2.1 SOFTWARE PACKAGE FOR GAUGE-ONLY PRECIPITATION ANALYSIS

Dongsoo Kim*
NOAA National Climatic Data Center, Asheville, NC

Dong-Jun Seo
NOAA National Weather Service Office of Hydrologic Development and University Corporation for Atmospheric Research, Boulder, CO

1. Introduction

The gauge-only analysis (referred to as gmosaic) is one of the gridded hourly quantitative precipitation estimates (QPE) generated by the NWS River Forecast Centers (RFCs). The National Climatic Data Center (NCDC), in collaboration with the Office of Hydrologic Development (OHD), adapted the operational version of the "gauge_only" module of the Multi-sensor Precipitation Estimator (MPE) for retrospective long-term hydro-climatic applications (Nelson et al. 2009). The gauge-only analysis algorithm uses a form of kriging (Seo, 1998) to spatially interpolate gauge observations onto the Hydrological Rainfall Analysis Project (HRAP) grid (~ 4km). The real-time version uses quality control including de-selection of multiple data in a grid-box. Hence, the retrospective version does not include the original MPE QC components but allows the user to test and use his/her own QC algorithm. Other modifications from the original version of the gauge-only analysis include;

- A new input data interface to allow inclusion of users' own gauge data,
- Off-line check of multiple gauges in a grid box,
- New routines for off-line gauge quality control to test sensitivity,
- User specification of the analysis domain (default is CONUS) and parameters, and
- Diagnostic output for investigations of top-of-the-hour observations and missing values.

The purpose of the gauge analysis package is to provide the community a tool test, evaluate and benchmark their advanced algorithms and applications against baseline for cost-effective research and development of new and improved algorithms and products. We hope that the package will help tackle outstanding scientific issues in gauge analysis such as i) treatment of anisotropy, ii) disaggregation of daily precipitation for subdaily analysis, iii) optimal merging with remotely sensed data.

* *Corresponding author address:* Dongsoo Kim, NOAA National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801; e-mail: Dongsoo.Kim@noaa.gov.

2. Functional Flow Diagram

The package is divided into two modules; preprocessing and analysis. The preprocessing module prepares input time-series data for analysis, which reduces computational load on the analysis module. First, user configures the analysis in HRAP coordinates. The CONUS domain, which is the default, is in 1121 by 881 in HRAP dimension. The "PrepHADS" in Figure 1 will subsets lat/lon locations of HADS gauge stations along with HRAP coordinates. The "Reformat" deselect station if it report off the-top-hour precipitation data or the number of qualified data are less than 5 hours. The "SuperOBS" averages multiple observations within a grid-box so that only single value represents the grid box. However, we preserve original values of multiple observations for future reference for sub-pixel variability study (e.g., Zhang et al. 2007).

The analysis module (Gauge_only) remains the same as the original code in Fortran 77.

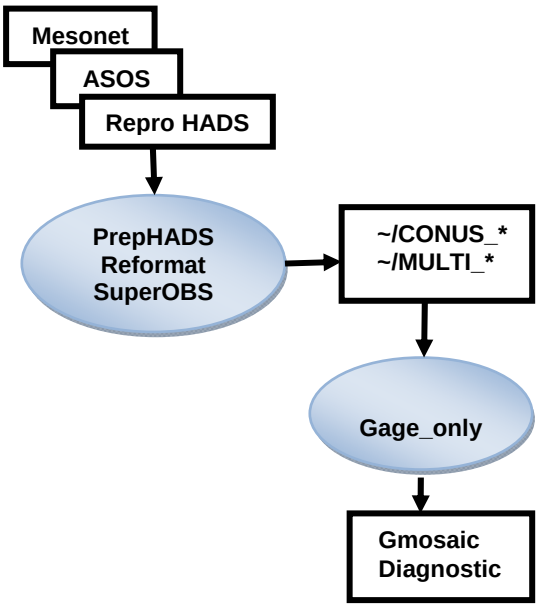

Fig. 1 Functional diagram of Gauge Analysis Package with two modules; preprocessing and analysis. The preprocessing module includes access of hourly precipitation data (PreHADS), reformat input data as latitude, longitude and precipitation value at the given hour (Reformat), and averaging of multiple data within a grid (SuperOBS).

3. GMOSAIC and Diagnostic Output

The analysis output is in the xmrg format used in the NWS (http://www.nws.noaa.gov/oh/hrl/misc/xmrg.pdf). In addition, the package produces a diagnostic file for review. An example in Table 1 lists the key information on analysis domain configuration, the period of analysis, the local directory name where input data resides, analysis parameters, the number of data ingested, the number of off-the-top-hour data and the number of missing values for each hour. Finally, it outputs the CPU time taken.

Figure 2 shows an example of the diagnostic output, including the numbers of input observations, data after deletion of missing values, and after deleting both missing and off-the-top-hour observations. The top figure is from the reprocessed HADS and ASOS data, and the bottom is from the real-time HADS data archived at UCAR and ASOS data. Note that the Repro HADS+ASOS (top) has a much more temporally-stable sample size than the realtime data.

