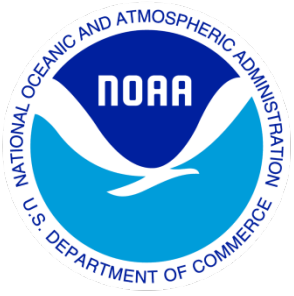
Climate Data Record (CDR) Program

Climate Algorithm Theoretical Basis Document (C-ATBD)

Bias-Corrected CMORPH High-Resolution Global Precipitation Estimates [Precipitation – CMORPH]



CDR Program Document Number: CDRP-ATBD-0812
Configuration Item Number: 01B-23
Revision 0 July 27, 2018

REVISION HISTORY

Rev.	Author	DSR No.	Description	Date
0	Pingping Xie, Robert Joyce, Shaorong Wu NOAA Climate Prediction Center	DSR-1017	Initial Submission Baselined in CDRP Library	07/27/2018

TABLE of CONTENTS

LIST of FIGURES

LIST of TABLES

1. Introduction

1.1 Purpose

The purpose of this document is to describe the algorithm submitted to the National Centers for Environmental Information (NCEI) by Dr. Pingping Xie / NOAA Climate Prediction Center (CPC) that will be used to create the Bias-Corrected CPC Morphing technique (CMORPH) Climate Data Record (CDR), using the Level 2 precipitation rate retrievals from passive microwave (PMW) measurements aboard multiple low earth orbit (LEO) satellites and infrared (IR) brightness temperature (TBB) data from geostationary (GEO) platforms. The actual algorithm is defined by the computer program (code) that accompanies this document. The intent of this document is to provide a guide to understanding that algorithm, from both a scientific perspective and in order to assist a software engineer or end-user performing an evaluation of the code.

1.2 Definitions

Bias-corrected CMORPH is generated on an 8kmx8km grid over the global domain (60^oS-60^oN) in 30-min intervals for an 18+ year period from January 1, 1998 to the present.

Inputs to the processing system include Level 2 retrievals of PMW measurements from sensors aboard all available LEO platforms, NESDIS daily snow maps, and IR TBB arrays from five GEO satellites:

PR_raw = raw precipitation retrievals from LEO PMW sensors

SNOW_surf = daily surface snow map

TBB_geo = TBB array from GEO satellites

First, raw precipitation retrievals (*PR_raw*) from individual PMW sensors are calibrated against those from the Tropical Rainfall Measurement Mission (TRMM) Microwave Imager (TMI). The calibrated precipitation retrievals (*PR_cal*) from individual PMW sensors are then combined in time and space to form a combined PMW retrieval field (*MWCOMB*):

PR_cal = calibrated precipitation retrievals from all LEO PMW sensors

MWCOMB = combine global fields of calibrated PMW precipitation retrievals

Daily snow cover maps (*SNOW_surf*) created by the NESDIS/STAR are utilized here to ensure the quality of MWCOMB over grid boxes covered by snow.

At the meantime, motion vectors of precipitating clouds ($\vec{V}$*_cld*) are derived through comparing the cloud patterns depicted on two consecutive *TBB_geo* images 30 minutes apart.

$\vec{V}\_cld$ = *motion vector of precipitating clouds*

Precipitating pixels depicted in the combined PMW retrievals (*MWCOMB*) are propagated from their respective measurement times to the target analysis time along the cloud motion vector ($\vec{V}\_cld$). This propagation is performed in both the forward and backward directions.

PR_fwd = *precipitation retrievals propagated forward*

PR_bwd = *precipitation retrievals propagated backward*

The raw CMORPH satellite precipitation estimates are defined as the weighted mean of the forward and backward propagated PMW retrievals, with the weight inversely proportional to the length of the propagation time.

CMORPH_raw = *raw CMORPH satellite precipitation estimates*

The last step of the processing is to remove the bias in the raw CMORPH. This is done through calibration against the CPC gauge-based analysis of daily precipitation over land and against the pentad Global Precipitation Climatology Project (GPCP) merged analysis over ocean.

CGA_dly = *CPC daily gauge analysis over land*

GPCP_pen = *pentad GPCP analysis over ocean*

CMORPH_crt = *bias-corrected CMORPH satellite precipitation estimates*

1.3 Referencing this Document

This document should be referenced as follows:

Bias-Corrected CMORPH - Climate Algorithm Theoretical Basis Document, NOAA Climate Data Record Program <CDRP-ATBD-0812 by CDRP Document Manager> Rev. X (YYYY). Available at http://www.ncdc.noaa.gov/cdr/operationalcdrs.html

1.4 Document Maintenance

Versions for the products and software described in this document are as follows:

Product Version	Software Version	CATBD Revision	Release Date	Remarks
1.0	1.0	1.0	2015-10-20	Initial Release

2. Observing Systems Overview

2.1 Products Generated

The products generated by this CDR package are the bias-corrected CMORPH integrated global satellite precipitation estimates. For this release, the CMORPH CDR precipitation data sets span an 18-year period from January 1, 1998 to the present and will be updated on a quasi real-time basis. The data set covers the globe from 60^{o}S to 60^{o}N.

Three sets of CMORPH satellite precipitation estimates are included in the products package, providing global precipitation fields at a combination of time-space resolutions including a) 30-minute / 8kmx8km, b) hourly / 0.25^{o}lat/lon; and c) daily / 0.25^{o}lat/lon, respectively.

2.2 Instrument Characteristics

Bias corrected CMORPH is created using information from five categories of space-borne and *in situ* measurements, including:

a) Level 2 precipitation rate retrievals from passive microwave (PMW) sensors aboard multiple low earth orbit (LEO) satellites. The retrievals are derived from PMW emission channels over ocean and from scattering channel measurements over land. Table 1 lists satellite PMW sensors whose measurements are used to derive the Level 2 precipitation retrievals used as inputs to the CMORPH integrated satellite precipitation estimates;

Table 1: PMW sensors whose measurements are used to derive L2 Precipitation Retrievals

PMW Sensor	Low Earth Orbit (LEO) Platform Carrying the Sensor
TRMM Microwave Imager(TMI)	Tropical Rainfall Measurement Mission (TRMM)
Advanced Microwave Scanning Radiometer (AMSR)	AQUA
Special Sensor Microwave Imager / Sounder (SSMIS)	Defense Meteorological Satellite Program (DMSP) F-16, F-17, F-18
Special Sensor Microwave Imager (SSM/I)	Defense Meteorological Satellite Program (DMSP) F-13, F-14, F-15
Microwave Humidity Sounder (MHS)	NOAA-18, NOAA-19, MetOp-A, MetOp-B
Advanced Microwave Sounding Unit (AMSU)	NOAA-15, NOAA-16, NOAA-17
Micro-Wave Radiation Imager (MWRI)	FunYun (FY) – 3B

b) Interactive Multisensor Snow and Ice mapping System (IMS) daily 4 km snow and sea ice maps over the northern hemisphere, constructed using satellite observations from visible, PMW and other channels;

c) Full-resolution (4kmx4km,30-min) global surface / cloud top temperature data defined through compositing infrared (IR) window channel measurements from five geostationary satellites, which currently including GOES-West, GOES-East, METEOSAT-10, METEOSAT-7, and HIMAWARI-8 as of March 2016;

d) CPC unified daily gauge analysis constructed through interpolation of quality controlled gauge reports of daily precipitation over land; and

e) GPCP merged analysis of pentad precipitation over ocean constructed on a 2.5^{o}lat/lon grid over the globe through combining information from gauge measurements and satellite estimates.

3. Algorithm Description

3.1 Algorithm Overview

Bias-corrected CMORPH high-resolution global precipitation CDR product is constructed in two steps through integrating information from multiple satellite and *in situ* based sources. The first step is to define integrated high-resolution global satellite precipitation estimates (the raw CMORPH). To this end, Level 2 precipitation rate retrievals from passive microwave (PMW) measurements aboard multiple low earth orbit (LEO) satellites are propagated from their respective observation times to the target analysis time along the cloud motion vectors derived from consecutive geostationary (GEO) infrared (IR) images. A description of the algorithm may be found in Joyce et al. (2004). In the second step, bias in the raw CMORPH is removed through Probability Density Function (PDF) matching against the CPC daily gauge analysis over land (Xie et al. 2010) and through adjustment against the pentad GPCP merged analysis over ocean (Xie et al. 2003).

