NOAA Technical Report NWS 54

Overview and Initial Evaluation of the Distributed Hydrologic Model Threshold Frequency (DHM-TF) Flash Flood Forecasting System

Brian A. Cosgrove
Edward Clark
Seann Reed
Victor Koren
Ziya Zhang
Zhengtao Cui
Michael Smith

Collaborators: Binghamton WFO, Baltimore/Washington WFO, Pittsburgh WFO, OHRFC, MARFC, NERFC, CBRFC, and Eastern Region

March 2012

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Weather Service

This page intentionally left blank

Office of Hydrologic Development (OHD)
SCIENCE TECHNICAL REPORT
For

Overview and Initial Evaluation of the Distributed Hydrologic Model Threshold Frequency (DHM-TF) Flash Flood Forecasting System

Brian A. Cosgrove[1], Edward Clark[2], Seann Reed[1], Victor Koren[1], Ziya Zhang[1,3], Zhengtao Cui[1,4] and Michael Smith[1]

Collaborators: Binghamton WFO, Baltimore/Washington WFO, Pittsburgh WFO, OHRFC, MARFC, NERFC, CBRFC, and Eastern Region

[1]NOAA, National Weather Service, Office of Hydrologic Development
[2]NOAA, National Weather Service, Office of Climate Water and Weather Services
[3]UCAR
[4]LEN Technologies

Final report approved March 23, 2012.

Table of Contents

ABSTRACT

A flash flood forecasting system has been developed which combines distributed modeling and statistical analyses to produce gridded forecasts of return periods. A distributed hydrologic model (DHM) coupled to a threshold frequency (TF) post-processor, DHM-TF, is currently being tested over Baltimore/Washington (LWX), Binghamton (BGM), and Pittsburgh (PBZ) Weather Forecast Office (WFO) domains, running in real-time on each WFO's server. It has verified well against National Weather Service (NWS) flash flood warning areas, local storm reports, and USGS gauge-based flood data. The development, prototype implementation, and evaluation of DHM-TF have been strongly collaborative efforts involving OHD, the PBZ, BGM, and LWX WFOs, OHRFC, MARFC, NERFC, and CBRFC. Comprehensive evaluation of real-time and retrospective DHM-TF simulations shows that, while not without shortcomings, the system is generally able to correctly depict the placement and timing of flash flood events, and offers several advantages over current operational NWS flash flood forecasting tools. These include a non-binary assessment of flood severity, high-resolution gridded output, routing of river flow, the inclusion of a snow model, and the potential for rapid updating and sub-HRAP resolution operations.

INTRODUCTION

Flash floods are a devastating natural disaster, causing millions of dollars of damage each year and putting many lives in danger. With the exception of excessive heat, flooding leads to more weather-related fatalities than any other cause. In 2010, the last year for which statistics were available, flash flooding caused 67 fatalities, 183 injuries, and $918 million dollars in damage (NWS, 2011). Just under half of the flood-related deaths were caused when victims were caught in vehicles and swept away, 21% died while in the water, and 20% died while camping. Given these statistics, accurate and timely predictions of flash floods are essential for the protection of life and property. Unfortunately, the nature of these events makes them quite difficult to monitor and predict. Flash floods feature a fast onset, less than six hours from the causative event (NWS, 2002), are local in scope, and depend greatly on fine scale weather and land surface conditions.

NWS forecasters use a variety of tools to monitor meteorological and environmental conditions for the possibility of flash flooding. In general, this includes the comparison of rainfall rates and rainfall volumes against Flash Flood Guidance. Flash Flood Guidance (FFG) is the amount of rainfall that would be needed to produce flooding over a given area. The product is derived from continuous lumped or distributed hydrologic models implemented at NWS River Forecast Centers. FFG is produced at least daily, and often multiple times per day. The Flash Flood Monitoring and Prediction (FFMP) software package is an operational tool within the Advanced Weather Interactive Processing System (AWIPS) that draws on FFG values to monitor flash flood conditions and facilitate WFO-based warning services. Storm-based flash flood warnings are products issued by the NWS that warn the public of impending or ongoing flash flood conditions that pose threat to life and property and that include a description of the specific geographic area of impact.

Monitoring efforts are valuable but often do not provide enough lead-time for affected parties to take the action needed to prevent loss of life and property. Hydrologic forecasts, which have the potential to increase warning lead-time, can be produced by standard lumped hydrological modeling. While model forecasts of raw river discharge (i.e., water flow in cms) can be informative, forecasts of flow frequency can provide forecasters with more actionable information by implicitly placing the current flow into historical context. Flow frequencies describe the rarity of any particular flow, and are calculated by comparing the current flow to a historical flow distribution. The frequency description is given in terms of the average number of years which can be expected to pass between floods of that particular magnitude (i.e., 100 year return period or recurrence interval) or in terms of the probability that the particular flow will be exceeded in any one year (i.e., one percent annual exceedence probability).

As useful as lumped models are, they are handicapped by the fact that they only provide information at basin outlets and cannot accurately represent the highly variable land surface and meteorological conditions that impact flash flooding. An alternative to lumped modeling is distributed modeling. Gridded distributed models operate on much finer space-time scales than lumped models and more effectively represent the variable nature of meteorological forcing and land surface parameters. They provide flood information, including flow frequency, at any grid point within the model domain. With this in mind, a method to use a distributed hydrologic model (DHM) in conjunction with a threshold frequency (TF) post processor (Reed et al., 2007) and NEXRAD precipitation data has been developed at NOAA's Office of Hydrologic Development (OHD). Precipitation forcings include Quantitative Precipitation Estimates (QPE) derived from the Multisensor Precipitation Estimator (MPE) and the High Resolution Precipitation Estimator (HPE) software packages, and Quantitative Precipitation Forecasts (QPF) derived from the High Resolution Precipitation Nowcaster (HPN) software package. This modeling approach is focused on improving flash flood prediction capabilities by increasing forecast accuracy and usability (Kitzmiller et al., 2008). It also seeks to improve upon the current NWS flash flood warning lead time goal of 38 minutes through use of channel routing and through leveraging one-hour HPN precipitation forecasts (NWS, 2011). Flash flood warning lead time has improved over the past several years (Figure 1), and DHM-TF forced by HPN output has the potential to improve lead times even more.

DHM-TF OVERVIEW

Operating on the Hydrologic Rainfall Analysis Project (HRAP) grid (Greene and Hudlow, 1982) at a 4km resolution and hourly time step, DHM-TF produces gridded flow forecasts, from which gridded frequency forecasts are derived using historical simulations. These frequency forecasts are then compared against flood threshold frequency grids to determine where flooding is occurring. In the absence of locally customized flood threshold grids, a uniform 2-year out-of-bank threshold value is used to indicate flooding (CITATION). That is, if DHM-TF simulates a return period value of 2 years or greater, flooding is taken to be occurring within that particular grid cell. DHM-TF relies on several hydrological modeling components to generate the required flow forecasts. These components, which include a gridded Sacramento Heat Transfer (SAC-HT) hydrological model, a gridded Snow17 snow model, overland and channel routing algorithms, and a statistical post processor, are part of the research version of OHD's Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM, Koren et al., 2004).

Hydrologic Modeling Components

The SAC-HT model (Koren et al., 2007) is an enhancement of the Sacramento Soil Moisture Accounting model (SAC, Burnash, 1973). It represents spatially heterogeneous runoff processes over river basins ranging from tens to a few thousand square kilometers. It accounts for processes in which the freeze and thaw of soil moisture can have significant effects on water balance and soil moisture dynamics. Routing is a key component of the DHM-TF flash flood forecast approach, and SAC-HT runoff in all DHM-TF simulations is first routed down conceptual hillslopes into a channel, and then routed within the channel to the outlet of the cell via kinematic wave routing (Koren et al., 2004). Whether or not this channelized flow is then routed downstream from cell-to-cell forms the basis of the two DHM-TF configurations utilized in this study.

1) **DHM-TF configured with standard cell-to-cell routing enabled**: Flash floods may occur near the causative rain event, or may occur downstream from the rainfall. The latter case is especially dangerous, as the lack of heavy rain in a particular area may provide residents or forecasters with a false sense of security. With cell-to-cell routing enabled, DHM-TF is able to represent the transport of water from channels in areas of heavy rainfall to downstream points, providing an accurate simulation of the potential for flash floods along an entire river network, and supporting downstream forecast analyses even when no QPF data is used. In this mode, streamflow within each grid cell is the result of locally generated within-cell runoff as well as channelized flow from upstream cells.
2) **DHM-TF configured without cell-to-cell routing (only local within-cell routing enabled)**: As valuable as cell-to-cell routing can be in many flood situations, the routing may mask the local flood signatures of small, fast-responding streams, especially if a cell contains a main-stem river channel. In such cells, the increase in the cell's overall streamflow due to flooding of the small stream will represent only a small percent of the total flow within the cell due to the large background flow supplied by the main river. To address this issue and better represent flooding on small steams, DHM-TF is run in this second unconnected "local" routing configuration. In this mode, the streamflow within each cell represents only locally generated runoff, with no contribution from upstream cells.

Statistical Processing

Distributed hydrologic models have the potential to provide valuable gridded flow information, and yet, as with other models, may be subject to biases which limit their applicability without calibration or post processing. To solve this problem, DHM-TF utilizes a threshold frequency post processing approach. Rather than assuming that the exact magnitudes of the simulated flows are correct, DHM-TF relies on the concept that the relative ranking of the flows are accurate. That is, even if the simulated flows are persistently biased, they will be consistent in that bias and thus can be correctly ranked against each other. It is this assumption which allows for the reliable conversion of flow values to return period values without need for accompanying observations. Reed et al. (2007) demonstrated the effectiveness of this inherent bias correction for a simulation in the Dutch Mills basin of Arkansas. In particular, they showed that although raw model flow values may be biased, the probabilities associated with these model flows are accurate and able to support the calculation of return periods.

The statistical package which accomplishes this task depends on a high-quality long-term simulation of flow, which in turn requires high-quality hourly precipitation data as input. Flow values from this long-term simulation are first passed through a routine which generates a grid of annual maximum peak flow values. A second routine fits the peaks to a log Pearson Type III (LPIII) distribution, and calculates a corresponding set of summary statistics which describe the distribution (IACOW, 1982). Specific thresholds of flood severity can then be associated with specific levels of probability within the distribution. Given historically derived probability distributions and associated thresholds for the level of severity, real-time hourly simulations of discharge can then be compared to the historical probability distribution to derive a level of severity (return period) for the specific real-time event. Forecasters can compare these grids of return periods to locally derived threshold frequencies (and associated return periods) or to a two-year out of bank rule-of-thumb to aid in warning decisions. Local threshold frequencies may be derived from several sources of information such as known flood frequencies at selected river locations or frequencies associated with culvert designs. An in-depth discussion on this process can be found in Reed et al. (2007)

The provision of a frequency and return period display provides additional information to a forecaster’s situational awareness. With standard hydrologic models, simulated discharge is converted to river stage at gaged locations through the application of a rating curve. These stages are associated with impacts for the river reach. The absence of stream gages at the spatial scale of the computational grid presents a challenge to the forecaster in predicting impacts from river or stream response. By converting the discharge to return period at all grid points within the model domain, DHM-TF successfully communicates the anticipated severity of the event by casting it in terms of what has previously occurred at both gauged and ungauged locations.

Taken together, the various components of the DHM-TF modeling approach produce flash flood forecasts which feature many advantages over traditional flash flood guidance. These include the ability to predict flash flooding at ungauged locations, a high resolution 4 km or less product (versus basin scale for standard flash flood guidance), a rapid update ability, production of non-binary flood severity information, and the output of verifiable small basin flow estimates.

NEXRAD Precipitation Data

Three NEXRAD-based precipitation products can be used as input to DHM-TF: MPE, HPE, and HPN. The MPE uses a combination of radar and gauge input data and is produced hourly within the AWIPS environment by each RFC on a 4 km grid (Kitzmiller et al., 2007). Rainfall estimates from Doppler radar, gauges, and satellites are automatically ingested, and bias correction factors are developed from a comparison of radar and gauge data. After automatic derivation of a gauge-only field, and a bias-corrected radar field, a blended radar/gauge product is produced through an automatic merging of the two fields. Since manual adjustments of input fields may occur repeatedly over several hours as additional gauge reports are received, the final MPE field, referred to as the 'best-estimate' QPE, may not be available for several hours (Kitzmiller et al., 2008). Thus, although the high quality of the MPE-derived QPE product makes it ideal for the long term baseline DHM-TF runs, the long lag times and slow updating characteristics of the product makes real-time use in flash flood forecasting impractical.

Although not offering the rigorous manual quality control that defines MPE, HPE features a lower latency time (less than 1 hour), a more rapid update (every 15 minutes), and a higher resolution (1 km), and is thus potentially well-suited for real-time, flash flood operations. Available within AWIPS, HPE leverages recent MPE gauge/radar bias information to automatically generate rainfall and rain-rate products statistically corrected for bias. The process also ingests a user-defined radar mask which determines how overlapping radars will be blended for each cell within the domain of interest.

While MPE- and HPE-derived QPE can be used by DHM-TF to bring model states up to the present, the most important aspect of the DHM-TF approach is its forecast capability which is powered by HPN-derived QPF data. Based on an updated extension of the Flash Flood Potential algorithm (Walton et al., 1985), the HPN process begins with the calculation of local motion vectors. These vectors are derived through a comparison of radar rain rates spaced 15 minutes apart, and are used to project current radar echoes forward in time out to two hours. Rain rates are then variably smoothed by a method based on the observed changes in echo structure over the past 15 minutes, as well as the current observed rain rate field (Walton et al., 1985; Kitzmiller et al., 2008).

DHM-TF EVALUATION

DHM-TF is being tested at the NWS Baltimore/Washington, Pittsburgh, and Binghamton Weather Forecast Offices (WFOs). The respective domains of 60,000 km^2, 89,000 km^2, and 57,500 km^2 cover, respectively, the LWX, PBZ, and BGM WFO County Warning Areas (CWA, Figure 2). DHM-TF is configured to use the SAC-HT model at all three locations and, additionally, the Snow17 model at the Binghamton WFO. While hydrologic and hydraulic routing options are available in other flood monitoring tools at NWS RFCs, only kinematic wave routing, which is available in HL-RDHM, is utilized in this study. No river regulation or control was implemented in the model setup, and so all simulated flows are natural. All simulations are performed in a free cycling fashion, with no data assimilation modifications (MODS) made to the model states. Both calibrated and uncalibrated Sacramento and Snow17 parameters were used in the Binghamton WFO simulations analyzed in this report, while only uncalibrated parameters were used for the other two WFOs. Uncalibrated parameters for the gridded SAC-HT model were taken from an *a priori* set of land surface parameters derived according to methods described by Koren et al. (2000) using the Natural Resources Conservation Service State Soil Geographic Database. These parameters were complemented by a percent impervious area dataset derived by Elvidge et al. (2004). Flow measurement data (cross sectional area and flow) at downstream gauges within the test domain were used to derive channel routing parameters, while values at upstream cells were derived using geomorphological relationships (Koren et al., 2004). Uncalibrated *a priori* Snow17 parameters were drawn from the CONUS data set of Mizukami and Koren (2008), and sourced from MARFC and NERFC lumped Snow17 data sets. Calibrated distributed SAC-HT and Snow17 parameters were derived through the use of the automated Stepwise Line Search (SLS) method of Kuzmin et al. (2008), in which *a priori* parameters are optimized through comparison with hourly streamflow from USGS observations.

