J4.6 ANALYSES OF HEAVY RAINFALL VARIATIONS IN THE U.S. FROM THE REPROCESSED HYDROMETEOROLOGICAL AUTOMATED DATA SYSTEM (HADS) DATA

Dongsoo Kim *
NOAA National Climatic Data Center, Asheville, North Carolina

1. INTRODUCTION

Recently, NOAA's National Climatic Data Center (NCDC) completed reprocessing of historical Hydrometeorological Automated Data System (HADS) precipitation data, which cover U.S. territories since July 1996. Earlier analyses of hourly precipitation data from the reprocessed HADS in the pilot domain, the Carolinas, provided improved quality controls with lfewer missing observations and better assignment of top-of-the-hour observations over real-time product (Kim and et al., 2006). The reprocessing of HADS data implemented several unique functions dealing with sub-hourly reports, accumulated precipitation (PC), and unrepresentative spikes and random errors in PC variables.

Each hourly precipitation product is generated two levels, the baseline level is an output of missing value recovery and level-1 is an output of spikes and random error correction. Each level produces associated quality flags which include the simple monthly statistics, namely, number of missing values, number of negative values, number of extreme precipitation events (1.0 in/hr), number of non-extreme precipitation events. Such simple statistics serve as a station qualifier for further analyses.

2. REPROCESSING VS. REAL-TIME PRODUCTION

Figure 1 is a schematic diagram of hourly precipitation production flow, done in the Office of Hydrologic Development (OHD), which deals with the time-sensitive delivery of data and the delayed production flow applied in the reprocessing. This reprocessing of HADS data, has only a one day latency. The diagram shows that some River Forecast Centers (RFC), some Weather Forecast Offices (WFO) and the National Center for Environmental Prediction (NCEP) receive hourly precipitation data which are derived from PC. NCDC receives the original format HADS data, and the delayed mode processing has more of a chance to capture late-arriving observations.

Table 1. Quality measures between reprocessed Level-1 product and real-time product during warm seasons (April through September), 2003-2005 in the Carolinas.

Year	Mon	N	Missing value (%)		Top-hour Obs (%)	
			Repro	RT	Repro	RT
2003	Apr	98	1	6	95	59
2003	May	99	1	8	95	60
2003	Jun	99	1	8	94	60
2003	Jul	99	1	7	94	58
2003	Aug	99	0	6	95	57
2003	Sep	103	1	9	94	56
2004	Apr	138	0	11	89	60
2004	May	141	1	8	91	61
2004	Jun	141	1	5	91	62
2004	Jul	143	0	4	90	61
2004	Aug	144	1	4	91	62
2004	Sep	141	1	5	92	62
2005	Apr	136	0	3	92	63
2005	May	136	0	3	91	62
2005	Jun	136	5*	2	91	62
2005	Jul	136	6*	3	91	63
2005	Aug	135	0	3	94	61
2005	Sep	135	0	2	89	58

The direct comparison between real-time and reprocessed hourly precipitation products is shown in Table 1. The number of HADS stations varies over time, but the percentage of missing values in the reprocessed product is much smaller than real-time product except for June and July of 2005, when a couple of days are missing in the original format HADS data due to a system failure at NCDC. Most of the recovered missing values were no-rain cases, but some were rain cases. As a result, monthly totals are always greater in the reprocessed product. For example, Figure 2 compares two time-series of hourly precipitation products at gauge station ARRN7, North Carolina, during September of 2004. All 44 missing values were recovered in the reprocessing including some rain events, so 2.03 inches of rain during the month was recovered. This was a 25% gain in the monthly total. Apparently, a single event was an extreme rain with 1.23 in/hr on 0700 UTC on September 28, 2004, due to Hurricane Jeanne.

