5.4 RETROSPECTIVE CASE STUDY OF THE IMPACT OF RAIN GAGE NETWORK REDUCTIONS ON NATIONAL WEATHER SERVICE RIVER FORECASTS IN THE SUSQUEHANNA RIVER BASIN

Peter Ahnert*, Kevin Hlywiak, Cody Moser, and Seann Reed
NOAA/NWS/Middle Atlantic River Forecast Center, State College PA

1.0 INTRODUCTION

One of the principal model forcings for operational river forecasts at the National Weather Service (NWS) Middle Atlantic River Forecast Center (MARFC) is 6-hour basin average precipitation. Precipitation is used by the lumped Continuous Antecedent Precipitation Index (ContAPI) rainfall-runoff model (Sitner et al. 1969; NWS 2013; Moser 2013) and the lumped energy balance snow model (SNOW17; Anderson 1973). High quality precipitation estimates are needed to produce accurate river flood forecasts.

The Susquehanna River Basin (SRB) river and rain gage networks were significantly enhanced during the period 1986 to 2011 through a multi-year program known as the Susquehanna Flood Forecast and Warning System (SFFWS). The 1985 SFFWS improvement plan (SRBC 1985) included $853K for 87 new and 20 upgraded rain gage reporting stations to “minimize the uncertainty in estimating runoff from large ungaged areas”. The system had an initial goal of increasing lead times by 6 to 10 hours and thereby reducing average annual flood damages by 15% or $12.4M per year (1985 dollars). Annual federal funding for the program between 1986 and 2011 fluctuated from a high of approximately $3.0M in 1987 to a low of approximately $700K in 1996. As of 2011, the SFFWS provided all or a portion of the funding for 73 high quality winter capable hourly rain gages. In 2011, a new draft SFFWS Strategic Plan for 2011-2016 recommended an additional 22 real-time hourly rain gages to ‘…fill gaps in coverage, and to support new or expanded forecast locations”, but none of these were installed (MHW Inc. 2010).

Beginning in FY2012, dedicated federal funding for the SFFWS was not included in the President’s budget or a congressional appropriation. Federal, state, and local government agencies, working closely with the Susquehanna River Basin Commission (SRBC), were able to arrange for continued funding of most of the stream gages and a few of the rain gages in the network by allocating funds from other sources. However, funding for 47 United States Geological Survey (USGS) Digital Collection Platform (DCP) rain gages in Pennsylvania, which cost from four to seven thousand dollars per year each to operate and maintain, could not be secured, and they had to be shut down in early 2013. On October 1, 2013 maintenance funding was cut for an additional 16 USGS DCP rain gages in the New York portion of the SRB and all 16 were shut down by the beginning of December. All 63 of these gages were high quality winter capable gages that reported rainfall at least hourly via satellite (Fig. 1). The remaining rain gage network used in MARFC operations as of December 1, 2013 included 76 winter capable gages that report hourly (USGS DCPs (7), Army Corps of Engineers DCPs (17), NWS/FAA ASOS gages (18), NWS Fischer & Porter gages (34)) as well as 112 non-winter capable gages that report hourly (AFWS gages (95), assorted DCPs (17)). The loss of the 63 gages represented a loss of 45% of the winter capable/hourly gages and loss of 25% of all hourly gages in the basin (Fig. 2). Other gages used are daily rainfall reports from NWS Cooperative Observer gages (75), and selected gages from the web based Community Collaborative Rain, Hail, and Snow (CoCoRaHS) network (61) (Fig. 2). These 136 stations report daily total rainfall and are available once a day at around 7AM local time.

This paper describes preliminary work done to assess the potential impact the loss of these 63 rain gages in Pennsylvania and New York could have on the accuracy and lead time of NWS flood forecasts in the SRB. First, the paper looks at the impact of gage reductions on gage network density in the basin and on calculations of Mean Areal Precipitation (MAP) for the sub-basins used for MARFC modeling and forecasting. To further examine the impacts these gage network cuts may have on operational river forecasts, retrospective hydrologic model simulations were made with and without the removed gages for several significant flood events that occurred prior to the loss of these gages.

*Corresponding author address: Peter R. Ahnert, NWS Middle Atlantic River Forecast Center, 328 Innovation Blvd. (STE 330), State College, PA 16803-6609; email; peter.ahnert@noaa.gov

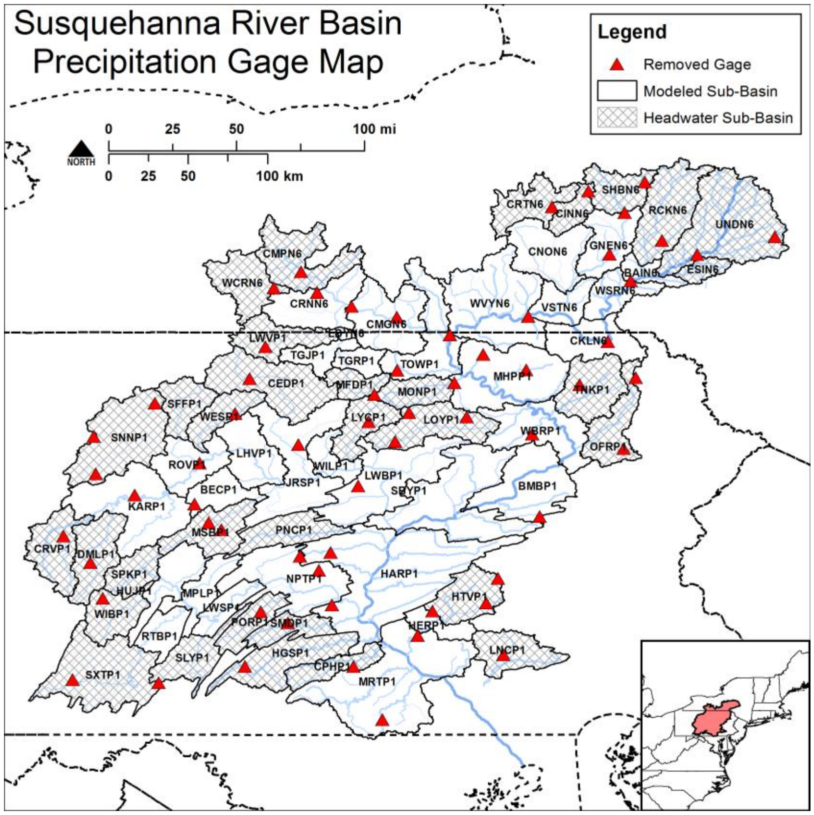

Figure 1. Basin/sub-basin map showing locations of the 63 discontinued USGS rain gages.

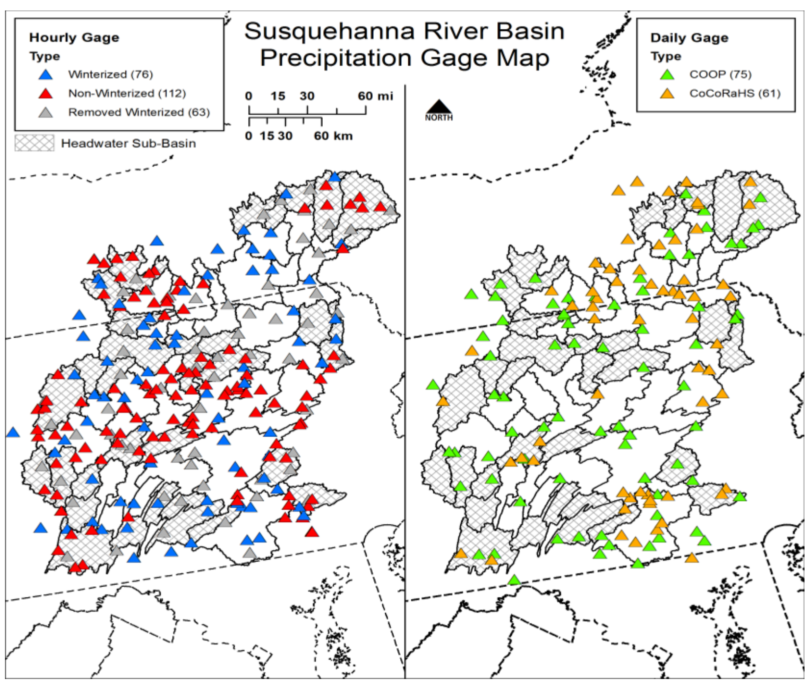

Figure 2. Location and type of MARFC (a)hourly rain gages and (b)daily rain gages.

1.1 Flood Cases

All three flood cases chosen for this study were recent floods that had a very significant impact within the Susquehanna basin. The first flood case chosen was caused by the remnants of Tropical Storm Lee in early September of 2011. Lee progressed slowly northward along the Appalachian Mountains and interacted with a quasi-stationary east-west frontal boundary producing extremely heavy rainfall across central Pennsylvania and New York. Storm total precipitation of 250 mm to 375 mm (10-15 inches) fell across portions of the SRB during this multi-day event resulting in record flooding at 12 forecast points, major flooding at 13 forecast points, moderate flooding at 13 forecast points, and minor flooding at 7 forecast points . Most flood crests occurred on 8 and 9 September. MARFC unofficially ranks this as the fourth largest flood in the Middle Atlantic region (MARFC 2013).

The second flood case chosen was the remnants of Hurricane Ivan in mid-September of 2004. MARFC unofficially ranks this as the sixth largest flood in the Middle Atlantic region (MARFC 2013). As Hurricane Ivan slammed into the Alabama/Mississippi coast, a cold front with its origins in Canada began to sweep across the Gulf Coast. As the dying tropical system became absorbed in the cold front on 17 September and formed a new low along the boundary, rainfall associated with the system spread from Maine to North Carolina with the heaviest rain falling over Pennsylvania. These heavy rains along with wet antecedent conditions caused by several preceding tropical storms resulted in a widespread flood disaster. Between 17 and 18 September, the New York portion of the SRB received 50 to 100 mm (2 to 4 inches) of rain while the Pennsylvania portion of the SRB received 75 to 175 mm (3 to 7 inches). Within the SRB, this event resulted in major flooding at 25 forecast points, moderate flooding at 15 forecast points, and minor flooding at 9 forecast points. Most flood crests occurred on 18 and 19 September (MARFC 2013).

The third case chosen was the late June 2006 flood. MARFC unofficially ranks this as the fifth largest flood in the Middle Atlantic region (MARFC 2013). Between 23 and 30 June, meridional flow over the U.S. provided the Northeast with very humid air. Meanwhile, a closed low over Hudson Bay pushed cooler air into the Mid-Atlantic, allowing a stationary front to form over the East coast of the U.S. bringing steady and locally very heavy rain into the Mid-Atlantic region between 23 and 27 June. Disturbances that progressed into the Mid-Atlantic from the upper low positioned in Canada brought even more rain to the region over 27 and 29 June, resulting in severe flooding. Since the majority of the rain accumulated due to lines of thunderstorms which filled the region, there were significant disparities in daily rainfall totals (even for gage locations usually deemed “nearby”). From 23 to 29 June, total precipitation amounts in many areas of the SRB ranged from 150 to 250 mm (6-10 inches). Within the SRB, this event resulted in major flooding at 16 points, moderate flooding at 13 points, and minor flooding at 3 points. Most flood crests occurred on 28 and 29 June (MARFC 2013).