Even with automatic and manual error correction procedures in place, a time-changing bias was found in the MPE fields used to force DHM-TF over the WFO domains. This bias stemmed from a truncation error within the NEXRAD precipitation processing scheme. Given the need for an accurate and internally consistent long-term flow simulation, a bias correction procedure was developed to account for this issue (Figure 3). In this procedure, monthly accumulations of Parameter-elevation Regressions on Independent Slopes Model (Daly et al., 1994) observation-based precipitation data are divided by monthly sums of RFC MPE data. The resulting monthly ratios form monthly correction factors that are applied to all hourly MPE data within each particular month. Application of this procedure to the MPE fields greatly reduces the inconsistency in the bias of resulting distributed model flow fields (Zhang et al., 2011).

Simulations conducted over the Binghamton, Baltimore/Washington, and Pittsburgh test areas fall into two categories: 1) retrospective, and 2) real-time. Long-term retrospective flow simulations serve as the baseline for conducting both specific retrospective case studies as well as real-time operations over the WFO domains. The simulations provide the annual maximum peak data which are used to construct the Log Pearson Type III distributions needed to convert flow values to return periods, and allow forecasted flows to be put into proper historical context. Currently 9-14 years in length, these simulations are forced by bias-corrected MPE data. Using previously saved restart files, they are extended once per year to generate new annual maximum peak data. Research has focused on several retrospective and real-time flash flood case studies over the three test site WFOs:

1) Pittsburgh WFO: 8/19/11
2) Baltimore/Washington WFO: 8/7/11

3) Binghamton WFO: 4/25/11-4/28/11, 5/3/11-5/4/11, 6/12/11-6/13/11, 8/27/11-8/29/11, 9/6/11-9/9/11

The verification of flash flood simulations is challenging due to sparse or delayed spotter information (i.e., flash flooding may occur but may not be reported, or may commence well before being reported), and due to the mismatch in format among 1) DHM-TF output (4km gridded return period information), 2) the storm-based flash flood warning polygons, and 3) point-type reported floods. Reporting location is also a complicating factor, as it is often difficult to obtain an exact latitude and longitude for reported flash flood events. For these reasons, and difficulties inherent in representing real-world fine-scale stream networks with a relatively coarse 4km connectivity file, a one-cell search radius was used when verifying DHM-TF return period values against flood observations. Additionally, it was necessary to take particular care in comparing the unregulated flows simulated by the model to USGS gauges and flood reports, which are often influenced by river controls. The preceding issues notwithstanding, it is still highly informative to compare DHM-TF output against available information, and such comparisons are detailed below.

Pittsburgh WFO

Case: August 19th, 2011 (small-scale urban flood), uncalibrated model parameters.

In this highly localized flood event which led to four deaths, over 2 inches of rain fell in downtown Pittsburgh, Pennsylvania in the span of an hour (Figure 4). The rainfall overwhelmed the city's drainage system and turned a major street into a 9-foot-deep river. The Pittsburgh WFO issued a flash flood warning covering this area at 20:24Z on the day of the event, and spotters reported flash flood conditions at 20:36Z and 20:45Z. Two challenging aspects of this flood event are the urbanized nature and fine scale topographical features of the region. One of the reported floods occurred at the junction of two 8.5 foot diameter underground storm sewer pipes and forced off 300lb manhole covers (Figure 5a), while the other flood occurred along a road on the side of a steep hill (Figure 5b). While DHM-TF cannot represent man-made drainage systems and operates on a scale that is too coarse to resolve the steep topography of the second reported flood, output from the automated run on the WFO PBZ server was examined for this case study to determine what utility the system would have offered forecasters in this type of situation. DHM-TF was run both with and without cell-to-cell routing, with results depicted in Figures 6a and 6b.

As shown in these figures, DHM-TF indicates flash flooding in the cell adjacent to both local storm reports. Given the one-cell search radius used for verification, this placement qualifies as a "hit". Also, the model's timing is excellent, with flooding indicated by the model at 21Z. Results are generally similar for the runs with and without cell-to-cell routing; however the unconnected simulation indicates slightly more widespread flooding. The overall magnitude of the flooding simulated by DHM-TF is low as judged against media descriptions of the flooding, an outcome most likely due to the lack of fine-scale topography and urban drainage networks in the model. Unfortunately there are no stream gauges within the area of flooding, so a direct comparison of observed and model results is not possible. DHM-TF did, however, match the return period of the USGS Sawmill Run stream gauge next to the area of flooding, with model and gauge each reporting a return period of < 2 years. In summary, even given the difficulty of the flooding situation and added complexities associated with urban infrastructure, DHM-TF was still able to accurately simulate the timing and placement of the flooding, and could have provided forecasters valuable input in the generation and issuance of flash flood warnings.

Baltimore/Washington WFO

Case: August 7th, 2011, (local upstream flash flood and within-bank downstream flash flood), uncalibrated with *a priori* model parameters.

The flash flooding that impacted the Virginia cities of Culpepper and Falmouth on August 7th, 2011 serves as a valuable case study as it occurred both in regions of heavy rain and in downstream regions that received little rain. In particular, the upstream region of flooding (Figure 7, yellow circle) around Culpepper received over six inches of rain during this event, while the downstream flood site near Falmouth received less than half an inch (Figure 7, top). Four instances of flash flooding were reported to the Baltimore/Washington WFO between 1:53Z and 2:30Z in the upstream area, and a flash flooding warning was in effect from 1:49Z until 16:45Z. Several headwaters are located in this area, and the location of this upstream flooding was well-simulated by DHM-TF with and without cell-to-cell routing (Figure 7, bottom). The magnitude of the flooding was also well-simulated by both instances of the model, with the 20+ year DHM-TF return periods corresponding well with the severity of the flooding as reported by the media (several car rescues and many road closures). It is also important to note that DHM-TF provided two hours of lead time in this case, with 2+ year return periods simulated in the Culpepper flood area starting at 00Z. This contrasts with the four minutes of lead time associated with the WFO's flash flood warning.

The downstream flood event near Falmouth was especially dangerous. It occurred 13 hours after the heavy upstream rainfall ended, at a time when large numbers of swimmers and hikers were in and along the Rappahannock River. One eyewitness report stated "The water was 2-foot low. All of a sudden, here comes a rush of water, and in 35 to 45 seconds it was 5-feet high and roaring." (8/7/11 Barry Beavon as reported by Michael Theis, Fredericksburg.Patch.com). No flood warning was active, and this quick surge of water led to the need for 18 water rescues around 19Z.

One of the striking features about this downstream flood, beyond the large lag time between the precipitation and the flood event, is that the river stayed within its banks, making this a within-bank flood. DHM-TF with cell-to-cell routing accurately predicted the timing, placement, and magnitude of this event, simulating a within-bank increase in flow reaching the Falmouth area at 19Z, right when the water rescues occurred. The modeled return period also matches that observed at the USGS Fredericksburg gauge (01668000), located near the water rescues, with both model and gauge reporting <2 year return period values during this event.

By contrast, DHM-TF without cell-to-cell routing was unable to accurately simulate this event. Without channel routing to carry the water downstream from Culpepper to Falmouth, the model was limited to depicting flooding only in the region of heavy precipitation (i.e., near Culpepper). This result highlights the vital importance of cell-to-cell routing in a localized precipitation event such as this, and demonstrates the added situational awareness that it offers forecasters.

Binghamton WFO

Case: June 13th, 2010, (localized urban and non-urban flash flood), calibrated model parameters.

As with the Pittsburgh flash flood case discussed above, the flooding which impacted the Scranton, Pennsylvania area on June 13th, 2010 was very localized in nature. The event was caused by two to four inches of rain which fell across a very limited area in a two hour time period (Figure 8). Although the west and south sides of Scranton were hardest hit, areas outside of the city and in the town of Old Forge experienced flooding as well. A flash flood warning was issued by the Binghamton WFO for this event at 17:34Z, and expired at 20:30Z. This provided over an hour of lead-time for flash flooding that was reported between 18:25Z and 23:00Z. DHM-TF output for this event with and without cell-to-cell routing is shown in Figure 9. While the cell-to-cell routing simulation failed to indicate any flooding during this event, the unconnected simulation provided a very accurate depiction of the spatial extent of the flooding in both Scranton and Old Forge. The unconnected simulation also displayed excellent timing, indicating an onset of flooding at 18Z, 25 minutes before the first report of flash flooding. The nearest USGS stream gauge available for verification activities is located just downstream from the Old Forge flood area. In this cell, both DHM-TF and the USGS stream gauge indicated return periods of less than two years (i.e., no flooding).

The performance difference between the connected and unconnected DHM-TF simulations likely lays in 1) the localized nature of the precipitation and 2) the type of flooding which occurred (urban and small stream). By virtue of cell-to-cell routing, discharge in each non-headwater cell in the connected DHM-TF simulation is the sum of within-cell generated discharge (usually relatively small) plus discharge generated upstream (the bulk of the flow). In a flood event such as this, where heavy precipitation only falls over a small area and upstream locations receive lesser amounts, channel discharge will increase relatively little compared to the pre-storm level. This will manifest as small increases in the return period of the flow, as was seen in the case study above. By contrast, when only locally generated discharge is considered, large increases in local runoff (from locally heavy precipitation) can translate into relatively large increases in within-cell discharge. In effect, the flood signal in the modeled channel was muted by the influx of "normal" flow from upstream in the connected DHM-TF simulation, but was isolated and allowed to emerge in the unconnected simulation. This isolation of local input along with alterations in the hillslope routing parameters is meant to represent the behavior of small streams in a DHM-TF cell, and in this case study, the bulk of the flooding occurred in small streams. These factors led to the superior performance of the DHM-TF run configured without cell-to-cell routing, and, combined with the results from the case studies above, indicate the value of executing both configurations of DHM-TF.

Case: May 3rd-4th, 2011, (null case), uncalibrated *a priori* model parameters.

DHM-TF has the capability to increase the lead-time and the timing and spatial accuracy of flash flood warnings, as well as to reduce false alarms where flash flood warnings are issued unnecessarily. It is the latter goal for which WFO BGM recommended this May 2011 case. From May 3rd to May 4th, close to 2 inches of rain fell over a broad swath extending from Pennsylvania through New York (Figure 10). Based on this rainfall and the flash flood guidance available at that time, WFO BGM issued three flash flood warnings for much of the area (Figure 10, blue outline). No flooding was reported in these areas during the time span covered by the warnings, and an examination of data from two USGS stream gauges (01532000 and 01553005) in the area of heaviest precipitation indicates streamflow return periods of less than 2 years. The three flash flood warnings are thus categorized as false alarms. Examining the output from DHM-TF simulations with and without cell-to-cell routing, the model correctly indicates a lack of flooding over the flash flood warning areas, depicting only small scattered rises in return period values which remain under 2 years in magnitude (Figure 11, cell-to-cell shown). In this case, use of DHM-TF would have enabled forecasters to avoid the issuance of unnecessary warnings.

Case: April 25th to April 28th 2011, (large scale and long duration flash flood events), calibrated and uncalibrated model parameters.

While many flash flood events are very localized, the flooding which impacted WFO BGM's domain in late April 2011 was widespread, and was the result of two separate precipitation events which struck the area from April 25th to the 28th. While the two precipitation events were separate and led to distinct rounds of flash flooding, it should be noted that the first event set the stage for the second event by saturating the soil before the second round of rain arrived. The Binghamton CWA was first impacted by rainfall associated with severe thunderstorms which erupted as a slow moving warm front passed through the area. It was next impacted by severe storms which developed as a storm system moved across the area out of the Great Lakes, and again as a cold front tracked eastward. Scattered areas received over five inches of rain, with widespread areas receiving over three inches (Figure 12). DHM-TF was run for the entire period with and without calibrated parameters, and model output is depicted in Figures 13, 14, and 15. While the preceding figures only display output from the calibrated simulations, Table 1 provides a summary of the performance of the WFO-issued flash flood warnings, all four model configurations (with and without cell-to-cell routing, with and without calibrated parameters), and the “flood union” of the DHM-TF calibrated simulations, where, if either simulation indicates flood conditions, the “Union” column was marked as a flood. Rather than focusing on return period values, the table focuses on a warn/no-warn concept. A flood event was considered 'warned' by the WFO if a WFO-issued flash flood warning was in effect at the time and location of the report, and 'warned' by DHM-TF if a return period of two years or greater was simulated with non-zero lead time within one grid cell of the reported flood.

From the table, it can be seen that even though DHM-TF operates on a relative ranking basis which can account for some types of biases, parameter calibration has a clear, positive impact on DHM-TF results. The second item of note is that over the three days of flooding, DHM-TF without cell-to-cell routing captured more of the flooding than did DHM-TF configured with cell-to-cell routing. A case-by-case investigation revealed that this superior performance tended to occur in areas of urban (non-stream) flooding, and areas where flooding occurred on a small stream residing in the same cell as a non-flooding large stream. Both are situations where the non-routing approach would be expected to have an advantage. The non-routing simulation also performed better in cells where the flood in the routed simulation was mistimed. Adding these performance advantages together, the calibrated DHM-TF simulation without cell-to-cell routing performed better than the WFO, catching all but five floods, compared to the six that were not warned for by the WFO. And, of those five misses, two were near-misses (1.8 and 1.9 year return periods), and one featured a one hour-delayed onset of flooding.

It should be noted that while each storm report is tagged with a specific point location (which is matched against a 4km by 4km HRAP cell for verification) as well as a polygon-type boundary, the storm report text often indicates flooding beyond the extent of the point and polygon locations. Examples include “scattered areas of flooding elsewhere in Luzerne County” and “many roads washed out”. Of the five DHM-TF misses noted above, three are associated with ambiguously worded storm reports, allowing for the possibility that the flooding actually matched nearby DHM-TF flood cells.