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

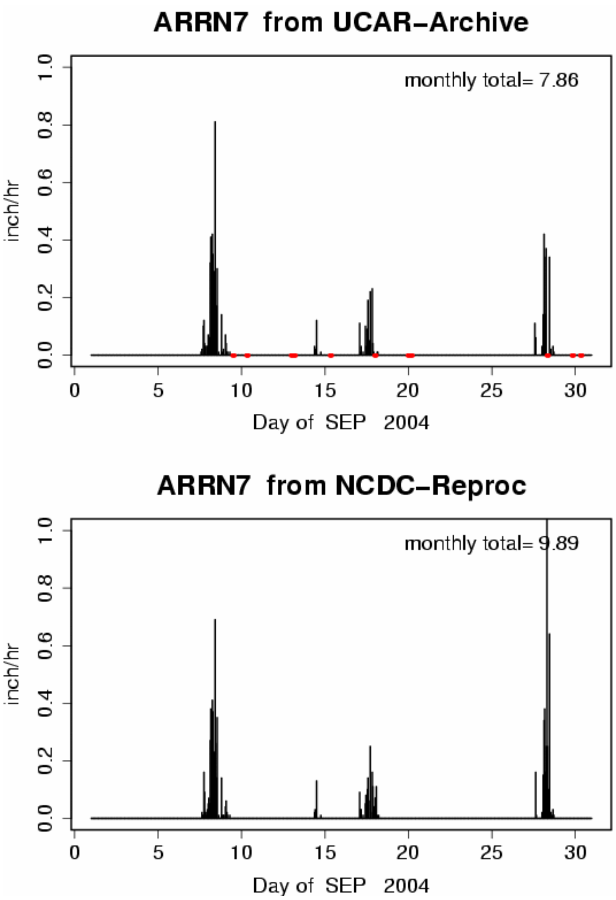

Figure 2. Two time-series of hourly precipitation at gauge station ARRN7, North Carolina, during September 2004. (Top) Real-time product archived in UCAR, (Bottom) Reprocessed product. All missing values in real-time product have been recovered, and the total monthly rainfall amount gained 2.03 inches.

3. COLLOCATED STATIONS

Some HADS stations are located near daily COOP stations, i.e. within a 0.5 deg distance. There are at most 20 of these collocated stations. Table 2a is a sample of collocation for HADS station ARRN7. The first row is a HADS station with a number of missing values in the number of hours (how many?), the rest are neighboring daily COOP stations in ascending order by distance (deg). The number of missing values for daily COOP stations is thus counted in days. Similar tables comparing HADS stations with neighboring HADS stations are also being produced as seen in Table 2b. This table will serve as a spatial consistency check on the monthly time scale.

PC at ARRN7 during September, 2004., both reprocessed HADS and COOP data, are in agreement on rain days even though rain amounts differ (Figure 3). The advantage of using hourly precipitation data is the ability to detect beginning hour of precipitation. The HADS-HADS collocation table identified the nearest HADS station COPN7. COPN7 reported 0.61 inches of

Table 2a. Neighboring daily COOP stations (11) with monthly total rainfall at the target HADS station ARRN7 for examination of hourly product.

Name	Miss	Total	Lat	Long	Dis(deg)
ARRN7	0	9.89	36.40	-80.56	0.0
315890	0	9.65	36.52	-80.62	0.12
319675	0	9.65	36.13	-80.55	0.27
445453	0	16.78	36.67	-80.45	0.28
314675	0	8.84	36.32	-80.27	0.31
317548	0	9.18	36.23	-80.30	0.31
312740	0	10.54	36.23	-80.85	0.34
312238	0	9.87	36.42	-80.20	0.36
448170	0	10.98	36.63	-80.27	0.37
443991	6	11.65	36.77	-80.73	0.40
449272	0	14.19	36.73	-80.28	0.43
443267	0	13.47	36.67	-80.92	0.44

Table 2b. Neighboring HADS stations with monthly total rainfall at the target HADS station.

Name	Miss	Total	Lat	Long	Dis(deg)
ARRN7	0	9.89	36.40	-80.56	0.0
COPN7	0	10.68	36.36	-80.69	0.13
ENNN7	0	10.82	36.13	-80.45	0.29
GAXV2	0	16.14	36.65	-80.98	0.48