2.0 REVIEW OF GAGE DENSITY REQUIREMENTS

The minimum acceptable rain gage network density depends on application and regional storm type characteristics (Cheng et al. 2008), as well as basin characteristics (shape, size, terrain, soils). Therefore, there is no definitive minimum rain gage density requirement for NWS river flood forecasting. The core NWS river forecast and warning operations are currently based primarily on lumped hydrologic models running on a 6-hour hour time-step. One of the principal inputs to the lumped model is 6-hour Mean Areal Precipitation (MAP) averaged over each sub-basin in the model. MARFC computes MAP values using the Thiessen polygon method (Thiessen 1911). During the warm season, MARFC incorporates 6-hour mean areal precipitation values (MAPX) that are based on Multi-sensor Precipitation Estimates (MPE) which combine radar rainfall estimates and rain gage observations. In the first phase of this study, gage-only analyses (MAP) were used to assess the impacts of gage reduction. Results from the first phase will be used to determine whether there will be additional value in analyzing multi-sensor radar-gage scenarios with and without the lost rainfall gages.

While trying to identify basins in the U.S. with existing high quality data networks for the International Model Parameter Estimation (MOPEX) experiment, Schaake et al. (2000) looked at the number of precipitation gages required for river forecasting applications. They used a practical estimate of gage density requirements developed by Schaake (1981) which was based on a study of observations from a very dense gage network (45 km^2 per gage). According to Schaake (1981), when the data time step is one-fourth of the basin lag time, the required number of gages N for a basin of area A (sq km), is

$$N = 0.6\ A^{0.3} \qquad (1)$$

According to Schaake (1981), the number of gages calculated using this equation should give MAP estimates for each time step that are accurate to within 20 percent 80 percent of the time during thunderstorm rainfall events (in the 20,000 km^2 Muskingum, OH river basin). The equation is reasonable to apply to basins between 200 and 20,000 km^2. For basins smaller than 200 km^2, he estimates a minimum of about three gages are needed for adequate noise filtering. Only two Susquehanna sub-basins are smaller than 250 km^2. For potential high quality MOPEX networks, Schaake et al. (2000) also included basins meeting a relaxed density requirement, which was half of the density (N/2) given by equation 1. Looking at National Climatic Data Center (NCDC) published daily climatological precipitation gages, Schaake et al. (2000) found that of 1,861 potential study basins examined in the U.S., only 16% met this relaxed rain gage density requirement.

Using the average size of MARFC modeled sub-basins in the SRB (1046 km^2), equation 1 states that 4.8 gages are required per sub-basin (1 gage per 218 km^2). The relaxed density requirement results in 2.4 gages per sub-basin (1 gage per 436 km^2). This suggests a good overall target gage density for SRB sub-basins is an average of between 2.4 and 4.8 gages per sub-basin.

Equation 1 is also conditioned on the rain gage data time interval being one-fourth of the basin lag time. Many of MARFC's headwater sub-basins have a lag time between 12 and 18 hours, which suggests the need for a rain gage time step of no more than 3 to 4.5 hours. The hourly gages satisfy this temporal requirement, while the daily gages do not.

Other studies (Eagleson 1967; Cheng et al. 2008; Awadallah 2012; Seo et al. 2013) have analyzed rain gage network designs to meet certain sets of requirements and have studied the relationship between rain gage density and the accuracy of basin average precipitation estimates for various spatial and temporal scales. Doviak (1983) compared some of the early studies by Huff (1970) in Illinois, Woodly et al. (1975) in Florida, and Hildebrand et al. (1979) in Montana, that looked at the accuracy of rain gage measurements for calculating average rainfall. Doviak found that for areas with appreciable air mass thunderstorms, a gage density of 1 gage per 143 km^2 should be adequate to obtain 6-hour rain gage estimates accurate to within 30% while also admitting that this spacing may be too dense for an economical measurement of rain over large areas.

Sharp et al. (1961) used 10 years of data from 39 rain gages in the 259 km^2 (100 mi^2) Sandstone Creek watershed in Oklahoma to study rainfall accuracy versus gage density. They found relatively small differences in mean storm rainfall over the watershed when using 39, 19, 13, or 10 gages. Differences rose sharply when five or fewer gages were used. In a later study in Oklahoma, Nicks (1965) found that a network of 10 gages adequately described the mean daily rainfall over the entire 2930 km^2 (1130 mi^2) Washita River basin study area (1 gage per 293 km^2). However, in contrast to these results, when examining the mean daily rainfall on small tributary catchments of 155 km^2 (60 mi^2) and 539 km^2 (208 mi^2), Nicks found that a gage density greater than 1 gage per 37 km^2 was needed to obtain adequate rainfall estimates.

In a case study in Bangalore, India, the Indian Space Research Organization studied the effect of rain gage density on the accuracy of rainfall (Mishra 2013). Looking at a 250 km^2 area, they found small errors when four to seven rain gages (>1 gage per 63 km^2) were used to represent the area. Reduction to three or fewer gages resulted in significant (>40%) error.

Eagleson (1967) developed a figure showing the number of gages needed for flood forecasting as a function of the desired percent error in the forecast of maximum discharge, catchment length and width, and effective storm radius. Applying Eagleson's figure to a forecast sub-basin size typical for the Susquehanna region (1046 km^2), and using an effective storm radius for convective storms of 16 km (10 miles), gives a minimum of 1.65 rain gages needed to be able to estimate catchment discharge within 10% (assuming precipitation estimation error is the only error source). This gives a minimum density of about 1 rain gage per 620 km^2 (242 mi^2). Eagleson's results indicate a density requirement somewhat less than equation 1.

Lebel et al. (1987) performed an experimental case study validation of the normalized error variance as a function of gage density using data from a very high density network located in the Gardon d'Anduze watershed in France. Their results show that for a catchment of 545 km^2, an hourly time step, and a scaled estimation variance of 5%, a gage density of 1 gage per 230 km^2 is needed. An extrapolation of their results indicates that for a 10% scaled estimation variance, a gage density of 1 gage per 300 km^2 is needed, which falls between the optimal and relaxed density values computed using equation 1.

The Sharp et al. (1961), Nicks (1965), and Mishra (2013) studies reveal the fact that when gages are used alone, the required gage density increases significantly for very small catchment sizes (≤250 km^2). Only two of MARFCs sub-basins are very small, Lindley on the Tioga River (LDYN6) which is 66 km^2 and Blanchard on Bald Eagle Creek (BCHP1) which is 183 km^2. Neither of these are modeled headwater basins. The smallest headwater sub-basin is East Sydney Dam on Ouleout Creek which is 267 km^2. Most MARFC sub-basins are between 500 km^2 and 1200 km^2.

For larger basins, Nicks (1965) results (1 gage per 293 km^2), Lebel et al. (1987) results (1 gage per 300 km^2), and Eagleson (1967) results (1 gage per 620 km^2) are in fair agreement with the relaxed requirement of equation 1 (N/2). It should be noted that equation 1 will yield a somewhat higher optimal-relaxed density range for smaller than average sub-basins in the Susquehanna and a somewhat lower optimal-relaxed density range for larger than average sub-basins.

In a recent study done by Seo et al. (2013) on the objective reduction of the SFFWS rain gage network via geostatistical analysis of uncertainty, it was demonstrated that keeping only the 20 most important gages could reduce the deterioration in precipitation data quality over the entire Susquehanna basin (not individual sub-basins) by 50% relative to removing all 73 funded gages. Seo et al. (2013) was completed before it was learned that funding cuts would result in the elimination of all but 10 of the SFFWS funded rain gages. This study did not recommend a minimum acceptable gage density, but it did identify the 20 most important rain gages for assessing the precipitation over the entire basin. Only two of the 20 rain gages deemed most important in Seo et al. (2013) are still operational.

In February 2012, MARFC also ranked the SFFWS gages based on the Seo et al. (2013) study, cost factors, gage quality, and forecaster experience. MARFC identified 20 gages that could be eliminated and 53 gages that it recommended be kept as part of the network. Unfortunately, only seven of the 53 are still operational as of December 2013.

Since many of the reviewed studies support the gage density thresholds derived by equation 1, they will be used here to represent the optimum gage density. Following Schaake et al. (2000), N/2 will be used to represent a relaxed gage density that should still provide good MAP estimates most of the time.

When and where good radar data for precipitation estimation are available, the needed gage density likely decreases. While some research has been done on the gage density required to produce accurate MPE products, we could not discern clear, quantitative guidance from our initial literature review. The number of rain gages required under a given radar umbrella to produce accurate rainfall estimates is a complex problem which was beyond the scope of this study.

3.0 GAGE DENSITY ANALYSES

The point density geoprocessing tool within the Esri ArcMap Spatial Analyst toolbox was used to calculate gridded gage density and the effective gage density for each sub-basin before and after the removal of the 63 rain gages. The tool calculates a magnitude per unit area from point features that fall within a neighborhood around each cell. For neighborhood shape and size, a search radius (18.2 km) resulting in a circle with an area (1,046 km^2) equal to the average size of our modeled Susquehanna River sub-basins was used.

To compute gage density, we first overlaid a grid with 4 km^2 grid spacing over the basin. For each grid cell, we searched the area within the circle of radius 18.2 km around each cell and counted the number of gages falling within the search radius (Fig. 3). That became the integer gage density assigned to the cell. We then averaged the integer gage density values for all cells whose centers were within the sub-basin to obtain the effective gage density for the sub-basin. For the Monroeton sub-basin example in Fig. 3, even though there is only one gage physically located in the basin, there are four additional gages nearby, resulting in an effective gage density of 2.87. As seen in Fig. 3, the advantage of this method is that nearby gages that are within a reasonable distance outside of the sub-basin boundaries are considered in the effective gage density calculation for the sub-basin.

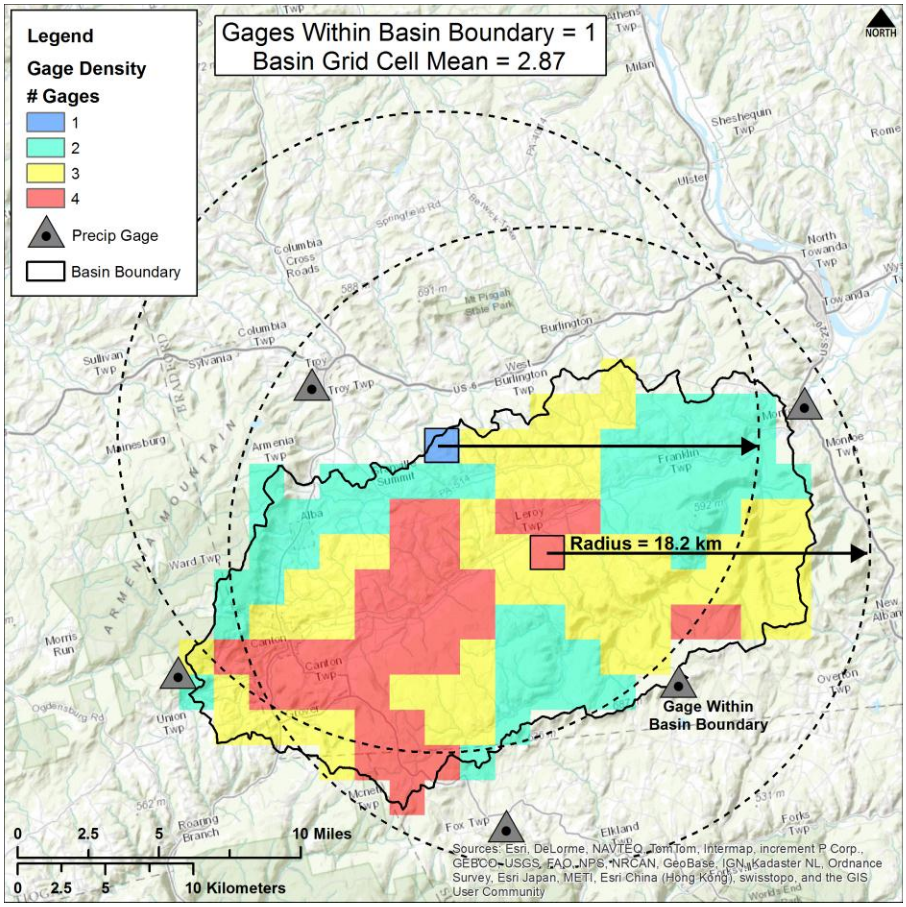

Figure 3. Illustration of how gage density was calculated for each sub-basin by first calculating the number of gages within a certain radius of influence of each grid cell in the basin and then averaging the gage density for all grid cells in the basin. Towanda Creek at Monroeton (MONP1) sub-basin is shown.