Turning to a graphical day-by-day examination of the flooding, Figure 13 shows the relatively limited flooding on the first day of the event, where only four local storm reports were recorded over two counties. Of these four, one event was missed by both the WFO BGM monitoring system and DHM-TF. The other three events were considered ‘hits’ in the analysis, with DHM-TF indicating flooding within a one cell search radius of the report. This calibrated non-channel-routing version of DHM-TF indicated flooding in areas without any storm reports. Given that a lack of storm report does not necessarily indicate a lack of flooding, it is difficult to determine the validity of the DHM-TF-indicated flooding.

To address this issue, return periods were computed for three USGS stream gauges (01518420, 01516350, and 01516500) in areas lacking local storm reports of flooding, but depicted by DHM-TF as having return periods greater than two years. Examining the data, DHM-TF’s calibrated run without cell-to-cell routing matched gauges 01518420 and 01516500, with each data source indicating return periods greater than two years. At gauge 01516350 however, the return period of the observed flow was calculated as less than 2 years, while DHM-TF produced a return period of just over 2 years. Additional verification was conducted with gauge 01534000 which indicated no flooding on the first and second days of rainfall, but displayed return periods greater than two years on the third day of the event. As with two of the three gauges discussed above, DHM-TF output verified well at this fourth location, indicating no out of bank flow.

The second day of rain led to increased levels of flooding, with 12 local storm reports. As indicated on Figure 14, five of these reports were not covered by a WFO warning. DHM-TF missed three of the 12 events, of which one was a near miss (1.9 year return period) and one featured flooding delayed by one hour. DHM-TF continued to indicate flooding in areas not containing local storm reports, and the same four USGS gauges used for verification on day one of the event were used for verification on day two. Model and observations coincide, with both data sources indicating greater than two year return periods at all three locations. Gauge 01534000 continued to indicate a lack of flooding, as did output from DHM-TF.

With saturated soils from two days of rainfall, the third day of the case study saw the most widespread flooding of the period. Figure 15 displays the southwest to northeast orientation of the flooding, with 21 reports of flash flooding located in a relatively narrow band. WFO BGM issued warnings for all 21 events, while the non-cell-to-cell routing, calibrated version of DHM-TF only missed one event. It is worth noting that on this day, the cell-to-cell routing configuration performed the best, catching all local storm reports of flooding. A comparison with the four USGS gauges yields mixed results. USGS-based return periods were greater than two years at 01534000 and 01516500, and were less than two years at the other two gauge locations. The calibrated DHM-TF simulation without cell-to-cell routing matched the USGS return periods at 01534000 and 01516350, but did not match at the 01518420 and 01516500 locations. Similarly, the calibrated simulation with cell-to-cell routing verified at two of the four gauge locations. The 50% verification rate is potentially deceiving though, as given the minor nature of the flooding at all four locations, differences of 1 year or less in the return period value made the difference between a match (USGS and DHM-TF both greater than or less than 2 year return period) and a miss (USGS and DHM-TF disagreeing on flood conditions).

DHM-TF performed very well over the three days of flooding, and many of the areas which lacked storm reports but which were depicted by DHM-TF as flooding were corroborated by USGS streamflow observations. Catching flood cases that were not warned for by the NWS, the calibrated, non-cell-to-cell routing DHM-TF configuration showed once again why it is important to not only run the cell-to-cell enabled configuration of DHM-TF. This strong performance notwithstanding, the union of output from DHM-TF simulations with and without cell-to-cell routing verified much better than either individual simulation (Table 1). Given this fact, along with the day-to-day variation of which model configuration performed the best, it may be most effective to use the union-type product. This would be especially true in situations where it is desirable to maximize flood detection.

Case: August 27th to August 28th 2011, (large scale and long duration flash flood events related to remnants of Hurricane Irene), calibrated model parameters.

As with the flooding in the April 2011 case described above, local storm reports associated with the remnants of Hurricane Irene were arrayed along a southwest to northeast axis (Figure 16) across the WFO BGM CWA. The flooding mainly occurred in regions which received over 3.5 inches of rain, although there were several reports in regions receiving less than 3 inches. Rain, which totaled over 6 inches in some areas, led to widespread flooding and one death. Hundreds of people were evacuated, and dozens were rescued. Damage from this event was estimated to top 20 million dollars in Delaware and Sullivan Counties alone, with week-long power outages.

Table 2 provides an overview of the performance of DHM-TF for this event, indicating which local storm reports were captured by DHM-TF, and which storm reports were missed. The table also indicates which flash floods were warned for by the NWS (29), and which floods were not warned for (3). The model performed very well for this multi-day event, with the cell-to-cell routing configuration of DHM-TF capturing all but two reported floods (which occurred in the same HRAP cell). These two misses occurred in locations different from those of the NWS misses. The non-routing configuration captured the vast majority of flood reports, but missed four flood events (at three locations), one of which overlapped with a NWS-miss location. The union of output from the two DHM-TF configurations did not provide a performance advantage, since the two floods that were missed by the cell-to-cell routing configuration were also missed by the non-routing configuration. In all of the missed flood events, return periods were elevated above normal, but did not reach the 2-year flood threshold. On some occasions, DHM-TF return periods receded back below the 2-year flood threshold before spotters reported flooding. Given the fact that an early indication of flooding still would have allowed issuance of a flash flood warning, and given the inexact nature of flood reporting (there is often a delay between start of flood and reporting of event) these instances were judged as captured events.

The visual summary of DHM-TF performance given in Figure 17 agrees well with the results from Table 2. The clusters of local storm reports generally coincide with the regions of highest return periods, and the southwest to northeast nature of the event is represented clearly by the model. A comparison of the cell-to-cell routing and non-routing output reveals that the non cell-to-cell routing configuration simulates less widespread flooding, with a flood-free region present in the northeast corner of the BGM WFO CWA. Since flash flooding was actually reported in this northeast region, it can be concluded that in this test case, cell-to-cell routing was vital to capturing the full range of flood events.

Both the routing-enabled and routing-disabled configurations of DHM-TF indicated flooding over the southeastern corner of the CWA, in an area devoid of any local storm reports. In order to determine if the simulated flooding was accurate (and just not reported), return periods were computed for this event at the two USGS gauge locations, 01428750 and 01439500, depicted in Figure 17. Attempts were made to include other gauging stations, but the degree of river regulation (which HL-

RDHM does not account for) prohibited the use of data from the region of highest return periods. Discharge at 01428750 peaked at 2760 CFS, which corresponds to a 3 year return period, while discharge at 01439500 peaked at a 17 year return period (5430 CFS). These values corroborate the flooding depicted by DHM-TF, indicating that a lack of observers is most likely to blame for the lack of storm reports in this region, and illustrating one of the challenges in verifying flood events.

DHM-TF depicted the flooding from Hurricane Irene very well, with the cell-to-cell routing configuration performing slightly better than the non-routing version. In particular, upstream flow contributions proved to be important at several storm report locations, and the heavy and widespread nature of the rain caused even large streams and rivers to flood, negating the need to isolate local runoff impacts as is done in the non-routing simulation. Nonetheless, both DHM-TF configurations would have provided valuable input into the flood forecasting and analysis process.

Case: September 6th to September 9th 2011, (large scale and long duration flash flood events related to remnants of Tropical Storm Lee), calibrated model parameters.

Following closely behind Hurricane Irene, the remnants of Tropical Storm Lee combined with flow off of the Atlantic Ocean to cause widespread and persistent rain over the BGM CWA. A large portion of the area received over 8 inches of rain, with some isolated regions receiving over 11 inches (Figure 18). Some of this rain fell over areas which were already very moist from having received rain during Hurricane Irene several days prior. This double rain region is delineated in Figure 18 by a black dashed line. Severe flooding resulted from the rain, and impacted much of central New York and northeastern PA. Flooding occurred in areas receiving large amounts of rain during Lee, as well as areas receiving relatively lesser amounts. The latter was made possible by the pre-moistened soil moisture conditions mentioned above.

Paralleling the analysis in Table 2, 58 local storm reports were compared against NWS warning polygons as well as return period output from routing-enabled and non-routing DHM-TF simulations (Table 3). The BGM-issued warnings performed very well, with only three flood events not covered by an active warning. Both configurations of DHM-TF also performed in an excellent fashion. The routing-enabled configuration only missed one flood event, while the non-routing version of the model only missed two events. None of the DHM-TF misses occurred at NWS-miss locations. In all three of these missed flood events, DHM-TF simulated return periods above the 2-year flood threshold value, but did so hours after the flood was reported. As in the Hurricane Irene case described above, some of the simulated floods ended before flooding was reported. Once again these were still counted as 'hits' owing to uncertainties in reporting flash floods and to the fact that the early simulated flood would have allowed for the issuance of a non-zero lead-time flash flood warning. Since the lone routing-enabled miss was caught by non-routing configuration of DHM-TF, the union of output from the two DHM-TF configurations verified perfectly.

A graphical analysis of DHM-TF's performance during this large-scale flood event is given in Figure 19. The two configurations produce relatively similar output, with regions of high return periods overlapping with the majority of storm reports. It is interesting to note the storm reports on the eastern half of the BGM WFO CWA. Referencing Figure 18, it can be seen that several of these occurred in areas receiving relatively little rainfall, and so owe their existence, in part, to the 3 inches+ of rain which moistened the ground during Hurricane Irene.

DHM-TF output was further validated through a comparison with return periods computed from three USGS gauges--the two gauges used for the Hurricane Irene case (01428750 and 01439500) along with an additional gauge located on the western edge of the simulation domain (01516500). These three gauges are situated in regions where flooding was not reported by spotters and where rainfall was relatively light. Nonetheless, stream gauge data at all three locations indicates that each surpassed the 2-year return period threshold for flood conditions. Both configurations of DHM-TF indicated flood conditions at 01439500 and 01516500, but failed to simulate flood conditions at 01428750. As the USGS-observed discharge at this latter location qualified as only a marginal flood (3 year return period), the return period miss by DHM-TF was not large in magnitude. As in previous cases, this gauge-based analysis indicates the caution that must be used in inferring the existence of flooding from the presence or absence of local storm reports.

Flooding from the remnants of Tropical Storm Lee resembled that of Hurricane Irene, in that it was widespread and resulted from heavy, prolonged rainfall that caused even large rivers to flood. This type of large river flooding negated the need to isolate the small stream flood signal from the main stem channel flow within each cell, since both rivers and streams were in flood. This eliminated the advantage of the non-routing DHM-TF configuration, which is aimed especially at situations where only the small streams within a particular cell are in flood. Additionally, the use of cell-to-cell routing proved important in capturing flood events dependent on upstream flow contributions. Nonetheless, both configurations of DHM-TF performed extremely well, were able to capture the timing and placement of flooding to a high degree of accuracy, and together led to a union of output with a 100% verification rate.

Model Enhancements: Calibration and Representation of Snow Pack

While the inclusion of snow pack modeling in DHM-TF operations adds a significant amount of data preparation and modeling overhead, experiments conducted over the WFO BGM domain illustrate the importance of representing snow pack processes in cold season flood simulations. A similar finding holds true for model calibration, a process that can be labor and resource intensive, but was found to greatly improve simulation of both discharge and return period values over the entire year. Figure 20 provides a visual example of these findings and displays four traces of model-produced discharge for the period February 2010 to April 2010. The flow that results from a simulation lacking a snow model and using uncalibrated parameters (red line) features a spurious early peak, and overly low subsequent peaks. This result stems from a heavy snowfall event that, given the lack of a snowpack model, was interpreted as rain within HL-RDHM and so resulted in an immediate increase in soil moisture and discharge. With no means of storing moisture in the form of snow for later release into the soil and channels as meltwater, the model produced a poor simulation of snowmelt-influenced runoff peaks in late March and April. The inclusion of the Snow17 model (magenta line) allows for a buildup of snow, removes the initial spurious peak, and leads to a greatly improved simulation. However, the simulation is still biased, and features overly large peaks for the snowmelt-influenced events. It is not until the SAC-HT and Snow17 models are calibrated that acceptable results are obtained (yellow line).

It is important to note that the preceding example was specially selected to illustrate the benefits that can result, in certain situations, from the use of calibrated parameters. In many cases, the benefit would not be as dramatic as that depicted in Figure 16. In fact, many test cases have been analyzed at the BGM, PBZ, and LWX WFOs, some of which are discussed in this report, in which satisfactory results emerge from uncalibrated simulations. They demonstrate that the *a priori* parameters used in the uncalibrated SAC-HT and Snow17 simulations function well in general, and provide a solid foundation for DHM-TF operations. It is not contradictory to state though, that if resources permit, calibration of these two models for use in DHM-TF would be very beneficial. This conclusion is supported by results from the DMIP2 Western Basin experiments which found that already skillful simulation of discharge and snowpack using a priori parameters improved after calibration (Smith et al., 2012).

Inclusion of a snow model in DHM-TF operations, however, is a different case. In regions of the country which receive large amounts of snowfall, accurate simulation of winter and spring discharge and return periods will be impossible without use of a snowpack model such as Snow17. Inclusion of this model in DHM-TF operations, therefore, is strongly recommended for areas that fall into this category.

Additional Considerations and Limitations

One concern in implementing DHM-TF is the relatively limited period of record when the MPE-derived QPE precipitation record (9-14 years) is available over each WFO CWA. If this precipitation record is not representative of the longer term climatology, there is a high likelihood that the discharge produced by HL-RDHM would also not reflect the long term record. Such a skew in the flow distribution would lead to the derivation of unrealistic return periods, and would reduce the utility of DHM-TF. One rule of thumb used by organizations such as the Federal Emergency Management Agency (FEMA), is that statistically computed return periods greater than approximately twice the length of the streamflow record should not be assumed to be accurate. Based on this assumption, DHM-TF return periods greater than roughly 25 years should be recast in a blanket fashion as simply ">25", with no further reliable information content as to the exact magnitude of the value. This upper return period limit will quickly grow as the precipitation (and streamflow) record increases.

In order to gauge the variability of log Pearson Type III-computed return periods as a function of record length, and thereby gain insight into the potential limitations of DHM-TF output, an initial investigation was made using streamflow data from six USGS gauges within the BGM WFO CWA. Each location features at least 73 years of hourly discharge data, supporting LPIII computation of return periods based on the most recent 13, 37, and 73 years of annual maximum peak flow data. The LPIII distribution was fit using the same method of moments approach as is used by DHM-TF (IACOW, 1982; Reed et al., 2007). The 13 year record length was chosen to represent the length of the MPE record over the majority of DHM-TF's area of operation, while 37 years was chosen to determine the impact of using only half of the available data record. Figure 21 displays return periods computed for 12 flood events using these three record lengths (some of the six gauges were used more than once). Variation with record length is minimal for the return period events of around 25 years or less, and generally increases as the calculated return period increases. This lends credence to the FEMA rule-of-thumb. While this sub-study is limited in both geographic extent and sample size, the results should provide increased confidence in DHM-TF's computation of return periods, even with the short available MPE record. In particular, the ability of DHM-TF to discriminate the range of flooding between marginal (2-year return period) and more serious (>25 year return period) does not appear to be hampered by the short period of record.

