A physically-based scheme for the retrieval of precipitation from cross-track sensors in the GPM constellation

UNIVERSITY OF MARYLAND

Chris Kidd[12], Toshi Matsui[12], Jiun-Dar Chern[12], Chris Kummerow[3], Dave Randel[3], Pierre Kirstetter[4] and Walt Petersen[5]

[1]UMD/ESSIC, [2]NASA/GSFC, [3]CSU, [4]NWC, [5]NASA/MSFC

Goddard Space Flight Center

Introduction

One of the proposed strengths of the current GPM GPROF precipitation retrieval scheme is the unification of estimates across all the sensors in the GPM constellation. This is currently achieved by sensor-dependent databases derived from the 'original' DPR-GMI database. However, since there are fundamental differences between the sensors in terms of channel availability and spatial resolution, the transfer of information from one database to another leads to a loss in fidelity.

It is therefore logical to, instead of deriving the other databases from the GMI database, generate databases for these sensors directly through comparisons with the DPR (or 'combined'). Consequently matchups were made between the GPM DPR-Ku, GMI and the MHS and SAPHIR instruments. The first 3 years of GPM provides 13,990 MHS cases and 2,439 SAPHIR cases with 22,727,663 and 10,365,732 matchups respectively: the matchups are confined to co-located (<2.5 km) footprints for all sensors within ± 5 minutes. Each matchup provides all L1C and L2A information from each sensor.

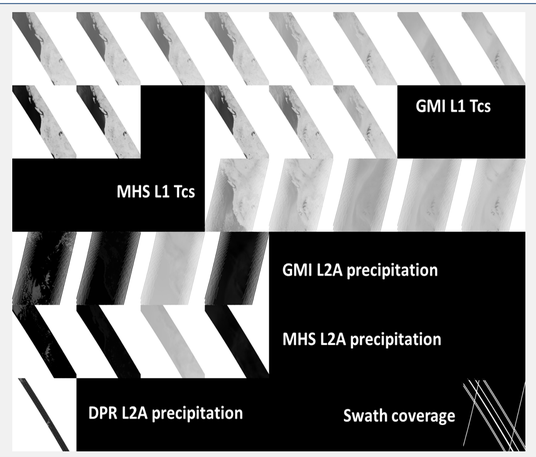


Figure 1: example of matchups between MHS, GMI and DPR

Each case is stored in an individual file with a header of the contributing filenames, together with the location of the coincident overpass location; this is followed by the matched fovs, one per line with time, lat/lon, Tcs (or precipitation rate), surface types, TPW and T2m. These files form the basis of the 'fat' database from which the operational database can be derived.

The 'thin' database, for simplicity and computational efficiency, contains the (calibrated) brightness temperatures together with the matched DPR-Ku near surface precipitation retrieval: the database is stored as a simple sequential text file. To maximise the speed of retrieval the database is accessed through an index file which contains the start/end location for each database bin, together the number of 'profiles', the mean Tcs of each channel and the Tc variance. One index was generated using the same TPW/T2m divisions used by GPROF, and another index using Tcs of the two most disparate channels (for SAPHIR, channels 1 & 6). The latter database needs no ancillary information.

Figure 2 below illustrates the production and retrieval flow of the scheme.

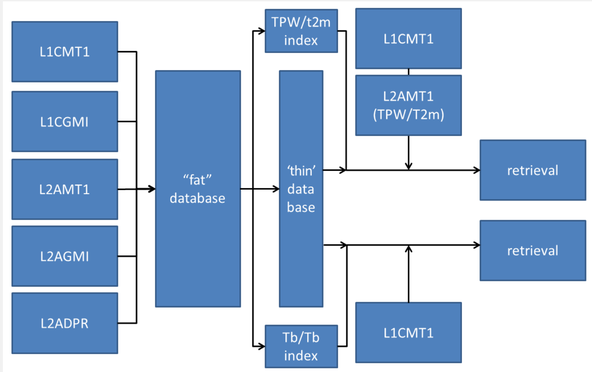


Figure 2: Flow diagram of SAPHIR (MT1), GMI and DPR matched databases and retrievals

The retrieval of precipitation using this scheme is straightforward, in this case for the retrieval of precipitation from SAPHIR:

i) the thin database of c. 10m values is read into memory, together with the index file;
ii) a list of orbits files is opened and each L1C orbit of Tcs is read in, then;
iii) For each fov the database bin is established (Tc1 & Tc6) and the observed Tcs are compared with the Tcs of the database.

