P1.24

Statistics of TRMM Data Archive and Distribution at the GES DAAC

Hualan Rui[1], Bill Teng[1], Long Chiu[2], George Serafino[3], Patricia Hrubiak[1], John Bonk[1]

GSFC Earth Sciences (GES), Data and Information Services Center (DISC)
Distributed Active Archive Center (DAAC)
NASA/Goddard Space Flight Center, Greenbelt Maryland, USA 20771

([1]Science Systems and Applications, Inc.; [2]George Mason University; [3]NASA)

1. Introduction

The Tropical Rainfall Measuring Mission (TRMM) is a joint mission of the National Aeronautics and Space Administration (NASA) and the National Space Development Agency (NASDA) of Japan to monitor and study tropical and subtropical rainfall systems. TRMM has been acquiring data from shortly after its launch on November 28, 1997 to the present. All TRMM standard products are processed by the TRMM Science Data and Information System (TSDIS) and archived and distributed by the GES DAAC. In addition to the standard products, the GES DAAC generates and/or maintains a set of derived TRMM products (e.g., satellite coincidence subsets, parameter subsets, resampled gridded subsets, GIS-compatible files) to facilitate use of TRMM data by the general public. TRMM data are reprocessed with improved science algorithms approximately once per year, currently at version 5. The average operating altitude for TRMM satellite was moved from 350 kilometers to 403 kilometers during the period from August 7 to 24, 2001, which will significantly extend the mission lifetime for TRMM.

2. TRMM Data Archive and Distribution System

The GES DAAC's architecture is shown in Fig. 1, which includes automated transfer, ingest, archive, and distribution.

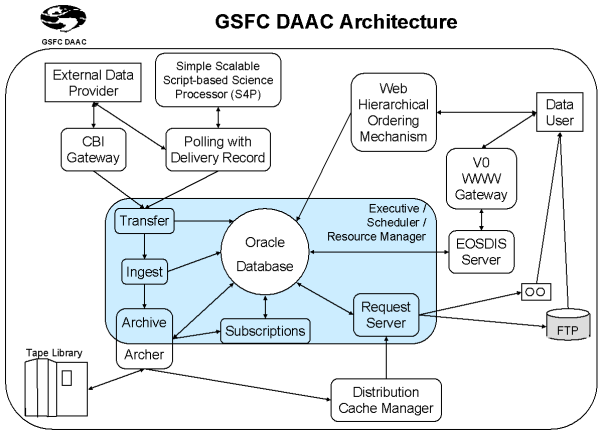

Fig. 1. Overview of GDAAC Architecture (courtesy of C. Lynnes)

The GES DAAC receives TRMM standard data products from the TSDIS. After data transfer, the metadata of each file are extracted and ingested, and data files are archived on tapes.

The GES DAAC Search and Order System, based on the Web Hierarchical Ordering Mechanism (WHOM), provides easy, friendly and quick access for searching, browsing, and ordering GES DAAC data. TRMM data, which can be accessed via the TRMM Search and Order System (http://lake.nascom.nasa.gov/data/dataset /TRMM/index.html), are updated on a daily basis automatically.

3. Statistics of TRMM Data Archive and Distribution

As of August 2001, there are 3224 registered TRMM users in the GES DAAC database. Of them, 657 users, from 32 different countries, have actually ordered TRMM data. Fig. 2 shows the profile of active TRMM users. There are 401 (60%) users from USA, and 86 (13%) users from Japan.

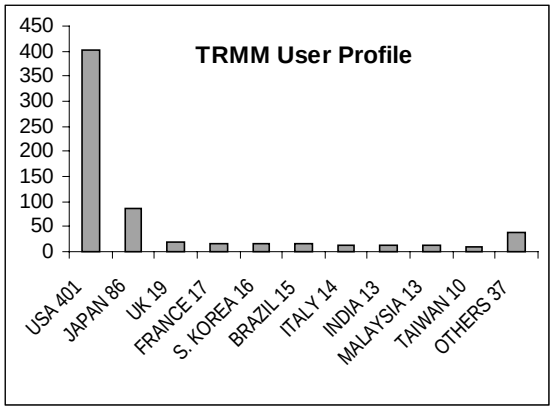

Fig.2. TRMM user profile

In order to better understand the data usage patterns and requirements of TRMM users, statistics of TRMM data archive and distribution are routinely derived from the GES DAAC database for each individual product, specific groups of data products, and the entire TRMM data set. Cumulative archive and distribution, along with the Utilization Ratio (UR), defined as the ratio of the number of distributed files to the number of archived files, are used to describe how

often a data product or a group of data products are being used by users.

The total cumulative distribution and archive of TRMM standard satellite products (as of August 2001) are 2,722,479 and 420,573, respectively, in terms of number of files; and 64.5 TB and 11.4 TB, respectively, in terms of data volumes. The UR of these TRMM standard satellite products is 6.5 (Fig. 3).