A further consideration is the use of the Log Pearson Type III distribution to compute discharge return periods. While no other distribution is as widely accepted in the United States as LPIII for this purpose, several alternative approaches do exist (e.g., normal distribution, log-normal distribution, Gumbel extreme value distribution; U.S. Dept. Agriculture, 2007). Consideration was given to altering the distribution used by DHM-TF; however, after consultation with staff from OHD's Hydrometeorology Design Studies Center, it was decided that the LPIII is well-suited for use within DHM-TF and that no changes need to be made at this time.

Forecast latency is another issue which impacts the utility of DHM-TF output. As currently implemented at the three test sites, DHM-TF is run up through the most recently available MPE output, and then run three additional hours into the future with zero precipitation. In this way, even though no QPF is used, the forecaster still gains a measure of insight into potential future downstream flood impacts through letting existing water route through the stream channel network. While this routing-based forecast is valuable, feedback from WFOs has indicated that the typical 45 minutes of lag time between the current time and the last DHM-TF return period map is too large and needs to be shortened for flash flood applications. As such, research is currently ongoing into the use of HPE and HPN data, which are produced more frequently, and with less lag time, than is MPE. In fact, HPN also includes QPF for the next hour, thereby providing further improvement of the DHM-TF lead time.

In parallel with the standard MPE-forced simulation, an MPE/HPE/HPN-forced real-time simulation was configured and activated at WFO BGM. This test simulation makes use of low-lag-time HPE precipitation to execute the model much closer to the current time, and features forecast-type return period output produced through use of 1-hour HPN forecasts. While the use of these two data sets addressed the two WFO concerns of lag time and forecast availability, examination of a September 30th, 2010 flood event revealed that DHM-TF output from this simulation did not verify as well as output from the standard MPE-forced simulation. Return period values were overly low and an unacceptably high number of flood events were missed. Examination of the forcing data indicates that it was to blame for the poor model performance. Figure 22 shows that signal artifacts were present in both the HPE and HPN data sets during this event. Further, precipitation values were much lower in these two data sets than in the MPE field at the bottom of the figure. When real-time discharge values based on this low-biased precipitation were compared against the multi-year MPE-based flow distribution in the LPIII processing module, current flows were under-ranked, and overly low return periods resulted. This mismatch in precipitation climatologies will need to be addressed in future research if HPE and HPN data is to be used in DHM-TF operations. Future research will also address the use of additional QPF data sources. This is slated occur as part of a proposed Weather Ready Nation DHM-TF Pilot Project involving MARFC, OHD, OCWWS-HSD, and NOHRSC.

SUMMARY

As detailed in the preceding results, DHM-TF performed very well across a wide range of flood test cases. Without the benefit of data assimilation MODS used in many NWS operational flood monitoring and forecasting tools, the model was able to accurately depict the timing and extent of flooding in both widespread tropical events and isolated convection-type events. DHM-TF also correctly identified the lack of flooding in a null case. Further, highlighting the strength its kinematic wave routing capabilities, the model performed well in a situation where the event was purely routing based. Future work may examine the relative performance advantages/disadvantages of the kinematic wave routing approach versus the hydrologic and hydraulic approaches available in other NWS flood monitoring and forecasting tools. While the exact role of DHM-TF in operations is still under discussion at the three WFOs, areas in which the tool will likely prove especially useful include the delineation of FFW polygon boundaries, the analysis of routing-dominated events, and the depiction of flood severity with greater clarity than that offered by current flood analysis tools. These strengths emerge in large part from its distributed model underpinnings, which give it the ability to produce 4km or finer forecasts at interior points in a fashion that lumped flash flood guidance cannot. Feedback from the WFOs has been very positive, with staff impressed at the accuracy of the model output. They have indicated that firmly believe the system has the potential to benefit their flood forecasting and analysis process, especially if the lag time of the system can be reduced. A sampling of WFO comments is provided in Appendix A.

Beyond illustrating the overall robust performance of DHM-TF, the case studies outlined in this report also demonstrate the necessity of executing DHM-TF both with and without cell-to-cell routing. DHM-TF without cell-to-cell routing tended to perform better in cases where non-stream flooding was reported, where the rainfall was localized in nature and/or cases where small stream flooding was reported in cells which also featured larger rivers. By contrast, routing stream flow downstream proved to be an advantage in widespread precipitation events (Lee and Irene) which caused both small and large stream flooding, and proved to be absolutely essential in the WFO LWX case where the flash flood impact was far downstream from the precipitation event. Given that this dual-configuration is recommended to capture the broadest range of flood events, there will inevitably be times when one configuration indicates flooding that the other does not. In these cases, factors to consider would include the characteristics of the ongoing precipitation event, the size and response-speed of the streams within the area of interest, and the performance history of the two model configurations at that particular location. In the absence of other guidance or determining factors, the best practice may be to err on the side of caution and use the union of the two outputs; that is, consider a flood event as occurring if either DHM-TF configuration indicates flood conditions. Forecaster education, training and experience will likely facilitate this type of expert level analysis that is vital to developing a robust concept of operations for any new approaches to flash flood and areal flood modeling as they are incorporated into the decision support process.

Testing of DHM-TF is also ongoing at NSSL, a summary of which is provided in Appendix B. Analysis approaches complementary to those used in this report produced findings that mesh well with the results detailed in this paper. NSSL found that DHM-TF outperformed both the legacy FFG and gridded FFG (GFFG), and captured a large percentage of flood events. Taken together, the results from the OHD and NSSL studies present a very compelling case for the use of DHM-TF in NWS operations, and demonstrate the powerful set of advantages offered by distributed hydrologic modeling. DHM-TF outperformed current NWS operational flood analysis and forecasting tools and would have provided a better level of service.

REFERENCES

Anderson, E.A. (1973). “NWS River Forecast System-Snow Accumulation and Ablation Model”. Technical Memo. NOAA, Silver Spring, MD, pp. 217.

Burnash, R. J. C., Ferral, R. L., and McGuire, R. A. (1973) “A generalized streamflow simulation system - Conceptual modeling for digital computers”, Tech. Rep., Joint Federal and State River Forecast Center, U.S. NWS and California State Department of Water Resources, Sacramento, California, 204.

Daly, C., Neilson, R. P., and Phillips, D. L. (1994). “A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain”. J. Appl. Meteor., 33, 140-158.

Elvidge, C. D., Milesi, C., Dietz, J. B., Tuttle, B. T., Sutton, P. C., Nemani, R., and Vogelmann, J. E. (2004). “U.S. Constructed Area Approaches the Size of Ohio”, Eos Trans. AGU, 85(24), doi:10.1029/2004EO240001.

Greene, D. R. and M. D. Hudlow, (1982) “Hydrometeorologic grid mapping procedures”. AWRA Int. Symp. on Hydrometeor. June 13-17, Denver, CO. (available upon request from NWS/HL)

IACOW, 1982, Guidelines for determining flood flow frequency: Bulletin 17B of the Hydrology Subcommittee, US Geological Survey, Reston, VA.

Kitzmiller, D., Ding, F., Guan, S., Kondragunta, C. (2007). “Sources of NEXRAD and multisensor precipitation estimates and their functional characteristics”. In: World Environmental and Water Resources Congress 2007. Restoring our Natural Habitat. Amer. Soc. Civil Engineers, 10pp, (Preprint).

Kitzmiller, D., Ding, F., Guan, S., Riley, D., Fresch, M., Miller, D., Zhang, Y., Zhou, G. (2008). “Multisensor Precipitation Estimation in the NOAA National Weather Service: Recent Advances”. In: World Environmental and Water Resources Congress 2008. Amer. Soc. Civil Engineers, 10pp, (Preprint).

Koren, V., Schaake, J., Mitchell, K., Duan, Q.-Y., Chen, F., and Baker, J. M. (1999). “A parameterization of snowpack and frozen ground intended for NCEP weather and climate models”, JGR, vol. 104, No. D16, pp. 19,569-19,585.

Koren, V.I., Smith, M., Wang, D., and Zhang, Z. (2000). “Use of Soil Property Data in the Derivation of Conceptual Rainfall–runoff Model Parameters”, American Met. Soc. 15th Conference on Hydrology, Long Beach, CA, pp. 103–106.

Koren, V.I., Reed, S., Smith, M., Zhang, Z., and Seo, D.J. (2004). “Hydrology laboratory research modeling system (HL-RMS) of the US national weather service”. Journal of Hydrology 291, 297–318.

Koren, V., M. Smith, Z. Cui, and B. Cosgrove (2007), Physically-Based Modifications to the Sacramento Soil Moisture Accounting Model: Modeling the Effects of Frozen Ground on the Runoff Generation Process, NOAA Tech. Report 52, 43 pp., U.S. Dept. of Commerce, Silver Spring, MD. http://www.nws.noaa.gov/oh/hrl/hsmb/docs/hydrology/PBE_SAC-SMA/NOAA_Technical_Report_NWS_52.pdf

Kuzmin, V. et al., Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search, J. Hydrol. (2008), doi:10.1016/j.jhydrol.2008.02.001.

Mizukami, N., and V. Koren (2008), Methodology and evaluation of melt factor parameterization for distributed SNOW-17, Oral presentation, AGU Fall Meeting, December 2008, San Francisco.

National Weather Service (2002). "NWS Manual 10-950". [Available from http://www.nws.noaa.gov/directives/010/pd01009050a.pdf]

National Weather Service (2011) "Summary of 2010 Weather Events, Fatalities, Injuries, and Damage Costs". [Available from http://www.nws.noaa.gov/om/hazstats/sum10.pdf]

National Weather Service (2011). "GPRA Metrics Status Report - FY2011, National". [Available at: https://verification.nws.noaa.gov/services/gpra/view/glance.aspx?gpra=1&fy=2011&fyvs=&area=NO&format=full#top

Ortega, K. L., T. M. Smith, K. L. Manross, K. A. Scharfenberg, A. Witt, A. G. Kolodziej, and J. J. Gourley, 2009: Severe Hazards Analysis and Verification Experiment. Bull. Amer. Meteor. Soc., 90, 1519–1530.

Reed, S., Schaake, J., and Zhang, Z. (2007). "A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations", J. Hydrology, 337, 402–420.

U.S. Department of Agriculture, Natural Resources Conservation Service (2007). Stream Hydrology. National Engineering Handbook, part 654, chapter 5. Washington, DC.

Walton, M., E. Johnson, P. Ahnert, and Hudlow, M. (1985). "Proposed on-site Flash-Flood Potential system for NEXRAD". Preprints, 6th Conf. Hydrometeorology, Amer. Meteor. Soc., 122-129.

Zhang, Yu, Seann Reed, David Kitzmiller, 2011: Effects of Retrospective Gauge-Based Readjustment of Multisensor Precipitation Estimates on Hydrologic Simulations. J. Hydrometeor, 12, 429–443.

Location	Report Time (UTC)	NWS Warned	Remarks	Std Uncal.	Local Uncal.	Std. Cal.	Local Cal.	Union Cal.
931, 617	201104260250		Route 6	1.4	1.7	2.2		
929, 616	201104260300		Wetona road and Route 6	1.4	1.7			
922, 623	201104260300		Route 352	1.2	1.2	1.6	1.8	1.8
935, 619	201104260330		Route 6, roadway debris	1.7				
929, 616	201104270100		Many roads washed out				2.1	
937, 630	201104270100		Streams out of banks, Vestal/Johnson city	1.4		1.4		
939, 618	201104270130		Major flash flooding, closed many roads	1.8				
943, 647	201104270200		Water rescue in progress	1.3	1.5	1.7		
943, 649	201104270200		Water rescue in progress	1.3	1.4			
941, 619	201104270205		Creek threating to flood Laceyville homes					
937, 615	201104270300		Roads flooded and washed out					
939, 616	201104270300		Roads washed out and flooded					
941, 619	201104270300		Roads flooded and washed out					
949, 647	201104270330		Scattered county road washouts	1.1	1.2	1.3	1.6	1.6
938, 632	201104270120		Ponded water, blocked culverts, Johnson City	1.4	1.7	1.4		
950, 612	201104270417		Rt. 6 Plymouth Twp, scattered Luzerne County	1.4	1.6	1.8	1.9	1.9
929, 613	201104280800		Route 14 Flooded					
934, 628	201104280858		Water over bridge on Foster Valley Road					
938, 648	201104280907		Numerous roads flooded					
932, 618	201104280913		Burlington 8-10in water flowing, mudslides					
929, 616	201104280915		Many roads flooded with heavy inundation					
933, 632	201104280917		Route 38b flooded					
936, 629	201104280932		Evacuation due to flooding Gaskill Road					
936, 629	201104281026		Water rescue on Route 17c					
935, 634	201104281038		Rt 26 closed, houses cut off, fire station flood					
938, 646	201104281057		Canasawacta creek flooding, water rescues					
935, 640	201104281100		Flash flooding reported closing several roads					
941, 645	201104281231		Route 12 flooded by the Canasawacta Creek					
933, 625	201104281248		West River Road flooded near Tioga Downs					
936, 632	201104281300		Nanticoke Creek flooding homes				Early End	
942, 639	201104281306		Wilkins Brook flooding		Early End		1.5	
936, 634	201104281435		Area surrounding fire dept flooded		Early End	Early End	Early End	Early End
949, 640	201104281500		Route 206 flooded					
939, 649	201104281518		North Main Street Route 12 flooded					
941, 645	201104281600		Flash flooding, evacuation of Willard Court		Early End			
937, 631	201104281622		Route 26 still closed, Church of the Nazarene		Early End		Early End	
941, 645	201104281739		Red Mill Bridge collapsing into Canasawacta		Early End	Early End	Early End	Early End

Table 1. Summary of local storm reports of flash flooding over the BGM CWA during the 4/25/11 to 4/28/2011 time period. Table also depicts which events were warned for by the WFO, and which events were captured by DHM-TF. The “Union Cal.” column contains the union of the two calibrated DHM-TF simulations.

