Creation of Multisensor Precipitation Products from WSI NOWrad reflectivity data

Yu Zhang[1], David Kitzmiller[2], Dong-Jun Seo[3], Dongsoo Kim[4], and Robert Cifelli[5]

1. Physical Scientist, Office of Hydrologic Development, NOAA National Weather Service, 1325 East-west Hwy, Silver Spring, MD, 20910

2. Meteorologist, Office of Hydrologic Development, NOAA National Weather Service, 1325 East-west Hwy, Silver Spring, MD, 20910

3. Associate Professor, Dept of Civil Engineering, University of Texas at Arlington, Box 19308, Rm 248E Nedderman Hall, 416 Yates St, Arlington, TX 76019-03083.

4. Physical Scientist, NOAA National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801

5. Meteorologist, NOAA Oceanic and Atmospheric Research, Earth System Research Laboratory, 325 Broadway, Boulder, CO 80305-3337

Abstract

Radar and multisensor quantitative precipitation estimates (QPEs) have played a critical role in real-time hydrologic and weather predictions. The utility of this data set for hydrologic model calibration, however, is hampered by the limited data archive and the presence of data gaps. In this study, we investigate the use of a composite reflectivity data set, namely the WSI NOWrad data set, to create an hourly QPE that would complement the National Weather Service's Stage III archive by filling in the latter's gaps and potentially improving the latter's accuracy. We first perform an inventory analysis of Stage III and NOWrad products, and through which we find that as much as 17% of the Stage III data were missing for some of the regions, and some of these gaps can be filled using the NOWrad data. Then, we create two experimental products: the first one (NOWrad-RAW) is based on variable Z-R relationship derived for each month using monthly COOP station totals. The second product (NOWrad-BMO) is derived by bias correcting the NOWrad-RAW using hourly rain gauge products. We then assess the two products along with Stage III products against independent hourly gauge reports over two locations - Charlotte-Mecklenburg metropolitan area in North Carolina, and the Lower Colorado River drainage in central Texas. Our analyses over the two test sites reveal that 1) NOWrad-RAW product suffer from a negative overall and conditional bias, while the Stage III is closer to bias-neutral, and 2) NOWrad-RAW product, at least over the Charlotte-Mecklenburg area, can outperform the Stage III in terms of correlation with gauge data and skill in detecting light rainfall. Further analyses of the RAW and bias-corrected NOWrad suggest that negative conditional bias may be substantially improved, but there is a tendency to over-correct the bias for the summer months, possibly due to the presence of false detections. Our results show that NOWrad can be a viable source of high-resolution quantitative precipitation information and it indeed complements the current NWS archive in several respects. Possible mechanisms for further improving the accuracy of the NOWrad QPE are discussed.

Introduction

The US National Weather Service's Next-Generation Weather Radar (NEXRAD) network, since its deployment starting from the late 1980s, has yielded profound improvement in both the accuracy and definition of precipitation information over the nation. NEXRAD-based quantitative precipitation estimates (QPEs), which comprise a large suite of radar-only (Fulton et al. 1998) and radar-gauge multisensor products (Seo et al. 2010) , now serve multiple roles in the National Weather Service's river and flash flood forecast operations (Zhang et al. 2011, Kitzmiller et al. 2013). In real-time operations, the forecasters employ the QPEs directly to continuously update the states of hydrologic models or to monitor potential flash floods in conjunction with time-varying flash flood guidance. These QPEs also play a critical part in generating and validating precipitation forecast which serve as a forcing for forecasting river stage and soil moisture information. As an example, precipitation nowcast based on NEXRAD QPEs, such as the High-resolution Precipitation Nowcaster (Kitzmiller et al. 2013), is now in operation for flash flood prediction. The radar QPEs are also routinely used at the NWS National Centers for Environmental Predictions for validating the precipitation forecast issued by numerical weather prediction models (see related work in Marchok et al. 2007).

Outside these real-time applications, NEXRAD QPEs have also served as a forcing for calibrating hydrologic models (see Reed et al., 2004, Smith et al., 2012); and are potentially an important source of information of precipitation climatology (Baeck and Smith, 1995).. For these purposes, however, the limitations of the current NEXRAD QPE data are also apparent. The current operational national mosaic QPE produced at the NWS, namely the Stage IV data, has an archive that only goes back to 2002 (Lin and Mitchell 2005). For the earlier period of 1996-2001, radar-gauge multisensor QPE created via the legacy Stage III algorithm at NWS River Forecast Centers has been archived, as has the NCEP Stage II product suite. Yet, these archives include substantial data gaps, and the data quality was deemed relatively low, especially in the earlier years (Zhang et al. 2007, 2011). Table 1 summarizes the percentage of hours with missing data in the Stage III archive across the 12 RFCs in the Conterminous US for 1996-2001. This number varies greatly among RFCs and across years, and the average can be as high as 78% (CNRFC) and as low as 3% (NCRFC). These missing data present a critical challenge to the use of the data set for activities such as hydrologic calibration, where a continuous data record over relatively long periods is needed.

