Performance Progress Report - NA11NES4400003 (30 December 2011)

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: 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 proposal 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 would ensure continuation of these important data analysis sets.

This proposed activity will develop 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. The proposed work will organize, streamline, update and document all code needed for product production from the level of the satellite-calibrated radiances to the final merged products. Arrangements will be made to ensure continued access to the various input data sets. Testing and evaluation of output products from the new system will be a key part of this project, as will 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 first few months of the progress were partially spent organizing the Co-I's, re-evaluating the status of the component software and organizing the coming activity. The software status is generally good with all components working routinely, but with substantial human involvement, especially with regard to transferring data and intermediate products, the products that are generated by individual PI's. In general, we are very ready to begin process of developing the future operational system to generate GPCP products.

In addition, a number of the investigators (Huffman, Chiu, Ferraro and Adler) and others spent time diagnosing and implementing the change from SSMI to SSMIS as the baseline microwave input to the GPCP monthly product. This is a data switch that had to be addressed in order to continue the routine production of GPCP products and for incorporation of the new data into the plan for transfer to NCDC. This shift of key data type is a good example of changes that affect the routine production. In this case this switch shut down the routine production for over a year, while we came up with a

solution, tested it (and re-tested it) and implemented the solution. For two years we worked to focus on the two microwave input data sources (for ocean and land) to ensure the most continuity possible given the current on-going issues with SSMIS. We chose F17 to be the calibrating satellite for the closest time match to previous SSMI's. We also obtained the RSS data for the emission algorithm for use by Co-I Chiu. Several versions of each SSMIS product were evaluated and feedback was provided to converge on a reasonable final algorithm. A similar effort is taking place for the land product and the daily product (1DD) using an 85-GHz proxy version of GPROF 2004 that was created by Daniel Vila that works with SSMIS data. This work has also entailed working with the SSMIS navigation to create co-located Tb's. With these new algorithms implemented to be applied to SSMIS data we carefully compared an overlap period of SSMI and SSMIS Tb's and calculated rainfall distributions. After another iteration to produce a minimum difference between the two outcomes, the new technique was implemented and the GPCP data set was brought up to date. The user community was grateful for the update.

Another key result for this project is the hiring of Dr. Matthew Sapiano at ESSIC/U. of Maryland. Matt has a strong background with a number of the input data sets, the process and the GPCP products. He brings a strong set of scientific and software skills to the project and will be making major contributions to it.

Performance Progress Report - NA11NES4400003 (30 June 2012)

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 proposal 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 would ensure continuation of these important data analysis sets.

This proposed activity will develop 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. The proposed work will organize, streamline, update and document all code needed for product production from the level of the satellite-calibrated radiances to the final merged products. Arrangements will be made to ensure continued access to the various input data sets. Testing and evaluation of output products from the new system will be a key part of this project, as will 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

Good progress has been made in the first year with all aspects of the project shifting into the development phase. Dr. Mathew Sapiano has now joined the GPCP CDR team at ESSIC/U. of Maryland and has been working for several months to implement the main part of the project. Dr. Sapiano will be responsible for much of the implementation of the project including the design, coding and streamlining of the GPCP system and will ensure that NCDC standards are met for transition to operational processing.

2.2 Version 2.2 availability

Some remaining issues with GPCP V2.2 have been resolved prior to the incorporation into the CDR code. The OPI and SSMI periods (collectively 1979-2008) have been finalized. The final METH SSMIS are in process, due in the next week, and should differ very modestly from the provisional data. That delivery should return GPCP SG to routine processing, about two months after observation, upon receipt of the GPCC gauge analysis.

2.3 Establishment of development strategy

The project is to create a documented operational code package that can be run at NCDC with minimal human intervention. The current V2.2 processing codes will be streamlined and modified to comply with NCDC standards and incorporated into one of three code packages that will produce the monthly, pentad and daily GPCP V2.2 products. The code will be a mixture of Fortran and C code run by the Executive Control Package (ECP), which will be a UNIX shell script that calls the various elements of the code. The basic structure of each of the three ECPs is depicted in Figure 1. Data acquisition will be handled outside of the ECP since this will need to be done on a routine basis.

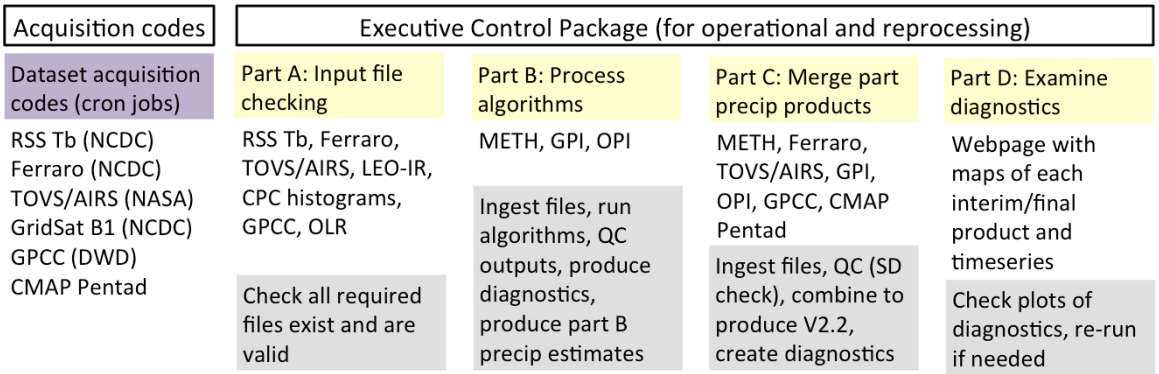

Figure 1. Schematic depicting major functions of the executive control package.