Local Storm Report and DHM-TF Output Summary

Location	Report Time (UTC)	NWS Warned	Remarks	Std. Cal.	Local Cal.	Union
936, 659	201108282300		french road in the kirkland is closed due to flooding.			
938, 630	201108281539		roads flooded		Early End	
938, 630	201108281630		choconut creek flooding. several roads flooded in the vestal area.		Early End	
938, 661	201108282000		flash flooding in the city of utica along the sauquoit creek. evacuations.			
939, 629	201108281726		west hill and juneberry roads closed due to water over bridges.		Early End	
939, 657	201108281845		flash flooding along the sauquoit creek from clayville to new hartford.			
940, 631	201108281321		water flowing across pierce creek rd			
940, 632	201108281744		numerous roads closed due to water over bridges.		Early End	
942, 631	201108281538		several roads and one bridge closed by flash flooding.			
943, 616	201108281358		20 homes surrounded by water. evacuations with 2 people trapped			
943, 631	201108281528		water over the roads			
944, 618	201108281500		dunlap grove bridge destroyed by floodwaters.			
944, 630	201108281515		trailers threatened by flooding. 20 homes affected. road closures			
945, 628	201108281534		10 trailer homes flooded in the new milford area. roads flooded.			
946, 608	201108281400		huntington creek washing over rt. 239		1.9	
944, 609	201108281931		kitchen creek over banks and flooding home	Early End	Early End	Early End
948, 635	201108281600		state of emergency, major flash flooding in deposit area. roads closed			
948, 654	201108281900		flash flooding reported around otsego county. at least 8 roads closed			
950, 658	201108281800		up to 50 people evacuated in cherry valley, 5 of rain fallen.			
951, 612	201108281228		minor flooding reported in ashley along the solomons creek.		Early End	
951, 612	201108281249		solomons creek rapidly rising. evacuations in south wilkes-barre.		Early End	
952, 610	201108281318		streams out of their banks	1.7	1.7	1.7
952, 610	201108281730		wapwallopen creek flooding roads	1.7	1.7	1.7
954, 645	201108281815		east branch delaware and small stream flooding. flood stage reached		1.6	
960, 645	201108281235		streams over roads and bridges			
961, 647	201108281438		batavia kill over rt. 36 near dimmock mountain rd.			
962, 635	201108281229		streams over roads or bridges			
962, 646	201108281328		main street evacuated, bridges washed out, catastrophic flooding.			
962, 646	201108281641		helicopter rescuing people off the roofs of buildings due to flooding.			
964, 640	201108281200		rt 19 covered with water. numerous mud slides. historic flooding			
965, 639	201108281517		water in buildings			
971, 632	201108281402		streams in buildings			

= NWS or DHM-TF Captured Flood   # = Missed Flood, For DHM-TF Max Storm Return Period Listed

Table 2. Summary of local storm reports of flash flooding over the BGM CWA during the 8/27/11 to 8/29/2011 time period. Table also depicts which events were warned for by the WFO, and which events were captured by DHM-TF. The "Union" column contains the union of the two calibrated DHM-TF simulations.

Local Storm Report and DHM-TF Output Summary

Location	Report Time (UTC)	NWS Warned	Remarks	Std. Cal.	Local Cal.	Union
921, 629	201109072100		road flooding on nys 14 between watkins glen/montour falls		62.2	
926, 634	201109072254		several roads flooding around ithaca.			
926, 634	201109072330		flooding on ellis hollow road at east hill plaza.			
926, 634	201109081515		yards flooded, water in basements in 100 block of park pl.		Early End	
927, 628	201109081143		flooding			
928, 629	201109081215		flooding			
928, 629	201109081354		east spencer road not passable from owl creek road east			
930, 630	201109081315		flooding 1 mile west of candor on route 96			
932, 625	201109071407		smith creek ro and sulpher springs rd have water over them			
933, 626	201109071514		east river rd flooded at nichols. houses flooded nr lounsberry.			
933, 627	201109080100		the engelbert farm on east river rd is several feet under water			
933, 629	201109071503		lisle rd near gary hunt rd water over the road.			
933, 632	201109071705		numerous roads and streams flooded. water over bridges.			
933, 661	201109072230		erie blvd flooded, 6 inches of water over 3rd street in rome	5.1		
934, 624	201109071536		route 187 flooded			
934, 628	201109071619		rds closed due to streams flooding, southern county			
934, 628	201109081249		rt 17 is closed from exit 70 johnson city to exit 60 waverly			
935, 618	201109071820		all the main roads through monroeton flooded and closed.			
935, 619	201109071802		state of emergency bradford cty. major flooding. evacuations.			
935, 619	201109071930		state of emergency bradford cty. major flooding. evacuations.			
935, 619	201109071930		route 6 is closed between wysox and wyalusing.			
935, 635	201109071723		leekville rd washout			
935, 640	201109071955		flash flooding of streams and creeks. road closures.			
936, 622	201109072045		flooding, roads closed in bradford county, towns unaccessible.			
936, 628	201109071411		water over streets and running through peoples yards.			
936, 628	201109081408		basements flooded, power outages, penn ave impassable			
936, 632	201109071921		route 26 just north of maine...3 ft of water was in a store			
936, 634	201109071935		nanticoke creek flooding roads and threatening homes			
936, 636	201109072300		multiple roads flooded in the town of barker.			
937, 627	201109071312		bolles hill rd has water over a bridge			
937, 627	201109071804		national guard being called in to rescue trapped residents.			
937, 630	201109071630		nanticoke creek is flooding neighborhood along river drive.			
937, 630	201109080338		route 434 vestal is flooded next to chucksters golf.			
937, 630	201109081529		enjoie golf course 50 percent flooded		Early End	
937, 630	201109081534		evacuations due to flooding		Early End	
938, 633	201109071500		west chenango rd flooded near rt 11			
938, 634	201109071454		brooks rd flooded near fox rd.			
938, 634	201109071925		in castle creek ... creek is flooding homes in the town			
938, 659	201109072342		a mobile home park in chadwicks is being evacuated.			
938, 661	201109080133		the saquoit creek is causing flooding along brookline drive			
939, 638	201109071512		county rd 32 in greene twp. closed due to water over the rd.			
939, 638	201109071530		streams over roads and bridges.			
940, 632	201109071431		numerous roads flooded in three cities area. deep ponding			
940, 632	201109071711		state of emergency broome cty. major flooding, roads closed			
940, 632	201109072000		i-88 east between exits 2,3 is closed due to a landslide			
940, 637	201109072051		page brook road is flooded at the chalker creek culvert.			
941, 623	201109071945		many roads flooded. rescues taking place			
941, 645	201109071945		willard court flooded. 30 to 40 people evacuated.			
941, 645	201109072136		state of emergency in chenango county. roads closed			
942, 639	201109071501		numerous roads flooded.			
943, 616	201109071605		streams over flowing banks. roads flooded.			
946, 618	201109071606		bowmans creek flooding. threatening a home.	Early End	Early End	
948, 605	201109071533		nescopek creek flooding.	Early End	Early End	
948, 645	201109071945		water rescues, roads closed, state of emergency otsego cty.			
949, 647	201109072025		numerous roads flooded. suny oneonta almost cut off on hill			
951, 612	201109070634		widespread flooding with some evacuations beginning			
960, 632	201109070930		water over route 52 impassable			
966, 636	201109071448		evacutions of 300 home near fallsburg.		6.3	

= NWS or DHM-TF Captured Flood

\# = Missed Flood, For DHM-TF Max Storm Return Period Listed

Table 3. Summary of local storm reports of flash flooding over the BGM CWA during the 9/7/11 to 9/10/2011 time period. Table also depicts which events were warned for by the WFO, and which events were captured by DHM-TF. The "Union" column contains the union of the two calibrated DHM-TF simulations.

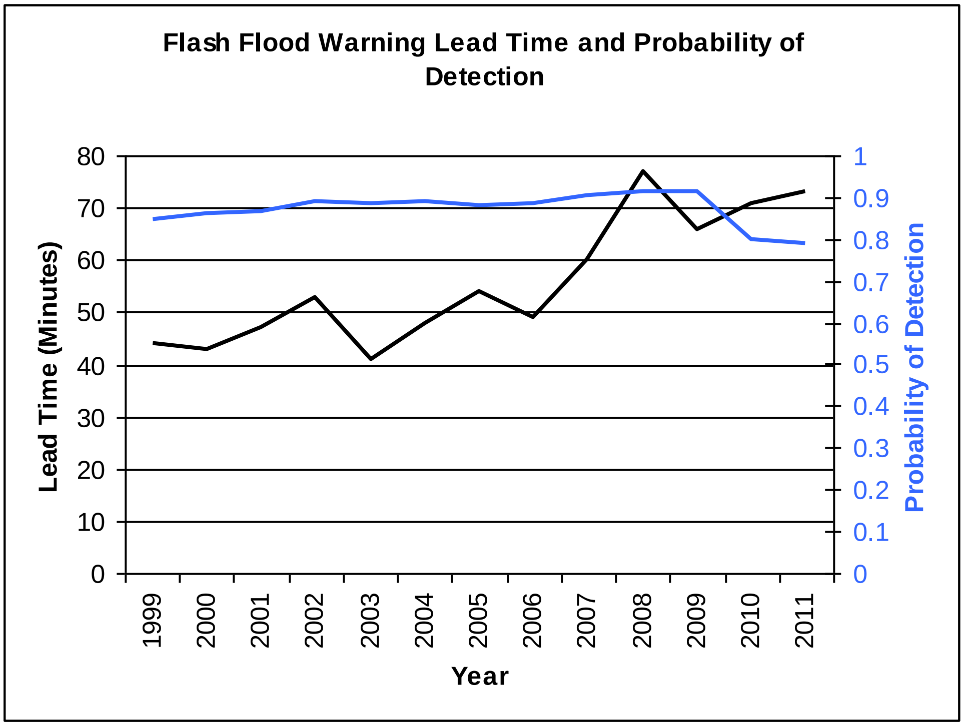


Figure 1. Average yearly National Weather Service flash flood warning lead-time and probability of detection from 1999 through 2011. The 2010 and 2011 values stem from the new storm-based verification system, while the earlier values are county-based statistics.

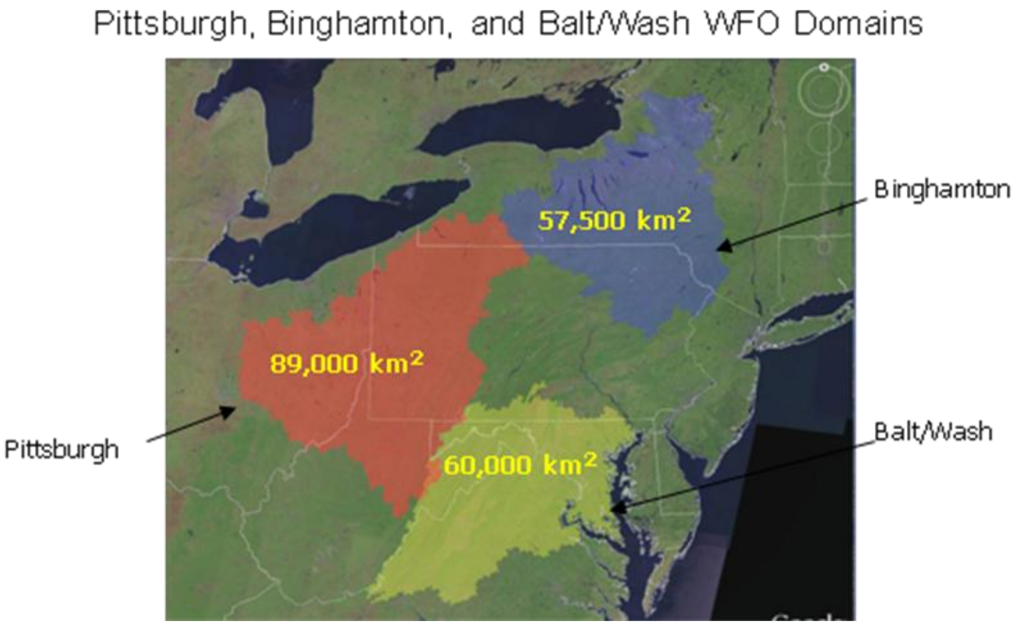


Figure 2. DHM-TF Case study domains centered on the Baltimore/Washington, Pittsburgh, and Binghamton WFO County Warning Areas, and featuring a $4km^2$ grid resolution.

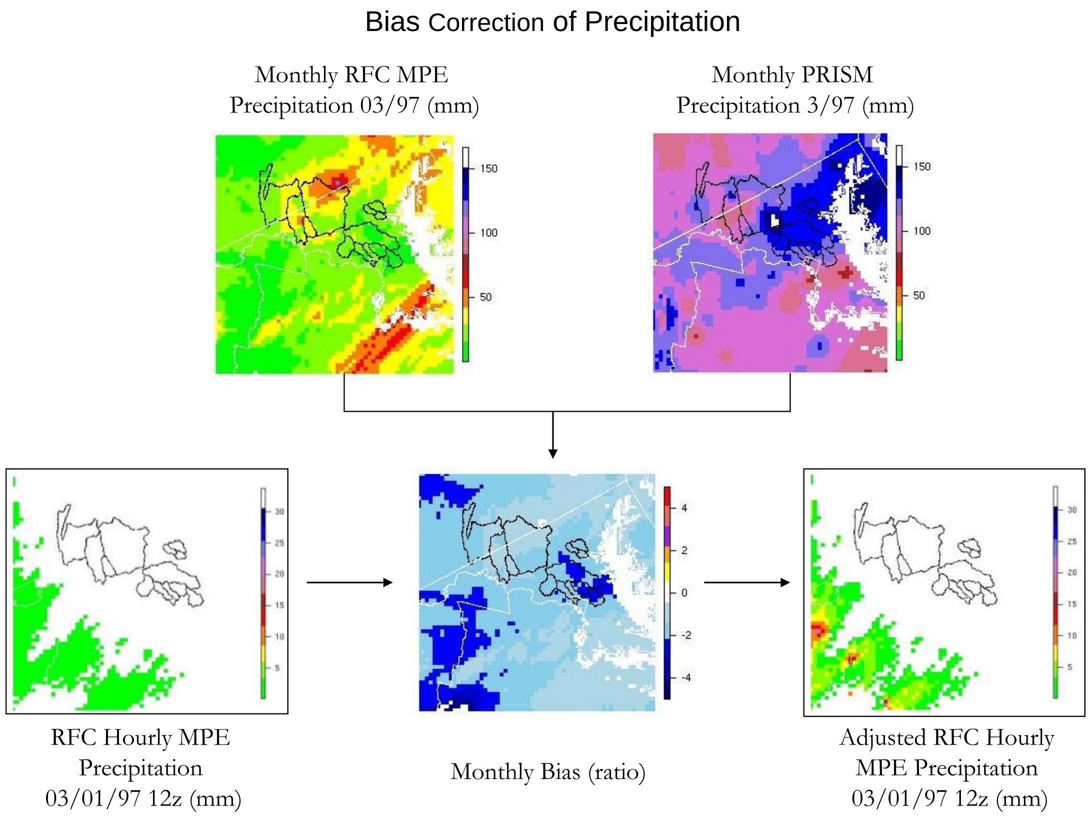


Figure 3. Overview of the bias correction procedure used to reduce the bias present in the MARFC and OHRFC MPE precipitation records (Adapted from Yu Zhang, personal communication).