3.2 Processing Outline

As illustrated in figure 1, the bias-corrected CMORPH processing is composed of two consecutive steps, i.e. a) constructing CMORPH integrated satellite precipitation estimates (the raw CMORPH), and b) removing bias in the raw CMORPH.

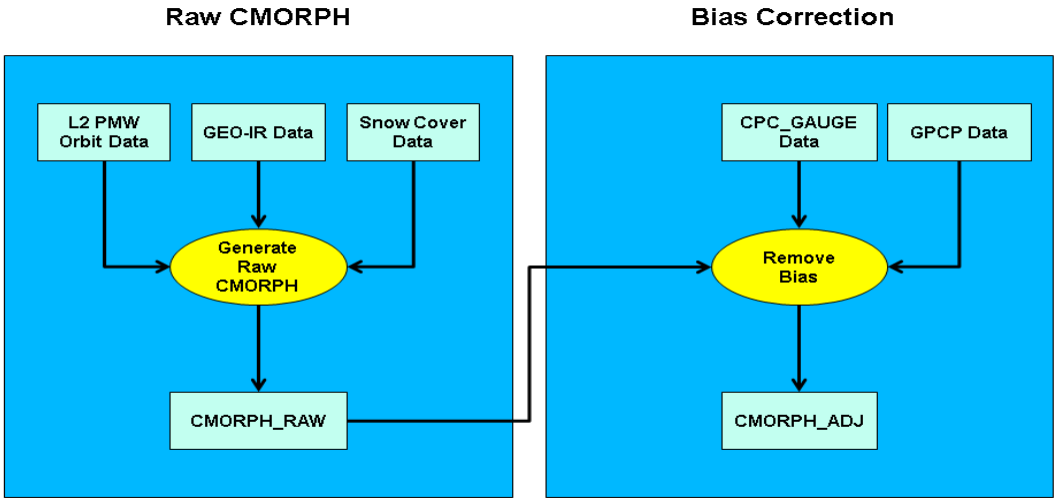


Figure 1: Overall flow chart of the bias-corrected CMORPH processing system. Blue rectangles indicate input / output data files while yellow ellipses represent individual processing procedures.

The purely satellite-based raw CMORPH is constructed through integrating precipitation information from multiple passive microwave (PMW) sensors aboard LEO

satellites (figure 2). Preprocessing is first performed for the Level 2 precipitation rate retrievals (*PR_raw*) from all available PMW sensors. Level 2 retrievals at respective satellite footprints are re-mapped to a common equal angle global grid system of 8kmx8km at the equator. Calibration is then performed through PDF matching against a common reference fields using temporally / spatially co-located data pairs. Retrievals from the Tropical Rainfall Measurements Mission (TRMM) Microwave Imager (TMI) are utilized as the calibrator in this version of the CMORPH CDR. This calibration ensures input PMW retrievals from individual sensors of close PDF characteristics. More details of this inter-satellite calibration may be found in section 3.4.4b).

The remapped and calibrated Level 2 PMW precipitation rate retrievals (*PR_cal*) are then composited into a single global field of precipitation estimates, called MWCOMB, at a time / space resolution of 30-minute / 8kmx8km. When PMW retrievals from multiple sensors are available over a specific grid box and for a specific 30-minute time slot, only the one with the highest quality is used to define the MWCOMB. The quality of PMW retrievals from various sensors is ranked, as shown in Table 2. This ranking is based on published studies by other scientists (e.g. Tian et al. 2014) as well as our own examinations (e.g. Joyce and Xie 2011). Evert time a new version of Level 2 precipitation retrievals are available from a new or existing sensor, a set of examination procedures will be implemented by the CMORPH developers at CPC to quantify the error of the L2 products through comparison against surface radar and gauge observations. Figure 3 illustrates an example of the MWCOMB for 00:00Z, August 1, 2014.

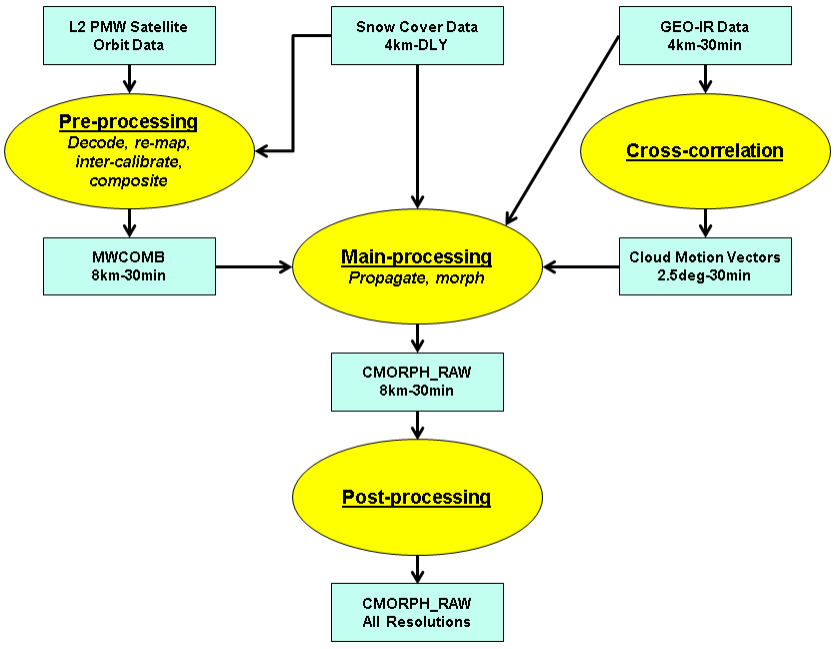


Figure 2: Flow chart of the raw CMORPH integrated satellite precipitation estimates.

Table 2: Quality Ranking of PMW Precipitation Retrievals from Various sensors Aboard Various Platforms

Quality Ranking	*PMW Sensor* *(In Decreasing Quality Order from Top to Bottom)*	*LEO Platforms* *(in Decreasing Quality Order from Left to Right)*
1	*TMI*	*TRMM*
2	*AMSR*	*AQUA*
3	*MWRI*	*FY-3B*
4	*SSMIS*	*F-18, F-17, F-16*
5	*SSMI*	*F-15, F-14, F-13*
6	*MHS*	*METOP-B, METOP-A, NOAA-19, NOAA-18*
7	*AMSU*	*NOAA-17, NOAA-16, NOAA-15*

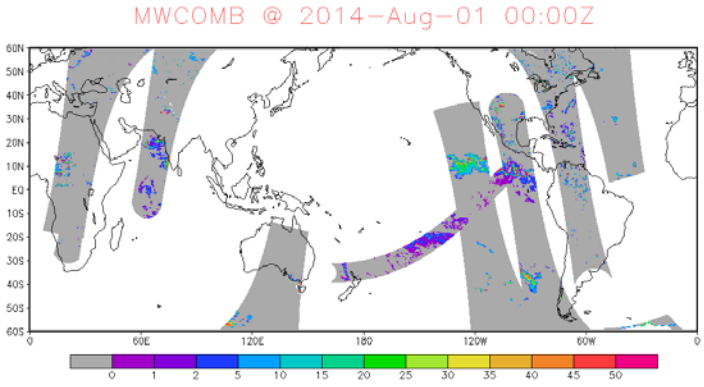


Figure 3: Composite calibrated passive microwave (PMW) precipitation retrievals (MWCOMB, mm/hr) for the 30minute time slot beginning at 00:00Z, August 1, 2014.

