice-phase-only precipitation (IP), and the percent of liquid/mixed-phase vs. total precipitation ($PERCENT$).

To make the four IMERG GIS variables consistent, the following relationships always hold true. The total-precipitation file is always equal to the sum of the liquid/mixed-phase

precipitation file and the liquid-equivalent ice-phase precipitation file. The liquid-precipitation percent file is always 100 times the ratio of the liquid/mixed-precipitation file and the total-precipitation file. These two relationships are expressed in Equation 1.

$$TP = LP + IP$$

$$PERCENT = 100\frac{LP}{TP}$$

(1)

The developer of the IMERG science algorithm recommends that the product method generally should be used for "long" periods of time (longer than 1 day) while the previously-mentioned 50%-cutoff method should be used for "short" periods of time (up to and including 1-day long). The rest of this section describes the mathematical details that implement these ideas. In the following equations, the values read from the input HDF5 IMERG file are rate (mm/h) and percent (0% to 100%). The notation for these read-in input values are TP_0 and $PERCENT_0$ if one input HDF5 file is used and TP_i and $PERCENT_i$ if more than one input HDF5 file is used. The output values written to the IMERG GIS files are total precipitation (*TP*), liquid/mixed-phase precipitation (*LP*), liquid-equivalent ice-phase precipitation (*IP*), and percent-liquid precipitation (*PERCENT*).

Equation 2 implements the 50%-threshold method. For 30-minute-duration Early and Late IMERG GIS accumulations, use Equation 2 with $k = 0.5$ hour, a value for k that converts a precipitation rate (mm/h) into a half-hour accumulation (mm). For 30-minute-duration Final IMERG GIS averages, $k = 1$ (unitless) in Equation 2 since both the input variable and output variables are precipitation rates (mm/h).

$$TP = k\,TP_0$$

$$LP = \begin{cases} TP & \text{if } PERCENT_0 \geq 50\% \\ \\ 0 & \text{if } PERCENT_0 < 50\% \end{cases}$$

$$IP = TP - LP \tag{2}$$

$$PERCENT = \begin{cases} 255 & \text{if } TP_0 = 0 \\ \\ 100 & \text{if } PERCENT_0 \geq 50\% \text{ and } TP_0 > 0 \\ \\ 0 & \text{if } PERCENT_0 < 50\% \text{ and } TP_0 > 0 \end{cases}$$

An "undefined" value of 255 is used for *PERCENT* when there is zero precipitation.

Equation 3 is used for the case of several 30-minute IMERG HDF5 files as input, but few enough that the 50%-threshold method is still applicable. In Equation 3, terms with subscript i are the input grids read from the ith individual 30-minute IMERG HDF5 file from among the n 30-minute IMERG HDF5 files read. Terms without a subscript are the output grids written to a 3-hour or 1-day IMERG GIS file. In Equation 3, $k = 0.5$ hour for Early IMERG GIS and Late IMERG GIS. In contrast, $k = 1/n$ (unitless) for Final IMERG. Final IMERG GIS files contain averages, and to calculate the average of n terms, one must divide the sum of the terms by n. For a 3-hour period, $n = 6$ and for a 1-day period, $n = 48$.

$$TP = k \sum_{i=1}^{n} TP_i \quad \text{for } n = 6 \text{ or } 48$$

$$b_i = \begin{cases} 1 & \text{if } PERCENT_i \geq 50\% \\ 0 & \text{if } PERCENT_i < 50\% \end{cases}$$

$$LP = k \sum_{i=1}^{n} b_i TP_i \quad \text{for } n = 6 \tag{3}$$

$$IP = TP - LP$$

$$PERCENT = \begin{cases} 255 & \text{if } TP = 0 \\ 100 \dfrac{LP}{TP} & \text{if } TP > 0 \end{cases}$$

Equation 4 is used when there are multiple 30-minute IMERG HDF5 files as input and the time period is long enough to use the product method rather than the 50%-threshold method. Equation 4 is only used for the 3-day, 7-day, and monthly accumulations calculated from Late IMERG HDF5 files. In Equation 4, n equals (48 d), where d is 3, 7, or 28–31, i.e., the number of days covered by the accumulation period. In Equation 4, $k = 0.5$ hours because Late IMERG GIS contains accumulations, as described earlier in this section.

$$TP = k \sum_{i=1}^{n} TP_i$$

$$LP = k \sum_{i=1}^{n} \frac{PERCENT_i}{100} TP_i$$

$$IP = TP - LP \tag{4}$$

$$PERCENT = \begin{cases} 255 & \text{if } TP = 0 \\ 100 \dfrac{LP}{TP} & \text{if } TP > 0 \end{cases}$$

14

Equation 5 is for the case that there is a single, monthly IMERG HDF5 file as the input. A month is considered a long time-period for the purpose of this calculation, so the product method is used rather than the 50%-threshold method. Only Final IMERG (not Early or Late IMERG) has a monthly HDF file, so Eq. 5 is used to create only the monthly Final IMERG GIS file.

$$TP = TP_0$$

$$LP = \frac{PERCENT_0}{100} TP_0$$

$$IP = TP - LP \tag{5}$$

$$PERCENT = 100\frac{LP}{TP} = PERCENT_0$$

3.4. Filenames of the GIS translation of Early and Late IMERG

The filename convention is different for IMERG GIS files created in realtime (i.e., Early and Late IMERG) than for IMERG GIS files created months later (i.e., Final IMERG).

Based on the way that realtime IMERG GIS is used, it makes sense that the end datetime of the accumulation period is part of the filename of Early or Late IMERG GIS files. Realtime users often want precipitation estimates that are as close as possible to the present. Realtime users can easily estimate how "fresh" an IMERG GIS file is by looking at the datetime stamp in the GIS filename. For Early and Late IMERG GIS accumulation from 30 minutes to 7 days long, the filenames state the start datetime and end time of the last 30-minute period included in that accumulation.

As soon as a new hour of Early or Late IMERG HDF5 file is created on the PPS Realtime System, PPS creates two 30-minute Early IMERG GIS accumulations and two 30-minute Late IMERG GIS accumulations. Every 3 hours, PPS also creates a set of Early and Late IMERG GIS accumulations for periods of 3 hours to 7 days. When the last Late IMERG HDF5 file of a month is created, PPS creates a Late IMERG GIS month accumulation. Table 2 describes the filename patterns for IMERG GIS files. Figure 2 is a schematic diagram of when the IMERG GIS files are created for various time periods near the end of a month.

