Performance Progress Report - NA11NES4400003 (31st December 2013)

The Global Precipitation Climatology Project (GPCP) Data Products—Transfer to Operations at NCDC

PRINCIPAL INVESTIGATOR: Robert F. Adler, Earth System Sciences Interdisciplinary Center ESSIC, University of Maryland—College Park (radler@umd.edu)
Co-Principal Investigator: George Huffman (SSAI/NASA Goddard)
Co-I: Mathew Sapiano (ESSIC/UMD)
Co-I: Ralph Ferraro (NOAA/NESDIS)
Co-I: Ping-Ping Xie (NOAA/NCEP/CPC)
Co-I: Long Chiu (George Mason University)
Co-I: Udo Schneider (Deutscher Wetterdienst [DWD])

1. Summary of Project

The objective of this work is to successfully transfer the routine production of Global Precipitation Climatology Project (GPCP) products to NCDC from the ad-hoc processing of the Co-Is listed above. The current monthly (1979-present), pentad (1979-present) and daily products (1997-present) have been developed by research groups over the last 15 years and are produced by a consortium of those groups, funded by various agencies. Transfer of the routine processing of the GPCP products to an operational entity will ensure continuation of these important data analysis sets.

This activity involves the development of a detailed strategy for transfer of scientific knowledge, satellite and other data source accesses, and processing code for successful implementation of an end-to-end processing system that would routinely produce the GPCP current (Version 2) products for archival and dissemination.

The current GPCP processing involves computation of individual intermediate products or data sets by a number of government and university entities and a merger of these products by another group. These codes are being organized, streamlined, updated and documented for product production from the level of the satellite-calibrated radiances to the final merged products. Testing and evaluation of output products from the new system is a key part of this project, as is an evaluation of the feasibility of reprocessing the 30-plus-year-record to make minor upgrades. Maintaining science quality of the output products will be a critical aspect and will require significant on-going effort.

2. Recent Progress

The code to produce the monthly 2.5° GPCP precipitation CDR has been finished and is currently in beta mode. This code has been used to produce GPCP data for the past couple of months and this has been compared to the official version of GPCP that is still being produced at NASA Goddard under the direction of George Huffman. The two datasets are in perfect agreement. We also note that routine production of GPCP with AIRS V6 precipitation from March 2013 has been resumed. In addition to the CDR, we have also developed an ICDR of GPCP precipitation that is available within 10 days of the end of each month, as shown in Figure 1 for December 2013. The GPCP CDR uses

the Global Precipitation Climatology Center (GPCC) Full Analysis/Monitoring product, the highest quality gauge analysis available from GPCC, which is which is produced with a two month lag so that all possible/reasonable gauge data can be collected. GPCC also produces a near real time First Guess product that is in excellent agreement with the Full Analysis/Monitoring products and this gauge data is used in the ICDR. The ICDR uses mostly the same input data as the CDR and uses many of the same routine. The ICDR has therefore been incorporated into the CDR processing code, with separate run codes and separate output and interim products directories. In recent weeks, we have been successful in securing agreement from data providers (with the help of NCDC) that the latency of the input data be reduced, so that the overall latency of the ICDR is currently within the target of 10 days. The December 2013 dataset was actually ready before this deadline on January 7th 2014.