In this study, we investigate the use of an alternative data source, the Weather Services International (WSI) NOWrad composite reflectivity data set, to retrospectively create a suite of gauge-radar QPE products that would complement the existing Stage III data set by filling the gaps and potentially mitigating some artifacts in the latter. In order to attain reliable QPEs, we devise a strategy which consists of two steps. The first step is to derive Z-R relationships using the monthly rainfall totals from the Cooperative Observer Program (COOP) gauges, and to create an hourly QPE product using the derived Z-R relationships. The second step involves applying the fusion mechanism of the operational Multisensor Precipitation Estimator (MPE) to bias-correct the products from Step 1 using hourly rain gauge reports. The former and latter products will be henceforth referred to as NOWrad-RAW and NOWrad-BMO products in this paper. We will then assess the accuracy of these data sets against hourly rain gauge data along with available archival Stage III data set to examine their strengths and weakness against the latter.

For the remainder of the paper, we first describe the data sets, including the NOWrad, gauge data used for merging and validation, and the archived Stage III data; then we present the methodology in creating the QPE data; and document the observations. The key outcomes and conclusions are summarized in the last section.

Data sets

Stage III

The Stage III data were retrieved from the archive at Office of Hydrologic Development (OHD) (http://dipper.nws.noaa.gov/hdsb/data/nexrad/nexrad.html). Stage III data were created using the legacy Stage III algorithm (Seo et al. 2010, Zhang et al. 2011, Kitzmiller et al. 2013), which takes the input, i.e., the hourly Digital Precipitation Array (DPA) data from the NEXRAD Precipitation Processing System (PPS; Fulton et al. 1998), and performs mean field bias correction to the DPAs of each radar using the algorithm developed by Smith and Krajewski (1991). The Stage III ceased to be operational after 2001 and was replaced by the Multisensor Precipitation Estimator (MPE; Seo et al. 2011) package, which consists of a wider suite of algorithms for performing bias correction and multisensor fusion. The Stage III data created in real time at the NWS RFCs were centrally archived at OHD. This archive, however, has numerous temporal gaps (Table 1) due to a variety of factors, as well as limited spatial coverage.

Stage III has been evaluated in a number of studies (e.g., Zhang et al. 2007, Hardegree et al. 2008, Zhang et al. 2011). A primary issue found in some of these studies is that Stage III tended to underestimate light rainfall, and this issue is related to a numerical truncation error in the NEXRAD PPS which was fixed in 2003 (Zhang et al. 2011).

NOWrad Composite Reflectivity

The NOWrad composite reflectivity data sets (Weather Services International Corporation 1995; National Center for Atmospheric Research and Global Hydrology Resources Center, 2014) were created by WSI by mosaicking the reflectivity from the Weather Surveillance Radar - 1988 Doppler (WSR-88Ds). These data have been used for similar studies (e.g., Germann and Zawadski 2002). Detailed documentation of this data set can be found at http://www2.mmm.ucar.edu/imagearchive/WSI/docs/GHRC_README.htm, and a brief description is provided here. According to this documentation, the instantaneous reflectivity data were manually quality assured, and a single value was selected for each 15-minute window over each grid box (0.0191 degree latitude and 0.0181 degree longitude, or nearly 2km by 2km). The reflectivity values were converted to reflectivity categories, with each category encompassing a 5-dBZ range (National Center for Atmospheric Research 2014). The data set is available for the Conterminous US on a 15-minute interval since 1996 till 2007. For our study, the data for 1996-2001 were acquired and processed as described below.

In comparison to the existing Stage III data, the NOWrad data set exhibit a number of attractive features. First, the data set is of higher spatial and temporal resolutions (~2km and 15 min) than the Stage III (~4km and hourly), and can be used to derive QPE products at these resolutions. Second, according to the NCAR website where the NOWrad data is distributed, the data set has gone through a degree of manual quality assurance (QA). Third, the data is present in many instances where Stage III is missing, and therefore it is useful for filling in the gaps of Stage III archive. Table 2 summarizes the number of missing data for each month in the NOWrad archive. Tables 3 and 4 provide the corresponding numbers in the Stage III archives in West Gulf River Forecast Center (WGRFC) and Southeast River Forecast Center (SERFC). NOWrad features a relatively small number of missing data over 1996-1998. For 1999-2001, however, a number of considerable gaps exist in the NOWrad data (Table 2). The Stage III archives in WGRFC and SERFC also consist of large gaps which, fortunately, do not coincide with those from NOWrad. For example, a large fraction of Stage III data is missing in the SERFC archive over the first 6 months of 1996 (Table 4), when NOWrad data are nearly complete (Table 2). As such, NOWrad presents a viable alternative data set for addressing the gaps in Stage III.