Table 2. The filename patterns for IMERG GIS files. [a]

Time-period boundary	Duration	Early and Late IMERG (i.e. "realtime" IMERG)	Final IMERG (i.e., "research" IMERG)
A 30-minute time period starting at 0000, 0030, 0100, ... 2300, 2330	30 minutes	3B-HHR-E.*A-S$hhmm$*.V.30min.X [b] 3B-HHR-L.*A-S$hhmm$*.V.30min.X [b]	3B-HHR-GIS.*A-S$hhmm$*.V.X [b]
Any period that begins at 0, 3, 6, 9, 12, 18, or 21 UTC	3 hours 1 day 3 hours 1 day 3 days 7 days	3B-HHR-E.*B-Shh30*.V.3hr.X 3B-HHR-E.*B-Shh30*.V.1day.X 3B-HHR-L.*B-Shh30*.V.3hr.X 3B-HHR-L.*B-Shh30*.V.1day.X 3B-HHR-L.*B-Shh30*.V.3day.X 3B-HHR-L.*B-Shh30*.V.7day.X	[not applicable]
Day boundary, 0000 UTC	1 day	3B-DAY-L.*A-S0000*.V.X [c]	3B-DAY-GIS.*A-S0000*.V.X
Month boundary, 0000 UTC	1 month	3B-MO-L.*A-S0000*.V.X	3B-MO-GIS.*A-S0000*.V.X

[a] In this table, "1 day" means "24 hours." The following symbols are used in this table:

    E refers to Early IMERG

    L refers to Late IMERG

    A = start date of the period covered by the GIS file, in yyyymmdd format

    B = start date of the last day in the multi-day period covered by the GIS file, in yyyymmdd format

    V = version number of the IMERG HDF5 data product such as "V05A"

    X = file extension of the GIS file, which can be tif, tfw, or zip, as discussed later in this section of the document

[b] For 30-minute-long periods, A=B, i.e., the start date of the last 30-minute portion of the period is the same as the start date of the full period. The symbol *hh* represents 02, 05, 08, 11, 14, 17, 20, or 23 hours. The symbol *mm* represent 00 or 30 minutes.

[c] For Late IMERG GIS, at 0 UTC each day, there are two copies of the 1-day accumulation. One copy follows the PPS standard naming convention that uses the start date/time of the period of coverage. The other copy uses the GIS realtime convention, specifically that the filename gives the last 30-minute granule within the period of coverage.
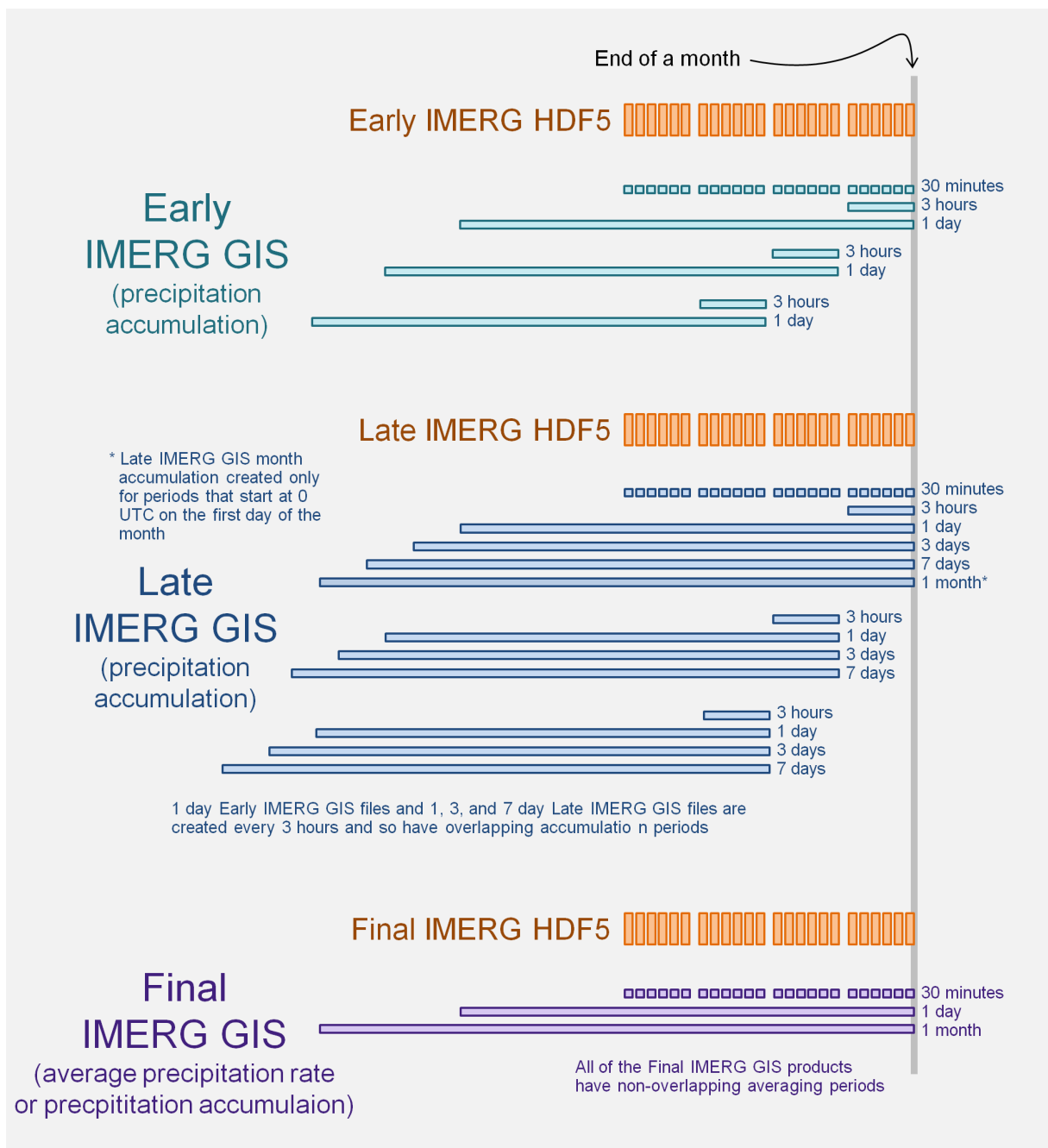
Figure 2. A schematic diagram of some of the IMERG GIS files that are created near the end of a month. The IMERG GIS files are shown in green, blue, and purple for Early, Late, and Final IMERG, respectively. The 30-minute input files (HDF5 files) are shown in orange. The horizontal axis represents time, with the vertical gray line indicating the end of the month, and observations earlier than that to the left of the vertical gray line.