The matching of the observed and database Tcs is a critical part of the retrieval process. GPROF currently uses a Bayesian scheme to do so, but this suffers from (potentially) including all possible data in each retrieval. Previous studies have suggested that similar, or improved retrievals are possible using just the nearest 'n' profiles, or a distance-weighted scheme (Figure 3).

A number of weighting functions have been tested (see below). Interestingly using a simple 2-channel Tc retrieval with the mean precipitation rate per bin provides a surprisingly good estimate of precipitation, particularly in high-precipitation regions, but less good in regions of low precipitation. Taking just the nearest profile (in Tc-space)provides a similarly good retrieval, although with greater noise due to the poorly-sampled heavy precipitation events (Figure 4).

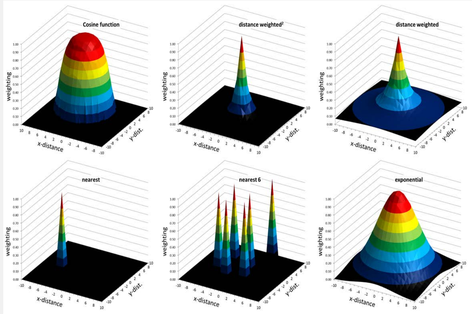


Figure 3: Proposed weighting functions of extracting precipitation from relevant 'profiles' in the database.

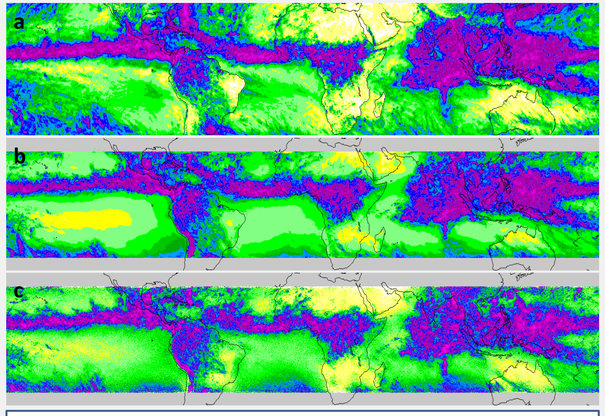
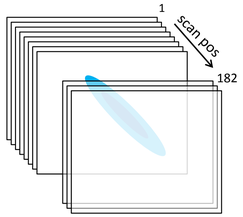
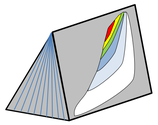


Figure 4: Monthly accumulated precipitation from (a) SSMIS (all sensors) using GPROF-V05, (b) SAPHIR using Tc:Tc database with no ancillary data and mean precipitation/bin, and (v) SAPHIR using the precipitation of nearest Tc match. It can be noted that (c) better matches the low-precipitation regions depicted in the SSMIS retrievals. However, note the anomalous SAPHIR retrievals over the Andes due to the dry, high-elevation.

Future directions:

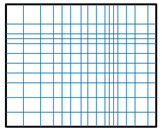
Cross track scan positions database: separation of databases by scan position to better represent the vagaries of variable Earth incident angles – and allow cross-scan position searching

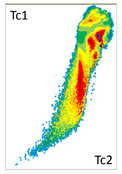




Cross database searches: with increasing precipitation – and water vapour the surface effects become increasing irrelevant, therefore it makes sense to combine the poorly-sampled bins with high intensity precipitation

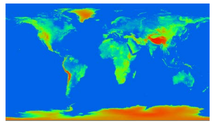
Variable 'mesh' databases: designed to better represent the information within each bin – those with large number of samples have finer database resolution, while those bins with few samples have a coarser database resolution.





Dynamic database searches: to better represent the database entries – some bins have little variation in the Tcs which require finer searches, while those bins with large variations require larger searches to provide similar profiles.

Use of static ancillary data: retrievals from sounders are affected over dry, high-elevation regions – thus the use of (static) altitude information is prudent to improve the accuracy of the retrievals



Conclusion

The current scheme here has shown great promise and will be further refined as the availability of more coincident GPM/constellation sensors become available. The scheme is to be implemented by the EUMETSAT CM-SAF for long term precipitation retrievals using the established FCDR Tb data series; critically, the static database and lack of reliance upon model information makes it an ideal technique to be used on such data climate data records.

Acknowledgements: This research is supported by NASA's PMM grant NNX16AE22G. Thanks also go to the NASA PPS for the provision of the data sets used in this study.