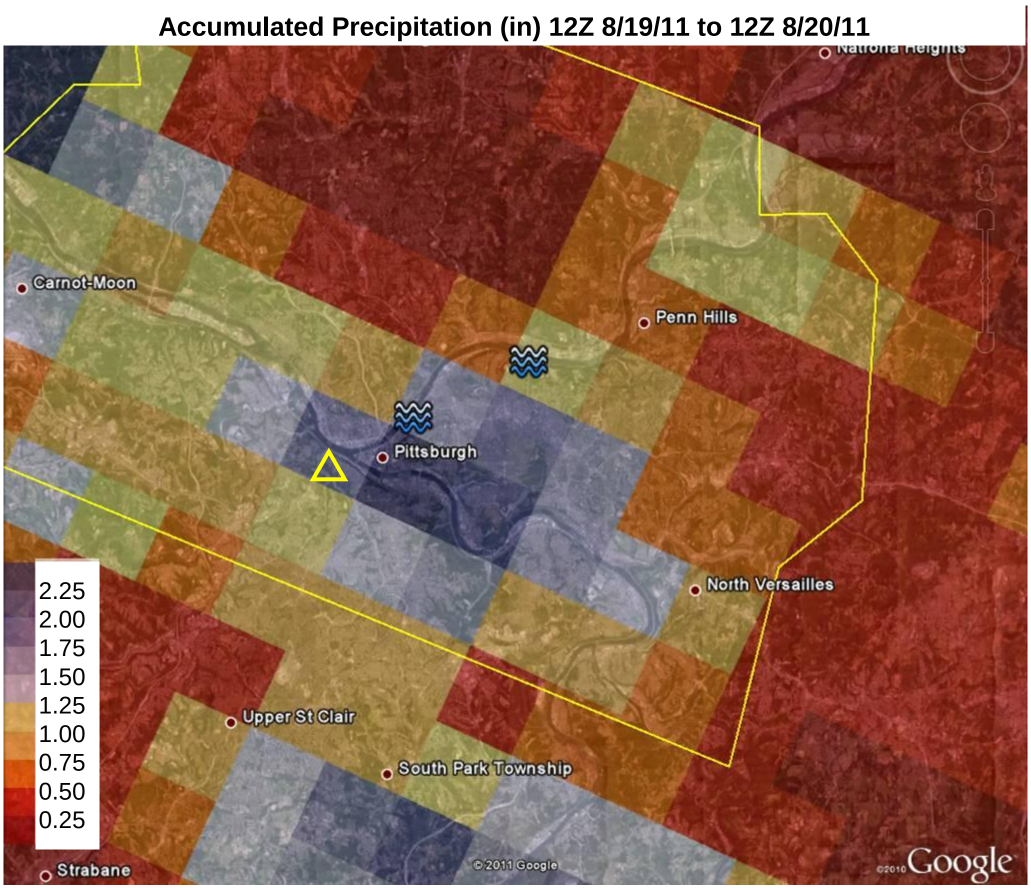


Figure 4. Accumulated 24-hr precipitation associated with flash flooding in Pittsburgh on August 20th, 2011. Yellow line represents flash flood warning issued by Pittsburgh WFO, blue wave symbols represent observed locations of flash flooding, and yellow triangle indicates location of USGS stream gauge used for verification.

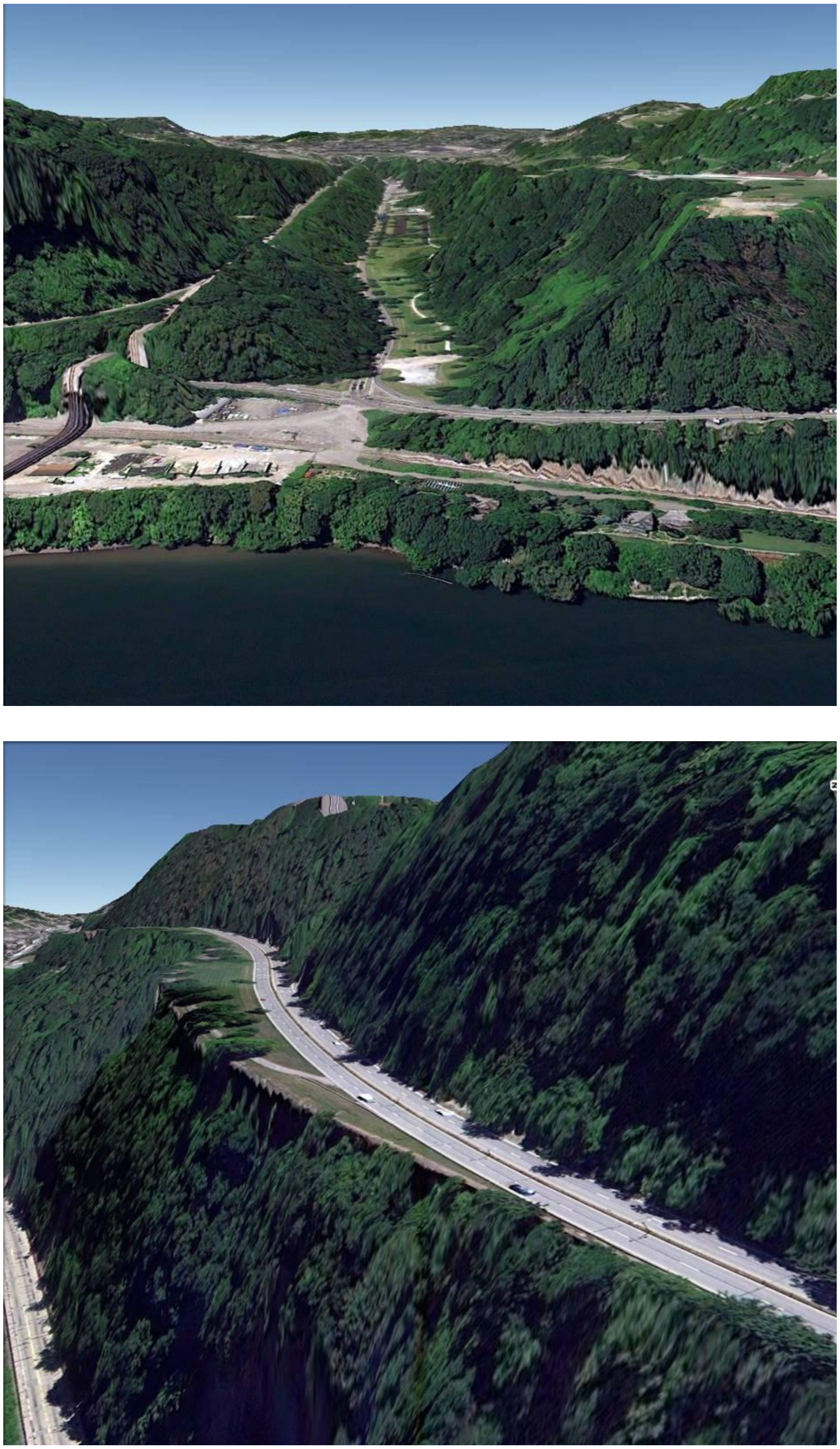


Figure 5. Locations of observed flash flooding in Pittsburgh on August 19th, 2011. Top: Roadway at bottom of valley along site of former stream and current junction of sewer system. Bottom: Roadway and park along steep hillside.

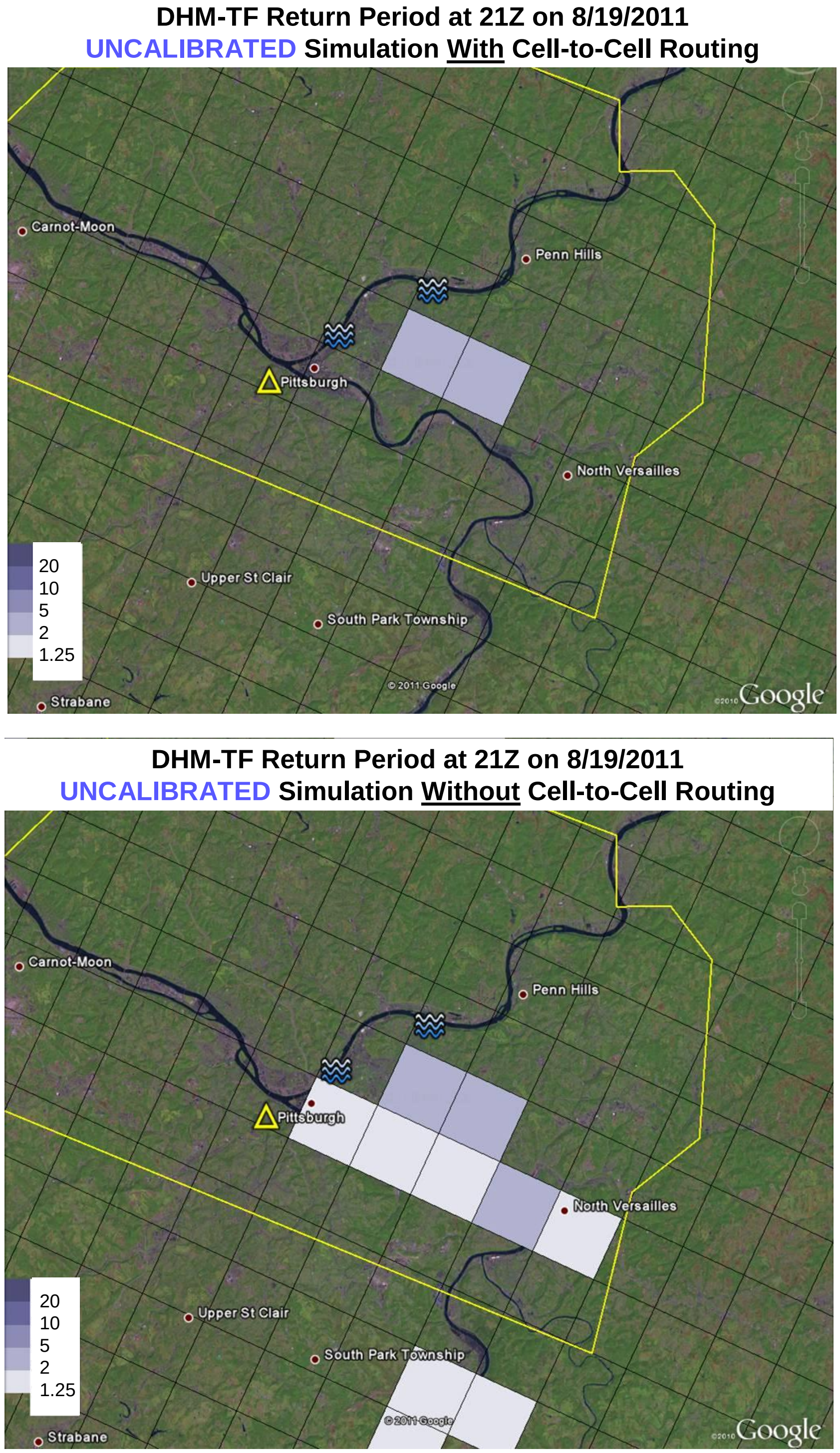
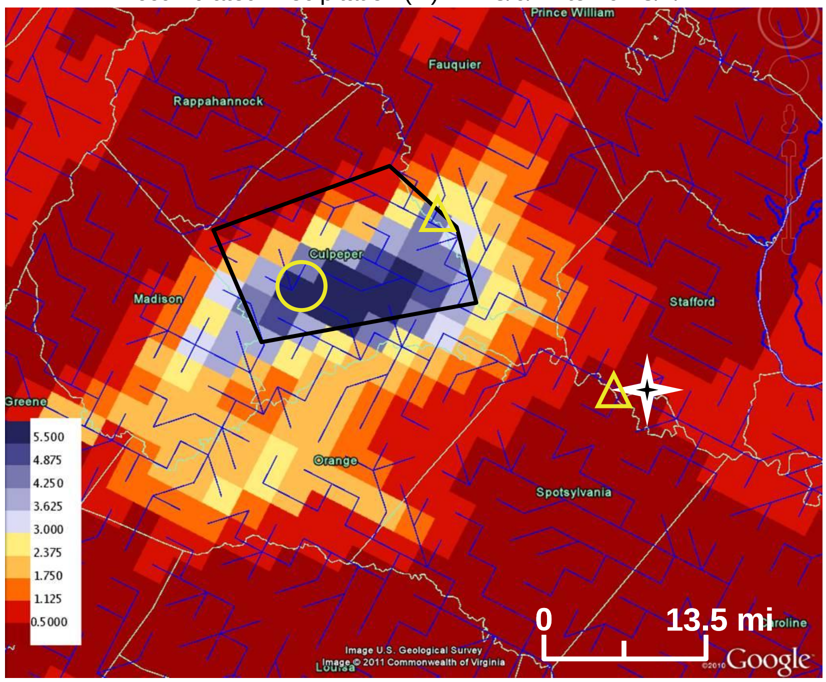
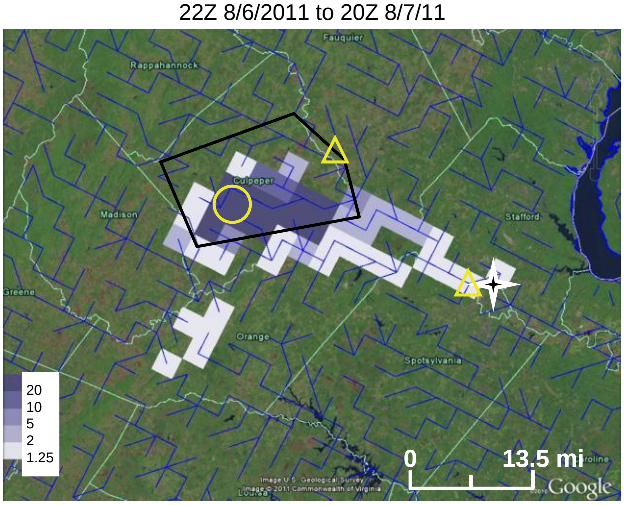


Figure 6. Plots showing DHM-TF return period (years) around time of flash flooding in downtown Pittsburgh. Top (5a)—Standard simulation using cell-to-cell routing. Bottom (5b)—Unconnected simulation using "local" routing. Yellow triangle indicates location of USGS stream gauge used for verification, yellow outline represents WFO flash flood warning area, and blue wave symbols indicate locations of observed flash flooding.