Consider a set of Early and Late IMERG GIS accumulations that all end at the same time. Their filenames would be identical except for the last few characters. In other words, the "root" of their filenames is the same, and only the suffix varies. As described in greater detail below, part of the suffix is the duration of the accumulation, part indicates the variable sorted in the GIS file, and part is the file extension.

The last (or only) IMERG HDF5 file in an Early or Late IMERG GIS accumulation determines the root of the filename. One could describe the filename of an IMERG HDF5 as consisting simply of a *Root* and a *fileExtension*. For example, the *Root* could be "3B-HHR-E.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V04A", while the *fileExtension* is ".RT-H5" for all realtime IMERG HDF5 files.

One takes the root of an IMERG HDF5 filename and creates an IMERG GIS filename by tacking on three items: a *Duration*, a *Variable*, and a *fileExtension*. Periods separate the parts of the filename when a part is not a blank string. The *Duration* is one of the following strings for Early or Late IMERG GIS files:

String	Meaning
"30min"	30 minutes
"3hr"	3 hours
"1d"	1 day (24 hours)
"3d"	3 days (72 hours)
"7d"	7 days (168 hours)

The *Variable* field is one of the following strings for Early and Late IMERG GIS files:

String	Meaning
[blank]	Total precipitation accumulation
"liquid"	Liquid-phase and mixed-phase precipitation accumulation
"ice"	Liquid-equivalent ice-phase precipitation accumulation
"liquidPercent"	Percent chance that the precipitation is liquid or mixed-phase in 30-minute to 1-day periods. For periods of 3 days or longer, the percent of precipitation that is liquid or mixed-phase.

The *fileExtension* field for an IMERG GIS file is one of the following strings:

String	Meaning
"tif"	The file is a GeoTIFF file
"tfw"	The file is an ESRI WorldFile
"zip"	The file is a zip compressed file
"txt"	For Early and Late IMERG only, a text file that is present when some of the expected 30-minute-duration HDF5 files were missing. In this situation, the

text file states how many 30-minute-duration HDF5 files were used to generate the GeoTIFF file.

A zip file can be uncompressed using the "unzip" Linux command or by double-clicking on the file in a Microsoft Windows system. The zip format is used to store groups of GIS files rather than the tar or gzip formats because all Microsoft Windows systems have native applications for uncompressing zip files while some Microsoft Windows systems do not have tar or gzip applications.

Along with the above-mentioned IMERG GIS files, there are two additional kinds of Late IMERG GIS files: day and month files that start on day or month boundaries. First, the Late IMERG 24-hour accumulation that covers 0000 UTC to 2359 UTC is placed in a zip file with a name that follows this pattern: 3B-DAY-L.*yyyymmdd.Vxxx.zip, where xxx is the version number such as "06A." Second, the Late IMERG accumulation for a full month is placed in a zip file with a name that follows this pattern: 3B-MO-L.*yyyymm01.Vxxx.zip.

3.5. Filenames for the GIS translation of Final IMERG

PPS creates overlapping accumulations for realtime IMERG GIS because realtime applications need up-to-date information, e.g., 24-hour accumulations updated more than once each day. However for research IMERG, there has not been a demand for overlapping GIS files. For this reason, there is only one daily research IMERG GIS file containing total precipitation rate, and its averaging period starts at 0000 UTC on each day.

For research IMERG GIS files, the three different averaging periods (30 minutes, 1 day, and 1 month) are indicated by the filename prefix rather than by adding a duration suffix to the filename. See Figure 2 for a schematic representation of the time periods covered by Final IMERG GIS files. Below is the pattern for the filenames for the three averaging periods that begin on 0000 UTC on 1 June 2014:

Filename root for GIS file   Duration of GIS file

3B-HHR-GIS.*.20140601-S000000-E002959   30 minutes

3B-DAY-GIS.*.20140601-S000000-E235959   1 day

 3B-MO-GIS.*.20140601-S000000-E235959   1 month

To view the total precipitation rate, download the *.tif and *.tfw files from the online archive. To view precipitation-phase information, download the *.zip file. Here are sample filenames:

3B-HHR-GIS.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.tif

3B-HHR-GIS.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.tfw

3B-HHR-GIS.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.zip

Inside the above-mentioned *.zip file are the following four GeoTIFF files for total precipitation rate and for various precipitation-phase related variables (and the associate WorldFiles that are not listed here):

3B-HHR.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.tif

3B-HHR.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.liquid.tif

3B-HHR.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.ice.tif

3B-HHR.MS.MRG.3IMERG.20140601-S000000-E002959.0000.V06A.liquidPercent.tif

Among these files, the first file is contains the total precipitation rate, the *.liquid.tif file contains the precipitation that reached the surface in liquid or mixed phase, the *.ice.tif file contains precipitation that reached the surface in ice phase, and the *.liquidPercent.tif extension indicates the fraction of the total precipitation that reached the surface in liquid or mixed phase.

3.6. Missing Data

A value of either 255 or 29999 is used to indicate missing data in IMERG GIS files depending on whether the data-type is 1 byte or 2 bytes. The two-byte IMERG GIS files that contain precipitation rate or precipitation accumulation use 29999 for the missing-data value in grid boxes for which the IMERG HDF5 files do not contain valid rainfall rates. The one-byte IMERG GIS files that contain the percentage of liquid precipitation use 255 for the missing value in grid boxes for which there are zero values or missing values for the total precipitation. The ratio of the liquid and mixed-phase precipitation to the total precipitation is undefined in this case (division by zero), so a missing data value is supplied in the percentage field of the IMERG GIS file.