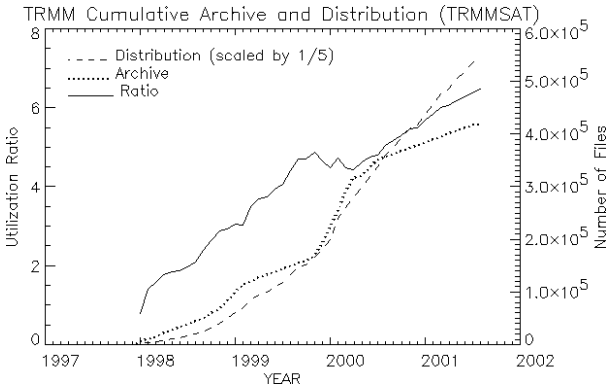

Fig. 3. Cumulative statistics of all TRMM standard satellite products, as of August 2001

3.1 Archive and Distribution Statistics of TRMM Orbital Products

There are nine TRMM orbital products (with level-1A excluded), four level -1 products (1B01, 1B11, 1B21, and 1C21) and five level-2 products (2A12, 2A21, 2A23, 2A25, and 2B31). The total cumulative distribution and archive are 2,621,957 and 342,641, respectively, in terms of number of files; and 62.5 TB and 9.1 TB, respectively, in terms of data volumes. The UR of orbital products is 7.7 (Fig. 4).

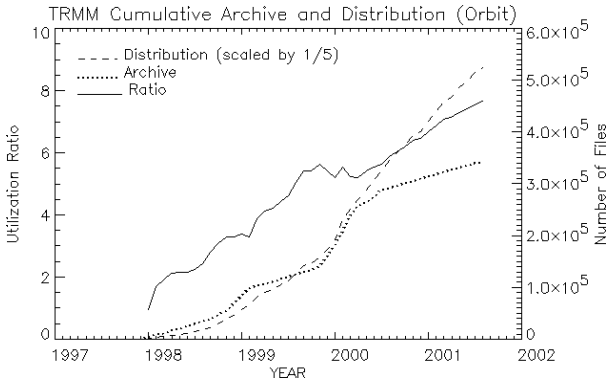

Fig. 4. Cumulative statistics of TRMM orbital products, as of August 2000

Orbital products can also be further grouped based on instruments: Visible/Infrared Scanner (VIRS), TRMM Microwave Imager (TMI) and Precipitation Radar (PR). The URs of these three groups are 8.5 for VIRS (1B01), 12.8 for TMI (1B11, 2A12), and 5.8 for PR (1B21, 1C21, 2A21, 2A25, and 2B31). The TMI data have a higher UR (Fig. 5), and the UR of the PR data is lower than the UR of total TRMM standard satellite products.

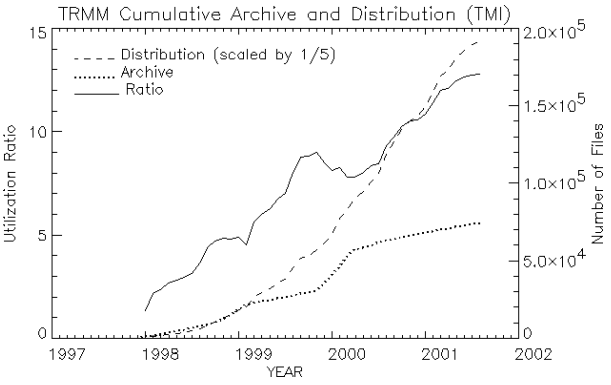

Fig. 5. Cumulative statistics of TRMM TMI products, as of August 2001

The URs of individual products were also examined. The product 1B11, TMI Brightness Temperature, has the highest UR, 16.1 (Fig. 6). This may be because of its usage in producing other products, such as SST. The product 1B21, PR Power, has the lowest UR, 2.4.

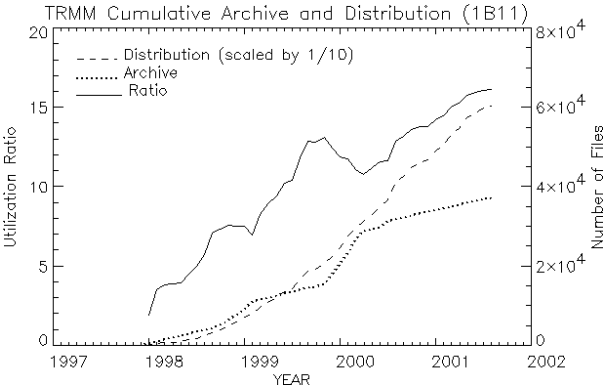

Fig. 6. Cumulative statistics of TRMM 1B11, as of August 2001

3.2 Archive and Distribution Statistics of TRMM Level-3 Products

There are seven TRMM level-3 gridded products, (3A11, 3A25, 3A26, 3A46, 3B31, 3B42, and 3B43). Their total cumulative distribution and archive are 34,969 and 2,140, respectively, in terms of number of files; and 52.7 GB and 2.6 GB, respectively, in terms of data volumes. The UR of gridded products is 16.3 (Fig. 7). The characteristics of this group are data on regular grid mesh, global (40S-40N) coverage, and small file sizes. Therefore, the UR values of TRMM level-3 data products are obviously higher than those of TRMM orbital products. Most of the gridded products have a UR above 10, with a few above 20. 3A25, Spaceborne Radar Rain, has the highest UR, 22.7. (Fig. 8). In fact, several gridded products can also be accessed via

anonymous ftp, the statistics of which are not included here. Therefore, the URs of level-3 products are actually higher.