Gauge data

We employ three gauge data sets in this study. The first one is NOAA Cooperative Observer Program (COOP) daily reports. These reports were quality assured and accumulated to create a monthly product. The second data set is the gauge network located in Charlotte-Mecklenburg metropolitan area for the period of 1997-2001. These are 5-minute reports aggregated onto hourly scales and quality assured at the National Climatic Data Center. The third one is the data set from the Lower Colorado River Authority (LCRA). The LCRA gauge data used in this study spans 2000-2001; it underwent both manual and automated QA procedures. The manual procedures helped identify and remove conspicuous temporal discontinuities, whereas the automated QA marked suspicious records through neighborhood check (Kondragunta and Shrestha. 2006).

Methodology

Our analyses consist of three phases. The first phase centers on the derivation of Z-R relationships and creation of an hourly QPE (NOWrad-RAW) from the NOWrad composite reflectivity grid. The second phase entails fusing the radar-only data with gauge observations using the mean field bias module of the Multisensor Precipitation Estimator to create a bias-corrected product, which we will hereafter refer to as the NOWrad-BMO (Bias-corrected

MOsaic) data set. The third phase is the validation phase. In this phase, we collect coincidental Stage III and NOWrad-MPE, and perform validation of both products against hourly rain gauge data, and compare the validation statistics to discern the comparative strengths of these data sets.

Generation of the NOWrad Radar-only QPE

The NOWrad-RAW product is generated following three steps. In the first step, the reflectivity categories for each grid box are translated back to reflectivity values. As the discretization of reflectivity value to NOWrad categories makes it impossible to fully recover the values from the latter, we use the centroid of each category as the representative reflectivity value. For example, for any pixel with reflectivity category "2", which corresponds to 5-10 dBZ range, we select 7.5dBZ as the approximate reflectivity value. To derive Z-R relationship for calculating rainfall rates from reflectivity, we devise a simple regression approach using monthly COOP precipitation gauging station reports as the ground truth and predictands. Fig. 1 shows the spatial distribution of the COOP stations used in the analysis. In the regression approach, the number of occurrences, or frequency, of each reflectivity category is calculated for each grid box. Then the coefficient and exponent in the Z-R relationship, a and b respectively, are optimized to maximize the correlation between the monthly totals from the gauge and the NOWrad-based product. The procedure is described below.

$$\widehat{R}_j = \sum_{i=1,n} \left(\frac{Z_j^i}{a}\right)^{\frac{1}{b}}$$

$$Find\ (a,b) to\ maximize\ Cor(\boldsymbol{R}, \widehat{\boldsymbol{R}})$$

where Z_j^i is the NOWrad reflectivity at pixel j and hour i; $\widehat{R}_j$ is the NOWrad-based monthly accumulation at pixel j; $\boldsymbol{R}$= (R_1,R_2..R_n) is the vector of observations at gauge locations (1,2..,n), and $\hat{\boldsymbol{R}}$ is the corresponding NOWrad-based estimates. Note that the hail cap is set at 50 dBZ and the 17.5 dBZ (centroid of NOWrad category 4) is considered the minimum reflectivity for precipitation signal (anything below would be considered as clutter).

Spatially uniform Z-R coefficients were derived for each month individually. The seasonal variations of the coefficients are shown in Fig. 2. The range of coefficient ‘a’ is between 280 and 360. It is in general lower in the summer and winter. The exponent ‘b’ varies between 1.5 and 3.0, and it is mostly lower in the summer (close to 1.7) and the highest in the winter. These results suggest that the summertime Z-R is in fact quite close to the convective one (a=300, b=1.4). Then, the rainfall rate from the WSI grid mesh was aggregated onto the Hydrologic Rainfall Analysis Project (HRAP) grid mesh, by simply averaging the values from the approximately four grid boxes embedded in each HRAP pixel. Subsequently, the rain rate over the four snapshots within an hour was added up to derive hourly accumulation with the assumption that the rate remained constant over each window. The resultant product is termed “NOWrad-RAW” that is roughly equivalent to the NWS Stage I product. It must be noted that this is not a purely radar-only product, as it did utilize monthly gauge data in deriving the Z-R relationship. Also, it should be heeded that the limited precision of the NOWrad reflectivity, which is a result of discretization, will necessarily limit the accuracy of precipitation estimates derived from it.

Creation of bias-corrected NOWrad QPE

For demonstration purpose, the bias-corrected NOWrad product is created for Texas for the year of 2000 using a module of the Multisensor Precipitation Estimator (MPE; Seo et al. 2010, Zhang et al. 2011, Kitzmiller et al. 2013). In order to maintain an independent gauge data set for validation, we used the gauge reports from the Hydrometeorological Automated Data System (HADS) in performing the analysis, and used the records from the Lower Colorado River Authority (LCRA) gauges as validation reference. Fig.3 shows the locations of HADS and LCRA gauges for this analysis. Note that we removed the LCRA gauges from the HADS archive, but some of the LCRA gauges may have been used in creating the Stage III product and therefore it is not necessarily an independent validation reference for Stage III.