The details of how IMERG GIS handles missing-data are based on the idea that IMERG GIS merely summarizes the precipitation estimates in either 30-minute-duration or monthly-duration IMERG HDF5 science files. The IMERG GIS algorithm itself is not a scientific algorithm, it is just a simple "accumulator" calculating sums and averages. The IMERG GIS algorithm avoids, to the extend possible, making scientific assumptions or using ancillary data.

There is only one option for handling missing-data when there is a single input file: any $0.1° \times 0.1°$ grid box with missing data-values in the input file is set to missing in the output IMERG GIS file. This is the case with 30-minute-duration and 1-month-duration IMERG GIS files.

A different calculation method is used to generate IMERG GIS files that have multiple input 30-minute-duration HDF5 files. A $0.1° \times 0.1°$ grid box in the output IMERG GIS file is given a missing-data value only if that grid box has missing data values in *all* of the input HDF5 files. If one or more of the input HDF5 files have a valid precipitation rate (≥ 0 mm h^{-1}) in a particular $0.1° \times 0.1°$ grid box, then those values are multiplied by 0.5 h and summed to create the accumulation (mm) written to that grid box in the output IMERG GIS file.

Obviously, this calculation method is likely to give lower-than-expected accumulation in grid boxes where some of the 30-minute input periods have missing-data values. A more unstable but less-biased calculation method would be to multiply the observed average precipitation rate by the ratio of the duration of the output accumulation period and the sum of the durations of the 30-minute input HDF files that contain non-missing data values for this 0.1°×0.1° grid box. A still more sophisticated algorithm would use climatological information, perhaps including seasonal and diurnal statistics to fill in greater than zero accumulation for the 30-minute periods that contain missing-data values in the input HDF5 files. The IMERG GIS algorithm, however, avoids making such scientific judgement calls.

3.7. Geographic boundaries

IMERG HDF5 files and IMERG GIS files have geographic boundaries of 90°S to 90°N latitude and 180°W to 180°E longitude. Many grid boxes near the poles are filled with the missing-data value. IMERG has a 0.1° spatial resolution. The southernmost grid boxes cover areas bounded by 90°S to 89.9°S latitude. The westernmost grid boxes cover areas bounded by 180°W to 179.9°W longitude.

4. How to Display the IMERG GIS Data Product

4.1. Displaying with ESRI ArcMap

This section describes how to display IMERG GIS files in the ESRI ArcMap GIS application. First, download the *.tif and *.tfw files from the online archive, as described in Section 2.1 of this document. Start ArcMap. Choose Add Data from the File menu or click on the Add Data button. Browse to the *.tif file of interest and click the OK button. ArcMap will locate the accompanying *.tfw file that you placed in the same directory. After selecting a *.tif file, click No if you see a pop-up window asking if you want to create data pyramids. Click OK if you see a pop-up window that warns that the spatial reference is unknown. The data will display correctly in ArcMap despite these warnings. After you dismiss these pop-up windows, the data will appear as a new layer listed on the left side of the ArcMap window.

Assuming that you are looking at a precipitation variable (instead of a percent variable), follow these steps. Adjust ArcMap's color table to improve how the precipitation data is displayed. To do so, right-click on the name of the layer to bring up the Layer Properties window. Click on the Symbology tab if it is not already selected. Change the color ramp from the default grayscale to a rainbow scheme or some other color ramp of your choice. Click the

checkbox next to Display Background Value, so that locations with zero precipitation have the background color showing.

In the Symbology tab of the Layer Properties window, scroll down to near the bottom in order to see the Stretch section. Figure 3 shows suggested settings for the Stretch section. These settings are only suggestions, but it is important to choose values for these fields carefully because they have a significant effect on how the data appears in ArcMap. First, change the Stretch Type from the default "PercentClip" to "Minimum-Maximum" (field A in Figure 3). Second, change the default Statistics field from "From Each Raster Data set" to "From Custom Settings (below)," as shown at field B of Figure 3.
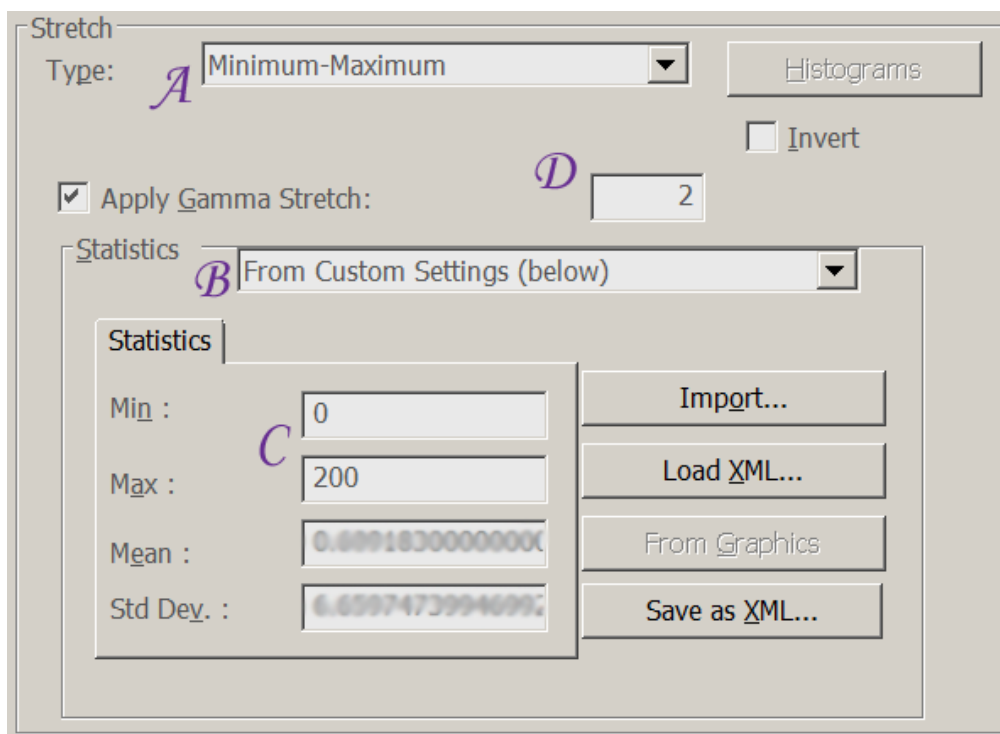
Figure 3. Suggested settings for the ArcMap Stretch window.

