Sorting Out the Variety of Precipitation Datasets

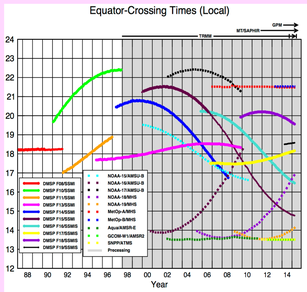
George J. Huffman
NASA Goddard Space Flight Center, Greenbelt, MD
george.j.huffman@nasa.gov

WHAT SATELLITES GIVE US PRECIPITATION ESTIMATES?

The original goal was 3-hourly observations, globally
- Original basis was sampling the diurnal cycle
- But also, "morphed" microwave loses skill outside ±90 minuntes

The current constellation includes:
- 6 polar-orbit passive microwave imagers
  - 4 SSMIS, AMSR-2, GMI
- 4 polar-orbit passive microwave sounders
  - 4 MHS
  - SAPHIR (asynchronous), ATMS (polar)

The future is "interesting"
- Legacy satellites are allowed to drift
  - apparently 06/18 (00/12) UTC is a stable (unstable) point
  - exact coverage is a complicated function of time
  - duplicate orbits aren't very useful for getting 3-hourly observations
- CGMS is studying placing a satellite in the 00/12 UTC gap in coverage
- Launch manifests are sparse



Ascending passes (F08 descending); satellites depicted above graph precess throughout the day.
Image by Eric Nelkin (SSAI), 19 May 2015, NASA/Goddard Space Flight Center, Greenbelt, MD.

WHAT GLOBAL PRECIPITATION DATASETS ARE THERE?

The CGMS International Precipitation Working Group (IPWG) lists
- Precipitation data sets that are
  - publicly available, quasi-operational, quasi-global, long-term
- 4 tables
  - Satellite-Gauge, Multi-Satellite, Single Satellite, Gauge
- http://www.isac.cnr.it/~ipwg/data/datasets.html
  - basic attributes and source (or contact)

Multi-Satellite with Gauge

Combination Data Sets with Gauge Data
Table 1. Summary of publicly available, quasi-operational, quasi-global precipitation estimates that are produced by combining input data from several sensor types, including satellite sensors and precipitation gauges. Where appropriate, the algorithms applied to the individual input data sets are mentioned. Solid shading indicates products that are not yet available. [Last updated 26 July 2013, G.J. Huffman]