The NOWrad-RAW product undergoes bias correction using the hourly gauge reports via the mean field bias module in MPE (Seo et al. 1999). The mean field bias module basically computes a single, spatially uniform bias factor for each region using collocated radar and gauge estimates over one or more time intervals. In our study, we defined the region as the entire Texas plus areas from adjacent states (OK and LA), and estimated the bias using all available reports from HADS stations. Multiplication of the radar-only field by the bias field yields the bias-corrected QPE.

Validation

The validation study is carried out over two locations, namely a) Charlotte-Mecklenburg metropolitan area in North Carolina, and b) Central Texas. The former site is located within the forecast domain of SERFC. For this site, reports from 70 tipping bucket gauges are used as validation reference (Fig. 4). For the latter, which is situated in the forecast domain of WGRFC, we use the LCRA gauge data as the reference (Fig. 3). For both sites, we also acquire Stage III data for hours when it was available. The validation period for a) is 1997-2001, whereas for b) is 2000-2001 as LCRA gauge data we acquired only cover this period. Our validation experiment focused on the hourly scale. We employ traditional metrics such as overall bias and correlation coefficient (CC), where bias is defined as the logarithm of the ratio of cumulative precipitation from radar to that from collocated gauges. In addition, we use quantile comparisons to depict the conditional bias. Moreover, since a key shortcoming of the Stage III data set is its inability to resolve light rainfall (Zhang et al. 2007, 2011), we calculate the probability of detection of false alarm ratio of light rainfall over the Charlotte-Mecklenburg area, where a relatively long archive of validation gauge data is available.

Observations

Charlotte-Mecklenburg

The validation results for Charlotte-Mecklenburg area are shown in Figs. 5-10.

Fig. 5 shows the scatter plots of CC and bias of hourly rainfall accumulations between NOWrad-RAW/Stage III and gauges calculated by considering all pairs between 1997 and 2001. It is clear that the NOWrad-RAW exhibits slightly higher correlation than Stage III (Fig. 5a), though it suffers from a more severely negative bias (Fig. 5b). Fig.6 shows the CC and bias computed from al hourly gauge-radar pairs for each month between 1997 and 2001 when NOWrad data were available. The following features are evident. First, for a majority of the stations, the CC is higher between NOWrad and gauge reports than that between Stage III and gauge data (Fig. 6a). Second, NOWrad-RAW contains fewer monthly totals with large bias than Stage III (Fig. 6b). A closer look at the bias of NOWrad-RAW and Stage III reveals that the bias is overall negative for both Stage III and for NOWrad, but the value of the latter being only slightly worse (mean = 0.865 and 0.861 for Stage III and NOWrad, respectively).

Quantile plots of Stage III and NOWrad versus gauge accumulations are shown in Fig. 7a and b for the entire record and for the summer, respectively. Apparently, Stage III data exhibit a positive conditional bias whereas the bias in NOWrad is overall negative. Such a feature is also evident in the summer months, except that the conditional bias is slowly diminishing for Stage III towards higher rainfall amounts, whereas for NOWrad the negative bias persists. Apparently, NOWrad suffers from an underrepresentation of high rainfall amounts, whereas Stage III experiences a slight overrepresentation. This difference is further explored through the time series plots of Fig. 8. The two events in July both point to overestimation of larger rainfall amounts by Stage III and underestimation by NOWrad-RAW.

As indicated earlier, one of the major issues of Stage III product is a lack of detection of light rainfall due to the truncation error. Figs. 9a and b explore the probability of detecting light rainfall (0.25 - 2mm/h) and the related false alarms, respectively. The probability of detection (POD) was calculated for Stage III and NOWrad-RAW by counting the fraction of instances with hourly precipitation from gauges falling within the range of (0.25mm, 2.00mm), whereas false alarm rate (FAR) was computed as the ratio of instances where NOWrad/Stage III indicates precipitation in the (0.25mm, 2.00mm) category whereas gauge reports zero (i.e., less than 0.25

mm). Notable features of Fig. 9a include: a) NOWrad-RAW indeed exhibits higher POD values over the entire time period than Stage III, and b) there is a conspicuous rising trend in the Stage III-based POD values, whereas those for NOWrad-RAW are relatively flat. The first feature is a clear indication that the NOWrad-RAW product, which is not affected by the truncation error of the NEXRAD PPS, most likely provides a more robust depiction of the light rain despite its overall and conditional negative bias. The second feature, i.e., the rising trend in the bias of Stage III, indicates that the Stage III products underwent incremental improvements in time, most likely due to the better use of gauge data in creating the Stage III. This feature is remarkably consistent with the observations of Zhang et al. (2011) that bias in the streamflow simulation results tended to improve over time using Stage III and MPE products from Mid-Atlantic River Forecast Center (MARFC). The key underlying cause of the latter feature, as stated in Zhang et al. (2011), is the incorporation of daily manual observations in creating the Stage III and MPE data sets.