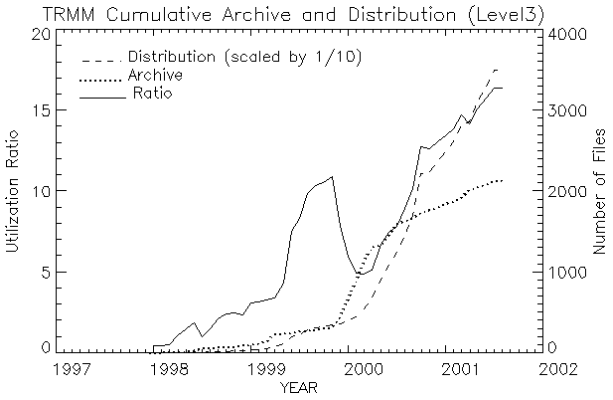

Fig. 7. Cumulative statistics of TRMM level-3 products, as of August 2001

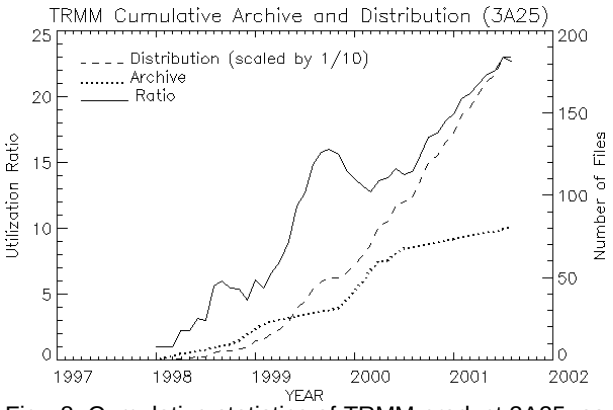

Fig. 8. Cumulative statistics of TRMM product 3A25, as of August 2001

There are some common characteristics in Fig. 3 to Fig. 8. All of the cumulative archive curves show two rapid increases in September 1998, and November 1999, representing the starting months of TRMM version 4 and version 5 reprocessing, respectively. The cumulative distribution curves also show corresponding increases with a short time delay. This trend shows TRMM users' preference for newly reprocessed data in their research.

Table 1. Total archive and distributions as of August 2001

Product/ Group	Distributed Files	Archived Files	Utilization Ratio
Satellite	2,722,479	420,573	6.47
Orbit	2,621,957	342,641	7.65
Level-3	34,969	2,140	16.34
VIRS	314,086	36,929	8.51
TMI	956,781	74,706	12.81
PR	1,351,090	231,006	5.85
1B11	602,501	37,348	16.13
3A25	1,841	81	22.73

Table 1 summarizes the total archive, distribution, and URs, as of August 2001, for the products and groups discussed above. Based on these archive and distribution statistics, we have set up several direct ftp sites for those products with higher URs, as well as other popular products with small file sizes.

4. Statistics of TRMM FTP Distribution

Several TRMM gridded products, value-added subset products, and TRMM ancillary data products are also available via the GES DAAC's anonymous FTP (ftp://lake.nascom.nasa.gov/TRMM/), in addition to TRMM Search and Order System (. These ftp sites provide TRMM users with more direct and faster access to TRMM data. Table 2 shows the FTP distributions for selected TRMM and its ancillary data products. One could see that the Gridded Orbital subsets G2A12 and G2B31, and their Georegional subsets are highly demanded.

Based on these FTP distribution statistics, we may make certain adjustments, such as suspending the G1B01 subsetting. An alternative VIRS L1B product, i.e., convert the 1B01 radiances to percent albedo for visible channels and to brightness temperatures for thermal channels as deemed more useful to VIRS users is being considered.

Table 2. FTP distributions (number of files) for selected TRMM data products, as of August 2001

Product	1999	2000	2001	Total
G1B01	23,125	240	0	23,365
G2A12	51,809	25787	8,404	86,000
G2B31	39,800	6,300	551	46,651
Rgn.Gridded*	13,988	2,557	10,067	26,612
3B42	0	0	40	40
3B43	0	0	28	28
3A44	1	59	1,060	1,120
3A45B	61	107	66	234

* Georegional subsets of Gridded orbital data G1B01, G2A12, and G2B31.

5. Conclusions

Overall, the UR has increased steadily as TRMM progressed, and that trend is continuing. As measured by the UR, the most frequently requested satellite orbital data products are TMI Brightness Temperature, and TMI Rain Profile, with URs above 12. In general, UR tends to increase with product level. Most of the level-3 products have a UR above 10, with a few above 20. FTP is by far the delivery mechanism of choice. These statistics not only help the GES DAAC to refine its data support strategy, but also are useful inputs to the design of future rain measuring satellite data support systems, such as the TRMM follow-on mission, proposed the Global Precipitation Mission (GPM).