Including gages outside each sub-basin results in a better representation of the effective gage density, but deviates from the density calculations used to derive equation 1. The resulting higher effective gage densities will increase the number of sub-basins that meet the optimal and relaxed density thresholds.

Using equation 1 and the average SRB sub-basin size of 1046 km^2, a gage density of approximately 2.5 gages or more meets the N/2 threshold . For comparison, we calculated the gage density within the SRB for four scenarios during the Lee flood; (1) 0600 UTC with the 63 gages, (2) 0600 UTC without the 63 gages, (3) 1200 UTC with the 63 gages, and (4) 1200 UTC without the 63 gages (Figs. 4 and 5). Fig. 4a and 4b represent the density of available gage data with and without the 63 discontinued gages at the 0600 UTC forecast preparation time during the Lee flood. As explained below in Section 5.0, the 0600 UTC forecast preparation time is important because it is the time at which the impacts from the loss of the 63 high quality hourly gages will likely be greatest.

At 0600 UTC during Lee, results show the expected significant increase in the areas with low gage density (density of less than 2.5 gages per 1046 km^2 when comparing the gage density plot with and without the 63 discontinued gages (Fig. 4a vs 4b). Of even greater concern is the increase in the areas with very poor gage coverage in Fig. 4b where the gage density falls to zero gages per 1046 km^2. In these areas, the rainfall estimates may be of poor quality even when good radar data are available to merge with the rain gage data. The areas with zero gages per 1046 km^2 shown in Fig. 4b are areas where the SFFWS should consider additional hourly gages whenever funding permits. The areas with five or more gages per 1046 km^2 shown in Fig. 4b are areas that future studies should examine for gages that might be relocated to areas with poor coverage.

Headwater sub-basins are identified in Fig. 4. In general, headwater basin flow modeling is more sensitive to rainfall/snowfall measurement uncertainty because there is no upstream routed flow component and stream response times are shorter. Maintaining a gage density of at least 2.5 gages per 1046 km^2 in headwaters should be a priority. Many of the headwater basins show grid cell density reductions of 1-2 gages over most of the basin. Fig. 4b shows that without the 63 gages, the point gage density over significant portions of 17 headwater basins falls into the 0-2 gage range. As a result, the MAP estimates for these sub-basins may be significantly degraded.

The available gage density improves for forecasts issued during the 1200 UTC forecast cycle because in addition to hourly data, the data from up to 136 daily reporting rain gages arrives at MARFC. MARFC time distributes the daily value from each gage into hourly and 6-hourly amounts using the time distribution of our hourly MPE. Fig. 5 shows the gridded gage density plot for Lee at 1200 UTC including all available daily and hourly rainfall reports before and after removal of the 63 discontinued rain gages. Of great concern is the increase in areas with no gage coverage in Fig. 5b where MAP rainfall estimates may be of poor quality. These are areas where the SFFWS could use additional daily or hourly gages.

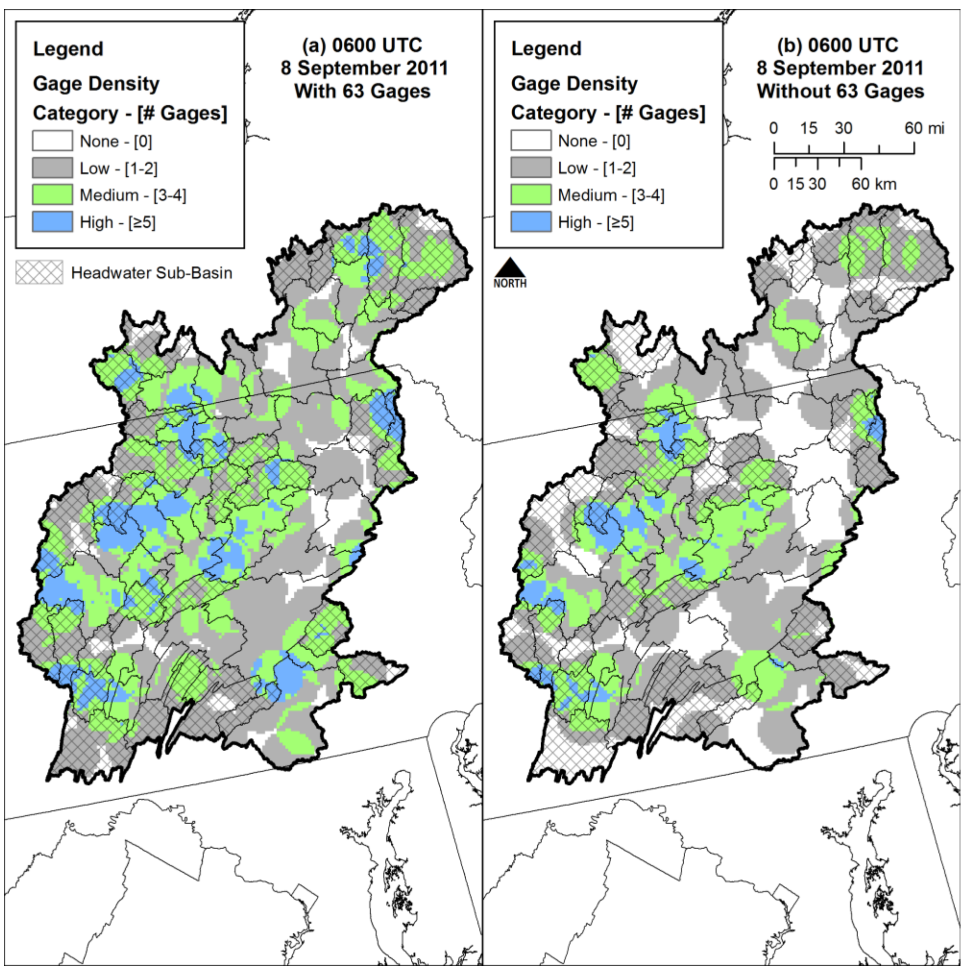

Figure 4. SRB gridded gage density plot for Lee at 0600 UTC; (a) with all available hourly gages and (b) with the 63 discontinued gages removed.

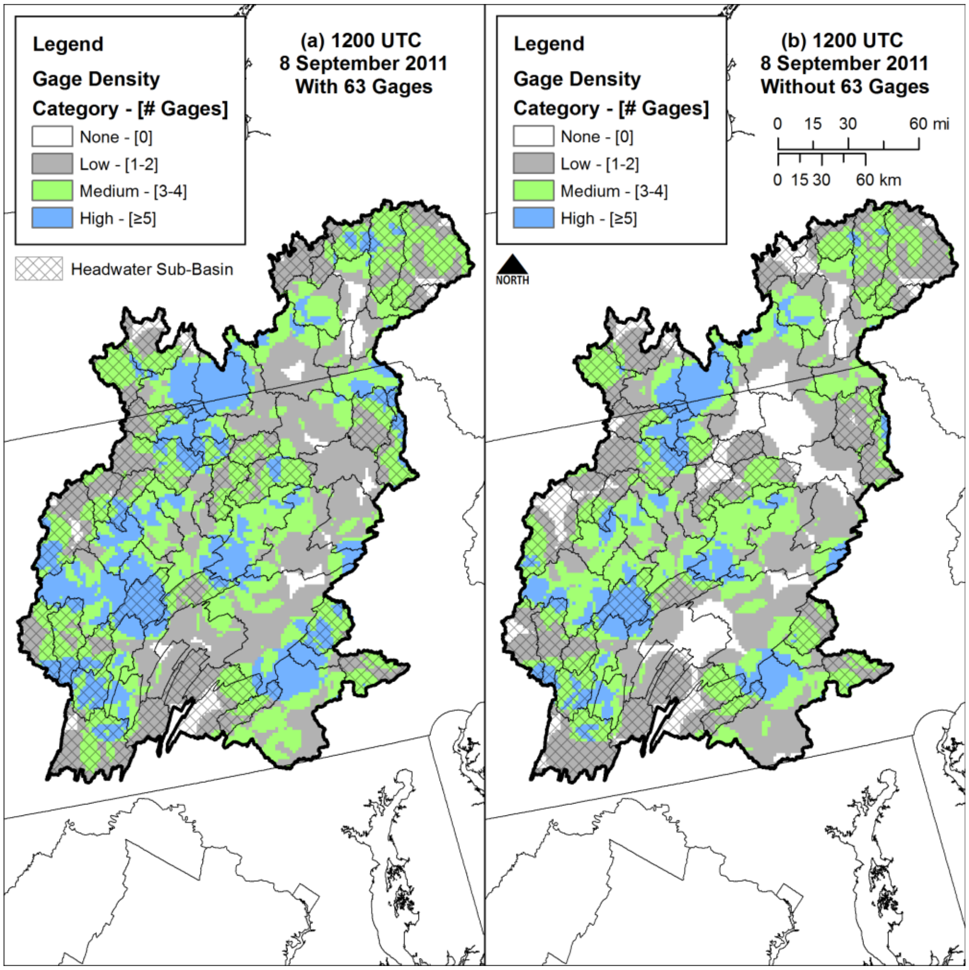

Figure 5. SRB gridded gage density plot for Lee at 1200 UTC; (a) all daily and hourly gages included and (b) with the 63 discontinued hourly gages removed.

For the four scenarios listed above, we also calculated the effective gage density versus the sub-basin area for all 66 modeled Susquehanna sub-basins and compared to N and N/2 from equation 1 (Fig. 6). The percent of sub-basins with poor (<N/2) gage densities increases from 14% to 37% when excluding the 63 gages for the 1200 UTC forecast cycle and from 31% to 71% when excluding the 63 gages for the 0600 UTC forecast cycle. Since the gage time step needed in order to satisfy equation 1 on headwater sub-basins is 3 to 4.5 hours, the results for 1200 UTC assume that the 24-hour gage data were accurately disaggregated down to hourly and 6-hourly time intervals using MPE data as per normal MARFC procedures.

The results for the Ivan flood (not shown) were found to be similar with the percent of forecast sub-basins with poor gage densities (not meeting the relaxed Schaake density (N/2) recommendations) increasing from 34% to 63% for the 0600 UTC case when excluding the 63 gages.

The density analysis results indicate that gage densities will likely be inadequate for reliable high quality gage-only MAP calculations in more than 50% of the modeled SRB sub-basins. Especially for fast responding headwater basins, these significantly lower rain gage densities will likely increase the uncertainty in the forecast and degrade the warning lead times and accuracy during some flood events.

