IPWG Algorithm Documentation for 3B40RT

George J. Huffman
17 February 2008

<NAME>

Real-Time TRMM HQ (3B40RT)

<ALGORITHM DESCRIPTION>

This algorithm provides a merger of TMI, AMSR-E, SSM/I, AMSU-B, and MHS microwave precipitation estimates into a "high-quality" (HQ) precipitation estimate. The TMI, AMSR-E, and SSM/I estimates are computed with version of the GPROF algorithm at the Precipitation Processing System at NASA/GSFC and the AMSU-B and MHS estimates are computed with the current NESDIS algorithm at NESDIS. Before merger all of the satellite estimates are calibrated to the TMI using algorithm-specific climatological global land and ocean histograms based on coincident time-space matched data.

Digital data:
ftp://trmmopen.nasa.gov/pub/merged/combinedMicro

Example GIF images and QuickTime movies:
http://trmm.gsfc.nasa.gov

Interactive Web-based display and analysis system:
http://lake.nascom.nasa.gov/tovas/

Detailed documentation and programming examples:
ftp://trmmopen.nasa.gov/pub/merged/ , 3B4XRT_doc.pdf and other files
ftp://trmmopen.nasa.gov/pub/merged/software

References:
Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. *J. Hydrometeor.*, **8**(1), 38-55.

Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin, 2009: The TRMM Multi-satellite Precipitation Analysis (TMPA). Chapter in *Satellite Applications for Surface Hydrology*, F. Hossain and M. Gebremichael, Eds. Springer Verlag, in revision.

<SPECTRAL INTERVALS AND APPLICABLE SATELLITES>

The input to HQ (3B40RT) consists of pre-computed precipitation estimates based on single-satellite passive-microwave data. At present we employ the GPROF to create the input estimates from conical-can passive microwave instruments and the NESDIS algorithm for sounders. In common with other physically-based algorithms, GPROF uses all available bands from modern conic-scan passive microwave imagers:

- 7 channels on SSM/I (active on DMSP F-13 and F-15),
- 9 channels on TMI and AMSR-E (active on TRMM and Aqua, respectively).

The NESDIS algorithm uses 5 “high-frequency” channels and 1 AMSU-A channels from modern cross-track-scanning sounders:

- AMSU-B (active on NOAA-16 and NOAA-17),
- MHS (active on NOAA-18 and MetOp).

<SPATIAL SCALE>

0.25°x0.25° latitude/longitude

<TEMPORAL SCALE>

3 hours

<ANCILLARY DATA>

Land/ocean surface type data

<ADDITIONAL COMMENTS>

<u>Introduction</u>

The HQ is the first stage of a system to produce the "TRMM and OtherData" estimates in real time. The system was developed as a testbed for developing concepts in merging quasi-global precipitation estimates and to take advantage of the increasing availability of input data sets in near-real time. The overall system is referred to as the real-time TRMM Multi-Satellite Precipitation Analysis (TMPA-RT). The TMPA-RT is run quasi-operationally on a best-effort basis at the Precipitation Processing System (PPS, formerly the TRMM Science Data and Information System, TSDIS) at NASA/GSFC, with on-going scientific development by the research team led by Dr. Robert Adler in the NASA/GSFC Laboratory for Atmospheres. Estimates are posted to the Web about 8 hours after observation time, although processing issues may delay or prevent this schedule. Due to the experimental nature of these estimates, users are encouraged to report their experiences with the data, and they should expect episodic upgrades or outages as the system develops.

<u>File Contents</u>

Each file starts with a header that is one 2-byte-integer row in length, or 2880 bytes. The header is ASCII in a "PARAMETER=VALUE" format that makes the file self-documenting (e.g., "algorithm_id=3B40RT").

Thereafter, 6 data fields appear:

precipitation	(2-byte integer)
precipitation_error	(2-byte integer)
total_pixels	(1-byte integer)
ambiguous_pixels	(1-byte integer; highly uncertain values)
rain_pixels	(1-byte integer)
source	(1-byte integer; the values are: 0 = no observation 1 = AMSU 2 = TMI 3 = AMSR 4 = SSMI 5 = SSMIS 6 = MHS 30 = AMSU&MHS avg. 31 = conical avg.)

All fields are 1440x720 grid boxes (0-360ºE,90ºN-S). The first grid box center is at (0.125ºE,89.875ºN). Files are produced every 3 hours on synoptic observation hours (00 UTC, 03 UTC, ..., 21 UTC) as an accumulation of all HQ swath data observed within +/-90 minutes of the nominal file time. Estimates are only computed for the band 70ºN-S.