Algorithm	Input data	Space/time grid	Areal coverage/ start date	Update frequency	Latency	Producer (Developer) URL
CAMS/OPI	CAMP-GPI, gauge	2.5°/monthly	Global/1979	Monthly	5 days	NOAA/NWS CPC (Xie) [1]
CMAP	OPI, SSM/I, SSMIS, GPI, MSU, gauge, model	2.5°/monthly	Global/1979 – Nov. 2011	Seasonal	3 months	NOAA/NWS CPC (Xie) [2]
	OPI, SSM/I, GPI, MSU, gauge, model	2.5°/pentad	Global/1979 – Nov. 2011	Seasonal	3 months	NOAA/NWS CPC (Xie) [3]
	OPI, SSM/I, GPI, gauge	2.5°/pentad-RT	Global/2000	Pentad	1 day	NOAA/NWS CPC (Xie) [4]
CMORPH V1.0 BIAS-CORRECTED	TMI, AMSR-E, SSM/I, SSMIS, AMSU, MHS, IR vectors, CPC Gauge, GPCP Pentad	0.25°/3-hourly	50°N-S/1998	Daily	18 hours	NOAA/CPC (Xie) [5]
CMORPH V1.0 BLENDED	TMI, AMSR-E, SSM/I, SSMIS, AMSU, MHS, IR vectors, daily gauge	0.25°/3-hourly	50°N-S regional/ 1998	Daily	18 hours	NOAA/CPC (Xie) [5]
GPCP One-Degree Daily (Version 1.2)	SSM/I & SSM/I5- F08+ (IR), TOVS, AIRS, GPCP, monthly	1°/daily	Global – 50°N-50°S/ Oct. 1997 – July. 2011	Monthly	3 months	NASA/GSFC 612 (Huffman) [6]
GPCP pentad (Version 1.1)	OPI, SSM/I, GPI, MSU, gauge, GPCP monthly	2.5°/5-day	Global/1979 - 2008	Seasonal	3 months	NOAA/NWS CPC (Xie) [7]
GPCP Version 2.2 Satellite-Gauge (SG)	OPI/OP-OPI, gauge 1979-7/87, 12/87 thereafter SSM/I & SSM/I5-A/2F (IR), gauge, TOVS, AIRS	2.5°/monthly	Global/1979 – 2010	Monthly	2 months	NASA/GSFC 612 (Huffman, Adler) [8]
IMERG Final Run V3	TMI, SSMI, SSMIS, AMSR-E, GMI, AMSR2, AMSU, MHS, SAPHIR, ATMS, IR, IR vectors, GPCP monthly	0.1°/half-hourly	Global – 60°N-S/ Mar. 10, 2014	Half-hour	3.5 months	NASA/GSFC PPS (Huffman) [14]
PERSIANN-CDR	GridSat-IRWIN, GPCP Monthly Precipitation	0.25°/daily	60°N-S/1980	Monthly	2 months	UC Irvine (Hsu) [9]
PR-OBS-05	PR-OBS-03, gauge, radar, NWP	5 km/3-hour	Europe/Jan. 2009	3 hour	0.5 hours	HSAF (Vocino) [10]
RFE	GPI, NOAA SSM/I, gauge	10 km/daily	Africa/Oct. 2000	Daily	6 hours	NOAA/NWS CPC (Xie) [11]
		10 km/daily	South Asia/April 2001	Daily	6 hours	NOAA/NWS CPC (Xie) [12]
TRMM Plus Other Data (3B43 Version 7)	TCI, TMI, SSMI, SSMIS, AMSR-E, AMSU, MHS, MW-VAR (IR), gauge	0.25°/monthly	Global – 50°N-S/Jan 1998	Monthly	2 months	NASA/GSFC PPS (Huffman, Adler) [13]
TRMM Plus Other Satellites (3B42 Version 7)	TCI, TMI, SSMI, SSMIS, AMSR-E, AMSU, MHS, MW-VAR (IR), V 7 3B43	0.25°/3-hourly	Global – 50°N-S/Jan 1998	Monthly	2 months	NASA/GSFC PPS (Huffman, Adler) [13]

[1] ftp://ftp.cpc.ncep.noaa.gov/precip/data-req/cams_opi_v0208/
[2] ftp://ftp.cpc.ncep.noaa.gov/precip/cmap/monthly/
[3] ftp://ftp.cpc.ncep.noaa.gov/precip/cmap/pentad/
[4] ftp://ftp.cpc.ncep.noaa.gov/precip/cmap/pentad_rt/
[5] ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0
[6] ftp://rsd.gsfc.nasa.gov/pub/1dd-v1.2/
[7] ftp://ftp.cpc.ncep.noaa.gov/precip/GPCP_PEN/
[8] ftp://precip.gsfc.nasa.gov/pub/gpcp-v2.2/psg/
[9] kuolin@uci.edu; Dr. Kuolin Hsu
[10] ftp://ftp.meteoam.it/from_archive/h05; ftp://ftp.meteoam.it/products/h05/h05_curr_mon_buf for last 60 days
[11] ftp://ftp.cpc.ncep.noaa.gov/fews/newalgo_est/
[12] ftp://ftp.cpc.ncep.noaa.gov/fews/S.Asia/
[13] http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&project=TRMM&dataGroup=Gridded
[14] ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/

Multi-Satellite

Satellite Combination Data Sets
Table 2. Summary of publicly available, quasi-operational, quasi-global precipitation estimates that are produced by combining input data from several satellite sensor types. Where appropriate, the algorithms applied to the individual input data sets are mentioned. The TCI is available as a separate product from the Goddard DISC, in addition to the 3G68 compilation. Hatched shading indicates a product being released in phases – see site for current availability. [Last updated 10 September 2013, G.J. Huffman]