Third, type in reasonable minimum and maximum data values in the GeoTIFF image that you want mapped to the bottom and top colors of the color table. These minimum and maximum fields are marked by C in Figure 3. Suggestions for what minimum and maximum values to enter in these fields are stated in column 3 of Table 3.

Table 3. For IMERG GIS precipitation fields, suggestions of values to enter for the minimum and maximum fields of the Stretch section of the Symbology Tab of ArcMap's Layer Properties window.

Latency [a]	Duration	Suggested min. and max. values to type in the color table definition	Data scale factor	Implied precipitation rate or precipitation accumulation after scaling is taken into account
		Precipitation Accumulation		
E, L, F	30 minutes	0 200	10	0 to 20 mm
E, L	3 hours	0 500	10	0 to 50 mm
E, L, F	1 day	0 1000	10	0 to 100 mm
L	3 days	0 2000	10	0 to 200 mm
L	7 days	0 5000	10	0 to 500 mm
L, F	1 month	0 1000	1	0 to 1000 mm
		Precipitation Rate		
F	30 minutes	0 200	10	0 to 20 mm/h
F	1 day	0 50	10	0 to 5 mm/h
F	1 month	0 to 1000	1000	0 to 1.0 mm/h
		Percent Liquid in Early, Late, and Final IMERG GIS		
E, L, F	various	0 to 100	1	0% to 100%

[a] **E** = Early IMERG GIS, **L** = Late IMERG GIS, and **F** = Final IMERG GIS

  To keep the color table labels consistent with the actual extent of the color table, scroll back up to the top of the Symbology tab of the Layer Properties window and edit the labels for the minimum and maximum of the color table. For example, if you were looking at a half-hour Final IMERG GIS file, change these labels to something like "0.0 mm/h (0)" and "20.0 mm/h (200)".

  Fourth, change the default gamma-correction factor for the colors so that the color table better shows the data in the GeoTIFF file. Check the box next to "Apply Gamma Stretch" if this box is not already checked off (point D in Figure 3). Then, enter a number between about 1 and 5 as the gamma correction value (2 is used in this example). Last, click the OK button at the bottom of the Layer Properties window to apply the changes that you have made.

  If you wish to display the percent-liquid IMERG GIS file, then follow the same instructions

as above for precipitation fields, except for the following differences. Near the bottom of the Symbology tab of the Properties window, set the minimum and maximum of the color table to 0 and 100. Near the top of the Symbology tab of the Properties window, do check off the Display Background Value field, but enter 255 instead of 0 for the value to be displayed as background. In the percent-liquid field, 255 indicates that zero precipitation was reported and therefore no percentage of liquid vs. total precipitation could be calculated.

Often, you may want to display multiple phases of precipitation using the same mapping of data to color table. To do so in ESRI ArcMap, repeat the display procedure outlined above for each IMERG GIS file. As you do so, make sure to use the same minimum and maximum values and the same gamma stretch factor. Figure 4 shows a screen capture of ArcMap displaying a set of four IMERG GIS files after following the above instructions. A separate shape file that contains country boundaries was also loaded into this ArcMap session.
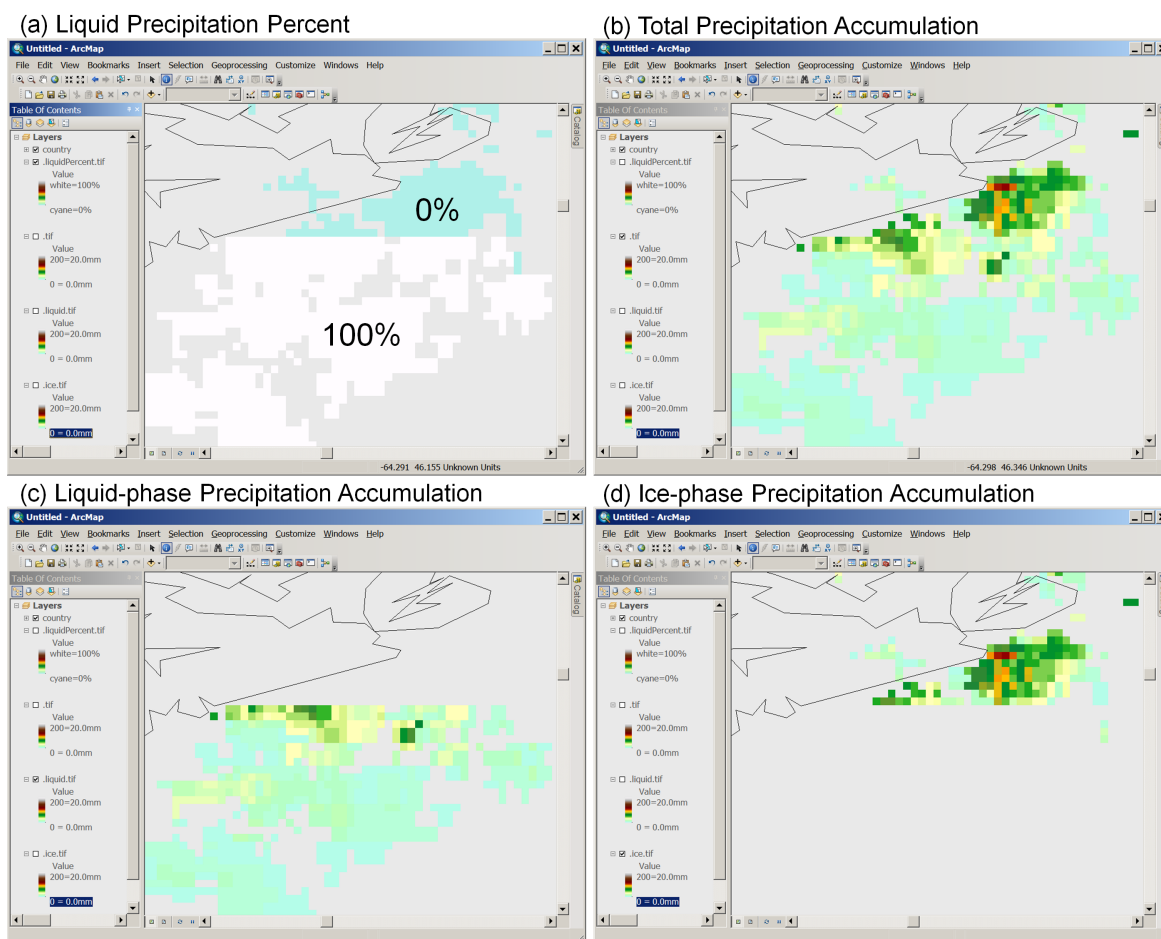
Figure 4. A storm observed off the coast of Nova Scotia, as seen in the 30-minute Late IMERG GIS files for 1930–2000 UTC on 10 April 2015. (a) Percent of precipitation-mass reaching the surface at liquid phase, with cyan indicating 0% liquid and white indicating 100% liquid. (b) Total surface precipitation accumulation during this 30-minute period. The color table maps 0