Note that we use the term "gridbox" to denote the values on Level 3 data (i.e., gridded data), while we use the term "pixel" to denote individual values of Level 2 data (i.e., instrument footprints). Thus, there can be many pixels contributing to a gridbox.

Both precipitation and random error are scaled by 100 before conversion to 2-byte integer. Thus, units are 0.01 mm/h. To recover the original floating-point values in mm/h, divide by 100. Missings are given the 2-byte-integer missing value, -31999. Currently the random error fields are all set to the 2-byte-integer missing value, -31999. This placeholder will be replaced with actual estimates as development proceeds.

The next 3 fields are in numbers of pixels. The variable ambiguous_pixels is the count of pixels for which the algorithm cannot determine whether the scene has valid or invalid data. It is a subset of the total_pixels and many, but not all, are included in raining_pixels. In general, a "high" fraction of ambiguous_pixels indicates that the grid box value is invalid.

The source field is coded to indicate which sensor provided each gridbox's estimate:

0 = no observation	1 = AMSU	2 = TMI	3 = AMSR	4 = SSMI
5 = SSMIS	6 = MHS	30 = AMSU&MHS avg.	31 = conical avg.	

The originating machine on which the data files are written is a Silicon Graphics, Inc. Unix workstation, which uses the "big-endian" IEEE 754-1985 representation of 4-byte floating-point unformatted binary numbers. Some CPUs, including PCs and DEC machines, might require a change of representation (i.e., byte swapping) before using the data. In some cases, the gunzip routine, used to uncompress the data, will change representations automatically.

Dataset Validation

The TMPA-RT is the NASA/GSFC precipitation group's testbed, produced on a best-effort basis, and should be considered experimental. Formal validation studies are underway, typically showing that the passive-microwave-based HQ (3B40RT) results are more accurate than the IR-based 3B41RT, but suffer more-frequent sampling limitations. The primary limitations on the HQ (3B40RT) are the sparse sampling by the collection of passive-microwave satellites and algorithm drop-outs in regions with icy or frozen surface. We encourage users to report successes and problems in applying these datasets to their particular applications.

Dataset Status

Beta testing began in early December 2001. An official (experimental) version was instituted in late January 2002, and several upgrades have been issued since that time. Data computed with versions earlier than 07Z 3 February 2005 are considered obsolete. In that upgrade, AMSR-E and AMSU-B estimates were introduced in 3B40RT and the calibrations for 3B41RT were recomputed every 3 hr (but still using an approximate trailing month of match-ups). The GPROF estimates for SSM/I were upgraded to correctly screen bad input values late on 9 March 2005. Upgraded AMSU-B and MHS estimates were introduced 31 May 2007 and 27 November 2007, respectively. As of 2 October 2008 a climatological adjustment to the then-current 3B42 V.6 was introduced, as well as detailed tracking of the satellite source for each estimate.

Users are strongly discouraged from using RT data from previous versions, namely those before 2 October 2008. Also, users should be aware that estimates from the current and previous versions overlap for the period 2 October 2008 into 17 February 2009, so they should ensure that they have the current version.

Users should anticipate a series of versions as the algorithm is developed further. The present areas of interest are: improving the HQ product by auditing out AMSU-B data that are deficient in precipitation coverage; moving to shorter-interval estimation periods to more accurately represent the time series of precipitation; and expanding coverage to higher latitudes.

Example Programs

The data fields are all written with C-language code as blocks of bytes, so there are no extraneous bytes in the files. Because the first two fields are 2-byte integers and the rest are 1-byte integers in each file (to save space), users must exercise care in using FORTRAN direct access to read the data. The FORTRAN example programs read all fields with a single OPEN. Alternatively, the files can be opened with different logical record sizes depending on whether one is reading 2-byte-integer or 1-byte-integer fields. Note as well that the units of the logical record size is not part of the FORTRAN 77 standard. On SGI machines it is in 4-byte words, but some other systems expect it in bytes. Also, to repeat an earlier comment, the originating machine on which the data files are written is a Silicon Graphics, Inc. Unix workstation. It uses the "big-endian" IEEE 754-1985 representation of 4-byte floating-point unformatted binary numbers, and some CPUs, such as PCs, might require a change of representation (i.e., byte swapping) before using the data.

The FTP site ftp://trmmopen.nasa.gov/pub/merged/software provides several example programs:

3B42RT.ctl	Example GrADS control file for 3B42RT
read3B4XRT.c	C example
read_header.f	FORTRAN header-read example
read_rt_file.f	FORTRAN single-read example
read_rt_file.pro	IDL example
read_rt_lines.f	FORTRAN line-by-line example

<COMPLETE NAME OF CONTACT PERSON>

George J. Huffman

<E-MAIL ADDRESS OF CONTACT PERSON>

george.j.huffman@nasa.gov