At the meantime, cloud motion vectors ($\vec{V}$_cld) are computed from consecutive GEO IR images in 30-minute intervals through the cross-correlation technique (*figure 2, right side*). The GEO IR images used here are the so-called even - odd satellite fields, which contains two global arrays of GEO IR data (*TBB_geo*) for the even (GOES-W, operational Meteosat coverage west of Greenwich, all of Indian Ocean METEOSAT coverage) and odd satellite configuration (GMS, GOES-E, operational Meteosat coverage east of Greenwich), respectively, to include the *TBB_geo* data from both satellites over the overlapping edges for improved definition of cloud motion vectors. Cloud motion vectors ($\vec{V}$_cld) are calculated for each grid of 2.5^{o}lat/lon interval using data over a 5^{o}lat/lon domain centering

at the target grid point. Spatial pattern correlation is computed between the *TBB_geo* array for the target analysis time and that for the previous time step at a combination of spatial displacements in both zonal and merional directions. The spatial displacement with the maximum correlation is assumed to be caused by the mean movement of the cloud systems over the region and is used to define the cloud motion vector ($\vec{V}$_cld) at the target grid point. The cloud motion vector defined on the 2.5°lat/lon grid system is then downscaled to the MWCOMB resolution of 8kmx8km through bi-linear interpolation. More details may be found in Section 3.4.4a). Figure 4 shows an example of the cloud motion vectors for 00:00Z, August 1, 2014.

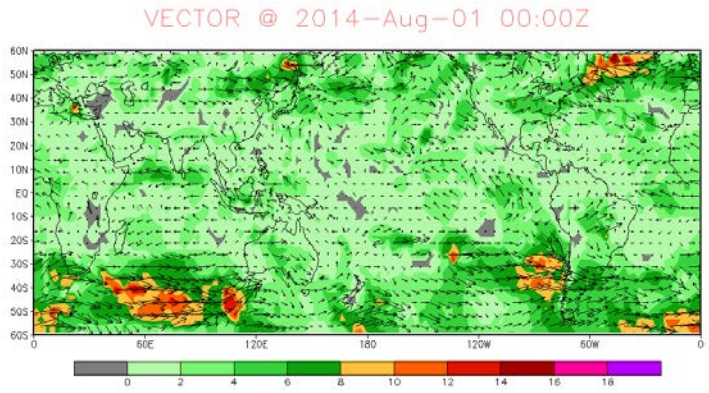


Figure 4: Motion vectors of precipitating clouds for 00:00Z, August 1, 2014. Arrows indicate the direction while the color shading represents the speed (in number of 8kmx8km grid boxes), respectively.

The PMW precipitation rate retrievals composited in the MWCOMB are then propagated from their respective observation times to the target analysis time (*figure 2, middle column*). The propagation is conducted in both the forward and backward directions to fully take advantage of the PMW observations before and after the target analysis time. The raw CMORPH integrated satellite precipitation estimation (*CMORPH_raw*) at a grid box of 8kmx8km is defined as the weighted mean of the PMW retrievals propagated there from the forward (*PR_fwd*) and backward directions (*PR_bwd*) with the weights inversely proportional to the length of the propagation time:

$$CMORPH\_raw = W\_fwd \bullet PR\_fwd + W\_bwd \bullet PR\_bwd \quad \text{<1>}$$

$$W\_fwd = T\_bwd / (T\_fwd + T\_bwd) \quad \text{<2>}$$

$$W\_bwd = T\_fwd / (T\_fwd + T\_bwd) \quad \text{<3>}$$

Where *T_fwd* and *T_bwd* are the lengths of the forward and backward propagation time, respectively.

Once the MWCOMB PMW retrievals are propagated and morphed onto an 8kmx8km grid and for a 30-minute time step, quality control is performed using the daily surface snow map produced by the NOAA National Environmental Satellite, Data, and Information Service (NESDIS) / Center for Satellite Applications and Research (STAR) and the full-resolution IR data to screen out suspicious precipitation estimates over grid boxes with snow or sea ice coverage (*figure 2, upper-right*).

Post-processing (*figure 2, bottom*) is finally implemented to create global fields of purely satellite based raw CMORPH precipitation estimates at an assortment of time /space resolutions, including 30-minute - 8km, hourly - 0.25°lat/lon, daily (00Z-00Z) - 0.25°lat/lon, and daily (gauge End of Day) – 0.25°lat/lon. Figure 5 shows an example of the raw CMORPH for 00:00Z, August 1, 2014.

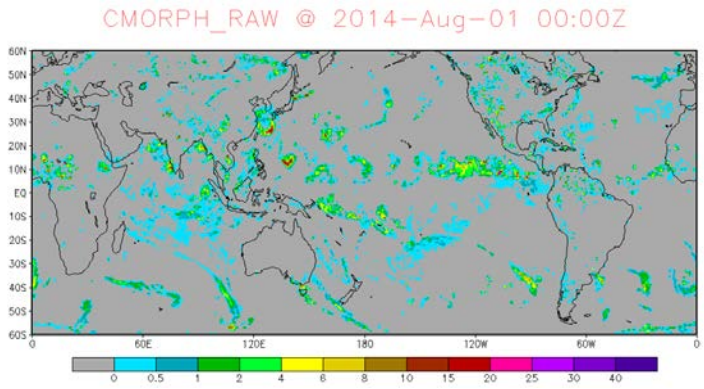


Figure 5: Raw CMORPH integrated satellite precipitation estimates (mm/hr) for 00:00Z, August 1, 2014.

The purely satellite based raw CMORPH described above contains bias passed through from the input Level 2 satellite PMW retrievals. The second big part of the CMORPH processing system is designed to remove that bias (figure 6).

The bias correction is carried out for the land and oceanic regions, respectively. Over land (*figure 6, left side*), the bias correction is performed through PDF matching against the gauge data: CPC gauge based analysis of daily precipitation (CGA_dly, Xie et al. 2010). Over ocean (*figure 6, right side*) the raw CMORPH is calibrated against the pentad GPCP merged analysis (*GPCP_pen*) of Xie et al. (2003).

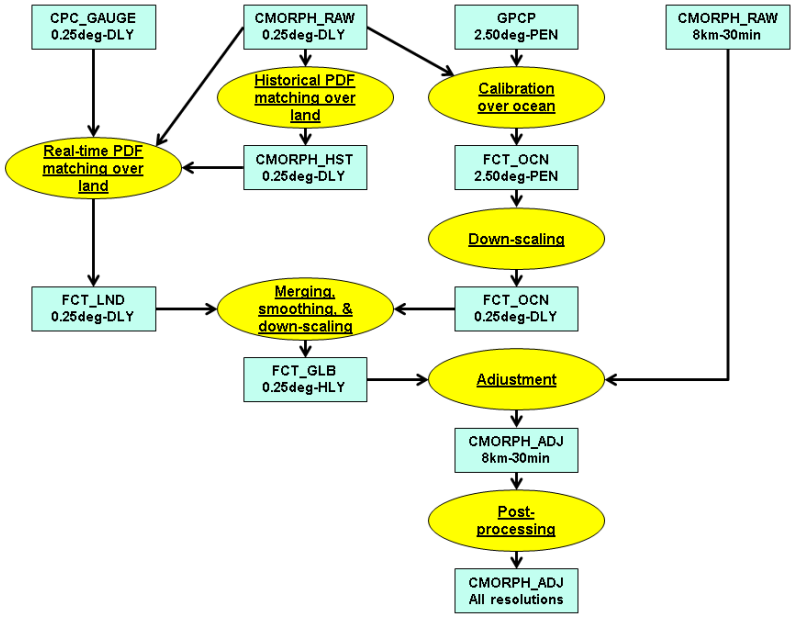


Figure 6: Flow chart for the CMORPH bias correction procedures.

The bias correction over land is implemented in two consecutive steps. First, bias correction is performed using PDF tables constructed using historical data. PDF tables are constructed for each calendar day and for each 0.25°lat/lon grid box over the global land using time-space co-located daily gauge and daily RAW CMORPH data over a region of 1°lat/lon centering at the target grid box and covering a 31-day period centering at the target date over the 15-year period from 1998 to 2012. Called Climatological Correction, the bias correction based on historical data is conducted through multiplying the raw CMORPH precipitation estimates with an adjustment factor, defined as the product of the correction ratio derived directly from the PDF matching and the under-estimation correction factor. More details on the construction of PDF tables using historical data and the definition of the under-estimation correction factor are included in Sections 3.4.5c) and d), respectively.

