Heading Toward Launch with the Integrated Multi-satellitE Retrievals for GPM (IMERG)

The GPM Multi-Satellite Team
George J. Huffman NASA/GSFC, Chair
David T. Bolvin SSAI and NASA/GSFC
Daniel Braithwaite Univ. of California Irvine
Kuolin Hsu Univ. of California Irvine
Robert Joyce Wyle Scientific and NOAA/NWS/CPC
Chris Kidd ESSIC and NASA/GSFC
Soroosh Sorooshian Univ. of California Irvine
Pingping Xie NOAA/NWS/CPC

Introduction
IMERG Design
Implementation
Future
Final Comments
1. INTRODUCTION

Individual precip estimates, present various
- periods of record
- observing times
- regions of coverage
- sensor-specific strengths and limitations

Combined datasets are critical to non-expert users

GPM wants to combine “all” precipitation-relevant satellites

Image by Eric Nelkin (SSAI), 20 April 2012, NASA/Goddard Space Flight Center, Greenbelt, MD.
1. INTRODUCTION

The GPM multi-satellite product goals:

- seek the longest, most detailed record of “global” precip
  - *don’t* use regional data sets
  - *do* use gauge data
- combine the input estimates into a “best” data set
  - *not* a Climate Data Record
  - relatively uniform input data

Image by Eric Nelkin (SSAI), 20 April 2012, NASA/Goddard Space Flight Center, Greenbelt, MD.
1. INTRODUCTION – Combination Concepts

The “good stuff” (microwave) is sparse
- 30 min has lots of gaps
- extra gaps due to snow in N. Hemi.
- 4 imagers (2 more getting ready), 3 sounders

IMERG is a unified U.S. algorithm that takes advantage of
- Kalman Filter CMORPH (lagrangian time interpolation) – NOAA
- PERSIANN with Cloud Classification System (IR) – U.C. Irvine
- TMPA (inter-satellite calibration, gauge combination) – NASA
- all three have received PMM support
2. IMERG DESIGN – Requirements/Goals

Resolution – 0.05° ~0.1° [i.e., roughly the resolution of microwave, IR footprints]

Time interval – 30 min. [i.e., the geo-satellite interval, then aggregated to 3 hr]

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

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

Product sequence – early sat. (~4 hr), late sat. (~12 hr), final sat.-gauge (~2 months after month) [more data in longer-latency products]

Sensor precipitation products intercalibrated to TRMM before launch, later to GPM

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 – liquid, mixed, solid

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
2. IMERGE DESIGN – Multiple Runs

Multiple runs serve different needs for timeliness
- more delay usually yields a better product
- pioneered in TMPA

**Early** – first approximation; flood, now-casting users
- current input data latencies
  PPS support ~4-hr delay
- truly operational users (3 hr) not well-addressed

**Late** – wait for full multi-satellite; crop, flood, drought analysts
- driver is the wait for microwave data for backward propagation
- expect delay of 12-18 hr

**Final** – after the best data are assembled; research users
- driver is precip gauge analysis
- GPCC gauge analysis is finished ~2 months after the month
2. IMERG DESIGN – Processing

Institutions are shown for module origins, but

- package will be an integrated system
- goal is single code system appropriate for all three runs
- “the devil is in the details”
2. IMERG DESIGN – Data Fields

Output dataset includes intermediate data fields
- users and developers require
  - traceability of processing, and
  - support for algorithm studies

0.1° global CED grid
- 3600x1800 = 6.2M boxes
- fields are 1-byte integer or or scaled 2-byte integer / 4-byte real
- but dataset compression means smaller disk files
- PPS will provide subsetting

“User” fields in italics, darker shading

<table>
<thead>
<tr>
<th>Half-hourly data file (early, late, final)</th>
<th>Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Calibrated multi-satellite precipitation</td>
<td>12 / 25</td>
</tr>
<tr>
<td>2  Uncalibrated multi-satellite precipitation</td>
<td>12 / 25</td>
</tr>
<tr>
<td>3  Calibrated multi-satellite precipitation error</td>
<td>12 / 25</td>
</tr>
<tr>
<td>4  PMW precipitation</td>
<td>12 / 25</td>
</tr>
<tr>
<td>5  PMW source 1 identifier</td>
<td>6</td>
</tr>
<tr>
<td>6  PMW source 1 time</td>
<td>6</td>
</tr>
<tr>
<td>7  PMW source 2 identifier</td>
<td>6</td>
</tr>
<tr>
<td>8  PMW source 2 time</td>
<td>6</td>
</tr>
<tr>
<td>9  IR precipitation</td>
<td>12 / 25</td>
</tr>
<tr>
<td>10 IR KF weight</td>
<td>6</td>
</tr>
<tr>
<td>11 Precipitation type (liquid/mixed/solid)</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monthly data file (final)</th>
<th>Size (MB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Satellite-Gauge precipitation</td>
<td>12 / 25</td>
</tr>
<tr>
<td>2  Satellite-Gauge precipitation error</td>
<td>12 / 25</td>
</tr>
<tr>
<td>3  Gauge relative weighting</td>
<td>6</td>
</tr>
<tr>
<td>4  Precipitation type (liquid/mixed/solid)</td>
<td>6</td>
</tr>
</tbody>
</table>
3. IMPLEMENTATION – Testing

“Baseline” Version 2 code delivered November 2011

“Launch-ready” code is due November 2012

Code will “freeze” in June 2013 for operational testing

Plan to bring up IMERG first in (more-flexible) PPS RT system
  • shake out bugs and conceptual problems
  • start quasi-operational production of “proxy” GPM data
  • likely we can release parallel products

Use lessons learned to upgrade the production code

PMM focus on validation is key
  • refine physical concepts
  • demonstrate level of confidence
3. IMPLEMENTATION – Transitioning from TRMM to GPM

IMERG will be computed at launch (February 2014) with TRMM-based coefficients.

About 6 months after launch expect to re-compute coefficients and run a fully GPM-based IMERG.

• compute the first-generation TRMM/GPM-based IMERG archive, 1998-present
• all runs will be recomputed for the entire data record
• when should we shut down the TMPA legacy code?

Contingency plan if TRMM ends before GPM is fully operational:

• institute climatological calibration coefficients for the legacy TMPA code and TRMM-based IMERG
• continue running
• particularly true for Early, Late
4. FUTURE – Outstanding Issues

High-quality estimates in snowy/icy regions

- not yet operational
- when snow estimates appear, we hope they will work with legacy sensors, at least back to the start of AMSU in 2000
4. FUTURE – What Next?

The clear goal for Day-1 is operational code meeting GPM deadlines; after that …

- **implement a high-latitude scheme**
  - develop high-latitude precip estimates
  - calibration schemes for high-latitude precip estimates
  - leo-IR–based displacement vectors
  - parallel observation-model combined product

- **use sub-monthly** (daily, pentad, or dekad) **gauge analyses**

- **refined precipitation type estimates**

- **alternative scheme for computing displacement vectors**

- **address cloud growth**

- **convective/stratiform classification**

- **address orographic enhancement**

- **error estimates**
  - bias and random
  - scale and weather regime dependence
  - user-friendly formats and cutting-edge science

- **intercalibrate across sensors with different capabilities**

- **revise precipitation gauge wind-loss corrections**
5. FINAL COMMENTS

The Day-1 GPM multi-satellite precipitation algorithm is planned as a unified U.S. algorithm.

IMERG will provide fine-scale estimates with three latencies for the entire TRMM/GPM era.

The system is planned to meet GPM requirements and to provide the hooks for future extensions.

There are still lots of interesting combination and science projects to address.

gorge.j.huffman@nasa.gov