While NOWrad data does exhibit higher skill in detecting light precipitation, it also suffers from a higher FAR. As shown in Fig. 9b, FAR from NOWrad is consistently higher than that from Stage III throughout the 1997-2001 period. The presence of false alarms for light rainfall in both radar products is unsurprising and can be attributed to limitations in both rain gauge and radar observing mechanisms. During light rainfall events, evaporation could reduce the water cumulated in tipping buckets and therefore result in undercatch by gauges which manifests as "false alarms". On the other hand, it is entirely possible that droplets observed by radar vanish prior to reaching ground due to subcloud evaporation.

It is also interesting to note that, the FAR from Stage III features a clear positive trend over the same period while POD experiences a substantial increase. It is unclear if the increasing FAR is indicative of deterioration in both products and this requires additional investigation, or simply that the limitation of gauges become more acute in light of improving radar QPE. In sum, as the NOWrad-Stage III difference in FAR appears to be narrower than that in POD, NOWrad-RAW data does appear to have an overall advantage over Stage III in terms of resolving light rainfall.

The differential ability of Stage III and NOWrad-RAW products in depicting rainfall intensity coverage is further illustrated in Fig. 10, where hourly accumulations ending 0 UTC on 30 July 2010 from the two data sets are contrasted. It is evident that the rainfall coverage from NOWrad is much broader and extensive; however, the NOWrad-RAW estimates are limited in dynamic range and are unable to resolve the intensity associated with the convective elements compared to the Stage III analysis.

Central Texas

The validation results of NOWrad-RAW and Stage III products over the LCRA domain are shown in Figs 11-16.

First presented are the scatter plots of hourly totals from Stage III, NOWrad versus gauge data lumped over the 18 months in 2000 and 2001 (Figs 11a and b). Between the two products, Stage III exhibit a positive overall bias where NOWrad a negative one, similar to the results for Charlotte-Mecklenburg, and the CC from Stage III with independent gauges is slightly higher than that of NOWrad.

Comparisons of CC and bias for individual months are shown in Figs 12a and b. For a majority of months, CC for NOWrad is in fact lower than that for Stage III, in direct contrast to the result obtained at Charlotte-Mecklenburg, where the opposite was found. As for bias, NOWrad still exhibits overall negative bias, with bias being negative for 12 out of 18 months, whereas bias for Stage III is mostly positive, with only 4 out of 18 months exhibiting negative bias. The absolute bias value (|bias -1|) from the two data sets are overall better (i.e., closer to neutral) for Stage III (for 7 out of 18 months).

One of the objectives of the work, as stated in mentioned in the methodology, is to appraise the possible incremental improvements of the NOWrad QPE through fusion with hourly gauge data. In our study, we apply the mean-field bias module of the MPE to the RAW NOWrad QPE for the year of 2000, for which we use the hourly reports from Hydrometeorological Automated Data System (HADS) gauges (Zhang et al. 2011) from the entire region as the basis for computing the bias and performing the correction. The mean field module calculates a spatially uniform bias factor over the entire WGRFC domain for each hour on the basis of both the current and past gauge-radar pairs, and it retrieves the so-called "best bias" factor (the non-zero bias value over the shortest time horizon) to correct the hourly rainfall accumulation for that hour. The results of this comparison are shown in Figs. 13-16.

The CC and bias for Stage III and the two NOWrad QPEs are shown in Figs. 13a and b for 10 months in 2000 when data from NOWrad were available. Evidently, the mean field bias correction has minor, and sometime slightly negative impacts on the CC: CC for NOWrad-BMO is only slightly higher than that for NOWrad-RAW over the earlier part of the year (up to June), and it is clearly worse for July, August and December (Fig. 13a). Stage III broadly outperforms both NOWrad products for most of the months in terms of CC. The comparisons of bias also yielded interesting results (Fig. 13b). First, it appears that Stage III exhibits a consistent seasonally-dependent bias which is negative in the winter but mostly positive in the warm season. The bias of NOWrad-RAW is overall lower than that of Stage III from January to July. For the summer, it ranges from being nearly neutral (June) to positive (July and August). Mean field bias correction had the effect of increasing the bias value, and in some cases significantly. For example, for January and April, bias correction rendered the bias of NOWrad close to that of Stage III. However, for the summer months, bias correction led to broad overestimation of total rainfall volume as indicated by the large positive bias in NOWrad-BMO for June, July and August.