In the second step, called real-time refinement, the PDF bias correction is repeated but with the PDF tables established using co-located CMORPH and daily gauge data for a 31-day period centering at the target date to account for the raw CMORPH bias variations not detected and removed by the climatological correction. PDF tables are constructed for each 0.25°lat/lon grid using data over a circular region centering at the target grid box. The data domain is expanded until sufficient number of data pairs is collected. The output of this land bias correction procedure is a daily array of adjustment factor defined on a grid of 0.25°lat/lon (*FCT_lnd*).

Over ocean, the raw CMORPH satellite precipitation is calibrated against the pentad GPCP merged analysis of precipitation (*GPCP_pen*). The raw CMORPH satellite estimates are first up-scaled from their original resolution (8kmx8km – 30-minute) to pentad and 2.5°lat/lon to match with that for the GPCP merged analysis. Adjustment factor for the raw CMORPH is first calculated for each pentad time step and for each 2.5°lat/lon grid as the ratio between the mean GPCP and mean raw CMORPH averaged over a circular region of 3 grid boxes in radius and a 19-pentad time period centering at the target grid box and target analysis time. The adjustment factor computed for each 2.5°lat/lon grid box and for each pentad period is then down-scaled to daily and 0.25°lat/lon resolution (*FCT_ocn*), assuming no changes in the adjustment factor inside the 2.5°lat/lon grid box and within a pentad time period.

The adjustment factors computed for land and ocean are then merged into a combined global field. Adjustment factors over the coastal grid boxes are first smothered to reduce discontinuities. The smoothing is performed over oceanic grid boxes within the distance of two 0.25°lat/lon grid boxes from any land points. The global field of 0.25°lat/lon and daily resolution is then down-scaled to create a global array of adjustment factor on an 8kmx8km and 30-min resolution (*FCT_glb*), again assuming no changes in the adjustment factor inside a 0.25°lat/lon – daily domain and applying smoothing crossing the grid box and time step boundaries.

The final bias corrected CMORPH (*CMORPH_adj*) is defined applying this adjustment factor (*FCT_glb*) to the raw CMORPH (*CMORPH_raw*):

$$CMORPH\_adj = FCT\_glb \bullet CMORPH\_raw \quad \ldots\ldots \quad <4>$$

The reference standard used in the bias correction over land is the CPC Unified Daily Gauge Analysis. Systematic difference of 5-10% exists between this CPC daily gauge analysis used here and the GPCP monthly merged analysis (Adler et al. 2003) used to adjust other CDR data sets such as the PERSIANN CDR (Hsu 2014). At least two factors attribute to this difference. First, a climatological adjustment is applied to the GPCP merged analysis to account for the under-catch of gauge measurements by wind effects while no such correction is conducted for the CPC gauge analysis. The other factor is the disagreements in the gauge station reports of daily and monthly precipitation used to define the CPC daily gauge analysis and the monthly gauge analysis created by the Global Precipitation Climatology Centre GPCC) of Germany and utilized to determine the magnitude of the GPCP merged analysis.

Figure 7 presents an example of the bias corrected CMORPH for 00:00Z, August, 2014.

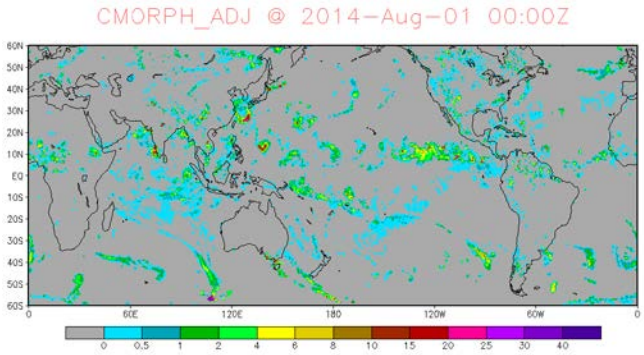


Figure 7: Bias corrected CMORPH integrated satellite precipitation estimates (mm/hr) for 00:00Z, August 1, 2014.

3.3 Algorithm Input

3.3.1 Primary Sensor Data

Bias corrected CMORPH is created using information from five categories of space-borne and *in situ* measurements, including:

a) Level 2 precipitation rate retrievals (*PR_raw*) from passive microwave (PMW) sensors aboard multiple low earth orbit (LEO) satellites. The retrievals are derived from PMW emission channels over ocean and from scatter channels over land. A full list of the PMW sensors whose measurements are utilized to derive precipitation estimates is given in Table 1.

   The Level 2 precipitation retrievals are accessible from NASA web site: ftp://trmmopen.gsfc.nasa.gov/pub/trmmdata/ByDate/V07/

While those from all other PMW sensors are available at the NOAA CLASS ftp site: http://www.class.ngdc.noaa.gov/saa/products/search?datatype_family=MSPPS_ORB

b) Interactive Multisensor Snow and Ice mapping System (IMS) daily 4 km snow and sea ice maps over the northern hemisphere. This data set is created by NOAA/NESDIS under the direction of NOAA Ice Center (NIC) and updated on a quasi real-time basis. The data files are available through ftp at: ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02156/4km/

c) Full-resolution (4kmx4km,30-min) global surface / cloud top temperature data defined through compositing infrared (IR) window channel measurements from five geostationary satellites. The data set is updated on a real-time basis at: ftp.cpc.ncep.noaa.gov/precip/global_full_res_IR

d) CPC unified daily gauge analysis constructed through interpolation of quality controlled gauge reports of daily precipitation. The data set is described in Xie et al. (2010) and available on a real-time basis from the CPC ftp site at: ftp.cpc.ncep.noaa.gov/precip/CPC_UNI_PRCP/GAUGE_GLB

e) GPCP merged analysis of pentad precipitation constructed on a 2.5^{o}lat/lon grid over the globe through combining information from gauge measurements and satellite estimates (Xie et al. 2003). The GPCP pentad analysis is currently available through ftp at the CPC ftp site: ftp.cpc.ncep.noaa.gov/precip/GPCP_PEN ftp.cpc.ncep.noaa.gov/precip/GPCP_PEN_RT

3.3.2 Ancillary Data

As described in Section 3.2, construction of the bias-corrected CMORPH high-resolution global precipitation analysis requires inputs of the following ancillary data sets:

a) Global land/sea masks at 8kmx8km and 0.25^{o}lat/lon grids. These land masks are used in both the creation of raw satellite CMORPH and the implementation of bias correction. The land mask data files are included as part of the bias-corrected CMORPH CDR processing package;

b) PDF tables established using historical data from 1998 to 2012. The PDF tables for each calendar are stored in a single data file with a GrADS control file describing the data structure. The historical PDF tables are included as part of the bias-corrected CMORPH CDR processing package;

A third set of ancillary data are PDF tables for the inter-calibration of PMW precipitation retrievals after the loss of the TRMM/TMI in April 2015:

c) PDF tables for calibration of PMW retrievals to the TMI using historical data. Detailed descriptions of these PDF tables can be found in Section 3.4.4b).

3.3.3 Derived Data

Not applicable.

3.3.4 Forward Models

Not applicable.

3.4 Theoretical Description

Not applicable.

3.4.1 Physical and Mathematical Description

Not applicable.

3.4.2 Data Merging Strategy

The bias –corrected CMORPH CDR is constructed through combining information from multiple sensors (IR, PMW, gauge observations et al) aboard multiple platforms (various GEO, LEO satellites as well as in situ measurements). Tempo-spatial homogeneity of the long-term precipitation data set is achieved through the following procedures:

- ***Inter-satellite calibration for the Level 2 retrievals from different sensors aboard different platforms***

  Level 2 retrievals of instantaneous precipitation rates derived from various PMW sensors aboard various LEO platforms are calibrated against a common reference field (TMI) before they are infused into the CMORPH morphing process.

- ***Calibration against long-term data sets***

  The raw CMORPH integrated satellite precipitation estimates are adjusted against the CPC daily gauge analysis over land and the GPCP pentad merged analysis over ocean to remove the bias and ensure long-term consistency.