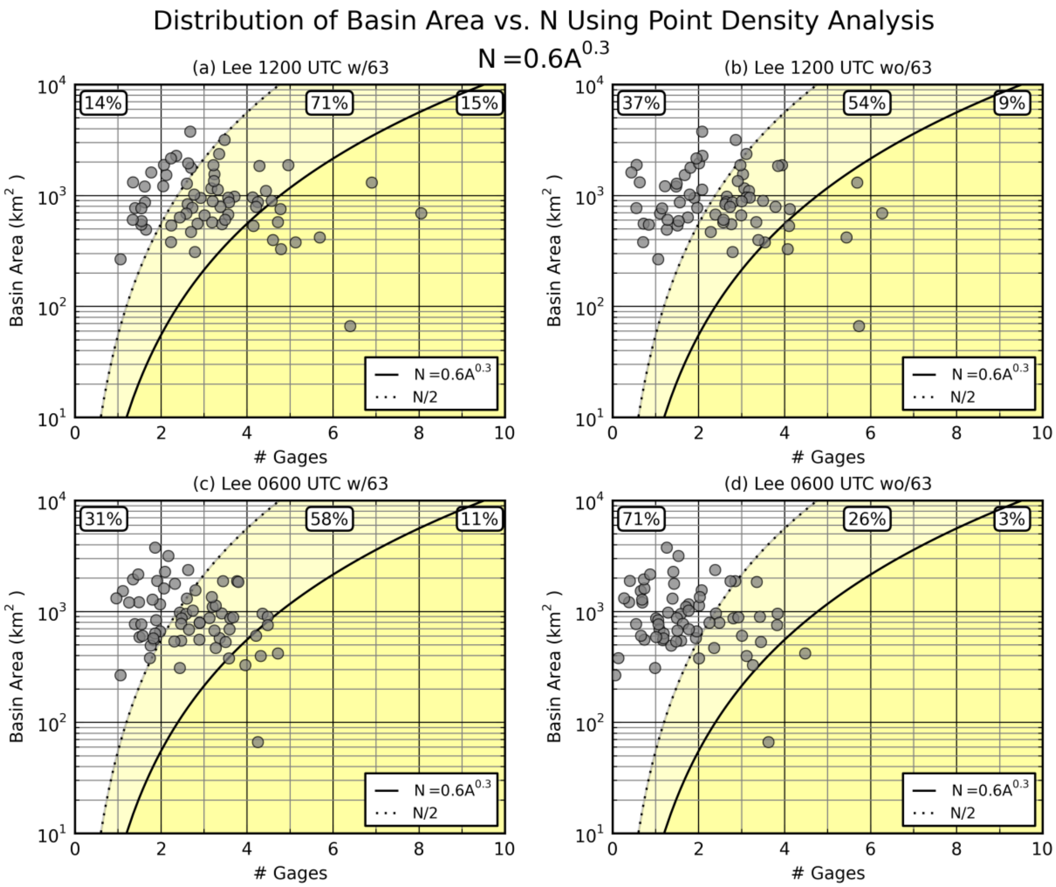

Figure 6. The effective gage density (# gages) versus sub-basin area is plotted for each of the 66 sub-basins in the SRB for (a) 1200 UTC forecast cycle (includes daily and hourly gages), (b) 1200 UTC forecast cycle without the 63 gages, (c) 0600 UTC forecast cycle (only hourly gages), and (d) 0600 UTC forecast cycle without the 63 gages. Plot is divided into 3 areas using Schaake et al. (2000) thresholds (N) and (N/2). Percent of sub-basins falling within each of these areas is shown.

4.0 MEAN AREAL PRECIPITATION ANALYSES

4.1 Lee Flood (September 2011)

The first flood case is the severe flooding from the remnants of Tropical Storm Lee. We calculated the storm total rainfall and the change in calculated storm total MAP values with and without the 63 rain gages at the 0600 UTC forecast preparation time on 8 September 2011 (Fig. 7). Removing the 63 rain gages changes the storm total MAP by anywhere between -44 mm and +32 mm. There were 10 out of 66 sub-basins (15%) where the absolute value of the MAP change was greater than 20 mm and 1 sub-basin where the change in MAP was greater than 40 mm. Comparing Fig. 7b to Fig. 4b, 13/19 sub-basins with MAP changes greater than 10mm have at least part of the basin with very poor gage coverage (0) after removal of the 63 gages. This shows that gage density reductions associated with the removal of the 63 gages can cause significant differences in sub-basin MAP values during major flood events.

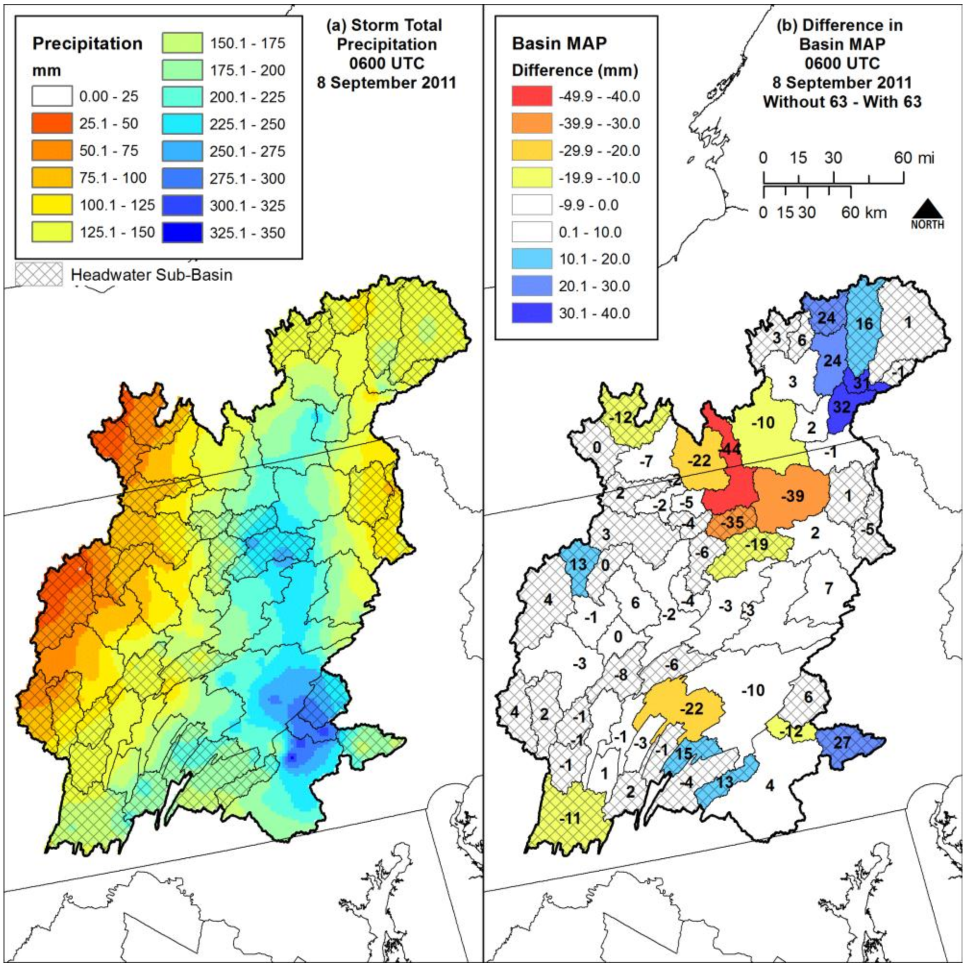

Figure 7. (a) Storm total precipitation up to 0600 UTC 8 September 2011 for Tropical Storm Lee and (b) change in calculated sub-basin MAP after removing 63 gages.

4.2 Ivan Flood (September 2004)

The second case of severe flooding was the Ivan Flood in September 2004. The storm total precipitation up to 0600 UTC on 18 September 2004 and the change in calculated storm total MAP value after the 63 gages were removed was calculated (Fig. 8). While in Lee the heaviest rainfall was across the eastern half of the SRB, in Ivan the heaviest rainfall totals were in the central and south-western portions of the SRB (Fig. 8a). The change in MAP values after removal of the 63 gages ranged from -25 mm to +14 mm (Fig. 8b). Three of the 66 sub-basins (5%) had changes in MAP greater than 20 mm. While six of the 10 sub-basins with MAP changes greater than 10mm are the same as during Lee, the sign of the MAP change is opposite in two of these sub-basins.

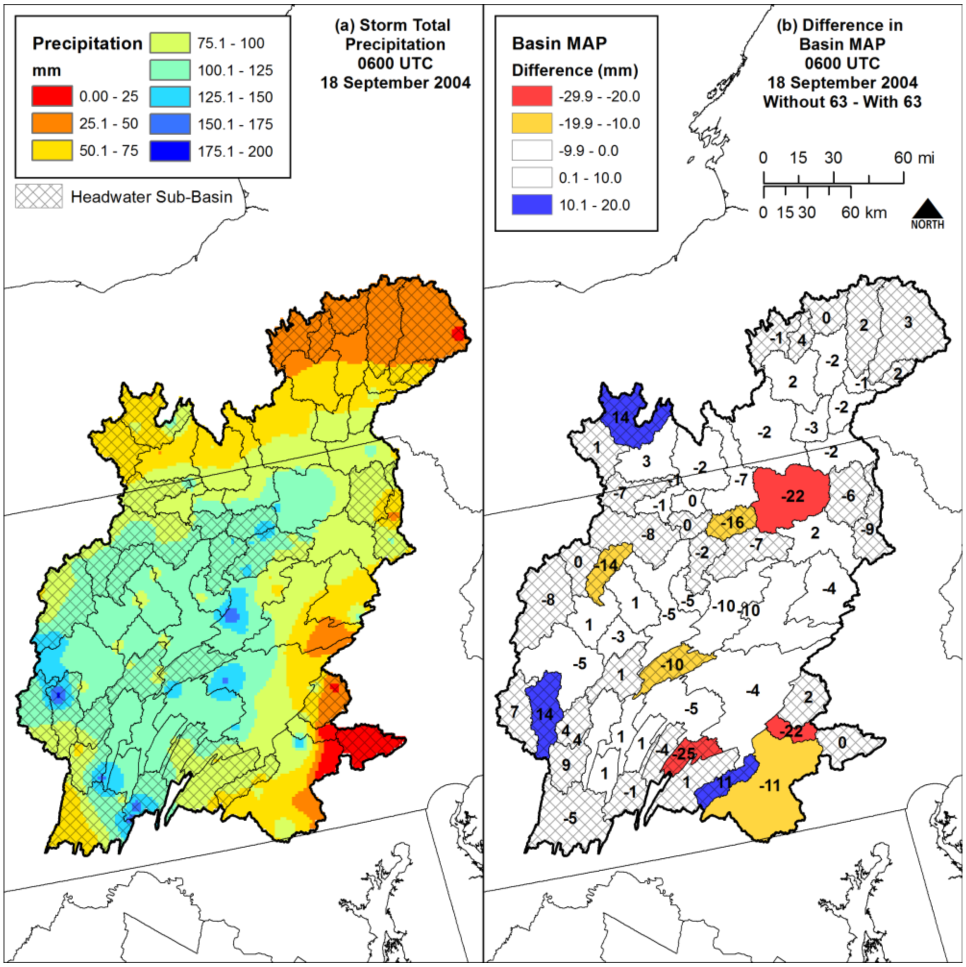

Figure 8. (a) Storm total precipitation up to 0600 UTC 18 September 2011 for Tropical Storm Ivan and (b) change in calculated sub-basin MAP after removing 63 gages.

4.3 June 2006 Flood

The third case of severe flooding was the June 2006 flood. Heavy rain causing river flooding fell from lines of convective storms falling over a several day period. The heaviest rain fell across the eastern third of the SRB where amounts ranged from 100 mm to 300 mm (Fig. 9a). The synoptic summary of the event mentions that gage rainfall measurements contained some big differences over short distances (MARFC 2013). These differences could also lead to some big MAP changes with the removal of the 63 gages.