Figs. 14a and b show the quantile-quantile plots of RAW and NOWrad-BMO, and Stage III data for the entire record (a) and for summer months only (b). As in Charlotte-Mecklenburg, Stage III exhibits a positive conditional bias where NOWrad-RAW features a negative one. After correcting for mean field bias, the quantiles are much closer to those based on gauges, and the effect is especially pronounced for the summer. These observations suggest that 1) a considerable number of large precipitation amounts reported by Stage III are not confirmed by the gauge reports; 2) NOWrad-RAW underrepresents the large rainfall amounts; and 3) mean field bias correction to a great extent mitigates the underrepresentation.

The observations from Figs 11 and 14 paint a complex picture of the inaccuracies of the three products. In particular, it appears that the NOWrad-RAW data sets may exhibit a positive overall bias but a negative conditional bias. There are several reasons underlying this apparent contradiction. First, NOWrad-RAW contains more false alarms than Stage III. The surplus rainfall volume contributed by these false alarms overweighs the deficit due to negative conditional bias at the heavy end of the precipitation spectrum, and thus leads to an overall positive volumetric bias in the summer months. Second, there are discrepancies between the reports from the analysis gauges (i.e., HADS) and the validation LCRA gauges, namely, NOWrad-RAW was biased low in reference to the former but the overall bias was indeed biased high against the latter. Such discrepancies may arise simply due to difference in the geographic coverage: HADS data were present for the entire WGRFC whereas the LCRA is concentrated over the central Texas. As mean field bias correction applies a uniform factor irrespective of rainfall magnitude, it effectively yielded elevated bias in NOWrad-BMO for the summer.

Figs 15 and 16 shows the comparisons of spatial rainfall patterns as derived from Stage III and two NOWrad products for 0900z on 1 August 2000 when heavy rainfall was reported near the Texas-Oklahoma border. The overall spatial patterns depicted from the three sources appear quite similar. However, a few distinctions are noticeable. First, precipitation areas for most of the storm cells as portrayed by NOWrad products are visibly larger. For this hour, Stage III features a cluster of disconnected rainy areas (Fig. 15a and 16a), whereas NOWrad products shows a contiguous umbrella over the southeastern corner (Fig. 15b and 16b), which is physically more realistic. Second, Stage III appears to miss the storm system along the AZ-NM border that shows up in the NOWrad data, possibly due to missing input from the New Mexico radars (Fig. 15a and b). Third, the rainfall amounts over the convective cells (marked by orange ovals) tend to be lower in the NOWrad-RAW product than in the Stage III product (Figs. 16b and a, respectively). NOWrad-BMO features much larger rainfall amounts over these locations, apparently a result of correcting for a negative mean field bias.

Summary and Conclusions

In this study we examined the existing gaps in the operational Stage III radar QPE archive, and assessed the feasibility of filling these gaps, as well as addressing the weakness of Stage III data, using the NOWrad composite reflectivity data.

Our inventory analysis indicates that there were significant gaps in the Stage III archive for some of the River Forecast Centers. These gaps are unevenly distributed in space and time. For example, Stage III were missing 77% of hours in 1996 in MARFC whereas ABRFC featured no missing data. NOWrad also features a number of large gaps. However, since the gaps in NOWrad and Stage III do not always overlap, many of the gaps in Stage III can indeed be filled using the QPE derived from the NOWrad data.

To investigate the efficacy of the NOWrad QPE, we derived NEXRAD- and MPE-like rainfall products from the NOWrad composite reflectivity data for the entire conterminous US, and evaluated the accuracy of these products against gauge observations over North Carolina and central Texas. In deriving the radar-only product, we employed a simple regression approach to estimate the Z-R relationship for each month based on monthly precipitation totals reported by COOP stations and the frequency of occurrence in each NOWrad reflectivity category. This product was further bias-corrected using the mean-field bias module of MPE.

Our assessment of these products and Stage III data in general pointed to mixed results in terms of quality. The NOWrad radar-only product, or NOWrad-RAW, in general exhibited a negative bias and tended to underrepresent precipitation totals during heavy events. On the other hand, however, the NOWrad-RAW product was shown to be more effective in resolving light precipitation (> 0.25 mm/h and < 2 mm/h) than the Stage III product, though at the expense of a higher false alarm rate. Our comparisons of NOWrad-RAW and Stage III products over the Charlotte-Mecklenburg metropolitan area in North Carolina also revealed that the detection skill of the latter product clearly improved during 1997-2001, consistent with the previous observation of Zhang et al. (2011). In addition, the NOWrad-RAW product was found to be more closely correlated with gauge data than the Stage III data set over Charlotte-Mecklenburg, but the correlation is slightly poorer than the latter over the central Texas area.