3.4.3 Numerical Strategy

Not applicable.

3.4.4 Calculations

As described in section 3.2, all computations involved in the bias-corrected CMORPH are quite straightforward. More detailed descriptions are given below for the numerical strategies applied for several key procedures:

a) Cross-correlation technique

Cross-correlation technique is applied to determine the motion vectors for the precipitating cloud systems from two consecutive geostationary IR images. The cloud motion vector is defined for each grid point of 2.5^o latitude / longitude interval. Pattern correlation between the IR data array of 8kmx8km resolution at the target time step and that for the previous time step is calculated using data over a 5^o longitude x 5^o latitude domain centering at the target grid point. This calculation is repeated with the center of the IR data array for the previous time step shifted in both the zonal and merional directions by various distances. Displacement resulting in the maximum pattern correlation is recorded and used to compute the cloud motion vectors.

To ensure reasonable computational efficiency, the cross-correlation calculation is executed in three steps. First, the pattern correlation is calculated for each grid point and for spatial displacement of every four grid boxes in both zonal and merional directions using IR data in every four grid boxes over the entire data domain of 5^o latitude / longitude. The rough estimation of the cloud motion vector is then refined through repeating the calculation for spatial displacement of every two grid boxes using IR data at one in every two grid boxes over a spatial domain of 2.5^{o}latitude / longitude. The final solution is achieved through once again repeat the pattern correlation calculation for spatial displacement of one grid box using data at all grid boxed over a spatial domain of 2.5^{o}latitude / longitude.

b) PDF matching for the inter-calibration of PMW Level 2 retrievals

As described in section 3.2, inter-calibration is performed for the Level 2 retrievals of instantaneous precipitation rates derived from PMW measurements of different sensors / platforms. This is done through matching the probability density function (PDF) of the precipitation rates derived from a PMW sensor against that from the TRMM Microwave Imagery (TMI). The TMI is flying on a precessing orbit, providing samplings overlapping with PMW sensors from all other LEO platforms. Retrievals from the TMI shows the best quality compared to those from other sensors.

PDF matching against the TMI retrievals is performed for retrievals from each of the PMW sensors, respectively. First of all, tempo-spatially collocated data pairs of the TMI and the target PMW retrievals are collected. A target PMW retrieval at an 8kmx8km grid box is considered collocated with an TMI retrieval at the same grid box in the same 30-min time slot or one time step before or after. PDF tables (histograms) are then constructed in 0.2 mm/hour interval for the retrievals from the target PMW sensor and the TMI. The PDF table for the target PMW sensor is adjusted, from the heavier precipitation end toward the lighter precipitation end, to match that for the TMI. Assume the total number of counts in class (j) of the PDF for the raw PMW sensor N_raw(j) is re-distributed into all m classes in the adjusted PDF tables:

$$N_{raw(j)} = \sum_{i=1}^{m} n\_adj(i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Where n_adj(i) is the number of counts distributed into class (i) in the adjusted PDF. The calibrated precipitation rate for class (j), PR_cal(j) is computed as the weighted mean of the precipitation rates associated with the adjusted PDF tables:

$$PR\_CAL(j) = \sum_{i=1}^{m} n\_adj(i) PR\_tmi(i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Due to the loss of TRMM/TMI measurements in April 2015, the PDF tables used to perform inter-calibration of PMW Level 2 precipitation rate retrievals as described above can no longer be updated. To continue the inter-calibration for PMW retrievals after April 2015, a set of PDF tables are established for PMW retrievals from each sensor using historical data. The PDF tables are constructed for each calendar month to account for seasonal variations, for every 10^{o}latitude band to reflect regional differences, and for land and ocean separately. These PMW inter-calibration tables based on historical data are included as part of the software package as described in Section 3.3.2c).

c) PDF matching for the bias correction

Bias correction is performed for the raw CMORPH satellite precipitation estimates. First, raw CMORPH satellite precipitation estimates are accumulated from their native resolution of 8kmx8km / 30-min to 0.25^{o}lat/lon / daily. Here, the end of a day is set as the same as that for the CPC daily gauge analysis which presents variations from country to country and is documented in Section 3.4.5b) (the Look-Up Tables). PDF tables are then constructed for each grid box of 0.25^{o}lat/lon over the land for each pentad using daily raw CMORPH and CPC gauge data over a 15-year period from 1998 to 2012.

In establishing PDF tables for each grid box of 0.25^{o}lat/lon and for each calendar date, raw CMORPH and CPC daily gauge data are collected over a 31-day period centering at the target date over the entire 15-year period and over a circular spatial domain of 2 grid boxes (0.5^{o}lat/lon) radius centering at the target grid box. The spatial domain is expanded to achieve at least 500 co-located data pairs with at least 300 cases of raw CMORPH reporting precipitation.

The collected raw CMORPH and the CPC gauge data are then ranked separately in descending order and grouped into 100 equal numbered classes. The adjustment factor is calculated for each class as the ratio between the mean value of CPC gauge data and that of the raw CMORPH averaged over 5 consecutive classes centering at the target class. The PDF bias correction table is therefore a 100-class

table composed of two columns: the 5-class mean raw CMORPH value and its associated adjustment factor.

3.4.5 Look-Up Table Description

Four sets of look-up tables are used in the implementation of the bias-corrected CMORPH CDR processing system, i.e. a) land-ocean masks, b) a global array of end of a day (EOD) definition for daily gauge data, c) PDF tables constructed using historical raw CMORPH and daily gauge analysis data, and d) Under-estimation Correction Factor tables.

a) Land-ocean masks

Many of the processing procedures inside the bias-corrected CMORPH CDR system apply different algorithms for different underlying surface types. Land-ocean masks are therefore needed as look-up tables fed into the system. In the current version of the bias-corrected CMORPH CDR system, two sets of land-ocean masks are included, at a spatial resolution of 8kmx8km and 0.25^{o}lat/lon, respectively:

- *Land-ocean mask at 8kmx8km resolution*
  - *CMORPH_V1.0_LAND_MASK_8km.lnx binary data file*
  - *CMORPH_V1.0_LAND_MASK_8km.lnx.ctl GrADS control file describing the data format*
- *Land-ocean mask at 0.25^{o}lat/lon resolution*
  - *CMORPH_V1.0_LAND_MASK_0.25deg.lnx binary data file*
  - *CMORPH_V1.0_LAND_MASK_0.25deg.lnx.ctl GrADS control file describing the data format*

Land-ocean masks are therefore needed Inside these mask data files, a grid box of land and ocean are designated as 1.0 and 0.0, respectively.

b) The Global EOD data file

Due to the different observation practices of the World Meteorological Organization (WMO) member countries, the beginning and ending times of a day vary from country to country. Reports of daily precipitation submitted by different countries therefore represent 24-hourly accumulations ending at different times. The definition of a day is here defined by the ending hour of the 24-hourly period, called End of Day (EOD). A gridded field of EOD time is constructed on a 0.25^{o}lat/lon for use in constructing daily precipitation estimates for the CMORPH satellite estimates.

- *CMORPH_V1.0_EOD_0.25deg_checked.lnx binary data file*
- *CMORPH_V1.0_EOD_0.25deg_checked..lnx.ctl GrADS control file describing the data format*

EOD is recoded as the World Standard Time hour, with 6.0 and 30.0 indicating a day ending at 06Z of the stamped date and the day next to the stamped date, respectively. Figure 8 display the global distribution of EOD.