The MAP difference plot shows MAP values changed from -26 mm to +19 mm upon removal of the 63 gages (Fig. 9b). Absolute changes in calculated MAP upon removing the 63 gages were greater than 20 mm in only one sub-basin. Of the 12 sub-basins with an absolute MAP change greater than 10 mm, four are different from sub-basins with greater than a 10 mm change in Lee or Ivan.

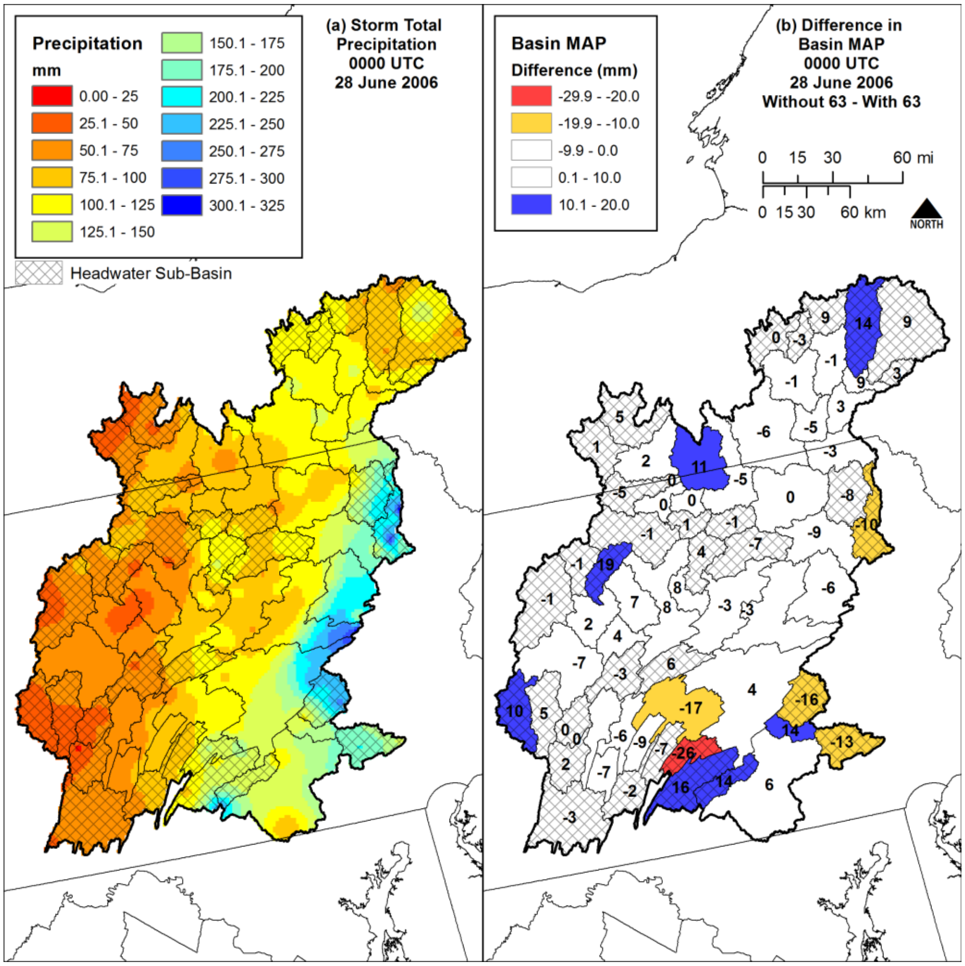

Figure 9. (a) Storm total precipitation up to 0600 UTC 28 June 2006 and (b) change in calculated sub-basin MAP after removing 63 gages.

Comparing the sub-basin change maps for all three events (Fig. 7b/8b/9b) shows that even though the same 63 gages were removed, the impact on the computed MAP can be significantly different from one event to another. For example, MAP changes at Campbell on the Cohocton River (the U-shaped basin at the top left of the maps were -12mm, +14mm, and +5mm, respectively for the three events. Furthermore, the sub-basins with changes in MAP greater than 10 mm varied with each event. The largest storm total MAP change resulting from the removal of the 63 gages was -44 mm in the Towanda sub-basin (TOWP1) basin during Lee, -25 mm in the Shermans Dale (SMDP1) basin during Ivan, and -26 mm in the Shermans Dale (SMDP1) basin during June 2006.

In summary, results show that while certain basins appear to be more susceptible to MAP changes due to the rain gage reductions, the impact on each sub-basin varies from event to event depending on the distribution of rainfall in and near the sub-basin. Based on the three events studied, most MAP values are not significantly impacted during an event. However, a few are, which will lead to an increase in MAP uncertainty during high impact events. If more cases are added to the MAP change analysis, statistics can be calculated to help identify and prioritize the basins in most need of additional gages.

5.0 STREAMFLOW ANALYSES

Retrospective hydrologic simulations for each of the three flood cases were run twice, (1) using all available gages; and (2) after removing the 63 discontinued gages, to illustrate the potential impacts of the gage reductions. The focus was on headwater forecast points where flows were not affected by an upstream (i.e., routed) flow component. We were most interested in looking at the potential impacts on the flood forecast and warning program and constrained our analyses to those headwater points that flooded during each event.

Our goal was to make the hydrologic re-simulations as representative as possible of simulations that were made operationally during the event. We used MARFC's operational hydrologic modeling system (Continuous API rainfall-runoff model). All three selected flood event cases occurred during the warm season, therefore, the SNOW-17 model and associated temperature data were not a factor. MARFC archives a copy of its model states (parameters and variables) every day, so the model parameters were initialized with the same values they had prior to the beginning of each flood event. To avoid any potential human bias or manipulation of the results, we did deviate from real-time operations by excluding any manual forecast modifications to the model parameters and variables. We also excluded radar rainfall estimates, which can be beneficial when used properly. Therefore, these 'raw' simulations are likely not as accurate as those MARFC would have produced operationally. Since we were most interested in the relative performance of the model simulations with and without the 63 gages, this decrease in accuracy should not impact the conclusions.

To simulate operational conditions at a particular forecast start time (1200 UTC, 1800 UTC, 0000 UTC, or 0600 UTC) during each of the three events, all available hourly and daily gage data were used in the 6-hourly MAP calculations at the previous 1200 UTC. Then, for the 1200-1800 UTC, 1800-0000 UTC, and 0000-0600 UTC time periods, only hourly data were used in the MAP calculations. As described earlier, more rainfall data are available at 1200 UTC because of the addition of the time distributed daily gage reports. At 1800 UTC, we have all of the daily data from six hours ago plus all of the hourly data (from only the hourly gages) for the 6-hour period from 1200-1800 UTC. At 0000 UTC, we have all of the daily data up to 12 hours ago plus hourly gage data for the past two 6-hour periods, 1200-1800 UTC and 1800-0000 UTC. At 0600 UTC, we have all of the daily data up to 18 hours ago and only the hourly data for the past three 6-hour periods, 1200-1800 UTC, 1800-0000 UTC, and 0000-0600 UTC. When the next 1200 UTC forecast cycle is reached, we again gain the benefit of up to 136 daily gage reports which can be time distributed back over the previous four 6-hourly periods. Therefore, in most cases, the risk for impacts from the recent gage reductions will be highest during the 0600 UTC forecast cycle and second highest during the 0000 UTC forecast cycle. The least risk from the gage reductions would be expected during the 1200 UTC forecast cycle.

A forecast time zero was chosen for each event based on the following criteria: (a) the majority of the heavy rain had fallen, (b) at least 6-12 hours prior to crests at most headwater points, and (c) preference was given to 0600 UTC or 0000 UTC. Based on these criteria the forecast time zero chosen was 0600 UTC 8 September 2011 for Lee, 0600 UTC 18 September 2004 for Ivan, and 0000Z UTC 28 June 2006 for the June 2006 event.

Each of the simulations also had to incorporate any future precipitation after time zero of the model simulation. During real-time RFC operations, a quantitative precipitation forecast (QPF) would be used. We eliminated additional uncertainty due to QPF in this study to isolate the effects of changes in the observed rain gage network. To do this we used observed 6-hourly MAP values computed using all gages as a surrogate for 'perfect QPF' for each of the simulations.

We then compared model simulated flows with and without the 63 gages to observed flows in an effort to evaluate the potential impacts these gage reductions may have on MARFC streamflow forecast accuracy during future floods. For the purposes of this study, crest flow simulations with error +/-20% or less were considered good, +/- 21-40% was a moderate over-forecast or under-forecast, and greater than +/-40% error was a severe over-forecast or under-forecast.

5.1 Lee Flood (September 2011)

Crest (flow) forecasts were compared to observed crests for each of the headwater forecast points that reached flood stage (Fig. 10). A +/-20% crest flow error was considered good considering the many potential error sources in a hydrologic simulation. During Lee, there was a slight tendency to over-forecast the flood crests. Below 1000 cubic meters per second (cms), removing the gages tended to increase the flow forecasts even more, leading to severe over-forecasting (>50%) at Rockdale (RCKN6) and Camp Hill (CPHP1). The forecast at Shermans Dale (SMDP1) also degraded to a moderate over-forecast when gages were removed. There was one case of severe under-forecasting (-45%) at Harpers Tavern (HTVP1-not labeled), but the gage cuts had no impact. Lancaster (LNCP1) and Sherburne (SHBN6), which were moderately under-forecast when the gages were included, improved significantly after removal of the gages. The root mean square error (RMSE) in the crest flow simulation for all headwater points reaching flood stage with the 63 gages included was 243 cms. When the 63 gages were removed, the RMSE increased to 264 cms (9% increase), indicating an overall decrease in headwater flood forecasting accuracy.

The forecast crest flows with and without 63 gages differ by more than 10% at seven locations. Including the 63 gages resulted in a better crest forecast at five of the seven locations. We further analyzed the event by plotting the hydrographs at several of the locations.

Lee resulted in record flooding at Loyalsockville on Loyalsock Cree, (LOYP1), where data go back to 1925. For LOYP1, both simulated hydrographs are one period (six hours) late with the crest forecast, but the simulation with the 63 gages resulted in a nearly perfect crest forecast, while removing the 63 gages resulted in a slightly lower (17%), but still good forecast crest flow (Fig. 11). This agrees with Fig. 7 which showed a 19 mm (0.75 inches) reduction in the storm total MAP in this sub-basin with the removal of the 63 gages. The 8% reduction in storm total MAP resulted in a 17% reduction in simulated crest flow. Both simulated crests and the observed crest were in the major flood category.