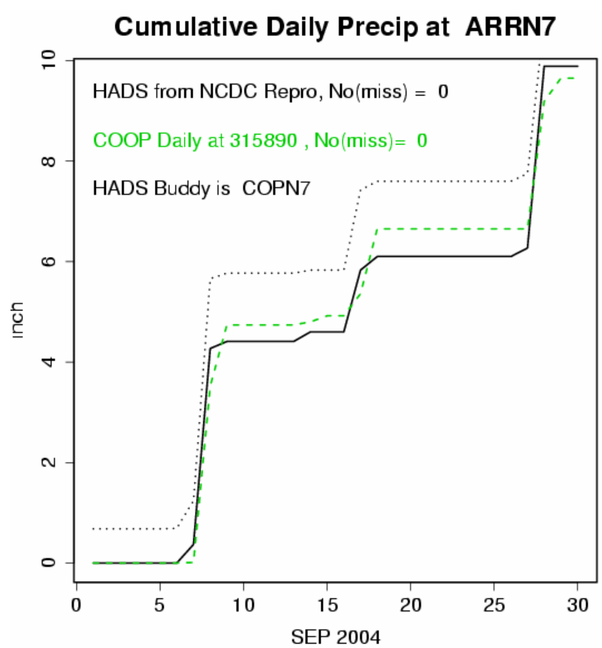

Figure 3. The cumulative daily precipitation at the HADS station ARRN7. The black solid line is accumulated precipitation from the reprocessed data, the black dot is the nearest station, COPN7, which is 0.13 deg away, and the green dash is from daily COOP which is 0.12 deg away.

rain at 0500 UTC, on September 1 while ARRN7 reported no-rain. This is a subject for further study to establish proper threshold values in spatial consistency checks.

4. INVENTORY OF LEVEL-1 PRODUCT

As mentioned in Section 2, there was difficulty retrieving HADS data from the archive on magnetic tapes at OHD (Cedrone, 2005, Personal communications). This problem resulted in an incomplete inventory of hourly precipitation product at NCDC. Table 3 lists of missing dates of historic HADS data beginning in July of 1996. The rescue effort of historic HADS precipitation data included contacting organizations such as The Western Region Climate Center and The University of Utah (Horel, 2006, Personal communications). Until the rescue effort is completed, the real-time hourly precipitation product archived at UCAR will be used.

5. ANALYSES OF EXTREME RAIN EVENTS

The hourly precipitation data provide an opportunity to analyze characteristics of sub-daily events, such as the diurnal cycle and extreme rain events on an hourly scale. The previous hourly precipitation data at NCDC (HPD; NCDC, 2003). Was used by Brooks and Stensrud (2000) to study heavy rain events climatology. That paper concluded that the HPD still missed most of the truly large precipitation events because of a relatively coarse station resolution of about 50km. If the reprocessed HADS proves to be of an equivalent quality, then the reprocessed HADS precipitation data will be an additional hydro-climate database.

Stations that report heavy rains are analyzed by comparing neighboring HADS and COOP stations to identify erroneous reports. Some patterns of heavy rainfall will be presented.

6. SUMMARY AND PLAN

The hourly precipitation product from historical HADS data is being compared with daily and hourly COOP precipitation to determine its quality. When rescue of HADS data with improved quality check is completed, it will triple the size of hourly precipitation data holding (DSI-3240) at NCDC.

Table 3. Table listing of missing dates in the reprocessed HADS hourly precipitation product at NCDC.

Year	Month	Date
1996	Jul Dec.	1 – 3, 31 Whole month
1997	Jan Aug	Whole month Whole month
1998	Jan June	Whole month 16 - 22
1999	Jan May June Dec.	3, 6 Whole month 22, 23 31
2000	Jan. – Apr. Aug. – Sep. Oct. Nov. Dec	Whole month Whole month 12 - 14, 30 - 31 5 Whole month
2001	Jan. – Sep. Nov.	Whole month 28
2002	Apr Nov	5 - 8 12
2003	Oct Nov	8 - 13 Whole month
2004	Jan Mar.	Whole month 19 - 26
2005	June July Dec	26 13, 19 31

7. ACKNOWLEDGMENT

The author acknowledges review by Anne Markel and Dr. John Bates and extends thanks to Drs. Seo, Kondragunta of OHD and Brian Nelson of NCDC for discussions on precipitation quality.

8. REFERENCES

Brooks, H. E. and D. J. Stensrud, 2000: Climatology of heavy rain events in the United States from hourly precipitation observations, *Mon. Wea. Rev*., Vol 128, 1194 – 1291.

Kim, D., B. Nelson and L. Cedrone, 2006: Reprocessing of historic hydrometeorological automated data system (HADS) precipitation data, Preprint, AMS 20th Conf on IOAS, 29 Jan. - 2 Feb., 2006, Atlanta, GA.

NCDC, 2003: Data documentation for data set 3240 (DSI-3240), National Climatic Data Center, 151 Patton Ave, Asheville, NC 28801-5001, USA.