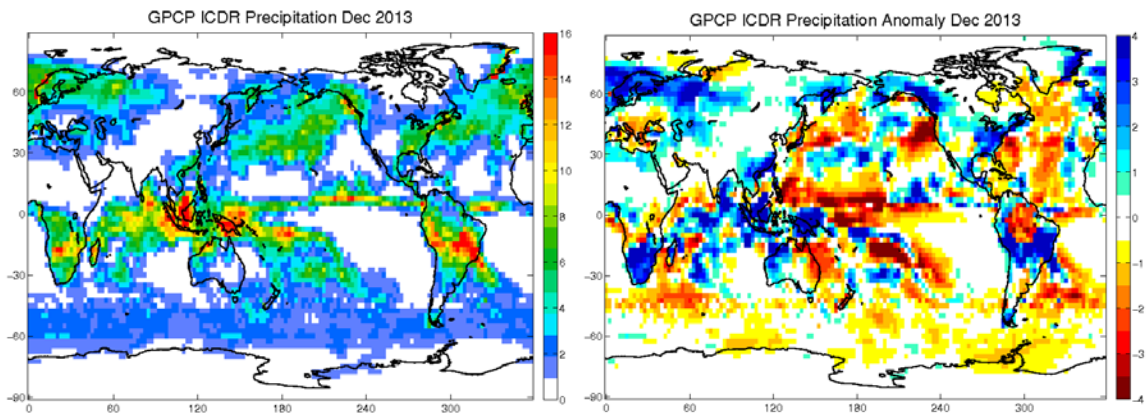

Figure 1. GPC ICDR precipitation and precipitation anomaly for December 2013.

2.1 Documentation

Our goal in this project is to provide a well-documented CDR package that can be run at NCDC. A key component of this documentation is the CATBD that is required by NCDC. A first draft of this document has been produced containing the necessary information for the 2.5 degree product in the required format. Similar, but separate documents will be produced for the One Degree Daily product and the Pentad product. Compared to most datasets, the science of the GPCP is well documented, with numerous scientific papers explaining the approach and the decisions made in the construction of the dataset. We have therefore spent some time producing User's guides that document how to run the code and that include information on how to debug the code etc. To that end, we have produced User's Guides for the CDR and the ICDR that outline how to install the code, how to obtain the data, how to run the code and how to run the diagnostics. We hope that these documents will aid in the transition to NCDC and we expect them to be a useful tool for future archive purposes.

2.2 Diagnostics

Diagnostic routines are an important tool for running the GPCPC CDR because of the complexity of the code and the large number of inputs required to run the code. In the past, such diagnostics were done manually with the dataset creator checking files existed and then plotting the output files. For the CDR, it is important that these diagnostics can

be standardized and easily run. New diagnostic routines have been written for the CDR code that check the inputs exist before running and plot the outputs of the CDR and the ICDR. The input diagnostics check each of the inputs exist by counting numbers of files and file sizes for the several inputs. The Ferraro data are supplied in annual files, so the diagnostic code reads the file and sums the number of observations for the month (a zero sum denotes missing data). Several Matlab routines have also been written to produce maps of the output and Interim fields (mean and anomaly) as well as time series of each of the Interim and output fields for both the CDR and the ICDR. A simple web interface has been established to allow for easy viewing of these diagnostics:
http://eagle1.umd.edu/GPCP_CDR_Diags/CDR_Diagnostics_Dec2013/index.htm

2.3 One Degree Daily (1DD) data

The next phase in the production of the GPCP suite is the modification of the code to produce the GPCP 1DD CDR. The 1DD code uses many elements from the 2.5 degree code, so it is expected that there will be some overlap between the two, with eth 2.5 degree code run first. The first part of the transition from NASA Goddard is the transfer of the GPROF 2004 algorithm, and this is the most complex part. The GPROF algorithm currently ingests the RSS binary brightness temperatures, whereas the CDR version will use the NCDC RSS NetCDF brightness temperatures. In addition, the algorithm code contains many unused input/output and processing options that are being removed as well as many lines of legacy code that must be removed. Progress is being made on all of these issues.

2.4 Plans for coming 6 months

There are several goals for the next 6 months. First, we plan to continue running the beta test of the CDR against the official NASA Goddard data product and to continue producing the ICDR as soon as the required data becomes available. The routine production of the ICDR allows for us to make other ICDR-based products available in near real time and we intend to start developing these products alongside the ICDR.

We will also continue work on the implementation of the 1DD CDR code and hope to have a beta version ready in the coming months.