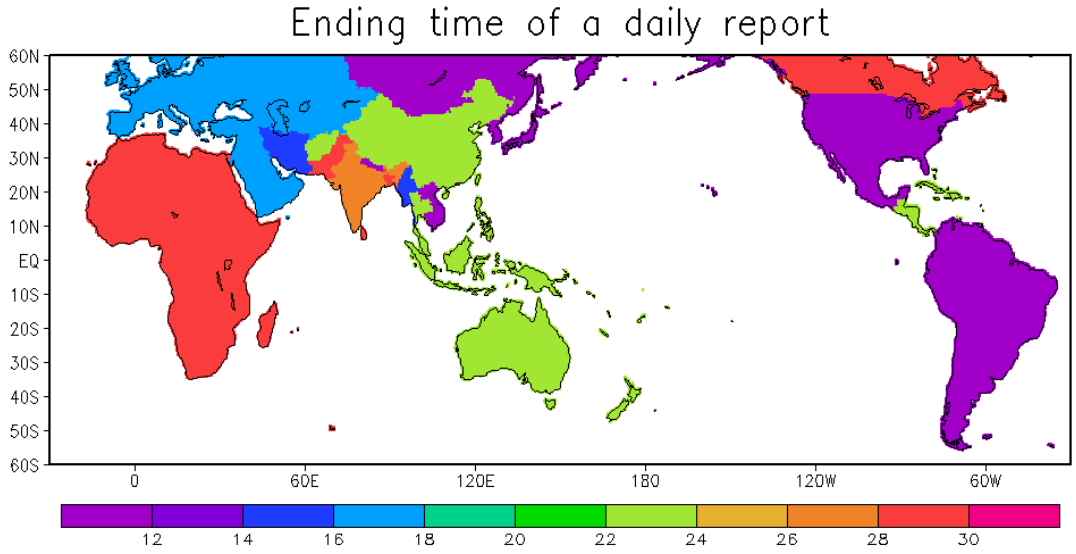


Figure 8: Global distribution of end of day (EOD) hour for the daily precipitation reports.

c) PDF tables constructed using historical data

As described in Section 3.2, the first step of the bias correction for the raw CMORPH is performed through a PDF table established using historical raw CMORPH and CPC daily gauge analysis data for a 15-year period from 1998 to 2012. To create the historical PDF tables, raw CMORPH satellite precipitation estimates are accumulated from their native resolution of 8kmx8km / 30-min to 0.25°lat/lon / daily. Here, the end of a day is set as the same as that for the CPC daily gauge analysis which present variations from country to country and is documented in the Look-Up Table b). PDF tables are then constructed for each grid box of 0.25°lat/lon over the land for each of the 365 calendar day using daily raw CMORPH and CPC gauge data over a 15-year period from 1998 to 2012.

In establishing PDF tables for each grid box of 0.25°lat/lon and for each calendar day, collocated raw CMORPH and CPC daily gauge data are collected over a 31-day period centering at the target calendar day over the entire 15-year period and over a circular spatial domain of 2 grid boxes (0.5°lat/lon) radius centering at the target grid box. The spatial domain is expanded to achieve at least 500 co-located data pairs with at least 300 cases of raw CMORPH reporting rain.

The collected raw CMORPH and the CPC gauge data are then ranked separately in descending order and grouped into 100 equal numbered classes. The adjustment factor is calculated for each class as the ratio between the mean value of CPC gauge data and that of the raw CMORPH averaged over 5 consecutive classes centering at the target class. The

historical PDF bias correction table is therefore a 100-class table composed of two columns: the 5-class mean raw CMORPH value and its associated adjustment factor.

The historical PDF tables for each calendar day is stored in a file, named:

CMORPH_BIAS_CRT_HS_V1.0_CPDF.lnx.xxx

where xxx indicates the Julian day for which the PDF tables are established.

Only the PDF tables of land grid boxes are included in the file, starting from the 0.25^{o}lat/lon grid box centering at [0.125^{o}E; 59.875^{o}S], ranging from west to the south, then from south to north. In total, there are xxxxx land grid boxes of 0.25^{o}lat/lon over the globe from 60^{o}S to 60^{o}N.

For each grid box of 0.25^{o}lat/lon, the PDF tables are output in the form of two real number arrays of 100 elements: the raw CMORPH precipitation rate and the adjustment factor associated with the precipitation rate.

d) Under-estimation Correction Factor for the Historical PDF Tables

The PDF matching technique for inter-calibrating precipitation rates is unable to make any correction when the raw CMORPH estimates showing zero precipitation. This would cause overall under-estimation of bias corrected CMORPH. To mitigate the effect, ratio between the CPC gauge analysis and the bias-corrected CMORPH is computed for each land grid box of 0.25^{o}lat/lon grid and for each calendar date using the data for the 15-year period from 1998 to 2012. This global field of ratio is applied to the historically corrected CMORPH estimates. The correction tables are stored in binary files named:

CMORPH_BIAS_CRT_HS_V1.0_STAT.lnx.xxx

where xxx indicates the Julian day for which the PDF tables are established.

3.4.6 Parameterization

Not applicable.

3.4.7 Algorithm Output

The bias-corrected CMORPH CDR processing system generates bias-corrected, integrated satellite precipitation estimates over the global domain for an extended period from January 1, 1998 to the present. The output precipitation fields are produced in a combination of three different time-space resolutions to accommodate user requirements of various backgrounds.

- *8kmx8km – 30-minute resolution*
    - *File name convention*

      CMORPH_V1.0_ADJ_8km-30min_yyyymmddhh

(each hourly files contain two global fields of 30-min precipitation rates in mm/hr)

- *GrADS control file describing the data format*

  CMORPH_V1.0_ADJ_8km-30min.ctl

- *0.25°lat/lon – hourly resolution*
  - *File name convention*

    CMORPH_V1.0_ADJ_0.25deg-HLY_yyyymmddhh
  - *GrADS control file describing the data format*

    CMORPH_V1.0_ADJ_0.25deg-HLY.ctl

- *0.25°lat/lon – daily resolution*
  - *File name convention*

    CMORPH_V1.0_ADJ_0.25deg-DLY_EOD_yyyymmdd
  - *GrADS control file describing the data format*

    CMORPH_V1.0_ADJ_0.25deg-DLY_EOD.ctl

Most of the inputs used in creating the CMPRPH CDR, including the Level 2 PMW retrievals, daily snow maps and the daily gauge analyses are available within 18 hours of the observations. While the CMORPH CDR is constructed as a post-process, a quasi real-time version should be able to be generated at a latency of ~18 hours.

4. Test Datasets and Outputs

4.1 Test Input Datasets

The Version 1 data set of the bias-corrected CMORPH is created for the entire data period from 1998 to the present. The input Level 2 PMW retrievals of instantaneous precipitation rates are those generated using the GPROF Version 2004. Level 2 data coverages for the PMW sensors aboard various platforms are shown in figure 9.

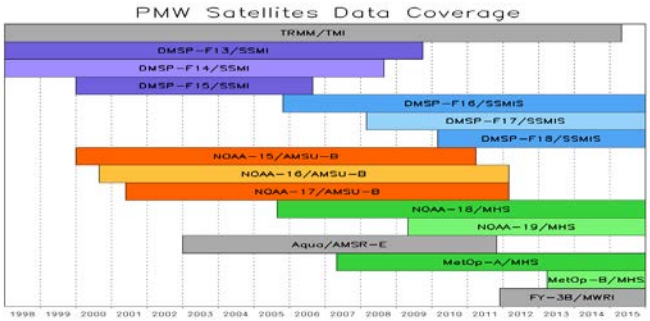


Figure 9: Level 2 retrieval data availability for various passive microwave sensors aboard various low earth orbit platforms.

4.2 Test Output Analysis

4.2.1 Reproducibility

Raw and bias-corrected CMORPH integrated satellite global precipitation estimates are generated for an 18-year period from January 1, 1998 to December 31, 2015. Figure 10 illustrate the December-January-February (DJF) and June-July-August (JJA) mean precipitation of the bias-corrected CMORPH averaged for an 18-year period from 1998 to 2015, while figure 11 shows the magnitude (shading) and phase (arrow) of the precipitation diurnal cycle derived from the bias-corrected CMORPH for JJA of 1998-2015 over the globe (top) and North America (bottom), respectively.