millimeters to cyan, and progressively higher rates to green, yellow, orange, and red, and white. White represents 20.0 millimeters of accumulation, although no IMERG grid boxes in this storm archive this much accumulation during these 30 minutes. (c) Liquid-phase surface precipitation accumulation displayed with the same color table as panel b. (d) Liquid-equivalent ice-phase surface precipitation accumulation displayed with the same color table as panel b.


4.2. Displaying with QGIS

It is possible to display the IMERG GIS data product using the QGIS (formerly Quantum GIS) application. QGIS is an open-source GIS application that can be downloaded at no cost from http://qgis.org. The use of QGIS, and ArcMap for that matter, in this document is not an endorsement. These two GIS applications are mentioned here merely to illustrate the kinds of steps that one may need to take to display the IMERG GIS files in a GIS application of your choice.

To begin, download the *.tif and *.tfw file from the online archive, as described in Section 2.1 of this document. Start QGIS. Select the Layer>AddLayer>AddRasterLayer menu option. Browse to the *.tif file that you want to display and click the Open button. QGIS will locate the accompanying *.tfw file that you placed in the same directory. If the Coordinate Reference System Selector window appears, click OK to dismiss it. Once QGIS reads the GeoTIFF file, an entry for this file will appear in the data set list on the left side of the QGIS window and the data itself will appear in the main display portion of the QGIS window.

Most likely the default color table shown by QGIS will be far from optimal and perhaps just a black block covering the globe.

This paragraph gives a quick way to improve the QGIS color table, and the rest of the section provides a more detailed way to improve the color table. These initial example assumes that one is looking at Early or Late IMERG GeoTIFF files, although the procedure is similar for looking at Late IMERG GeoTIFF Files. Begin by calling up the Layer Property window for this GeoTIFF File. To do so, find the name of this GeoTIFF File in the Layer Panel that usually appears on the left side of the QGIS window. On this filename, right click and select Properties from the pop-up menu that appears. In the Layer Properties window, select the "Style" tab from the vertical list along the left side of the Layer Properties window. Then do four things within the Style tab. First, find the pull-down menu next to the Render Type label and select "Singleband pseudocolor". Second, click the box next to the "Invert" label so that a X symbol appears in that box. Third, type in an appropriate maximum value in the field next to the "Max" label. For 30-minute to 7-day accumulations (which are scaled by 10 in the GeoTIFF file), start off with a maximum of 10000; for 1-month accumulations (which are not scaled in the GeoTIFF file), start off with a maximum of 1000; and for percent liquid, start off with a maximum of 255 (since they are stored as unsigned 1-byte integers in the GeoTIFF file). Fourth, click the Apply button at the bottom of the Layer Properties window. You could see an improved display of the data in the QGIS window.
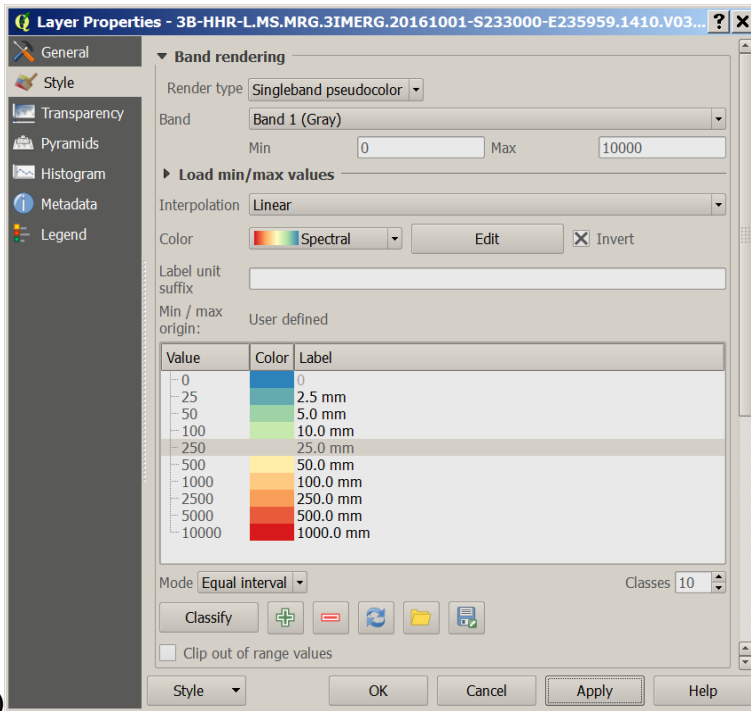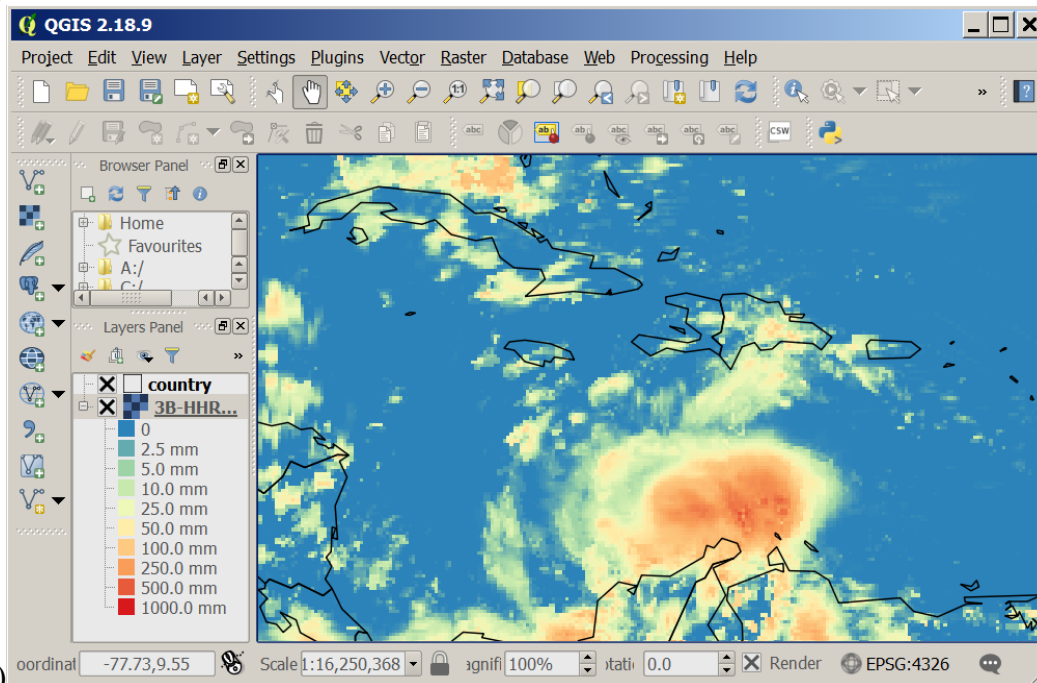
Figure 5. Screen captures of QGIS displaying IMERG GIS files. (a) The Style tab of the QGIS Layer Properties window after an unequally-spaced color table is defined for displaying precipitation scaled by 10 and stored as a 2-byte integer in an Early or Late IMERG GIS accumulation for a period of 30 minutes to 7 days. (b) The QGIS main window with a 1-day Late IMERG GIS accumulation for 8 October 2016 showing Hurricane Matthew in the Caribbean.