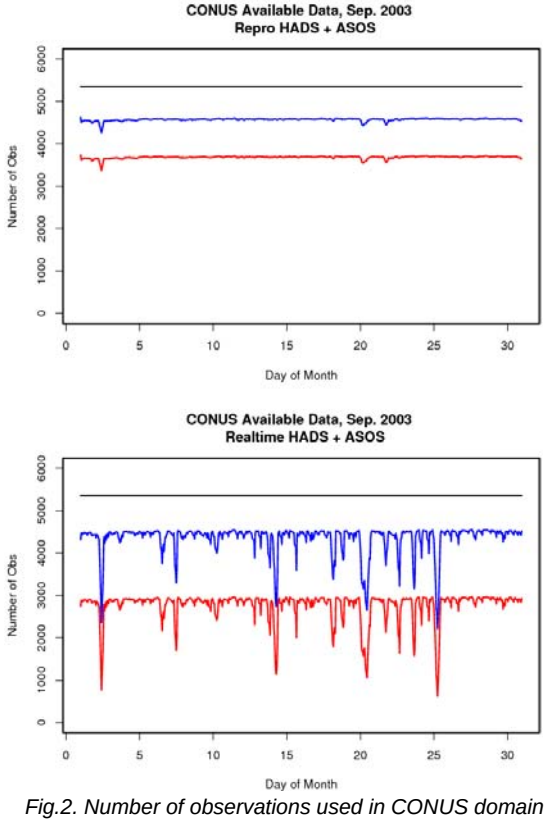

Fig.2. Number of observations used in CONUS domain Gmosaic analysis. Black lines are number of stations ingested, blue lines are number of data after deleting missing values, and red lines are after deleting both missing values and off-the top-of-the-hour data. The red lines are actual data used in Gmosaic. (Top) HADS data are from the NCDC's reprocessed HADS, (Bottom) HADS data are realtime HADS from the UCAR archive.

Figure 3a is the Gmosaic field generated with reprocessed HADS data valid at 0700 UTC 12 Sep. 2003. Figure 3b is from the Stage 2 Gmosaic archive of UCAR. Stage 2 is the NCEP's rendition of CONUS analysis based on the real-time Gmosaic fields from the NWS River Forecast Center (RFC). These figures show sensitivity of the analysis to latency of the input data in southern Texas. For reference, Figure 3c shows the CONUS mosaic of the NEXRAD digital precipitation array (DPA) data.

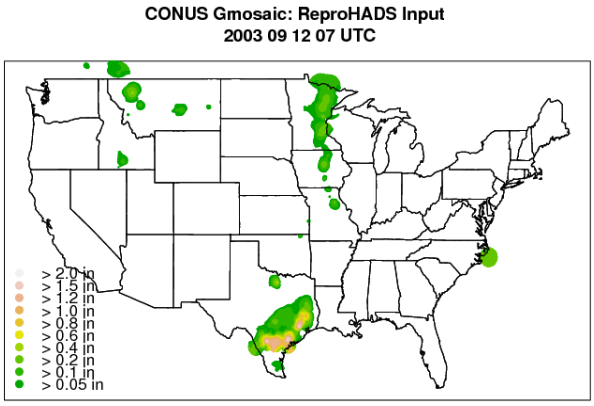

Fig. 3a CONUS domain Gmosaic with Reprocessed HADS and ASOS data valid at 07 UTC 12 Sep. 2003

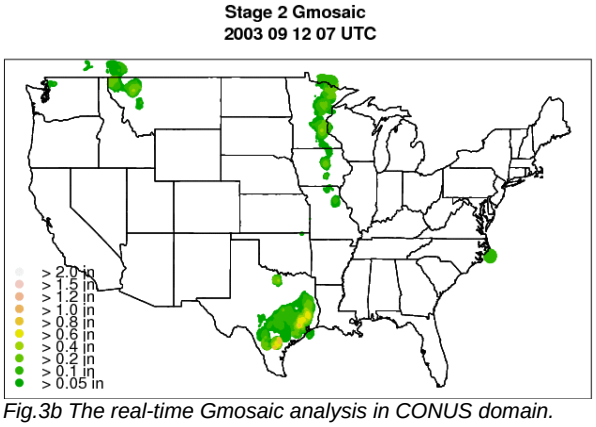

Fig.3b The real-time Gmosaic analysis in CONUS domain. This is from the archive at UCAR.