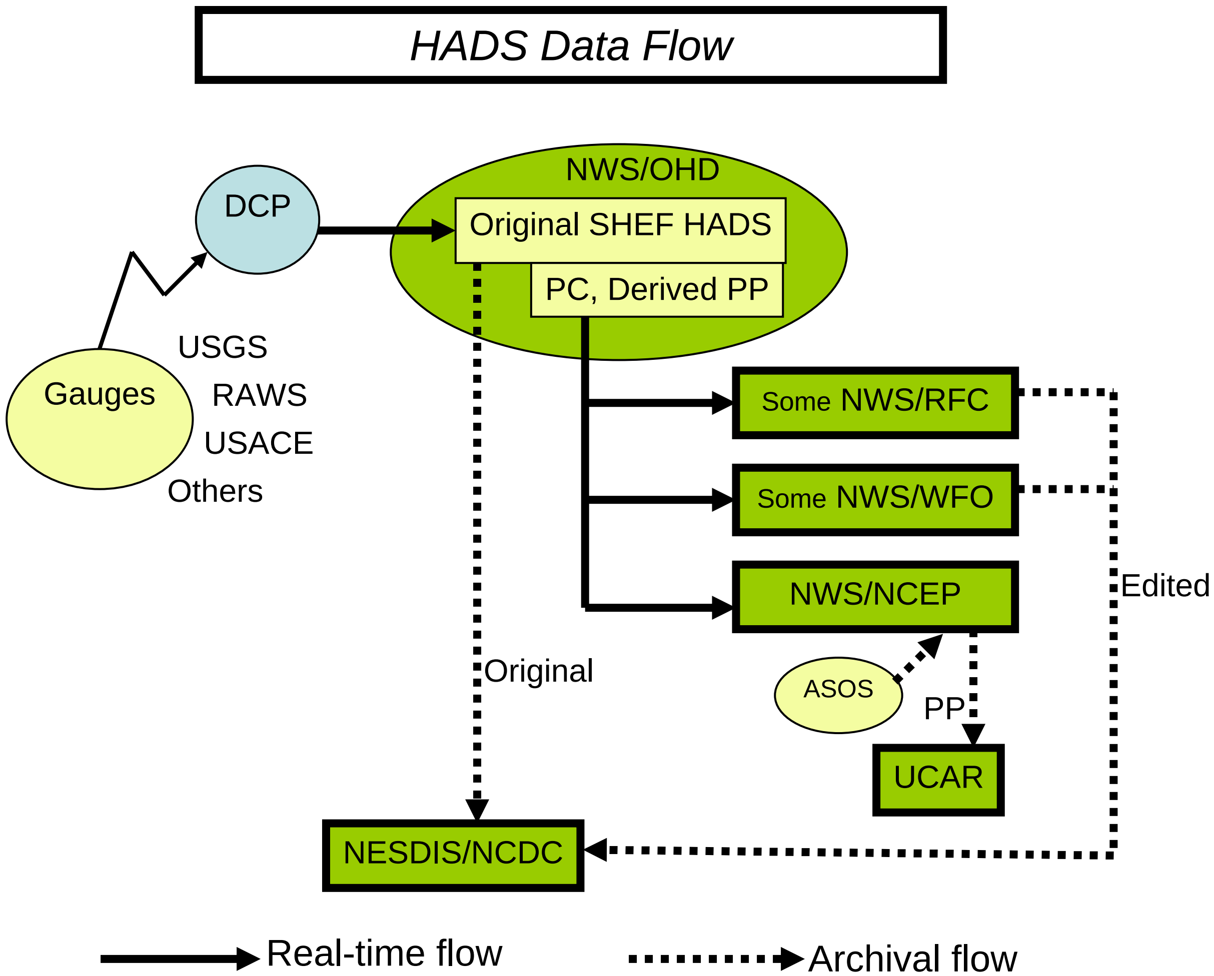

Figure 1. A schematic diagram of HADS real-time suite of data and product dissemination. The HADS Program Office converts raw gauge data into Standardized Hydrometeorological Exchange Format(SHEF), produces hourly precipitation, and disseminates it to some RFCs, WFOs and NCEP in real-time. The HADS office pushes original SHEF-format HADS data to NCDC once a day. All surface observations and manually edited data are permanently archived at NCDC under the Service Records Retention System (SRRS).