Accumulated Precipitation (in) 21Z 8/6/11 to 20Z 8/7/11



Max Return Period (yrs) Simulation w/ Cell-to-Cell Routing
22Z 8/6/2011 to 20Z 8/7/11



Max Return Period (yrs) Simulation w/o Cell-to-Cell Routing
22Z 8/6/2011 to 20Z 8/7/11

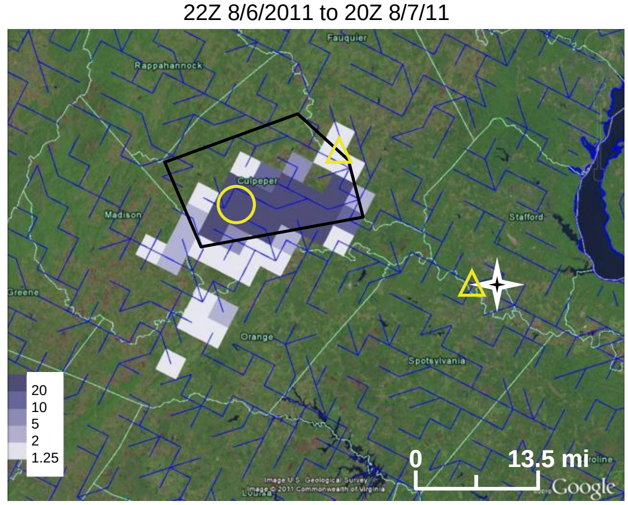


Figure 7. Overview of the August 6th-7th 2011 WFO LWX flood event. Top: Accumulated precipitation (inches) during flood period, Bottom: Maximum return period during the event as simulated by DHM-TF with (left) and without (right) cell-to-cell routing. Black border depicts WFO LWX flash flood warning, yellow triangles indicate USGS stream gauges, yellow circle depicts location of local storm reports of flash flooding, and white star indicates location of rescued swimmers.

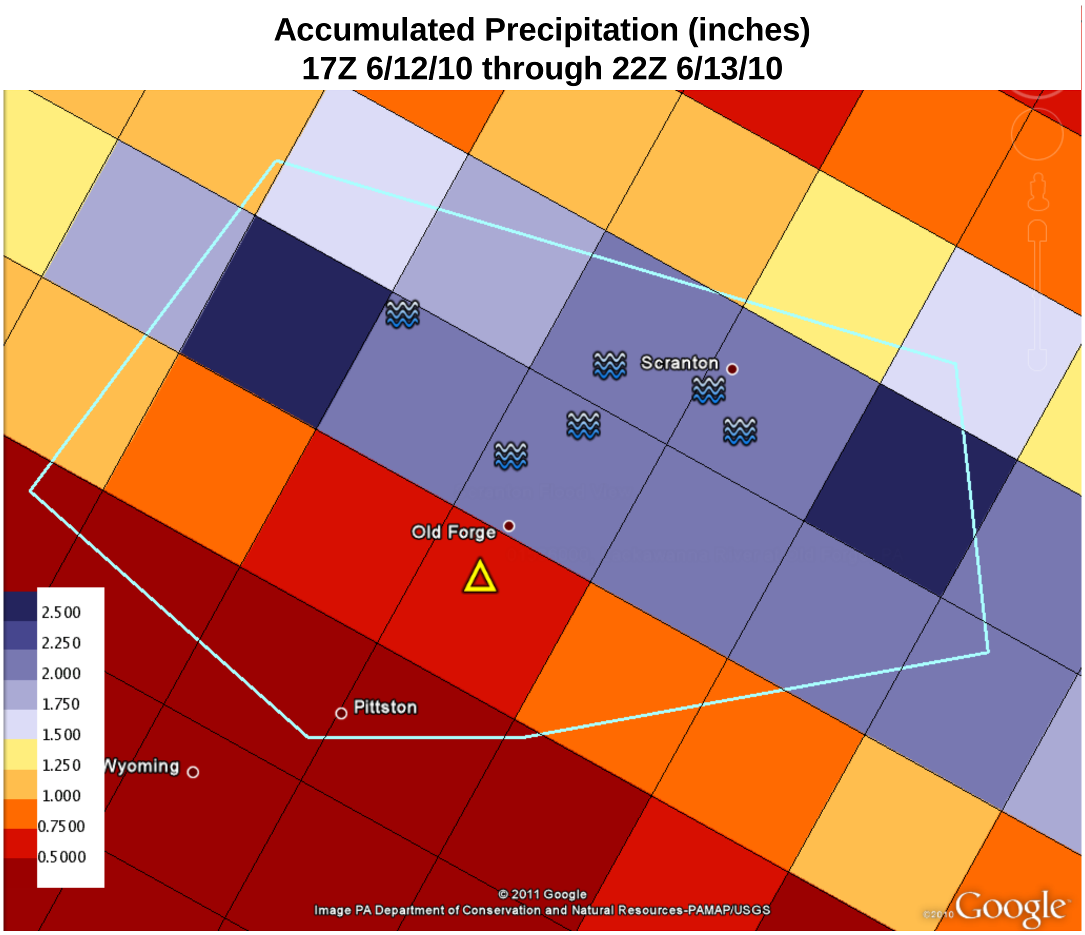


Figure 8. Accumulated precipitation associated with flash flooding in Scranton and Old Forge from 17Z June 12th to 22Z June 13th, 2010. Blue line represents flash flood warning issued by Binghamton WFO, blue wave symbols represent observed locations of flash flooding, and yellow triangle indicates location of USGS stream gauge used for verification.

DHM-TF Max Return Period (Analysis w/Cell-to-Cell Routing)
18Z 6/12/10 through 22Z 6/13/10

DHM-TF Max Return Period (Analysis w/o Cell-to-Cell Routing)
18Z 6/12/10 through 22Z 6/13/10

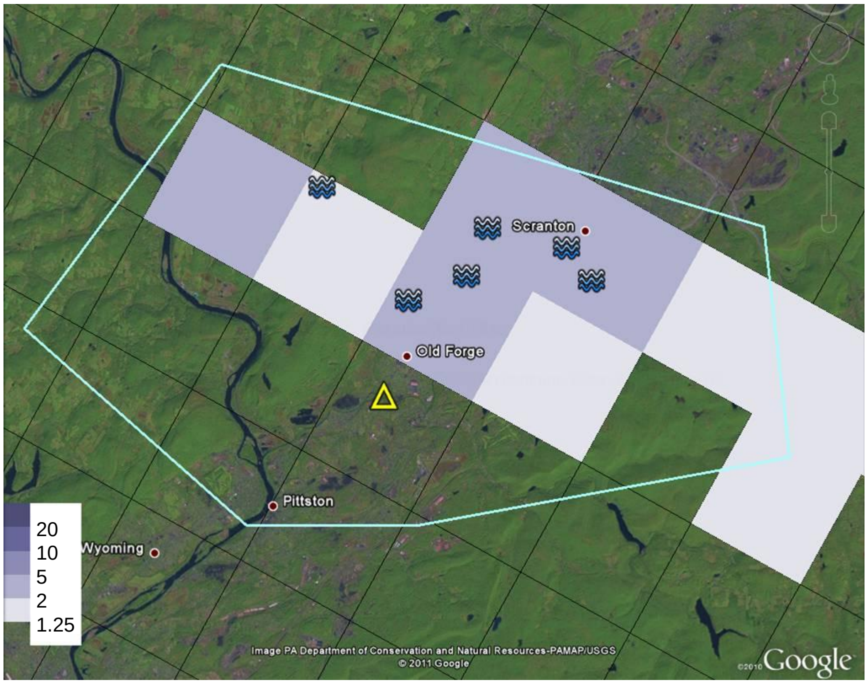


Figure 9. Plots of maximum return period values during 6/13/10 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangle indicates location of USGS gauge 1536000 used in verification, and blue wave symbols indicate observed locations of flash flooding.

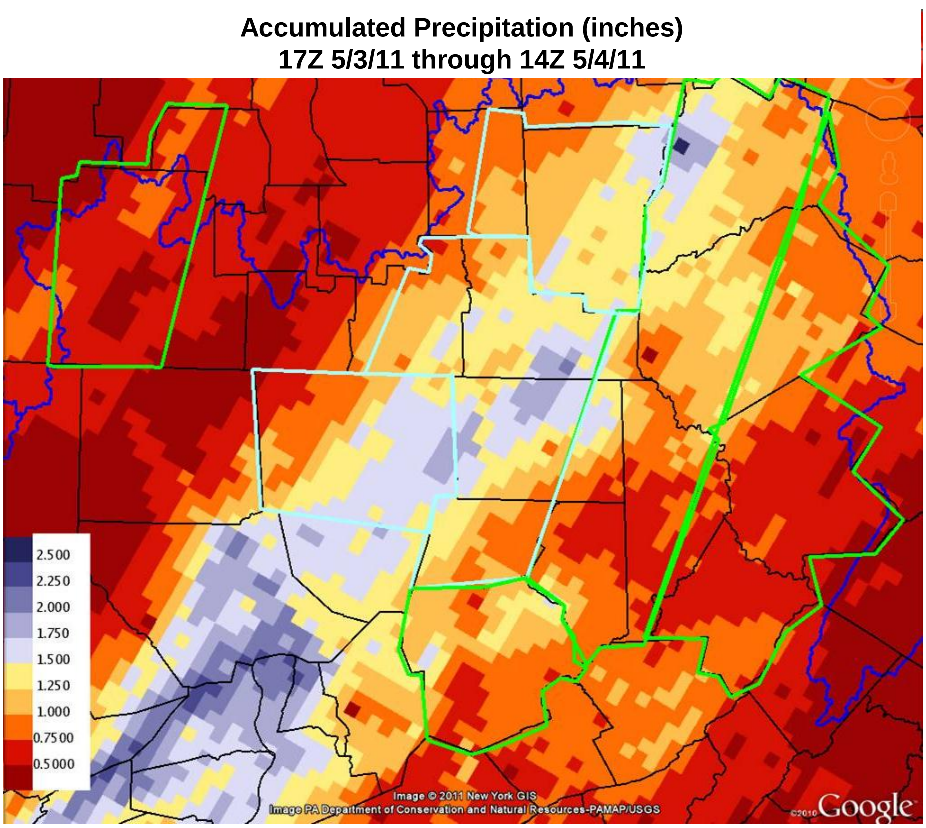


Figure 10. Accumulated precipitation from 17Z May 3rd to 14Z May 4th, 2011 which prompted issuance of flash flood warnings (light blue outlines) by WFO BGM. Dark blue line represents RFC boundary, green line represents areal flood advisory, and yellow line indicates boundary of WFO BGM.

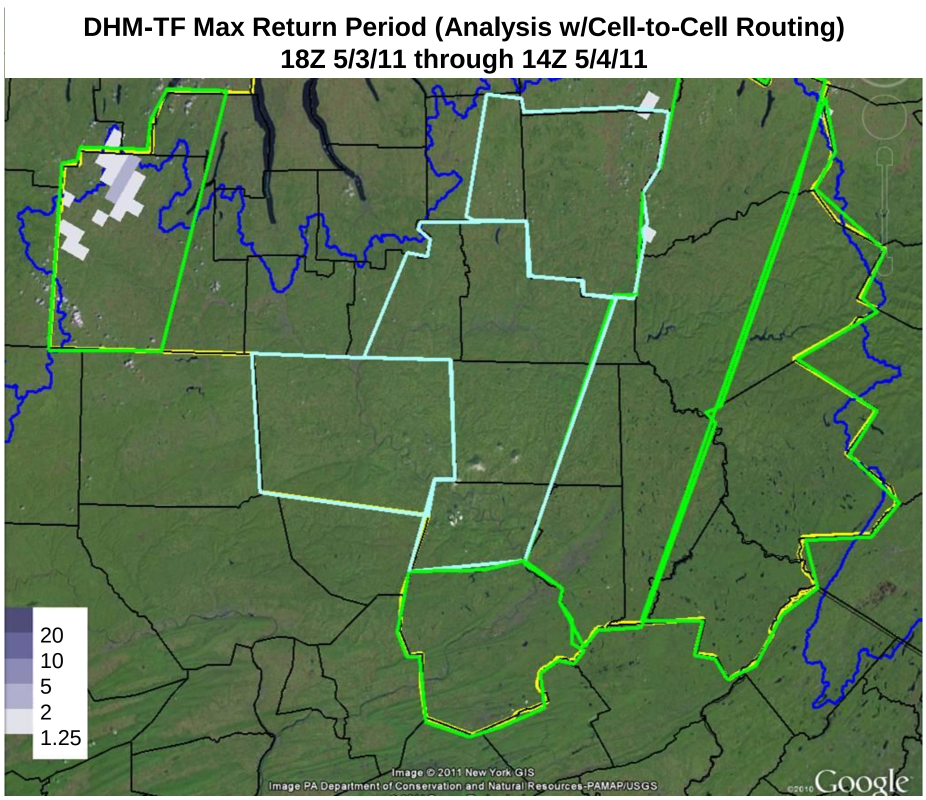


Figure 11. Plot of maximum return period values from 18Z 5/3/11 to 14Z 5/4/11 generated by DHM-TF with cell-to-cell routing enabled. Dark blue line represents RFC boundary, green line represents areal flood advisory, and yellow line indicates boundary of WFO BGM.