One can, optionally, create a more refined color table by manually selecting the limits of each color in the color table. To do so, first change the Mode from "Continuous" to "Equal Interval" so that it becomes possible to increase the number of colors in the color table. Increase the Classes field to 10 colors. At this point, one could click the Apply button to preview what the display will look like. From this point on, this example will be for displaying 30-minute to 7-day accumulations, although a similar procedure could be followed for monthly accumulations or for percent liquid data.

You will notice that most locations will have precipitation near the low end of the color table. Precipitation rates or accumulations are most numerous near the zero end of the color table with a few widely-scattered large values because precipitation is approximately logarithmically distributed. In other words, precipitation is usually light (clustered near ~1 mm/h) but a few places on Earth at any given time might have precipitation that is one or two orders of magnitude higher (10 to 100 mm/h). The color table will better match the data set if most of the graduations are located near the lower end of the scale, as described here. Double-click on the number to the left of the second lightest precipitation rate (the rate one step heavier than the zero color) and change the number to 25. Double-click on the next lowest entry in the color table and change it to 50. Continue this process, choosing values of 100, 250, 500, 1000, 2500, 5000, and 10000 for the rest of the color table entries. After making these changes, the Style tab of the Layer Properties window will look something like Figure 5. Click the Apply button to see the updated, unequally-spaced color table.

To speed up the process the next time that you want to display an IMERG GIS file, select the Style menu at the bottom-left of the Layer Properties window and choose the Save Style menu item. In the file-selector window that pops up, type in a filename and click the Save button to dismiss the file-selector window. You may note that the Style menu at the bottom-left of the Layer Properties window has a Load Style menu item. This menu item is the one that you will select in the future to reload this color table that you just created and saved. In the Layer Property window, select the OK button to dismiss the Layer Property window. To view the color-table legend in the main QGIS window, click on the plus symbol to the left of the filename in the list of open files on the left side of the QGIS window.

Table 4 states the physical meaning of the color table cutoffs suggested in the previous paragraph (25, 50, ... 5000, 10000). The table also provides the physical meaning of these cutoffs, depending on what kind of IMERG GIS file is being displayed. If you wish, you may type in these physical meanings as color-table annotations so that they appear to the right of each color of the color table. To do so, double-click on the color table or use another means to open the Layer Properties window. Then, type the text labels you wish in the field to the right of each color (Figure 5a). Click the Apply button and then the OK button to apply the labels (Figure 5b) and dismiss the Layer Properties window.

Table 4. For IMERG GIS files, the physical meaning of the suggested color-table cutoffs for precipitation variables displayed in QGIS.

	Physical meaning			
Value in the *.tif file	Precipitation accumulation		Precipitation rate	
	30-minute to 7-day	Monthly	30-minute or 1-day	Monthly
0	0 mm	0 mm	0 mm/h	0 mm/h
25	2.5 mm	25 mm	2.5 mm/h	0.025 mm/h
50	5.0 mm	50 mm	5.0 mm/h	0.050 mm/h
100	10.0 mm	100 mm	10.0 mm/h	0.100 mm/h
250	25.0 mm	250 mm	25.0 mm/h	0.250 mm/h
500	50.0 mm	500 mm	50.0 mm/h	0.500 mm/h
1000	100.0 mm	1000 mm	100.0 mm/h	1.000 mm/h
2500	250.0 mm	2.5 meters	250.0 mm/h	2.500 mm/h
5000	500.0 mm	5.0 meters	500.0 mm/h	5.000 mm/h
10000	1000.0 mm	10.0 meters	1000.0 mm/h	10.000 mm/h

5. The GIS Translation of TMPA, the TRMM Analog to IMERG

5.1. The TMPA GIS format

So far, this document has described the IMERG data product. This section describes a legacy data product called TMPA that was first created in 1998 (Huffman et al. 2007, 2010). TMPA is the multi-satellite product of the Tropical Rainfall Measurement Mission (TRMM) satellite in the same way that IMERG is the multi-satellite product associated with the GPM core satellite. TMPA stands for TRMM Multi-satellite Precipitation Analysis, and this data product is also known as the TRMM 3B42 algorithm. As of 2015, several hundred journal articles have been written that mention TMPA or 3B42 in their title or abstract according to the article database called the "Web of Knowledge" (http://webofknowledge.com). The standard TMPA

files are written in the HDF4 format for the research product and in a custom binary format for the realtime product. The TMPA GIS files are written as GeoTIFF files and WorldFiles.