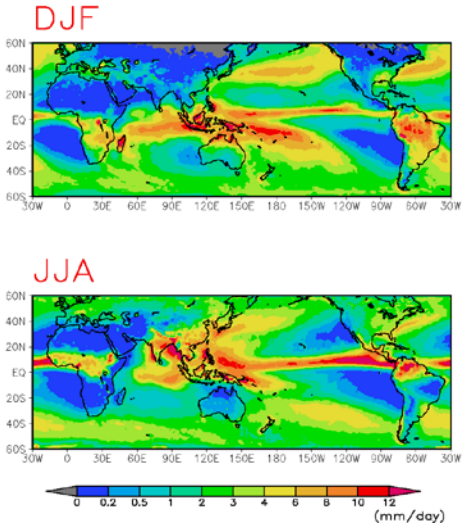


Figure 10: December-January-February (DJF) and June-July-August (JJA) mean precipitation (mm/day) derived from bias-corrected CMORPH for an 18-year period from 1998 to 2015.

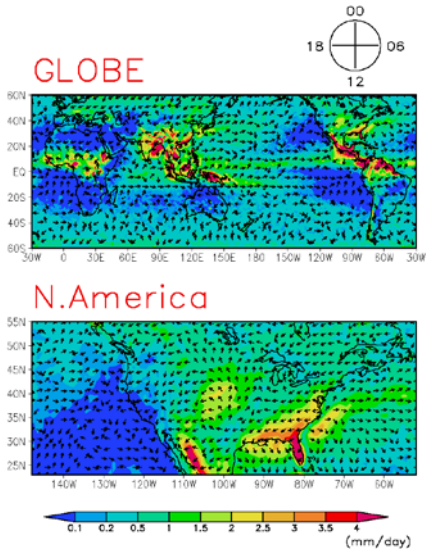


Figure 11: Magnitude (mm/day, shaing) and phase (timing of maximum, arrow) of JJA precipitation derived from JJA mean hourly precipitation for 1998-2015 over the globe (top) and North America (bottom). Local time of maximum precipitation is shown by the direction of the arrows as indicated by the circle on the top-right corner of the figure.

4.2.2 Precision and Accuracy

Performance of the bias-corrected CMORPH is examined through comparison against *in situ* observations of precipitation. Figure 12 shows the 1998-2015 18-year mean land precipitation derived from the CPC gauge analysis, the raw and bias-corrected CMORPH. While the raw CMORPH presents similar spatial patterns of annual mean precipitation, it tends to over-/under-estimates the magnitude of precipitation over the extra-tropics / tropics. The bias correction succeeded in reducing the bias in the raw CMORPH, especially over the tropics, while annual precipitation in the bias-corrected CMORPH still exhibits under-estimation due to the poor performance of the input PMW Level 2 precipitation retrievals in detecting and quantifying cold season precipitation (Xie and Joyce et al. 2013).

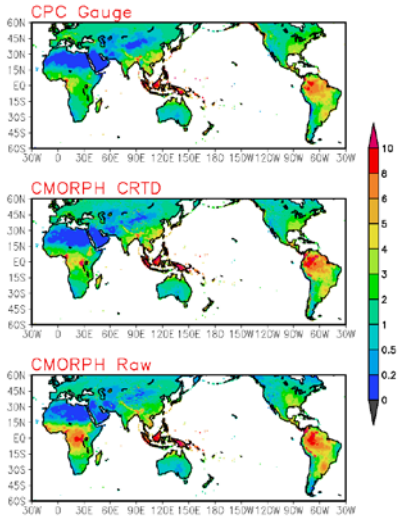


Figure 12: 1998-2015 18-year mean precipitation (mm/day) derived from (top) CPC unified gauge analysis, (middle) bias-corrected CMORPH, and (bottom) raw CMORPH.

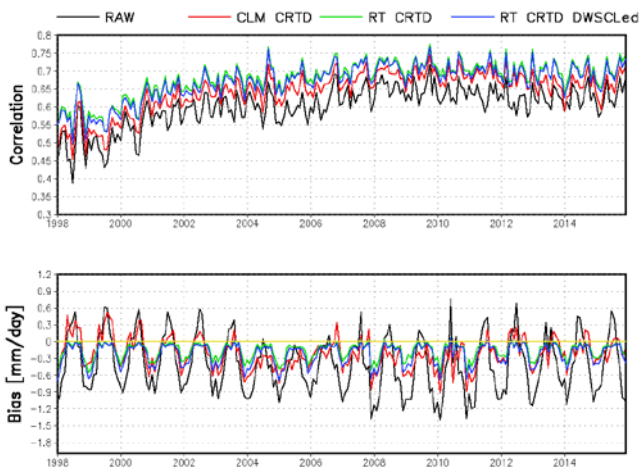


Figure 13: Time series of the correlation (top) and the bias (bottom) between the CMORPH satellite estimates and the CPC gauge analysis of daily precipitation over the global land. The comparison statistics is computed using data over the 0.25olat/lon boxes where / when at least one reporting gauge is available. Results for the raw, climatologically corrected (CLM_CRTD), real-time refined (RT_CRTD), and real-time refined / down-scaled (RT_CRTD DWSCLed) CMORPH are plotted in black, red, green, and blue lines, respectively.

Performance of the CMORPH satellite estimates improves with time as number of the LEO platforms with PMW sensors increases, especially during the earlier half of the data period when the number of PMW sensors increases from 3 in 1998 to 7 in 2004 (figure 13).

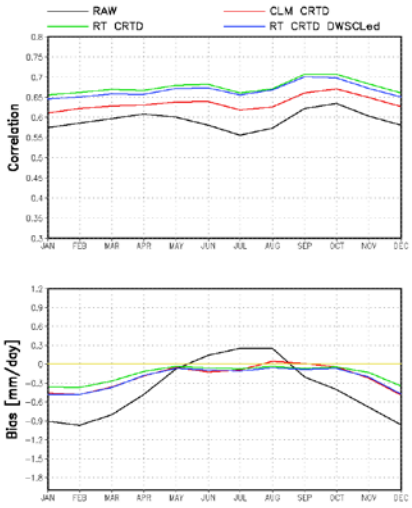


Figure 14: Correlation (top) and bias (bottom) of the CMORPH satellite precipitation estimates as a function of season. Statistics is computed through comparison against the CPC unified daily gauge analysis over 0.25^{o}lat/lon grid boxes for an 18-year period from 1998 to 2015. Only data over a 0.25^{o}lat/lon grid box where / when at least one reporting gauge is available are included in the calculation. Results for the raw, climatologically corrected (CLM_CRTD), real-time refined (RT_CRTD), and real-time refined / down-scaled (RT_CRTD DWSCLed) CMORPH are plotted in black, red, green, and blue lines, respectively

The PDF matching technique described in section 3.2 succeeded in removing the bias in the raw CMORPH over the land areas during warm season (figure 14). Appling real-time refinement (blue lines) to the climatological correction (green lines) apparently improved the performance of the bias corrected CMORPH. Over the cold seasons, bias remains, although its magnitude is reduced substantially. The imperfect performance of the bias correction procedure is caused by the limited capability of the input PMW Level 2 precipitation retrievals for cold season precipitation. The PDF technique is unable to raise a zero value retrievals to a non-zero estimation.

4.2.3 Error Budget

Not applicable to this version.

5. Practical Considerations

5.1 Numerical Computation Considerations

The bias corrected CMORPH CDR processing is conducted step by step as indicated in the flow charts. Computation is quite straightforward with improved efficiency with a machine of faster CPU and larger memory.

5.2 Programming and Procedural Considerations

The system is running on a Linux based system. All of the programs are written in Fortran 77, Fortran 90, and C language while scripts in C shell.

5.3 Quality Assessment and Diagnostics

Quality of the bias-corrected CMORPH CDR may be assessed through comparison against ground truth including gauge measurements / analyses and radar observations and through inter-comparison with other products. Anomalous values may be checked out through these comparisons and through manual monitoring of global maps.

5.4 Exception Handling

A key step for the bias corrected CMORPH CDR is to propagate the PMW retrievals of instantaneous precipitation rates with the cloud motion vectors derived from the consecutive GEO IR images. The GEO IR images, however, are not available over certain regions and for a very small fraction of time periods. Cloud motion vectors are computed from monthly files of hourly, 0.5 degree, 700-mb winds produced by the National Center for Environmental Prediction (NCEP) reanalysis (Kalnay et al. 1996).

5.5 Algorithm Validation

Not Applicable to this version.