Each of the three ECPs will possess the same basic elements with file checking, integrity checks and diagnostic outputs at every level. The first part of the code will be to check that all input datasets are available and to produce text file summaries of missing data for external plotting. In practice this initial step will be treat Tb and rain rates separately, and processing will be halted if errors or missing files are encountered. The second part of the ECP will be to run any algorithms to produce precipitation estimates. The third part of the ECP will be to merge precipitation estimates, some of which are the result of the second step and some of which will be obtained as rain rates in the first step. This step will produce the final estimate as well as files containing the interim products. The final part of the ECP will be to produce diagnostics. This is a crucial part of the process, since it is anticipated that the NCDC personnel that operationally run the code are not assumed to be expert in precipitation datasets. We anticipate producing diagnostics of all products that will be made available to the GPCP developer group via a web interface and that can be used to diagnose specific problems as they arise.

2.4 Implementation of Passive Microwave codes

A schematic of the planned processing of the monthly data is shown in Figure 2. A plan was developed for the production of the ECP and necessary packages to begin work on a subset of this code that can function as a prototype, but that can also help expose issues that might be encountered. It was decided to start by producing a simplified version of the ECP that produces the monthly merged PMW product. This will involve the following steps:

1. Check RSS Tb files and produce diagnostics
2. Run METH algorithm to produce rain rates
3. Check rain rates from METH and Ferraro algorithms and produce diagnostics
4. Combine METH and Ferraro rain rates to obtain merged PMW estimate
5. Output diagnostics

This part of the code was chosen since it includes some of the more challenging aspects of the plan that involve combining inputs and code currently being run by various GPCP team members. This will also test several important issues that will need to be overcome at NCDC and so it we expect to be able to provide these to NCDC for comment and perhaps to be run operationally in their environment. These issues include the timely availability of suitable RSS Tb data at NCDC, the availability of Ferraro estimates and mechanisms required to link these data within NCDC.

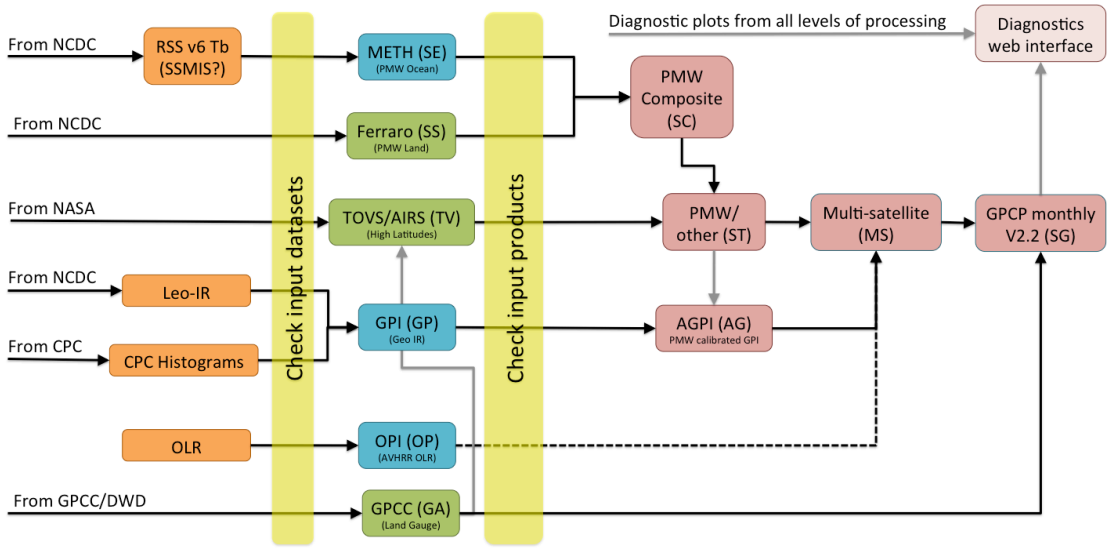

Figure 2. Schematic depicting major sections of monthly executive control package.

2.5 Availability of RSS SSMIS Tb data

RSS Tb data are an important input for GPCP as these data are used by the METH algorithm to produce estimates over the ocean. Our proposal assumes availability of the RSS SSM/I FCDR (currently available at NCDC for 1987-2008 for the GPCP calibrators F08, F10, F13) for all re-processing of GPCP. Additionally, GPCP V2.2 uses RSS V6 SSMIS Tb to continue the record past the SSM/I period (which essentially ends with F13 in October 2009, although for simplicity and to avoid late-life artifacts, the record is only used through December 2008). It is not clear that RSS V6 SSMIS Tb will be available at NCDC. This issue needs to be addressed in the coming months.