Algorithm	Input data	Space/time grid	Areal coverage/ start date	Update interval	Latency	Producer (Developer) [URL]
AIRS AIRX2SSD	AIRS, AMSU, HSB sounding retrievals	Level 2/ swath/orbit segments	Global/Sept. 2002	Daily	1 day	NASA/GSFC GES DISC (Susskind) [1]
AIRS AIRX3STD, AIRX3STB, AIRX3STM	AIRS, AMSU, HSB sounding retrievals	1°/daily, 1°/8-day, 1°/monthly	Global/Sept. 2002	Daily, 8-day, monthly	1 day	NASA/GSFC GES DISC (Susskind) [1]
CMORPH	TMI, AMSR-E, SSMI, SSMIS, AMSU, MHS, IR vectors	8 km/30-min	50°N-S/1998	Daily	18 hours	NOAA/CPC (Xie) [2]
CMORPH V1.0 RAW	TMI, AMSR-E, SSMI, SSMIS, AMSU, MHS, IR vectors	0.25°/3-hourly	50°N-S/1998	Daily	18 hours	NOAA/CPC (Xie) [3]
GSMaP NRT	TMI, AMSR-E, SSMI, SSMIS, AMSU, IR vectors	0.1°/hourly	60°N-S/Oct. 2007	1 hour	4 hours	JAXA (Kachi & Kubota) [4]
GSMaP MWR	TMI, AMSR-E, AMSR, SSMI, IR vectors	0.25°/hourly, daily,monthly	60°N-S/1998-2006	–	–	JAXA (Aonashi & Kubota) [5]
GSMaP MVK	TMI, AMSR-E, AMSR, SSMI, SSMIS, AMSU, IR vectors	0.1°/hourly	60°N-S/2000 (currently 2003-2008 data available)	Monthly	Reprocess now; will become operational	JAXA (Ushio) [4]
GSMaP MVK+	TMI, AMSR-E, AMSR, SSMI, AMSU, IR vectors	0.1°/hourly	60°N-S/2003-2006	–	–	JAXA (Ushio) [5]
IMERG Early Run V3	TMI, SSMI, SSMIS, AMSR-E, GMI, AMSR2, AMSU, MHS, SAPHIR, ATMS, IR, IR vectors	0.1°/half-hourly	Global – 60°N-S/ Mar. 10, 2015	Half-hour	5 hours	NASA/GSFC PPS (Huffman) [16]
IMERG Late Run V3	TMI, SSMI, SSMIS, AMSR-E, GMI, AMSR2, AMSU, MHS, SAPHIR, ATMS, IR, IR vectors	0.1°/half-hourly	Global – 60°N-S/ April, 2015	Half-hour	16 hours	NASA/GSFC PPS (Huffman) [17]
MPE	Meteosat 7,8,9,10 IR and SSM/I, SSMIS	MFG: original pixels/30-min	Indian Oc. 8°W-122°E,65°N-S	30 min	10 min	EUMETSAT [8]
		MSG: original pixels/15-min	Europe/Africa 79°W-E,81°N-S	15 min	10 min	
NRL Real Time	SSM/I- & F16/SSMIS-cal IR (Prob.-Matching Method)	0.25°/hourly	40°N-S/ July 2000	Hourly	3 hours	NRL Monterey (Turk) [7]
PERSIANN	ITMI, AMSR-E, SSMI, SSMIS, AMSU, MHS)-cal. IR	0.25°/30-min	60°N-S/March 2000	Hourly	1 day	UC Irvine (Hsu) [8]
PERSIANN-CCS	Neural Net GEO-IR	0.04°/30-min	60°N-S/2003	30 min	1 hour	UC Irvine (Hsu) [9]
PR-OBS-03	SSMIS, AMSU/MHS, MSG-IR	5 km/15-min	Europe/Jan. 2009	15 min	15 min	HSAF (Melfi, Cattani) [10]
PR-OBS-04	SSMIS, AMSU/MHS, MSG-IR	8 km/30-min	Europe/Nov. 2011	1 hour	3 hours	HSAF (Melfi, Cattani) [10]
TCI (3G68)	PR, TMI	0.5°/hourly	Global – 37°N-S/ Dec. 1997	Daily	4 days	NASA/GSFC PPS (Haddad) [11]
TOVS	HIRS, MSU sounding retrievals	1°/daily	Global/1979-April 2005	Daily	1 month	NASA/GSFC 610 (Susskind) [12]
TRMM Real-Time HQ Version 7 (3B40RT)	TMI, SSMI, SSMIS, AMSR-E, AMSU, MHS	0.25°/3-hourly	Global – 70°N-S/ Mar. 2000	3 hours	9 hours	NASA/GSFC PPS (Huffman, Adler) [13]
TRMM Real-Time VAR Version 7 (3B41RT)	MW-VAR (IR)	0.25°/hourly	Global – 50°N-S/ Mar. 2000	1 hour	9 hours	NASA/GSFC PPS (Huffman, Adler) [14]
TRMM Real-Time HQ/VAR Version 7 (3B42RT)	HQ, MW-VAR (IR)	0.25°/3-hourly	Global – 50°N-S/ Mar. 2000	3 hours	9 hours	NASA/GSFC PPS (Huffman, Adler) [15]