5.6 Processing Environment and Resources

- Work station, OS, and compliers

Currently the CMORPH CDR is executed on a Work Station in Red Hat Enterprise Linux 5 (RHEL5).

The input Level 2 PMW retrieval (the orbit data) files are in Hierarchical Data Format (HDF). Most of the preprocessing programs to decode the HDF files are written in C language, except those for the files from the SSMIS that are coded in Fortran. Library packages needed included hdf-4.2.6, hdf5-1.8.8, HDF-EOS2.18v1, netcdf-4.1.3, zlib-1.2.5, and jpeg-6b.

The C- and Fortran- programs are compiled by the 'cc' and the 'gfortran' command under RHEL5.

- Hard Disk Usage

The Level 2 retrievals from individual PMW sensors (the orbit data) are first decoded from the HDF files and output into ASCII data files. The extracted precipitation rate retrievals are then mapped, calibrated, and outputted into binary data files of gridded fields.

The Orbit data and raw data files are compressed after being used to generate gridded data files. The gridded data files, however, need to be kept uncompressed until the MWCOMB files are created.

Typical volumes for data files involved in various steps of defining raw CMORPH for a monthly period are listed below in table 3.

Table 3: Typical Data Volumes in Processing CMORPH for a Monthly Period

Satellite / PMW Sensors	Orbit Data(GB) (compressed)	ASCII Data(GB) (compressed)	Gridded Data(GB) (uncompressed)
TMI	9	1	12
AMSR	4.5	2	24
NOAA (each satellite)	2	0.4	48
NOAA (combined)	------	------	48
DMSP (each satellite)	9	2	48
DMSP (combined)	-----	-----	48
MetOp (each satellite)	1.5	0.4	48
MWCOMB	-----	-----	48
Merged IR			48
Morphing Results			72
Final Raw CMORPH			36

6. Assumptions and Limitations

6.1 Algorithm Performance

The bias corrected CMORPH CDR algorithm is developed under the following assumptions:

a) The L2 PMW retrievals are capable of detecting and quantifying instantaneous precipitation rates over the entire spectrum of precipitation intensity, for all the seasons and over the entire target domain of 60°S to 60°N;

b) The motion vectors derived from the consecutive GEO IR images through the cross-correlation technique represent the movement of precipitating cloud systems;

c) Precipitating cloud systems persist throughout the time period of propagation with no changes in intensity;

d) CPC gauge based analysis accurately represents the overall magnitude; and

e) Bias structure for the raw CMORPH as defined through comparison against CPC daily gauge analysis and pentad GPCP is the same for the raw CMORPH at 30-min and 8kmx8km time scales.

Limitations of the bias-corrected CMORPH include:

a) Incomplete global coverage (60°S-60°N);

b) Less than desirable performance for snowfall and cold season rainfall;

c) Under-estimation of heavy rainfall events, and

d) Inaccurate magnitude where / when the CPC gauge analysis is not accurate.

6.2 Sensor Performance

Performance of the bias-corrected CMORPH relies heavily on the quality of the input data, including the Level 2 PMW precipitation retrievals, the gauge-based analysis of daily precipitation, and the GPCP pentad merged analysis.

a) Currently operational PMW retrievals used as inputs to the bias-corrected CMORPH CDR are unable of generating precipitation estimates over snow/ice covered surface, tend to under-estimate heavy rainfall, miss light rainfall, and present poor performance for cold season rainfall and snowfall (Ebert et al. 2007, Xie and Joyce 2014);

b) The CPC gauge-based analysis is derived from reports of gauge-measured daily precipitation which tend to show negative bias compared to monthly reports over certain regions; and

c) The GPCP pentad merged analysis, defined through calibration against the monthly GPCP analysis, contains uncertainty in its magnitude over global oceans.

7. Future Enhancements

The CMORPH CDR described in this document is created using the currently operational algorithm. It covers a quasi global domain from 60°S to 60°N and performs quite well in depicting precipitation during warm seasons and over tropical, sub-tropical and mid-latitudes. A newer version of the algorithm, called the second generation CMORPH, is developed and is being tested at NOAA Climate Prediction Center (CPC) to generate integrated satellite precipitation estimates at a refined spatial resolution (5kmx5km) covering the entire globe from the South Pole to the North Pole with improved quality especially in depicting the snowfall (Joyce and Xie 2011, Xie and Joyce 2014).

7.1 Enhancement 1

The current version CMORPH covers the globe from 60°S to 60°N, largely limited by the use of GEO IR data to derive the cloud motion vectors. In the second generation CMORPH, the cloud motion vectors are first computed from individual sources including a) GEO IR based precipitation estimates, b) CFS reanalysis precipitation simulations, and c) LEO PMW and IR based precipitation fields. Vectors defined from these individual sources are then combined through a 2DVAR technique to form global fields of precipitating cloud motion vectors with optimized accuracy (Xie et al. 2015). The integrated CMORPH satellite precipitation estimates are then constructed through morphing the PMW and IR based precipitation retrievals under a Kalman Filter (KF) framework (Joyce and Xie 2011).

7.2 Enhancement 2

Representation of snowfall rate in the current version CMORPH is less than desirable, due to the poor capability of the input PMW retrievals from the currently operational products (Xie and Joyce 2014). Recently, research groups with NESDIS and NASA have developed test products with much improved quality in capturing falling snow during cold seasons (Meng et al. 2011). We have successfully infused their snowfall test products into our second generation CMORPH system and demonstrated improved performance of the CMORPH.

8. References

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Neikin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979 – Present*). J. Hyrometeo.*, **4**, 1147 – 1167.

Hsu, K.-L., 2014: Climate Algorithm Theoretical Basis Document (C-ATBD) for the PERSIANN Precipitation CDR. Available at : http://gis.ncdc.noaa.gov/all-records/catalog/search/resource/details.page?id=gov.noaa.ncdc:C00854

Joyce, R.J., J.E.Janowiak, P.A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487 – 503.

Joyce, R.J., and P. Xie, 2011: Kalman Filter Based CMORPH, *J. Hydrometeor.*, 12, 1547 – 1563, DOI: 10.1175/JHM-D-11-022.1.

Tang, L., Y. Tian, and X. Lin (2014), Validation of precipitation retrievals over land from satellite-based passive microwave sensors, *J. Geophys. Res. Atmos*., **119**, 4546–4567, doi:10.1002/ 2013JD020933.

Xie, P., M. Chen, and W. Shi, 2010: CPC Unified Gauge-Based Analysis of Global Daily Precipitation. Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc.,

Xie, P., J.E. Janowiak, P.A. Arkin, R. Adler, A. Gruber, R. Ferraro, G. J. Huffman, and S. Curtis, 2003: GPCP Pentad Precipitation Analyses: An Experimental Dataset Based on Gauge Observations and Satellite Estimates. J. Climate, 16, 2197 – 2214.

Appendix A. Acronyms and Abbreviations

AMSU	Advanced Microwave Sounding Unit
AMSR	Advanced Microwave Scanning Radiometer
ATMS	Advanced Technology Microwave Sounder
CDR	Climate Data Record
CMORPH	CPC Morphing Technique
CPC	Climate Prediction
DMSP	Defense Meteorological Satellite Program
EOD	End of Day
FY	Fun Yun (Name of Chinese Meteorological Satellites)
GEO	Geostationary
GMS	Geostationary Meteorological Satellite
GOES	Geostationary Operational Environmental Satellite
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
IMS	Interactive Multisensor Snow and Ice Mapping System
IR	Infrared
LEO	Low Earth Orbit
METEOSAT	Meteorological Satellites of the European Space Agency
MHS	Microwave Humidity Sounder
MWCOMB	MicroWave based precipitation rate retrievals COMBined
MWRI	MicroWave Radiation Imager
NCEI	National Centers for Environmental Information
NESDIS	National Environmental Data and Information Service
NOAA	National Oceanic and Atmospheric Administration
PDF	Probability Density Function
PERSIANN	Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks
PMW	Passive Microwave
SSM/I	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager / Sounder
STAR	NESDIS Center for Satellite Applications and Research
TMI	TRMM Microwave Imager
TRMM	Tropical Rainfall Measurement Mission