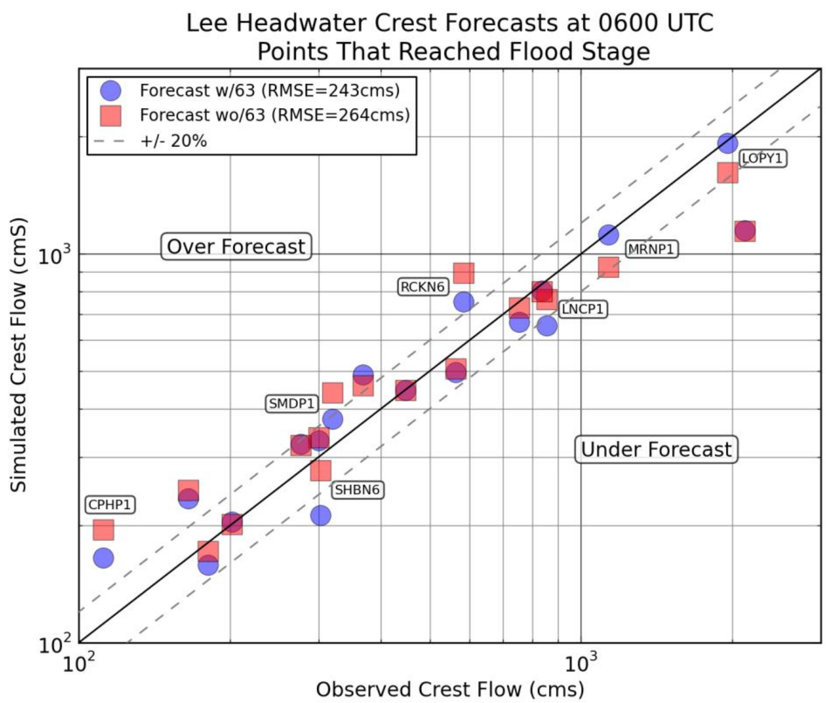

Figure 10. Simulated versus observed crest flows with the 63 gages (blue circles) and without the 63 gages (red squares) during Lee for the retrospective simulation at 0600 UTC 8 September 2011. Headwater forecast points that reached flood stage are shown. Basins with the largest crest flow simulation differences are labeled.

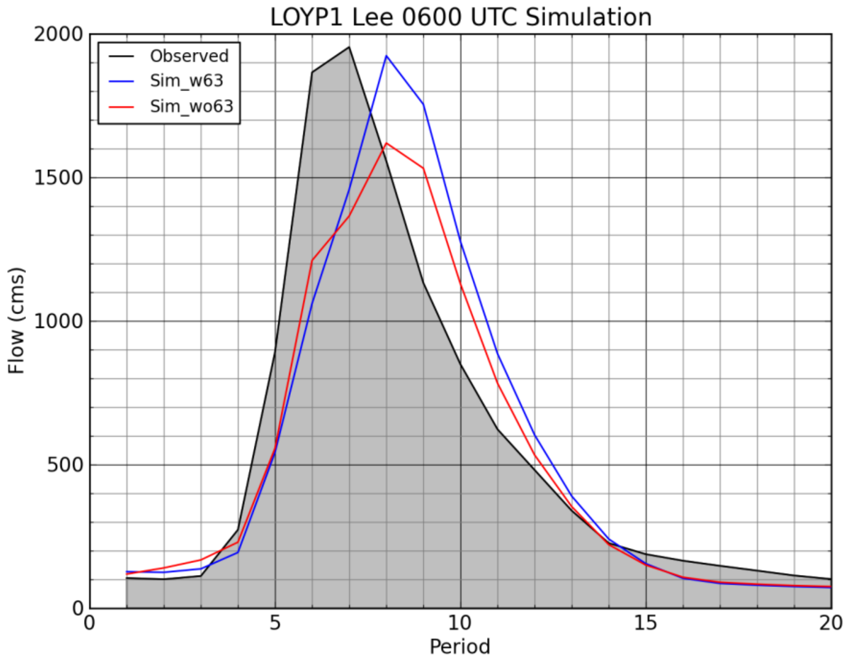

Figure 11. Observed (black line) and retrospectively simulated (blue line=with 63 gages, red line=without 63 gages) streamflow hydrographs for Loyalsockville, PA on Loyalsock Creek for the Lee flood. X-axis time zero is 0600 UTC 8 September 2011 and time intervals are 6-hour periods.

Monroeton on Towanda creek (MRNP1 or MONP1). also had record flooding (data go back to 1914). The storm total MAP change upon removal of the 63 gages was -35 mm (1.4 inches), a change of -16%. The hydrograph (not shown) shows results very similar to LOYP1. The simulation with 63 gages was very close to the observed crest flow and the simulation without the 63 gages was slightly under-forecast (18%), but still considered accurate. Both simulations and the observed crests were in the major flood category.

The second highest flood crest of record occurred during Lee at Lancaster on the Conestoga River (LNCP1), where records go back to 1928, . LNCP1 storm total MAP up to 0600 UTC increased by 27 mm (1.1 inches) or 13% when the 63 gages were removed from the network. The simulated crest is moderately under-forecast (31%) with the 63 gages (Fig. 12). Removing the 63 gages results in a higher MAP, higher simulated flows, and a slightly under-forecast (14%). The observed crest and both simulated crests were in the major flood category.

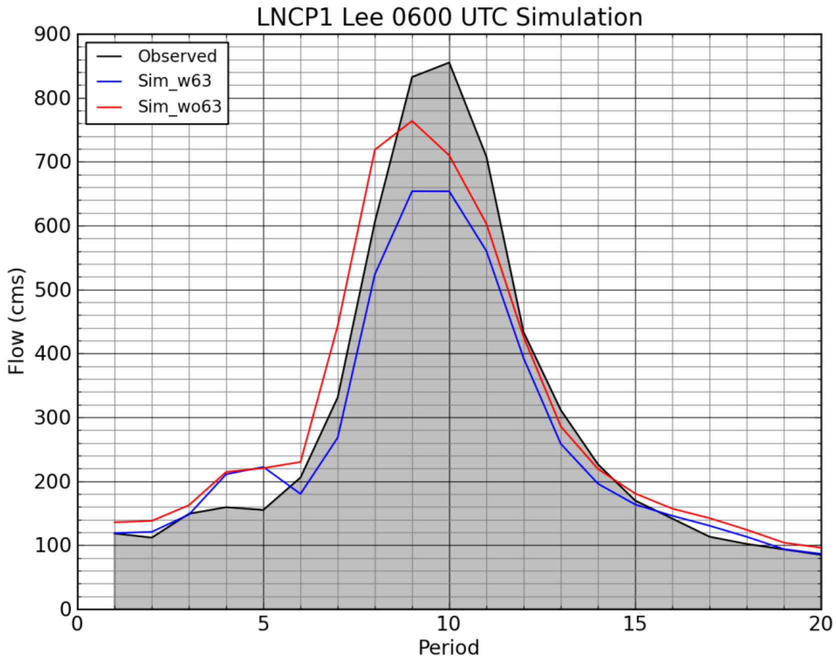

Figure 12. Observed (black line) and retrospectively simulated (blue line=with 63 gages, red line=without 63 gages) streamflow hydrographs for Lancaster, PA on the Conestoga River. X-axis time zero is 0600 UTC 8 September 2011 and time intervals are 6-hour periods.

Lee caused record flooding at Rockdale on the Unadilla River (RCKN6), where records go back to 1929. As of 0600 UTC, the storm total MAP was 16 mm (0.63 inches) or just 9% higher for the Rockdale sub-basin after removing the 63 gages. Crest simulations were one period (six hours) late, and both simulations over-forecast the crest (Fig. 13). The simulation with the 63 gages resulted in a moderate crest over-forecast of 29%, and simulation without the gages resulted in a severe crest over-forecast of 53%. This is a case where a seemingly small 9% increase in the storm total MAP at RCKN6 was amplified by basin and model characteristics into a 24% increase in the crest flow over-forecast. Both simulated crests and the observed crest were in the major flood category.

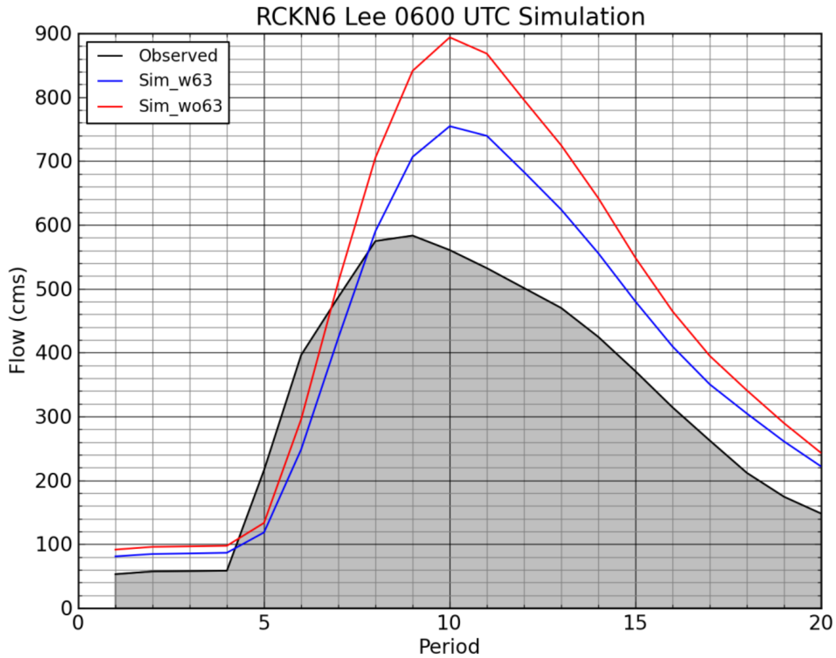

Figure 13. Observed (black line) and retrospectively simulated (blue line=with 63 gages, red line=without 63 gages) streamflow hydrographs for Rockdale, NY on the Unadilla River. X-axis time zero is 0600 UTC 8 September 2011 and time intervals are 6-hour periods.

The results from the final three points with greater than a 10% change in crest flow during this event were also mixed. West of the RCKN6 sub-basin in New York is the Sherburne sub-basin on the Chenango River (SHBN6), Lee caused record flooding at SHNB6, where records go back to 1938. Storm total MAP calculations showed a 24 mm (17%) increase in calculated MAP after removal of the 63 gages. The simulated crest with the 63 gages was moderately under-forecast (29%), while the forecast without the gages was considered good (-7.5%). The simulation with the 63 gages crested in the moderate flood category, while the simulated crest without the 63 gages and the observed crest were both in the major flood category.

Shermans Dale (SMDP1), a small basin west of Harrisburg PA, had a double flood crest with the first crest exceeding the moderate flood threshold. This was not an unusual flood for SMDP1 (25^{th} highest flood crest on record since 1929). The simulation with the 63 gages was considered good (+18%). The storm total MAP increased by 15 mm (8%) when the 63 gages were removed. The increased MAP resulted in a moderate over-forecast of 38% (Fig. 14). Both simulated crests and the observed crest were in the moderate flood category. SMDP1 is also of interest because of the second and third peaks. The small rise between periods 10 and 13 was over-forecast primarily because the simulated flows at the beginning of the rise were too high. During actual RFC operations, a forecaster would have likely made a manual adjustment to the model simulation just prior to this rise, leading to a much improved forecast.