The potential of further improving the NOWrad QPE product through blending with hourly rain gauge data was also analyzed with a focus on using the gauge data as the reference to bias-correct the NOWrad-RAW data. The bias-correction was done using the MPE mean field bias module, which calculates a spatially uniform bias factor for each hour using pairs of positive radar and gauge estimates. The application of this mean field bias to the NOWrad-RAW data for 2000 in general elevated the precipitation amounts. For some of the months (mostly in the cool season), such an increase helped mitigate the negative bias, whereas for the summer, this magnified and thereby worsened the positive overall bias. Yet, the bias correction did have a broad benefit - it increased the number of instances of heavy rainfall and rendered the NOWrad-based distribution of hourly rainfall amounts close to that from the gauge reports.

The findings from our investigations point to the fact that the NOWrad composite reflectivity data, despite a number of shortcomings, can be a viable source of high-resolution quantitative precipitation information and it indeed complements the current NWS archive in several respects. In particular, the wide availability of this data set during 1996-2000 allows it to fill the large number of gaps in Stage III archive. Moreover, as demonstrated in the study, the NOWrad-based QPE can exhibit better correlation with gauge data, as it does in the Charlotte-Mecklenburg area, and it is conspicuously better in detecting light rainfall over the earlier years. As such, we are ready to create the NOWrad-BMO QPE product on the grid mesh of HRAP for the entire CONUS. Yet, it must be realized that a number of challenges remain to be addressed before the NOWrad QPE becomes a fully reliable archive for, say, calibration of distributed hydrologic models. First, the regional variation of the performance of NOWrad versus Stage III data needs to be more thoroughly depicted. In this study, the NOWrad-RAW QPE shows better correlation over the Charlotte-Mecklenburg area, whereas its results are not as encouraging in Texas. This knowledge of regionally-dependent strengths would help determine when and where the NOWrad QPE can be used in place of Stage III. Second, the negative overall and conditional bias need to be further mitigated. The lack of representation of heavy precipitation would clearly be detrimental to the resolution of historical flood events through retrospective hydrologic model simulations. Mean-field bias correction, as done in this study, was shown to be useful, but not necessarily robust for removing conditional bias. In addition, applying a uniform bias to a large area spanning multiple radar coverage is clearly subject to uncertainty. the most important among which is the variation in bias among radars arising from differing Z-R relationship, and calibration differences. A couple of additional steps can be taken to address this. First, our method of identifying Z-R could be improved to reduce the negative conditional bias. It is possible to use objective functions other than CC that would seek Z-R relationships that yield expanded range in the precipitation values. Second, histogram-matching could be considered as an alternative to mean field bias correction. In addition, the spatially varying bias may also be further mitigated. Biases likely to differ among WSR-88Ds given the differences in radar calibration and Z-R relationship, and these differences cannot be accounted for the mean-field bias correction approach. It is possible that further gains can be realized by partitioning the area into subdomains (e.g., effective coverage of each radar) and applying the mean field bias module to each of subdomains individually. In addition, the MPE does offer a local bias correction module for mitigating spatially non-uniform bias (Seo and Breidenbach. 2002). The performance of this module, as shown in Habib et al. 2013, can be variable depending on the abundance of gauge data and may be inferior to that of mean field bias module. Nevertheless, further studies are warranted to assess the potential benefit of integrating this module, as well as the multisensor blending modules (Seo 1998), as a function of the density of hourly rain gauge reports from the HADS. In this respect, other means of bias correction may also be explored. For instance, the use of PRISM monthly gridded rainfall total to bias-correct the radar QPE on a pixel-by-pixel basis, which has been by Zhang et al. (2011) to be beneficial to water budget analysis, may be tested and compared with the MPE-based bias correction. Finally, the strengths of other archives, such as the NCEP Stage II radar-only and bias-corrected radar QPEs, need to carefully analyzed in order to design a better archive that combines the strengths of multiple data sources. These will be left to future work.

Acknowledgements

The NOWrad reflectivity data used in this study were provided by the Mesoscale Meteorology

Division of the National Center for Atmospheric Research, sponsored by the National Science Foundation. We gratefully acknowledge the assistance provided by David Ahijevych of NCAR, and information on NOWrad history provided by Carrie Gillespie and staff of WSI and The Weather Company.

References

Baeck, M.L. and Smith, J.A. (1995). Climatological analysis of manually digitized radar data for the United States east of the Rocky Mountains. *Water Resources Research* 31(12), 3033-3049.

Fulton, R. A., J. P. Breidenbach, D. J. Seo, D. A. Miller, and O'Bannon, T. (1998): The WSR-88D rainfall algorithm. *Wea. Forecasting*, 13(2) , 377–395.

Germann, U., and Zawadski, I. (2002): Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Wea. Rev.*, 130(12**)**, 2859-2873.

Habib, E., Qin, L., Seo D J, and Ciach, G (2013): Independent Assessment of Incremental Complexity in the NWS Multi-sensor Precipitation Estimator Algorithms. *J. Hydrol. Eng.*, 18:143-155.