Performance Progress Report - NA11NES4400003 (30th June 2014)

The Global Precipitation Climatology Project (GPCP) Data Products—Transfer to Operations at NCDC

PRINCIPAL INVESTIGATOR: Robert F. Adler, Earth System Sciences Interdisciplinary Center ESSIC, University of Maryland—College Park (radler@umd.edu)
Co-Principal Investigator: George Huffman (NASA Goddard)
Co-I: Mathew Sapiano (ESSIC/UMD)
Co-I: Ralph Ferraro (NOAA/NESDIS)
Co-I: Ping-Ping Xie (NOAA/NCEP/CPC)
Co-I: Long Chiu (George Mason University)
Co-I: Udo Schneider (Deutscher Wetterdienst [DWD])

1. Summary of Project

The objective of this work is to successfully transfer the routine production of Global Precipitation Climatology Project (GPCP) products to NCDC from the ad-hoc processing of the Co-Is listed above. The current monthly (1979-present), pentad (1979-present) and daily products (1997-present) have been developed by research groups over the last 15 years and are produced by a consortium of those groups, funded by various agencies. Transfer of the routine processing of the GPCP products to an operational entity will ensure continuation of these important data analysis sets.

This activity involves the development of a detailed strategy for transfer of scientific knowledge, satellite and other data source accesses, and processing code for successful implementation of an end-to-end processing system that would routinely produce the GPCP current (Version 2) products for archival and dissemination. Initial data processing with the new system will take place at University of Maryland CICS.

The current GPCP processing involves computation of individual intermediate products or data sets by a number of government and university entities and a merger of these products by another group. These codes are being organized, streamlined, updated and documented for product production from the level of the satellite-calibrated radiances to the final merged products. Testing and evaluation of output products from the new system is a key part of this project, as is an evaluation of the feasibility of reprocessing the 30-plus-year-record to make minor upgrades. Maintaining science quality of the output products will be a critical aspect and will require significant on-going effort.

2. Recent Progress

Testing of the monthly CDR code has been going well, and the diagnostics for this product have been further enhanced with the development of a new webpage: http://eagle1.umd.edu/GPCP_ICDR/index.htm

The CDR code has been ported to a linux environment and works with minimal need for extra libraries (several external libraries were required in the original code that complicated compilation and compatibility). There remains some work to do to reach full compliance with NCDC's coding standards, such as the removal of goto loops and

other such programming constructs. We have started to do this type of streamlining of the code and progress is good.

2.1 Progress with the ICDR

The GPCP CDR uses the Global Precipitation Climatology Center (GPCC) Full Analysis/Monitoring product, the highest quality gauge analysis available from GPCC, which is which is produced with a two month lag so that all possible/reasonable gauge data can be collected. As reported in December 2013, we have developed an Interim CDR (ICDR) that makes use of the near real time GPCC First Guess product that allows us to produce data within a few days of the end of the month. The ICDR uses mostly the same input data and routines as the CDR and is therefore produced as part of the CDR code by a modified version of the monthly CDR driver code. The initial goal was to make the ICDR available within 10 days of the end of the month and this has been exceeded every month over the past 6-months thanks to timely delivery of the IR data by Pingping Xie at CPC and of the SSMIS data by RSS. A minor bug was found and fixed in the ICDR and it has been added to the new diagnostic page:
http://eagle1.umd.edu/GPCP_ICDR/index.htm

2.2 GMU METH issue

Ocean emission data in the CDR are produced by GMU using the METH algorithm. We have previously reported that METH data produced using the NCDC RSS Tb give slightly different results to those produced using the same algorithm applied to RSS binary Tb. It was found that the CDR differs by no more than 0.1 K over approximately half of the pixels, a difference that should be too negligible to produce any difference in the algorithm. We have been working with our collaborators at GMU to isolate the source of the discrepancy and continue to study the issue. Preliminary tests indicate that there may be some issues with the convergence at some pixels, so that a small perturbation leads the algorithm to give a different answer. We are working to properly quantify this issue and to apply a fix to the CDR