[1] http://disc.sci.gsfc.nasa.gov/AIRS
[2] http://www.cpc.ncep.noaa.gov/products/janowiak/cmorph_description.html
[3] ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0
[4] http://sharaku.eorc.jaxa.jp/GSMaP/
[5] http://sharaku.eorc.jaxa.jp/GSMaP_crest/
[6] http://www.eumetsat.int/Home/Main/DataProducts/Atmosphere/index.htm?l=en
[7] song.yang@nrlmry.navy.mil; Dr. Song Yang
[8] http://chrs.web.uci.edu/persiann/data.html
[9] http://hydis.eng.uci.edu/gwad/
[10] http://hsaf.meteoam.it/user-registration.php
[11] ftp://pps.gsfc.nasa.gov/pub/trmmdata/3G/3G68/
[12] joel.susskind-1@nasa.gov; Dr. Joel Susskind
[13] ftp://trmmopen.nascom.nasa.gov/pub/merged/combinedMicro/
[14] ftp://trmmopen.nascom.nasa.gov/pub/merged/calibratedIR/
[15] ftp://trmmopen.nascom.nasa.gov/pub/merged/mergeIRMicro/
[16] ftp://jsimpson.pps.eosdis.nasa.gov/NRTPUB/imerg/early/
[17] ftp://jsimpson.pps.eosdis.nasa.gov/NRTPUB/imerg/late/

WHAT DATASET AT GODDARD WORKS BEST FOR ME?

GPCP Satellite-Gauge (SG)
- Long record, 1979–present, 2-month latency
- Relatively coarse (2.5° lat./lon.) global grid
- Includes gauge data

TMPA production
- Long record, 1998–present, 2-month latency
- Relatively fine (0.25° lat./lon.), 50°N-S
- Includes gauge data

IMERG Final
- April 2014–present, 3.5-month latency
- Relatively fine (0.1° lat./lon.), 60°N-S
- Includes gauge data

IMERG Early
- March 2015–present, 5-hour latency
- Relatively fine (0.1° lat./lon.), 60°N-S
- Extrapolated "morphing"

GPCP One-Degree Daily (1DD)
- Long record, October 1996–present, 2-month latency
- Relatively coarse (1° lat./lon.) global grid
- Consistent with GPCP SG

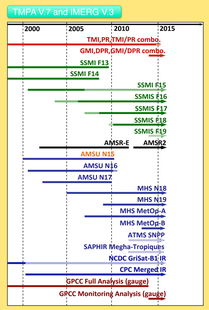
TMPA Real-Time
- Long record, February 2000–present, 8-hour latency
- Relatively fine (0.25° lat./lon.), 50°N-S
- Includes gauge data

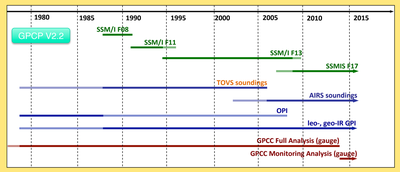
IMERG Late
- April 2015–present, 16-hour latency
- Relatively fine (0.1° lat./lon.), 60°N-S
- Interpolated "morphing"

Each SG tends to follow the calibrators
- Over land – the GPCC gauge analysis
- Over ocean – satellite calibrator

Real-time products show various biases compared to (post-event) validation

Periods of record not used in the datasets are shown in lighter color





THE NEW CONCEPT – IMERG REQUIREMENTS/GOALS

Resolution – 0.1° [i.e., roughly the resolution of microwave, IR footprints]