Hardegree, S.P., Van Vactor, S.S., Levinson, D.H., and Winstral A. H.(2008): Evaluation of NEXRAD Radar Precipitation Products for Natural Resource Applications. *Rangeland Ecology and Management* 61:339-345.

Kitzmiller, D., Miller, D., Fulton, R., and Ding, F. (2013): Radar and multisensor precipitation estimation techniques in National Weather Service hydrologic operations." *J. Hydrol. Eng.*, 18(2), 133–142.

Kondragunta, C. R., and Shrestha, K. (2006): Automated real-time operational rain gauge quality-control tools in NWS hydrologic operations. Preprints, 20th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc., P2.4.

Lin, Y. and Mitchell, K.E. (2005). The NCEP Stage II/IV hourly precipitation analyses: Development and applications. 19th Conference on Hydrology, San Diego, CA, Amer. Meteor. Soc., 1.2

Marchok, T., R. Rogers, and Tuleya, R. (2007): Validation Schemes for Tropical Cyclone Quantitative Precipitation Forecasts: Evaluation of Operational Models for U.S. Landfalling Cases. *Wea. Forecasting*, 22(4), 726–746.

National Center for Atmospheric Research, (2014): US Composite Reflectivity (2km by 2km) from National Weather Service Radars. [http://locust.mmm.ucar.edu/imagearchive1/WSI/docs/GHRC_README.htm, last accessed 29 January 2014]

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and DMIP Participants, (2004):. Overall distributed model Intercomparison project results, *J. of Hydrol.*, Vol. 298(1-4), 27-60.

Seo. D., Real-time estimation of rainfall fields using radar rainfall and rain gage data, (1998) *J. Hydrol.*

208 (1-2), 23-52.

Seo D., J. P. Breidenbach, and Johnson, E.R.(1999): Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol*., 223(3-4), 131–147.

Seo D., and Breidenbach, J. P. (2002): Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor*., 3(2), 93–111.

Seo D., A. Seed, and Delrieu, G. (2010): Radar and multisensor rainfall estimation for hydrologic applications. Rainfall: State of the Science, *Geophys. Monogr*., Vol. 191, Amer. Geophys. Union, 79-104.

Smith, J.A., and Krajewski, W.F., (1991): Estimation of mean field bias of radar rainfall estimates. *J. Appl. Meteorol.* 30(4), 397–411.

Smith, M. B., and Coauthors, (2012): Results of the DMIP 2 Oklahoma experiments. *J. Hydrology* 418-419(8), 17-48.

Weather Services International, (1995): *Data Services Catalog*. WSI, Billerica MA.

Zhang, Y., T. Adams, and Bonta, J. V. (2007): Subpixel-scale rainfall variability and the effects on separation of radar and gauge rainfall errors. *J. Hydrometeor*., 8(6), 1348–1363.

Zhang Y., S. Reed, and Kitzmiller, D. (2011): Effects of retrospective gauge-based readjustment of multisensor precipitation estimates on hydrologic simulations. *J. Hydrometeor*., 12(3), 429–443.

Table 1: Percentage of Hours with Missing Hourly QPE Data in NWS Stage III Archive

	1996	1997	1998	1999	2000	2001	Mean
NERFC	73	1	0	3	6	17	17
MARFC	77	0	96	6	11	0	32
SERFC	36	0	0	0	1	12	8
OHRFC	23	0	0	0	17	17	9
LMRFC	49	2	2	4	8	5	12
NCRFC	0	0	0	3	8	9	3
MBRFC	0	0	0	8	15	14	6
ABRFC	0	0	0	0	1	21	4
WGRFC	0	0	1	3	34	13	8
NWRFC	55	2	0	10	3	17	14
CBRFC	71	10	1	5	15	2	17
CNRFC	100	100	100	100	53	13	78

Table 2: Number of hours with missing NOWrad reflectivity data

Month\Year	1996	1997	1998	1999	2000	2001
1				170		312
2				59		168
3						
4						24
5						
6						24
7						528
8					696	
9					720	24
10					720	0
11					720	24
12					120	0

Table 3: Number of hours with missing Stage III QPE data for WGRFC

Month\Year	1996	1997	1998	1999	2000	2001
1	4	27	8	744	4	744
2	3	4	7	672	223	672
3	4	8	3	6	744	744
4	8	15	5	0	10	720
5	14	5	7	63	11	115
6	13	19	1	2	11	4
7	10	12	16	3	1	7
8	7	1	6	2	0	2
9	19	9	2	5	1	703
10	10	6	1	38	0	1
11	4	0	4	3	57	5
12	19	31	3	13	7	1

Table 3: Number of hours with missing Stage III QPE data for SERFC

Month\Year	1996	1997	1998	1999	2000	2001
1	507			4		0
2	528		1			0
3	546		1			
4	621					720
5	648		2			
6	309					
7						290
8					744	0
9					0	0
10					0	0
11					0	0
12		2		18	17	0