2.3 Potential shift starting in 2009 due to switch to SSMIS Tb

The GPCP monthly product uses microwave estimates from SSM/I from 1987 to the end of 2008, with Tb from the SSM/I aboard F13 used form 1995. In January 2009, Tb from the SSMIS mounted on the F17 satellite are used since the feed of data from F13 effectively ends in mid 2009. It was recently noticed that the might be a discontinuity in the microwave record created by the switch to SSMIS, with the primary shift being observed over ocean (Figure 1). The shift is around -.36 in GPCP, but is -0.4 in the METH data that is input to GPCP. We have carried out multiple comparisons with other data, including comparing with monthly TMI estimates. The TMI is of similar construction to the SSMI and data are derived from similar algorithms as the ones used for SSMI and SSMIS. The comparison with TMI appear to show a shift when SSMIS data is used, which might suggest that the shift is caused by non-physical mechanisms, although it is important to note that the shift is about half the apparent algorithm error. Work is continuing on this problem.

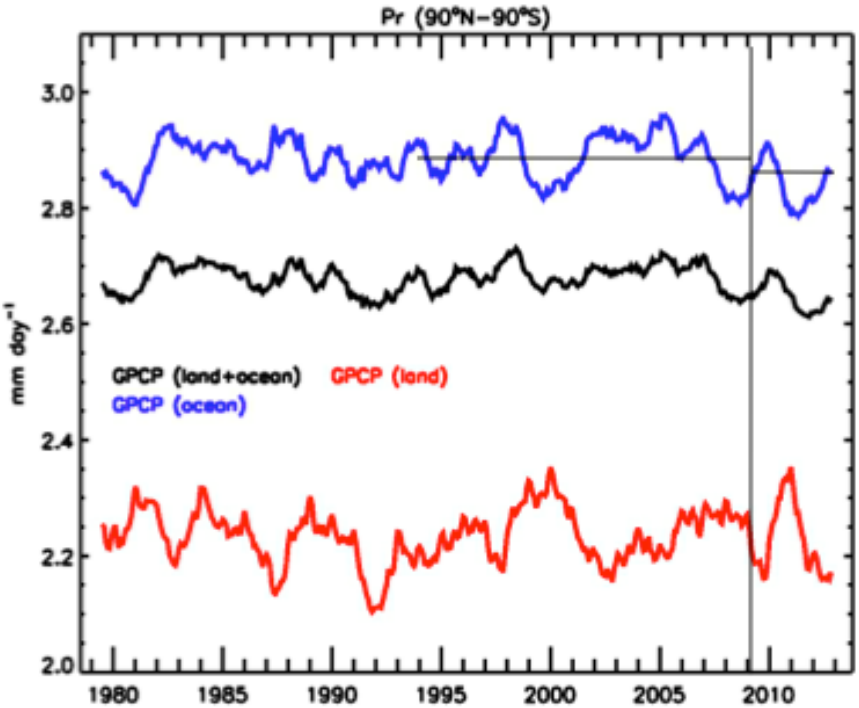

Figure 1. Monthly mean global GPCP precipitation time series for ocean only, land+ocean and land only from 1979 to 2013.

2.4 Plans for coming 6 months

The main task for the next 6 months is to continue improving and streamlining the monthly CDR to conform to NCDC coding standards. We will continue to produce this each month and compare to the official NASA Goddard data product. In addition, we continue to produce the ICDR for each month and make available to the community. When available, the ICDR will be compared with the CDR to assess the level of fidelity to the CDR product. We are also starting to work on the transition of the daily and pentad codes. We will work with our collaborator, Dr Pingping Xie, to handover the code to produce the pentad GPCP. The Daily product is a fairly complex piece of code and we are starting to look at this and understand what steps need to be taken to successfully transition this code to a CDR.