**IPWG Algorithm Documentation for 3B41RT**

George J. Huffman
17 February 2009

<NAME>

Real-Time TRMM VAR (3B41RT)

<ALGORITHM DESCRIPTION>

This algorithm provides precipitation estimates from geostationary infrared (IR) observations using spatially and temporally varying calibration by the HQ (3B40RT). The algorithm is a probability-matched threshold approach that ensures that the histogram of gridbox-average IR precipitation rates matches the histogram of gridbox-average HQ precipitation rates locally. As such, the colder an IR pixel is than the zero-precipitation threshold brightness temperature, the higher the rainrate it receives. We refer to this as the variable-rainrate (VAR) infrared algorithm.

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

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 VAR (3B41RT) consists of the TRMM real-time HQ merged passive microwave precipitation estimates and the National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC) merged global geosynchronous 11-micron infrared (geo-IR) brightness temperatures. The latter are provided half-hourly on a 4x4-km-equivalent Cylindrical Equidistant Grid for the latitude belt 60°N-S based on merging all available images from:
* GOES-E,
* GOES-W,
* MTSAT, previously GMS, GOES-9,
* METEOSAT 5, and
* METEOSAT 7.

The approach is equally applicable to other gridded "high-quality" precipitation estimates and/or other sources of gridded 11-micron IR data. It would be key to ensure that enough coincident samples were available to the calibration that the coefficients were stable.

<SPATIAL SCALE>

0.25°x0.25° latitude/longitude

<TEMPORAL SCALE>

1 hour

<ANCILLARY DATA>

None

<ADDITIONAL COMMENTS>

Introduction

The VAR is the second stage of a system to produce the "TRMM and Other Data" 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.

File Contents

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=3B41RT").

Thereafter three data fields follow. All the fields are on a 0.25° lat./long. grid that increments most rapidly to the east (from the Prime Meridian) and then to the south (from the northern edge). Gridbox edges are on multiples of 0.25° The data fields are written as binary data in big-endian byte order. The data fields are:

| | |
|---|---|
| precipitation | (2-byte integer) |
| precipitation_error | (2-byte integer) |
| total_pixels | (1-byte integer) |

All fields are 1440x480 gridboxes (0-360°E, 60°N-S). The first grid box center is at (0.125°E, 59.875°N). Files are produced every hour from the on-hour IR image (except for the half-past image for GMS/GOES9/MTSAT), with fill-in by the half-past image (except for GMS, where the on-hour image is used for fill-in). Valid estimates are only provided in the band 50°N-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 final field is in number of pixels.

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 infrared results are designed to emulate the microwave results as closely as possible, so known deficiencies in the microwave

will likely be reflected in the infrared as well. In addition, it is well-known that infrared algorithms of the kind used here have large random errors at the fine time and space scales provided. However, we expect the infrared estimates to match the histogram of microwave estimates, so that user-specified averaging should yield approximately unbiased results. 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