The simulation at SMDP1 also shows an artifact of our methodology. Since we use actual observed rainfall in place of QPF beyond time zero, the two curves gradually converge after period 10 since both simulations use the same precipitation forecast input. The third rise between periods 15 and 18 illustrates that while the previous two crests were over-forecast by the model, the next crest may not necessarily also be over-forecast. This model behavior presents many possibilities including; (1) the highest rainfall in the basin may not have been captured by the gages, and/or (2) the heavy rainfall fell in in less than 6 hours, and/or (3) the highest rainfall fell right near the sub-basin outlet. Alternatively, the first two rain events may have resulted in observed antecedent soil

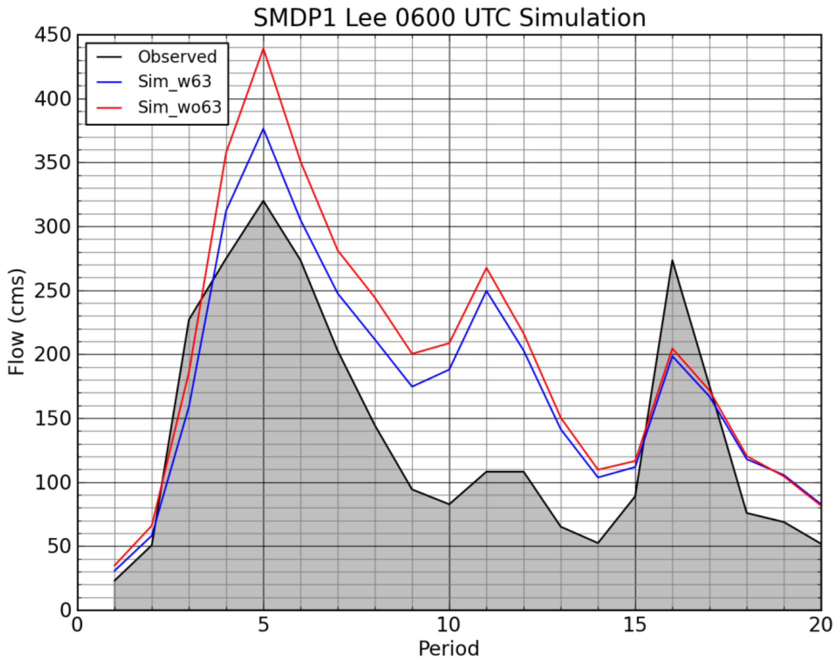

Figure 14. Observed (black line) and retrospectively simulated (blue line=with 63 gages, red line=without 63 gages) streamflow hydrographs for Shermans Dale, PA on Sherman Creek. X-axis time zero is 0600 UTC 8 September 2011 and time intervals are 6-hour periods.

moisture considerably wetter than what was being simulated. All of these possibilities relate to complexities forecasters take into account when preparing forecasts.

The last headwater forecast point we examined was the moderate flood event at Camp Hill on the Yellow Breeches Creek (CPHP1) south of Harrisburg, PA. Lee caused the 10^{th} highest crest at Camp Hill since records began in 1909. The storm total MAP increased by 13 mm (7%) when the 63 rain gages were removed. Both simulations severely over-forecasted the crest. With the 63 gages, the over-forecast was 48% and without the gages the over-forecast was significantly worse (75%). A 7% increase in the storm total MAP resulted in a 27% increase in the crest flow over-forecast. Despite the significant severe over-forecast, both of the simulated crests and the observed crest fell within the moderate flood category.

Analysis of the hydrographs from Lee clearly show that large differences in the simulated crest flow forecast can occur as a result of the removal of the 63 rain gages from the gaging network. The change in crest flow caused by the change in the gaging network was more than 10% at 7 of 18 (39%) headwater forecast points that flooded. In five of seven (71%) cases where the change was more than 10%, the simulated crest forecast was significantly worse when the 63 gages were removed. For this event, the change in the gaging network had an impact on MAP (rain gage-only) based crest flow forecasts at numerous locations. In some cases the antecedent conditions and modeled basin characteristics will lead to amplification of the MAP error into a higher percent crest flow error. The increased uncertainty in the rainfall amounts and resulting increased uncertainty in the flow forecasts may make warning decisions more difficult, resulting in lost lead time, more false alarms, and more missed warnings. The level of impact may depend on the quality of the radar rainfall estimates during the event. If radar rainfall estimates are accurate enough to be combined with the rain gage data into high quality MPE, the impacts on flood warnings may be minimal.

On the other hand, in those areas and during those periods of time where radar rainfall estimates are of poor quality and cannot be used, degradation in warning lead time and accuracy is much more likely than it was before the removal of these 63 gages from the network.

5.2 Ivan Flood (September 2004)

Retrospective hydrologic simulations with the 63 gages and without the 63 gages were made for headwater points that flooded during Ivan (Fig. 15). The forecast time zero was 0600 UTC 18 September 2004. During Ivan, most of the simulations (8 of 13) under-forecast the observed crest flows by more than 20% when the 63 gages were included. After removal of the 63 gages, many of the simulations (7 of 13) still under-forecast the crest flows by more than 20%. For the three sub-basins where removal of the 63 gages resulted in more than a 10% change in the crest flow simulation, the results are again mixed. The simulation with the 63 gages was 16% better at Shermans Dale (SMDP1), 15% worse at Williamsburg (WIBP1), and 11% worse at Spruce Creek (SPKP1). For the five highest flow cases, the results were also mixed. The removal of the 63 gages resulted in an increase in the under-forecast error in two sub-basins (SAXP1-Saxton , SNNP1-Sinnemahoning), decreased under-forecast error in one sub-basin (HUNP1-Huntingdon), and little change in two sub-basins (LOYP1-Loyalsockville, SLYP1-Shirleysburg). The RMSE in the crest flow simulation for all headwater points reaching flood stage with the 63 gages included was 225 cms. When the 63 gages were removed the RMSE was 238 cms (6% increase), indicating an overall decrease in headwater flood forecasting accuracy without the gages. This is similar to the Lee results where the increase in RMSE was 9%. Results indicate that while the impacts of the gage network reductions on individual sub-basins vary from event to event, there will likely be a few sub-basins significantly impacted during a major flood event, and the overall impact will likely be decreased forecast accuracy.

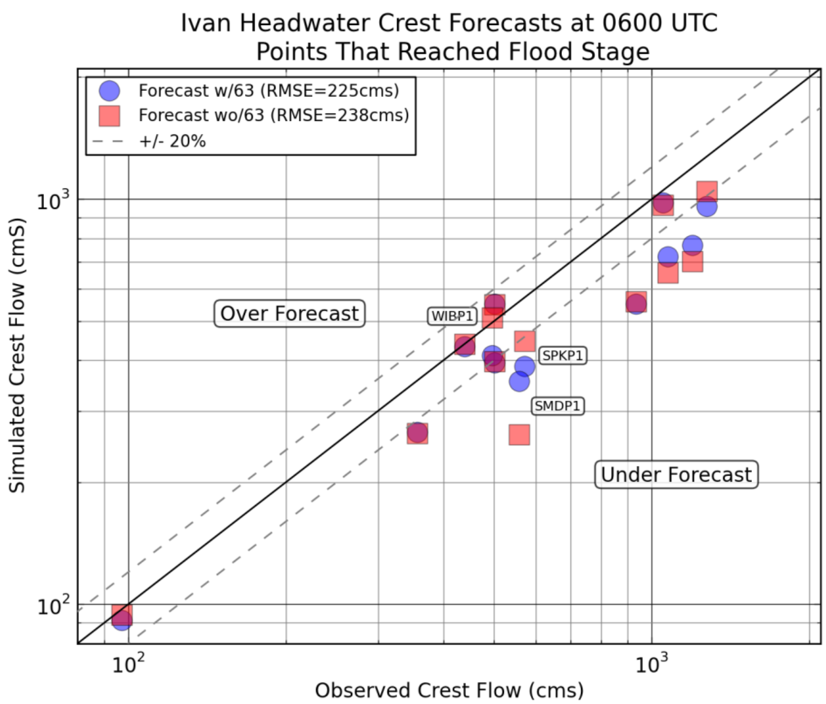

Figure 15. Simulated versus observed crest flows with the 63 gages (blue circles) and without the 63 gages (red squares) during Ivan for the restrospective simulation at 0600 UTC on 18 September 2004. Headwater forecast points that reached flood stage are shown. Basins with the largest crest flow simulation differences are labeled

5.3 June 2006 Flood

The chosen forecast time zero was 0000 UTC 28 June 2006 for the June 2006 flood event0000 UTC. 0000 UTC was chosen instead of 0600 UTC, because most of the rain had fallen as of 0000 UTC. As in Ivan, many of the simulations (6 of 12) under-forecast the observed crest flows by more than 20% when the 63 gages were included (Fig. 16). After removal of the 63 gages, two additional sub-basin crest flows were under-forecast by more than 20%. Removal of the 63 gages resulted in more than a 10% change in the crest flow simulation at four locations (HTVP1-Harper Tavern, RCKN6-Rockdale, SMDP1-Shemans Dale, and CPHP1-Camp Hill). For two of these cases (SMDP1 and CPHP1) the simulation was significantly worse (29% and 31%, respectively) when the gages were removed. At HTVP1, the simulation was slightly worse (15%) when gages were removed, and at RCKN6 the simulation was slightly better (14%) when the gages were removed. For the locations with the two highest flows, the removal of the 63 gages resulted in approximately a 7% increase in the moderate under-forecast at Tunkhannock (TNKP1) and an 8% decrease in the moderate under-forecast at Unadilla (UNDN6). The RMSE in the crest flow simulations for the 12 headwater points that flooded was 169 cms with the 63 gages and 171 cms without the 63 gages. At the four locations where the simulated flow changed by more than 10%, three had better crest flow forecasts when the 63 gages were included. In two of these cases, the crest flow error increased by approximately 30% when the 63 gages were excluded.

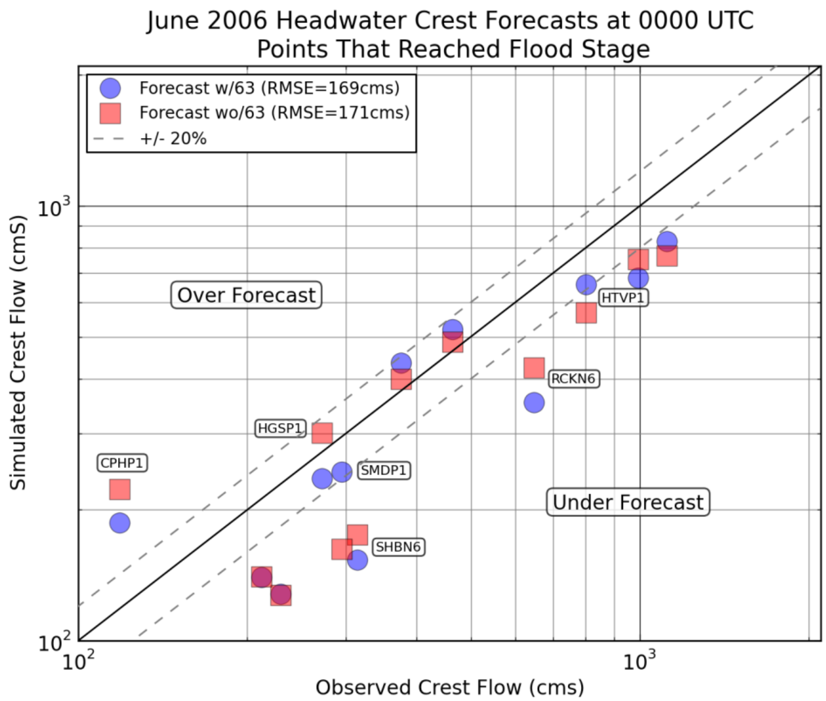

Figure 16. Simulated versus observed crest flows with the 63 gages (blue circles) and without the 63 gages (red squares) during June 2006 flood event for the restrospective simulation at 0000 UTC 28 June 2006. Headwater forecast points that reached flood stage are shown. Basins with the largest crest flow simulation differences are labeled

5.4 Discussion and Summary of Streamflow Analyses

While the RMSE for the three events show that the simulations with the 63 gages were more accurate, the mixed results when looking at individual sub-basins was unexpected. A likely explanation could be the following. Assume a well calibrated model. With perfect MAP values, the model might be expected to over-simulate the observed crest about 50% of the time and under-simulate about 50% of the time. Let's assume for a moment the model is over-simulating. Now, let's degrade the rainfall estimate for sub-basins by randomly removing some of the gages around each sub-basin. In some sub-basins, the gages that were removed will be the ones with the higher rainfall amounts. In other sub-basins, the nearby gages that are removed will be the ones with lower rainfall amounts. When we substitute gages that are farther away in the sub-basin MAP calculation, we will increase the MAP about 50% of the time and decrease the MAP about 50% of the time. If the original simulation was over-forecast, the change in MAP will often improve the forecast if the MAP decreases (about 50% of the time) and degrade the forecast if the MAP increases (about 50% of the time). The opposite holds true if the original simulation as under-forecast. Hence, mixed results are likely for most of the sub-basins.