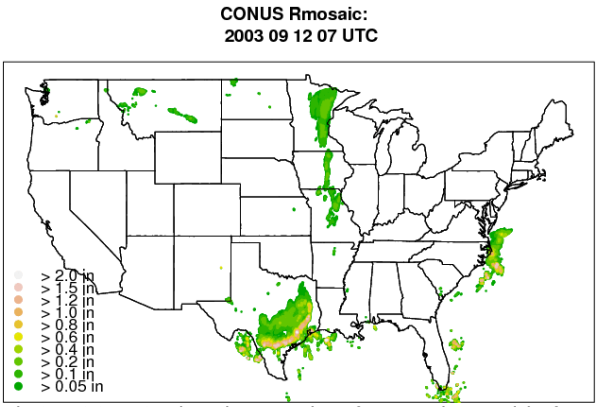

Fig.3c CONUS domain mosaic of DPA (Rmosaic) from NCDC's radar archive.

4. Resources

a. Reprocessed HADS data

The HADS hourly precipitation data have been reprocessed and are available for 2002 – 2008 in CONUS domain http://eclipse.ncdc.noaa.gov/hads/ (Kim et al. 2009) and associated quality code is available via anonymous ftp from eclipse.ncdc.noaa.gov:

```
cd /pub/Repro_HADS/Ver1.0
cd ${year}
get HD1.${year}.$[month}.tar.gz
get QG1.${year}.$[month}.tar.gz
quit
```

Table 2 describes the quality code in the QG1 file for further rectification of HD1 file.

b. Integrated Surface Database (ISD)

The ISD-Lite is a simplified version to access the global hourly surface data from which precipitation data can be obtained:

ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-lite

c. NEXRAD Digital Precipitation Array (DPA)

The WSR-88D (NEXRAD) data and display resources are available free of charge. Data inventories and a display toolkit (Ansari et al. 2009) is available at the following address:

http://www.ncdc.noaa.gov/oa/radar/radardata.html

d. Gauge Analysis Package

The manual of installing gauge analysis along with the source codes is available via anonymous ftp from:

```
eclipse.ncdc.noaa.gov
cd /pub/Repro_HADS/Gmosaic
mget *
```

5. Acknowledgment

Authors acknowledge Yu Zhang and Bill Hankins for testing the beta version of the package.

6. References

Ansari, S., C. Hutchins, S. Del Greco, N. Stroumentova, and M. Phillips, 2009: The weather and climate toolkit. *Preprints, 25th AMS Conf. on Interactive Information and Preocessing Systems (IIPS) for Meteorology, oceanography, and Hydrology*, 10 – 16, 2009, Phoenix, AZ.

Kim, D., B. Nelson and D. J. Seo, 2009: Characteristics of reprocessed hydrometeorological automated data system (HADS) hourly precipitation data, *Weather and Forecasting,* **25**, 1287-1296.

Kondragunta, C. and K. Shrestha, 2006: Automated real-time operational rain gauge quality controls in NWS hydrologic operations. *Preprints, 20th AMS Conf. on Hydrology*, 29 Jan. – 2 Feb. 2006, Atlanta, GA.

Nelson B., D. J. Seo and D. Kim, 2009: Multisensor precipitation reanalysis, submitted to *J. Hydrometeorology*.

Seo, D. J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gauge data. *J. Hydrology*, **208**, 37-52.

Zhang, Y., T. Adams and J. Bonta, 2007: Subpixel-scale rainfall variability and the effect s on separation of radar and gauge rainfall errors. *J. Hydrometeorology,* **8**, 1348-1363.

Table 1. Sample of diagnostic output

```
Thu Feb 19 09:01:41 EST 2009
 Domain: CONUS
 Grid size : 1121 x 881 = 987601
 Begin time:  2003090100
 Ending time: 2003093023
 Input Data: /san1/CONUS_GM0/Ingest/
 itype: SOE
 isearch: Double Heap
 rain_min in mm:  0.00000000E+00
 dist_min in km:  0.100000001
 number of nearest gauges (cool/warm):  20 20
 radi(cool/warm) in km:   100.0000000  60.0000000
 rngi(cool/warm) in km:   100.0000000  60.0000000
 rngc(cool/warm) in km:  100.0000000  60.0000000
 ====================
 CONUS_2003_09_01_00  nOBS/off-top-hr/missing val: 5249 876 839
 CONUS_2003_09_01_01  nOBS/off-top-hr/missing val: 5249 876 817

                ...

CONUS_2003_09_30_22  nOBS/off-top-hr/missing val: 5249 876 799
CONUS_2003_09_30_23  nOBS/off-top-hr/missing val: 5249 876 812
 CPU taken (min) :  147.092575
```

Table 2. Description of quality code associated with each datum, and suggested action for rectification

Quality Description	Quality code	Suggested Action
False light precip.	0	$P_i \leftarrow 0.0$
	1	$P_i \leftarrow 0.0$ w/risk
Outliers	2	$P_i \leftarrow$ -9
	3	$P_i \leftarrow$ -9 w/risk
Under-estimate	4	$P_i \leftarrow$ -9 w/risk
False zero precip.	5	$P_i \leftarrow$ -9 or *w/radar*
	6	*Compare w/radar*
Recoverable missing value	7	$P_i \leftarrow 0.0$
Not enough sample	8	*Compare w/radar*
No suggestion	9	No action