Time interval – 30 min. [i.e., the geo-satellite interval]

Spatial domain – global, initially covering 60°N-60°S

Time domain – 1998-present; later explore entire DMSP era (1987-present)

Multiple product latencies – early sat. (~5 hr), late sat. (~16 hr), final sat.-gauge (~3.5 months after month) [more data in longer-latency products]

Sensor precipitation products intercalibrated to GPM GMI/DPR combined (2BCMB)

Global, monthly gauge analyses including retrospective product – explore use in submonthly-to-daily and near-real-time products

Error estimates – still open for definition

Embedded data fields showing how the estimates were computed

Precipitation type estimates – probability of liquid

Operationally feasible, robust to data drop-outs and (strongly) changing constellation

Output in HDF5 v1.8 – compatible with NetCDF4

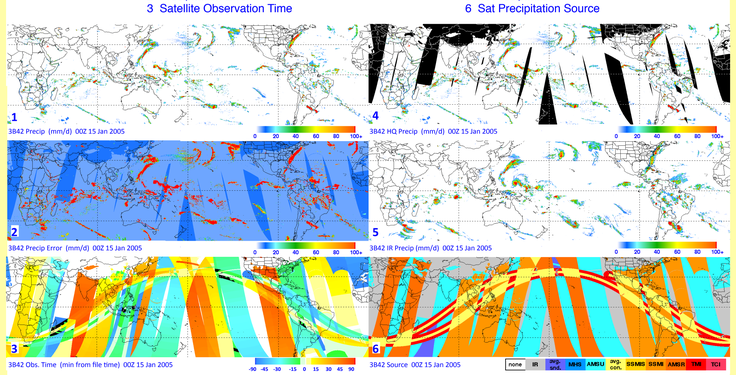
Archiving and reprocessing for near- and post-RT products

IMERG Data Fields

	Half-hourly data file (Early, Late, Final)
1	*[multi-sat.] precipitationCal*
2	[multi-sat.] precipitationUncal
3	*[multi-sat. precip] randomError*
4	[PMW] HQprecipitation
5	[PMW] HQprecipSource [identifier]
6	[PMW] HQobservationTime
7	IRprecipitation
8	IRkalmanFilterWeight
9	*probabilityLiquidPrecipitation [phase]*
	Monthly data file (Final)
1	*[sat.-gauge] precipitation*
2	*[sat.-gauge precip] randomError*
3	GaugeRelativeWeighting
4	*probabilityLiquidPrecipitation [phase]*

TMPA V.7 PRODUCTION 3-HOURLY OUTPUT FIELDS

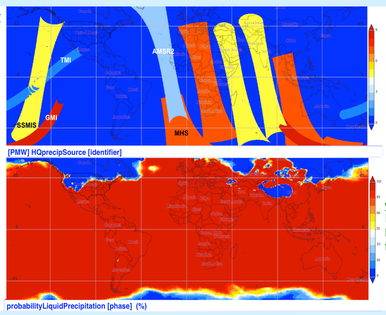
1 Precipitation
2 Precipitation Error
3 Satellite Observation Time
4 HQ Precipitation
5 IR Precipitation
6 Sat Precipitation Source



SELECTED IMERG HALF-HOURLY OUTPUT FIELDS
1430 – 1500Z 3 April 2014

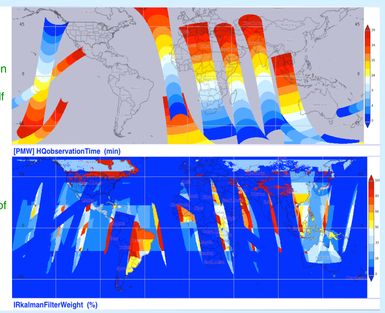
PMW sensor contributing the data, selected as imager first, then sounder, then closest to center time

probability that precipitation phase is liquid; diagnostic computed from ancillary data