At some time in the future, NASA will discontinue the production of TMPA (both HDF4 and GIS, research and realtime) after allowing time for TMPA users to switch over to using the IMERG data product. It is possible that NASA will discontinue TMPA production in the second half of 2019.

In several respects, TMPA is limited in comparison to the newer IMERG data product. First, only total precipitation is stored in TMPA GIS files, while IMERG GIS contains both precipitation and precipitation-phase. Second, TMPA's 0.25° spatial resolution is coarser than the 0.1° resolution of IMERG. For research TMPA, the geographic boundaries are 50°S to 50°N latitude and 180°W to 180°E longitude (Huffman and Bolvin 2013). For realtime TMPA, the geographic boundaries are 60°S to 60°N latitude and 180°W to 180°E longitude (Huffman 2003). Third, the time resolution of TMPA is 3 hours, which is coarser than the 30-minute resolution of IMERG.

 For realtime TMPA GIS files, PPS creates a precipitation accumulation for the original binary file's 3-hour period and for a 1-day, 3-day, and 7-day period that ends with this 3-hour period. TMPA HDF or binary files have their 3-hour averaging period centered on one of the following times each day: 0, 3, 6, 9, 12, 18, or 21 UTC. The datetime stamp expressed in the TMPA filename is this center time (not the start time of the 3-hour period). Stated another way, the TMPA averaging period includes the UTC time stated in the filename plus or minus 90 minutes. For example, the TMPA file for 6 UTC covers the 3-hour period of 0430 UTC to 0730 UTC.

To obtain Realtime TMPA GIS files, visit the following URL: ftp://trmmopen.gsfc.nasa.gov/pub/gis/ . Realtime TMPA files are generated typically within about 8 hours following the time of the earliest observation in the file. The GIS version of research TMPA can be obtained from this URL: ftp://trmmopen.gsfc.nasa.gov/pub/trmmdata/GIS/ . Research TMPA files are available approximately 3 months after the satellite observations are taken to allow calibration data sets to be generated from other instruments.

The units of Realtime TMPA GIS files are 0.1 millimeters of accumulation, i.e., precipitation accumulation multiplied by 10. The units of Research TMPA GIS files are 0.1 millimeter per hour, i.e., precipitation rate multiplied by 10.

5.2. Filenames for TMPA GIS files

For research TMPA, a sample HDF4 filename is 3B42.20140501.00.7.HDF. The analogous research TMPA GIS files would be called 3B42.20140501.00.7.{tif|tfw}. These filenames can be parsed as the algorithm (3B42), the date in yyyymmdd format (year=yyyy, month=mm, day of month=dd) and the UTC hour in the middle of the 3-hour period for which the precipitation

estimate applies, the data version (7 in this case), and the file format ("HDF" for HDF4 files, "tif" for GeoTIFF files, and "tfw" for ESRI WorldFiles).

For realtime TMPA, an example filename for the original binary product is 3B42RT.2014060100.7.bin. For each realtime TMPA file, PPS generates four TMPA GIS files, one for each of the four accumulation periods listed in Table 5.

Table 5. Sample filenames for realtime TMPA GIS.

Name of TMPA GIS GeoTIFF file	Covers a period ending 1 June 2014 0130 UTC and that has a duration of...
3B42RT.20140601.00.07.3hr.tif	3 hours [a]
3B42RT.20140601.00.07.1day.tif	1 day
3B42RT.20140601.00.07.3day.tif	3 days
3B42RT.20140601.00.07.7day.tif	7 days

[a] The 3-hour TMPA GIS file covers the same 3-hour time interval as the original, binary realtime TMPA file. The 3-hour TMPA GIS file stores precipitation accumulation during these 3 hours while the original, binary realtime TMPA file stores precipitation rate (mm/h) averaged over this 3-hour period. For example, the 3-hour TMPA GIS file would report a 30 millimeter accumulation in 3 hours for a grid box if the original, binary TMPA file reported a 10 mm/h precipitation rate for that grid box.

6. References

Hong, Y., R. F. Adler, A. Negri, and G. J. Huffman, 2007: Flood and landslide applications of near real-time satellite rainfall products. *Nat. Hazards*, **43**, 285−294.

Huffman, G. J., D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E. J. Nelkin, S. Sorooshian, J. Tan, and P. Xie, 2019a: *NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD)*. https://pps.gsfc.nasa.gov/Documents/IMERG_ATBD_V06.pdf .

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, and J. Tan, 2019b: *Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation*. White paper. https://pps.gsfc.nasa.gov/Documents/IMERG_doc_190313.pdf

Huffman, G. J., and D. T. Bolvin, 2013: *TRMM and other data precipitation set documentation*. File specification TMPA HDF4. ftp://precip.gsfc.nasa.gov/pub/trmmdocs/TMPA_3B43_doc.pdf .

Huffman, G. J., 2003: *README for Accessing Realtime TMPA from the Experimental TRMM Real-Time Multi-Satellite Precipitation Analysis (TMPA-RT) Data Set.* File specification for Realtime TMPA binary files. http://www.eol.ucar.edu/projects/name/documentation/Realtime TMPA_README .

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. *J. Hydrometeor.*, **8**, 38−55. ftp://meso.gsfc.nasa.gov/agnes/huffman/papers/TMPA_jhm_07.pdf.gz .

Huffman, G. J., R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-satellite Precipitation Analysis (TMPA). Chapter 1 of *Satellite Rainfall Applications for Surface Hydrology*. F. Hossain and M. Gebremichael, Eds. Springer Verlag, ISBN: 978-90-481-2914-0, 3−22.

Wikipedia, 2015: "World file." Entry in online encyclopedia. http://en.wikipedia.org/wiki/World_file .