A third case exists when the original simulation for a sub-basin with accurate MAP values results in an accurate simulation of the observed crest. In this case, more accurate MAP values result in a more accurate streamflow forecast, and degrading the rain gage network results in less accurate MAP values. If the change in MAP from the degraded network is significant, the resulting crests will be either over-simulated or under-simulated. But in both cases, the resulting simulations will be less accurate than the original. Because of this third case, one would expect to see an overall improvement in forecasts with improved MAP input if a enough flood events were examined. The RMSE from the three flood events appear to confirm this, although the 43 sampled sub-basin crests may not provide enough samples to conclusively prove this.

Both simulations (with the 63 gages and without the 63 gages) were compared to the observed crest for all headwater points that flooded during the three flood events (Fig. 17). Results show that for the majority of crest simulations (29/43 or 67%) the removal of the 63 gages did not have a significant impact on the crest flow forecast. When the 63 gages were removed the crest flow forecasts were worse 21% of the time and better only 12% of the time. The ratio of better to worse forecasts was four to one for the most significant crest changes (those >20%).

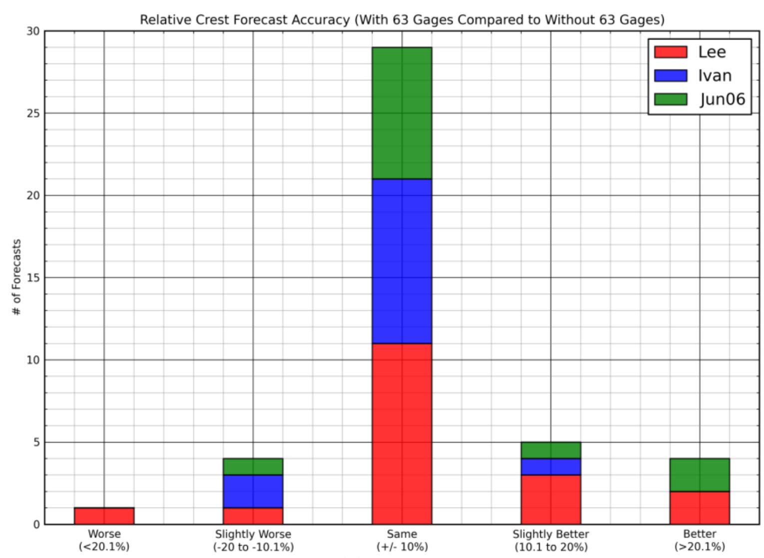

Figure 17. Summary of results from three case studies at headwater locations that reached flood stage. The percent error in the simulated crest flow with the 63 gages included in the network is compared to the percent error without the 63 gages. Better category - crest flow forecasts were more than 20% better with the 63 gages. Worse category - crest flow forecasts were more than 20% worse with the 63 gages included.

6.0 CONCLUSIONS

The potential impacts from recent reductions in the SRB hourly rain gage network were examined from three perspectives: (1) the impacts on sub-basin gage densities, (2) MAP calculations, and (3) flood crest simulations. A better understanding was gained about how the reduced gage network will impact hydrologic forecast and warnings within the SRB.

A GIS based gage density analysis method was introduced which includes gages within a reasonable distance outside of a basin boundary and uses geospatial point density averaging. The technique is adjustable to account for various sub-basin sizes or an appropriate gage radius of influence and is a valuable tool to evaluate potential/actual gage network changes (reductions or additions). To examine potential/actual gage network reductions, MAP analyses for recent heavy rainfall events were useful for evaluating the impact on rainfall estimates. Finally, while it required the most effort, the retrospective simulations and flood crest analyses were valuable in quantifying the potential impacts on hydrologic forecasts and should be an essential part of any rain gage network reduction analysis when the necessary data and tools are available.

The gage density analyses showed that removal of the 63 hourly gages increased areas where the gage density is less than the recommended density for a high quality rain gage network. The resulting gage density will likely be inadequate for high quality gage-only MAP calculations in more the 50% of the Susquehanna sub-basins. The density maps and graphs in this paper will help determine where to add additional gages or relocate gages when funding permits. Additional gages for headwater basins with the poorest coverage should be a high priority.

The impacts from the loss of gages on river forecasts may be more significant in the cold season (1 November to 31 March), due to the fewer number of winter-capable gages and the reduced quality of radar precipitation estimates. However, the impacts on the Flash Flood warning program may be more significant during the warm season, since heavy rainfall from small/slow moving convective storms may be missed when there are fewer hourly gages. Radar-based rainfall products will help compensate for the loss of gages but the quality of MPE products is also degraded with a reduced rain gage network. We do not currently have precise information about gage density requirements for Multi-sensor Precipitation Estimates (MPE) over a wide range of hydrometeorological conditions.

The MAP analyses for the three events show how variable the impact from the removal of the 63 gages is from one event to another and from one sub-basin to another. The impact varies significantly depending on the distribution of rainfall over the gages in and around each sub-basin. During widespread heavy rainfall/flood events, there will likely be at least a few sub-basins with large storm total MAP changes (> 20mm) caused by the elimination of the 63 gages. This will lead to an increase in MAP uncertainty during high impact events.

The retrospective streamflow analyses of 43 headwater flood crests shows that for the majority (67%) of the sub-basin crests, the change in the SRB gage network changed the forecast crest flow by 10% or less. This affirms that the SFFWS will likely continue to provide the same high quality forecast and warnings, even without these gages, most of the time.

Due to hydrologic uncertainties in modeling and observation errors, a small number of simulations actually improved with the loss of gages; however, on average, simulated floods were degraded in headwater basins without the 63 gages. In 9% of the simulated flood forecasts, the crest flow accuracy was degraded by more than 20%. Hydrologic simulations did highlight basins which are more sensitive to the rain gage network changes. These sensitivities are consistent with our network density analysis.

During any widespread flood event, there will likely be a few sub-basins where the degradation in the MAP estimates and crest flow simulations caused by the elimination of the 63 gages notably increases the river forecast error and/or uncertainty. Much of the time, the impacts will likely be inconsequential. However, once in a while, the increased forecast error and/or uncertainty will make it more difficult to make river flood warning and evacuation decisions. In these instances, the resulting lost warning lead time, reduced warning accuracy, and/or delay in critical evacuation decisions could lead to increased flood damages and increased risk of lives lost. These conclusions must be tempered by the fact that three events is a limited sample size. Nevertheless, the sample did consist of 43 headwater flood crest cases.

7.0 FUTURE WORK

The study confirms the value of gage density, MAP, and retrospective streamflow simulations for evaluating the impacts from this gage network change. Future work may include additional flood cases to increase the sample size.

Future work is recommended to determine if and how much further gage density requirements can be relaxed when good quality radar data are available for use in precipitation estimation. Retrospective streamflow simulations using MPE computed with and without the 63 gages may help address this question.

Results and methods from this study will be useful in planning future precipitation gaging changes in the SRB that may be necessary due to funding fluctuations.

8.0 ACKNOWLEDGEMENTS

The authors wish to thank Caroline Woods of the Library Reference staff at the NOAA Library in Silver Spring for her assistance with the literature review for this paper. We would also like to thank John Schaake for providing reference materials. Finally, we'd like to thank Brian Miretzky, NWS Eastern Region Scientific Service Division for his thoughtful review and comments.

9.0 REFERENCES

Anderson, E., 1973: National Weather Service river forecast system - snow accumulation and ablation model. Silver Spring MD. NOAA/NWS.

Awadallah, A. G., 2012: Selecting optimum locations of rainfall stations using kriging and entropy. International J. of Civil & Environmental Engineering, **12**, 36-41.

Cheng, K. S., Y. Lin, J. Liou, 2008: Rain-gauge network evaluation and augmentation using geostatistics. Hydrological Processes, **22**, 2554-2564.

Doviak, R. J., 1983: A survey of radar rain measurment techniques. J. Appl. Meteor., **22**, 832-849.

Eagleson, P., 1967: Optimum density of rainfall networks. Water Resources Research, **3**, 1021-1033.

Hildebrand, P. N., M. Towery, M.R. Snell, 1979: Measurement of convective mean rainfall over small areas using high-density raingages and radar. J. Appl. Meteor., **18**, 1316-1326.

Huff, F., 1970: Sampling errors in measurements of mean precipitation. J. Appl Meteor., **9**, 35-44.

Lebel, T., G. Bastin, C. Obled, J.D. Creutin, 1987: On the accuracy of areal rainfall estimation: a case study. Water Resources Research, **23**, No. 11, pp 2123-2134.

MARFC, 2013: On-line river flood climatology. http://www.erh.noaa.gov/marfc/Rivers/FloodClimo/Rankings/AllFloodsRank.php. State College: Middle Atlantic River Forecast Center.

MHW, Inc., 2010: Susquehanna flood forecast and warning system improvements plan - phase I. Harrisburg, PA, Susquehanna River Basin Commission.

Mishra, K., 2013: Effect of rain gauge density over the accuracy of rainfall, a case study over Bangalore, India. http://www.springerplus.com/content/pdf/2193-1801-2-311.pdf, SpringerPlus, **2**:311.

Moser, C. L., S.T. Kroczynski, K.P. Hlywiak, 2013: Comparison of the SAC-SMA and API-CONT hydrologic models at several Susquehanna river headwater basins, NWS Eastern Region Tech. Attach. 2013-01, Bohemia NY, NWS Eastern Region.

Nicks, A., 1965: Field evaluation of rain gage network design principles. IAHS Hydrological Sciences J., **67/68**, 82-96.

NWS, 2013: NWSRFS user's manual. Silver Spring MD, NWS Office of Hydrology, http:/www.nws.noaa.gov/oh/hrl/nwsrfs/users_manual/part2/_pdf/23apicont.pdf.

Schaake, J. C., 1981: Raingage network density requirement for hydrologic modeling. Silver Spring, MD, OHD Report, http://www.nws.noaa.gov/oh/mopex/raingage%20density%20requirement.htm

Schaake, J. C., Q. Duan, M. Smith, V. Koren, 2000: Criteria to select basins for hydrologic model development and testing. 15th Conf. on Hydrology, AMS, P1.8.

Seo D.J., R. S. Siddique, P.R. Ahnert, 2013: On the objective reduction of rain gauge network via geostatistical analysis of uncertainty in radar-gauge precipitation estimation. ASCE J. of Hydrologic Engineering. accepted manuscript available on-line.

Sharp, A. W., W.J. Owen, A.E. Gibbs, 1961: A comparison of methods for estimating precipitation on watersheds. Forty-Second Annual meeting of the Amer. Geophys. Union. AGU.

Sitner, W. S., C.E. Schauss, J.C. Monro, 1969: Continuous hydrograph synthesis with an API-type hydrologic model. Water Resources Research, **5**, 1007-1022.

SRBC, 1985: Proposed flood forecasting system improvement program. Harrisburg PA, Susquehanna River Basin Commission.

Thiessen, A. , 1911; Precipitation averages for large areas. Mon. Wea. Rev., **30**, 1082-1084.

Woodly, W. L., A.R. Olsen, A.Herndon, V.Wiggert, 1975; Comparison of gage and radar methods of convective rain measurement. J. Appl. Meteor., **14**, 909-928.