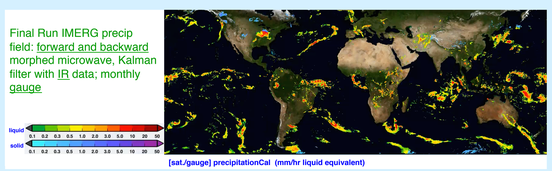


PMW sensor observation time after start of half hour

weighting of IR in the Kalman filter step

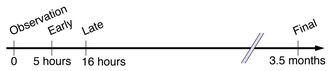


Final Run IMERG precip field: forward and backward morphed microwave, Kalman filter with IR data; monthly gauge



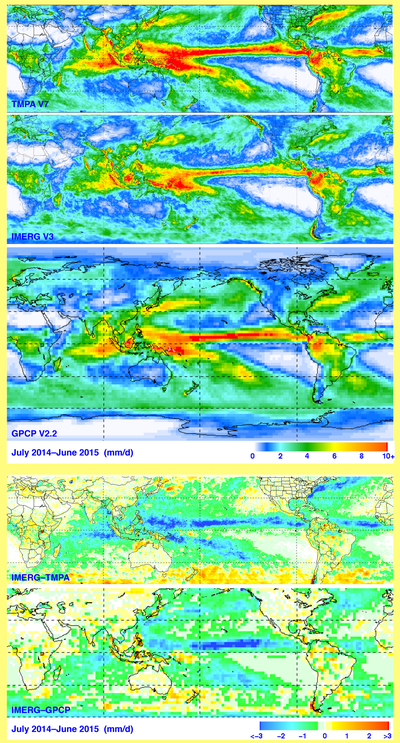
IMERG RUN STRATEGY

Different users have different requirements for latency and accuracy
- "Early" – forward-only morphing 6 hours after observation time for users needing a quick answer (flash flooding, landslides)
- "Late" – a complete forward/backward morphing 16 hours after observation time for "next-day" users (agricultural forecasting)
- "Final" – combined multi-satellite and gauge analyses 3 months after observation for archival science research data (climate statistics, model validation)
- Terminology based on NWS weather map analysis cycle



Studies underway for reducing near-real-time latency
- Can we get the data faster?
- Can we make the satellite-by-satellite computations faster?
- Can we make reasonable compromises in use of input data without compromising IMERG quality "too much"?

PRELIMINARY IMERG VALIDATION



Averages for July 2014–June 2015
- All land areas incorporate the same (GPCC) gauge analysis
- TMPA is likely 5-7% high over oceans
- IMERG is "Day 1", not yet using full GPM-based calibrations

Near-Real-Time Behavior
- Oceans are substantially the same as for post-real-time
- TMPA-RT tends to perform more poorly outside the TRMM coverage area (35°N-S)
- IMERG Early and Late should perform better at high latitudes because
  - the GPM Core Observatory covers out to 65°N-S
  - the algorithms that compute input precip data estimates are better
- First validations for (Day-1) IMERG Early and Late show competitive performance in many situations
- In recent excessive rain situations in South Carolina and Texas, IMERG Early and Late have overestimated, particularly at high rates

SCHEDULE AND FINAL REMARKS

GPCP is currently in Version 2.2
- Version 2.3 will come out in early 2016
  - retrospectively processed to 1979
  - it will include an "Interim" product computed about 10 days after the month
- Version 3 is in development
  - finer scale, more modern algorithms

TMPA is currently in Version 7
- Production should continue until IMERG is retrospectively processed
  - the notional schedule is continuation through Spring 2017
  - failures in input data could push termination earlier

IMERG is currently in Version 3
- Version 4 will come out in Spring 2016
  - March 2014–present
  - first fully GPM-based calibrations

IMERG will be retrospectively processed back to the start of TRMM (January 1998)
- The first "Interim" dataset covering the TRMMM era (1998-22014) will be computed in early 2017
- The Version 5 product, covering the entire period starting with 1998, will be computed in early 2018
  - this also represents the TRMM Version 8 final processing
  - the schedule is susceptible to slip

Real Time IMERG Near-Real-Time Data Animation – *http://svs.gsfc.nasa.gov/cgi-bin/details.cgi?aid=4285*

IMERG Email List – contact David Bolvin at *david.t.bolvin@nasa.gov* to be added

GPCP data – *http://precip.gsfc.nasa.gov*

TMPA and IMERG data – *http://pmm.nasa.gov/data-access/*