Accumulated Precipitation (inches) 17Z 4/25/11 through 18Z 4/28/11

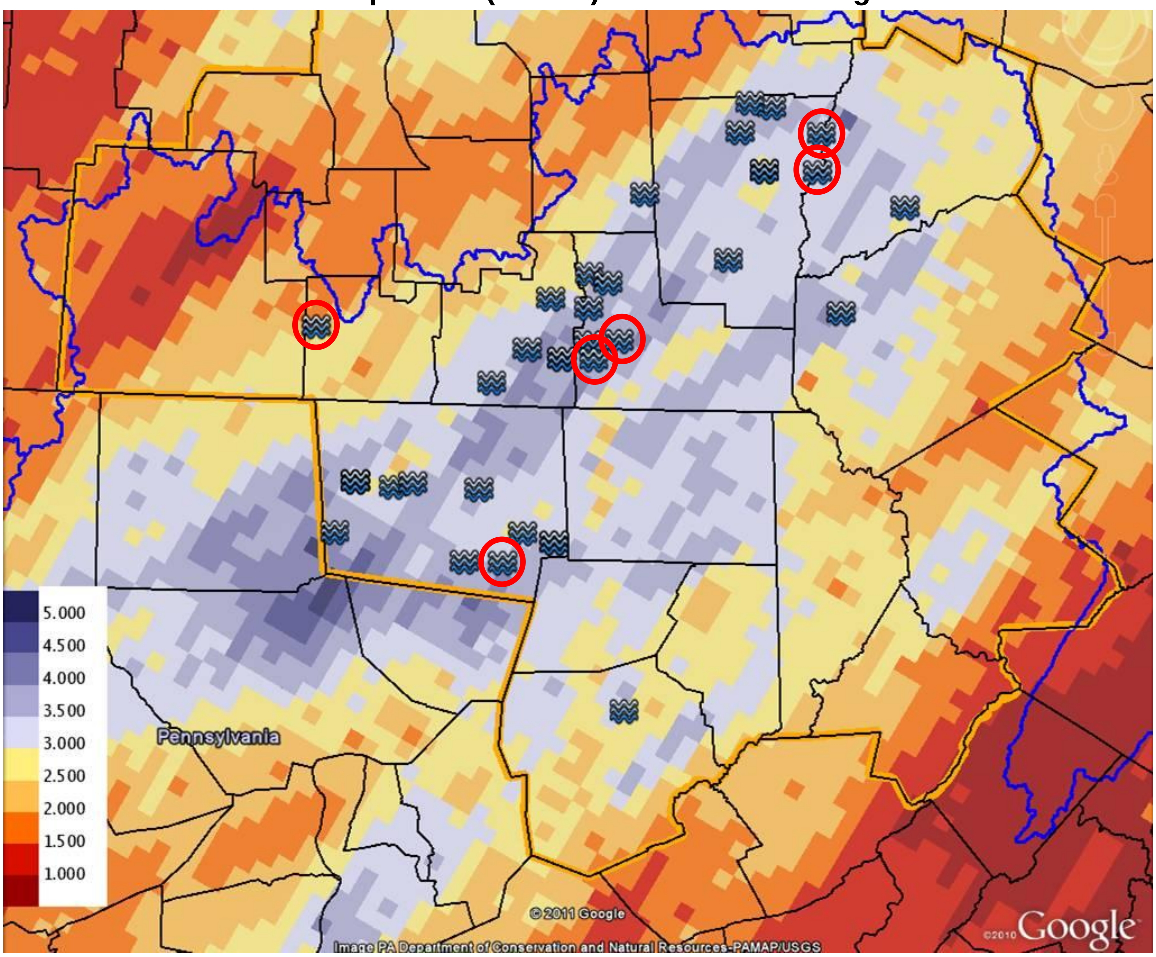
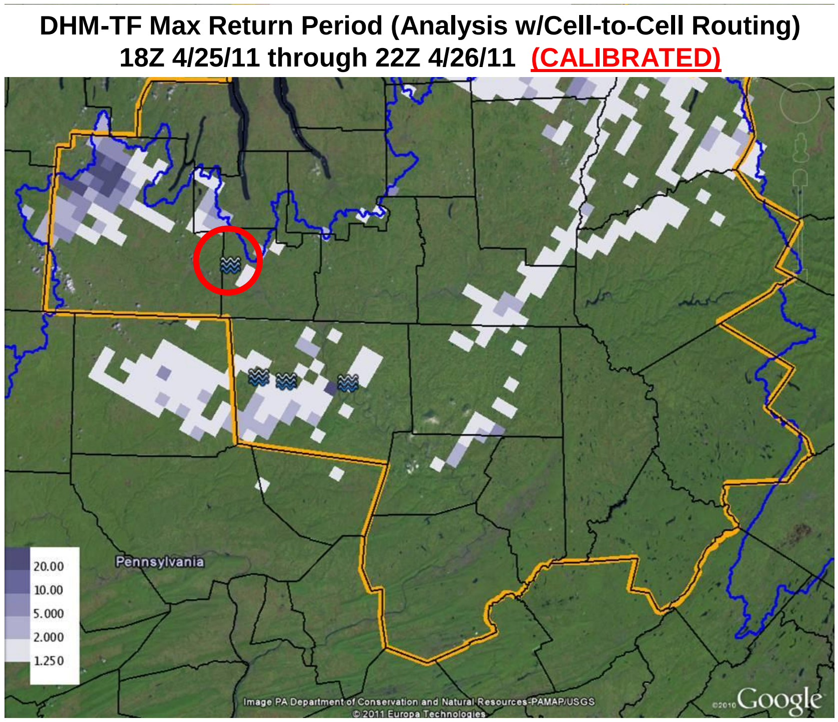


Figure 12. Accumulated precipitation associated with flash flooding in the WFO BGM county warning area from 17Z April 25th to 18Z April 28th, 2011. The blue line represents the RFC boundary, the orange line represents the WFO boundary, the blue wave symbols represent observed locations of flash flooding, and the red circles represent unwarned storm reports.



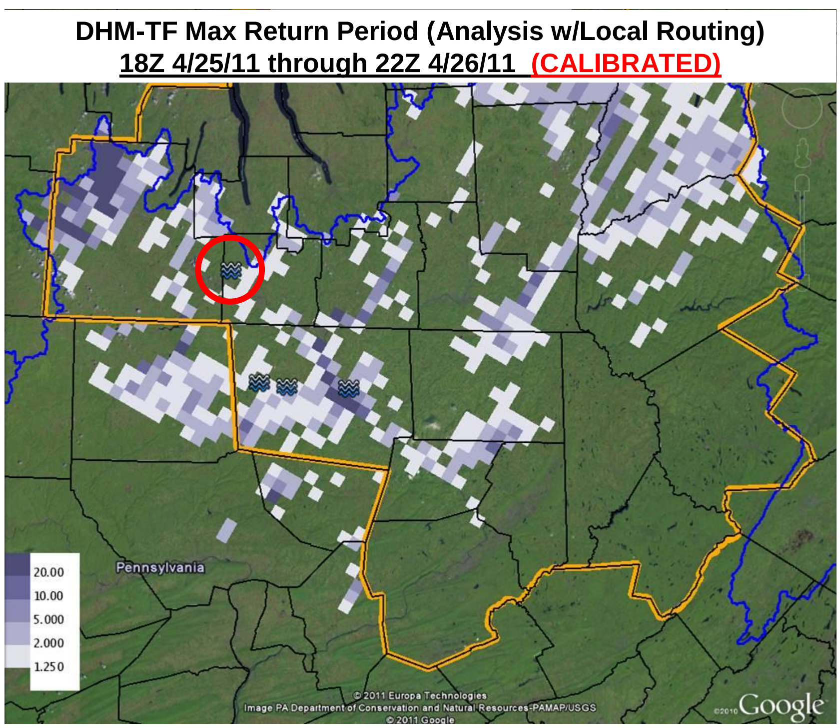


Figure 13. Plots of maximum return period values during 4/25-4/26 2011 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangles indicate locations of USGS gauges used in verification, the red circle indicates reports of flooding that were not warned for, and blue wave symbols represent observed locations of flash flooding.

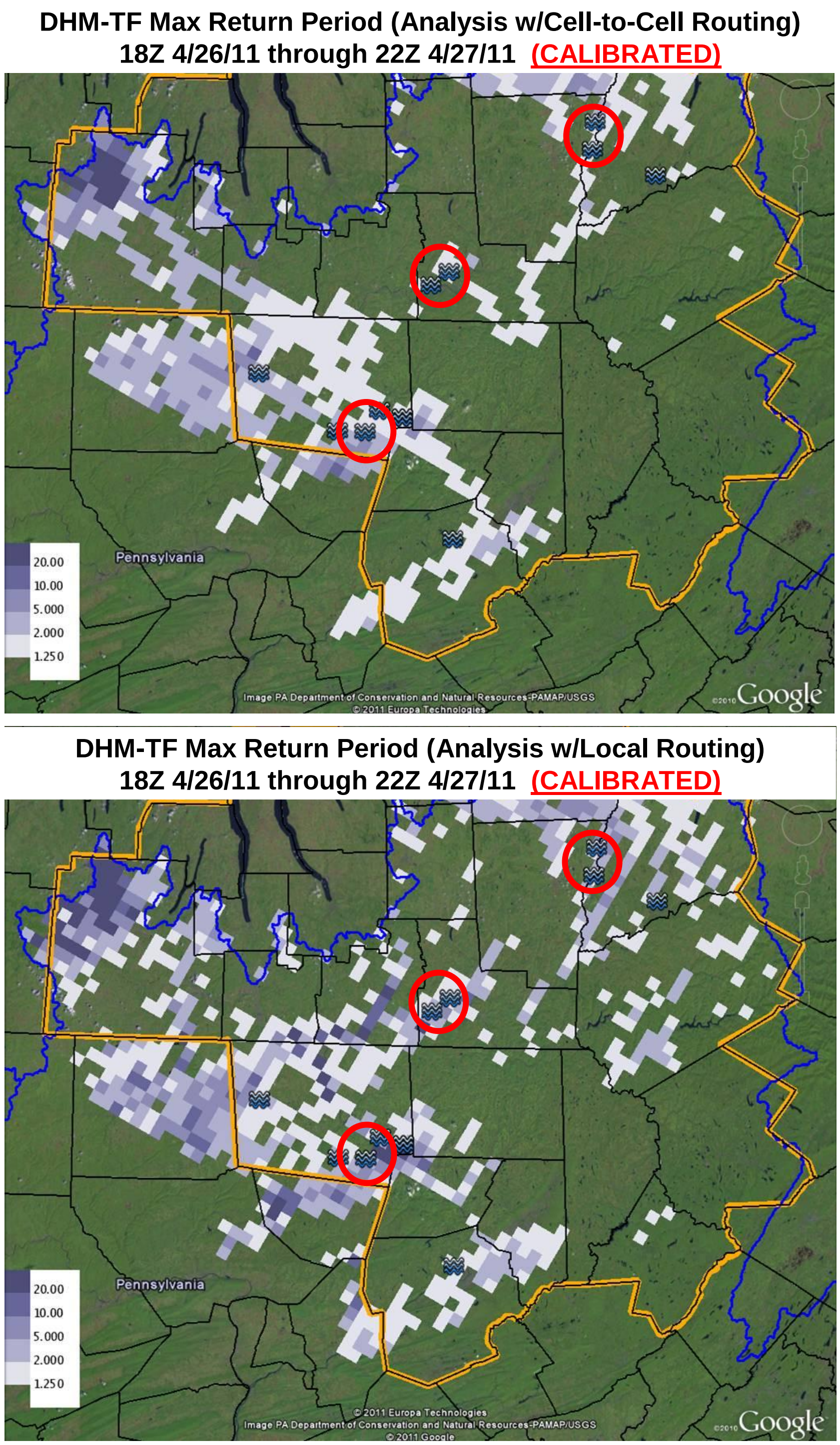


Figure 14. Plots of maximum return period values during 4/26-4/27 2011 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangles indicate locations of USGS gauges used in verification, the red circles indicate reports of flooding that were not warned for, and blue wave symbols represent observed locations of flash flooding.

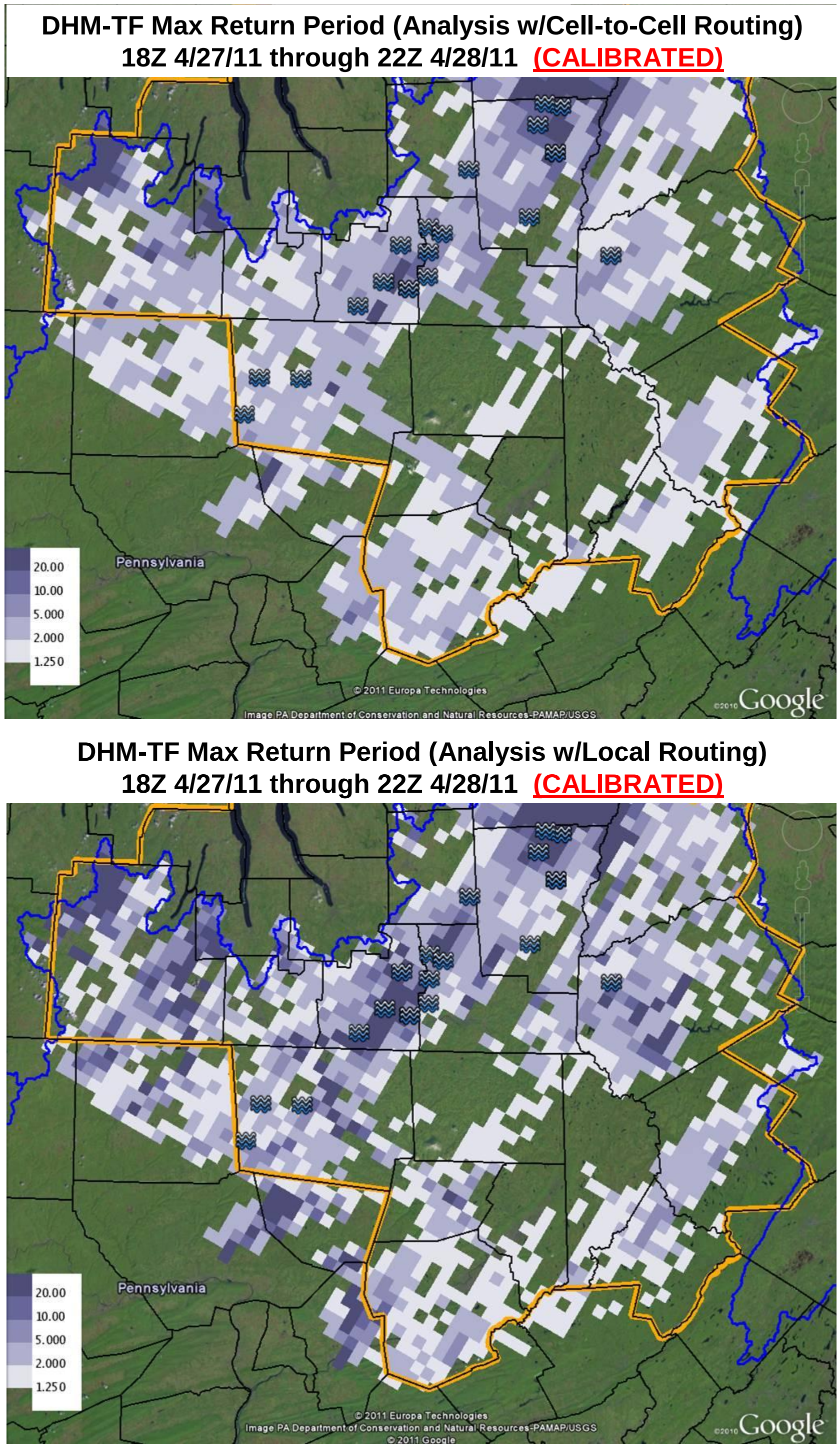


Figure 15. Plots of maximum return period values during 4/27-4/28 2011 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangles indicate locations of USGS gauges used in verification, the red circle indicates reports of flooding that were not warned for, and blue wave symbols represent observed locations of flash flooding.

Accumulated Precipitation (inches)
11Z 8/27/11 through 12Z 8/29/11

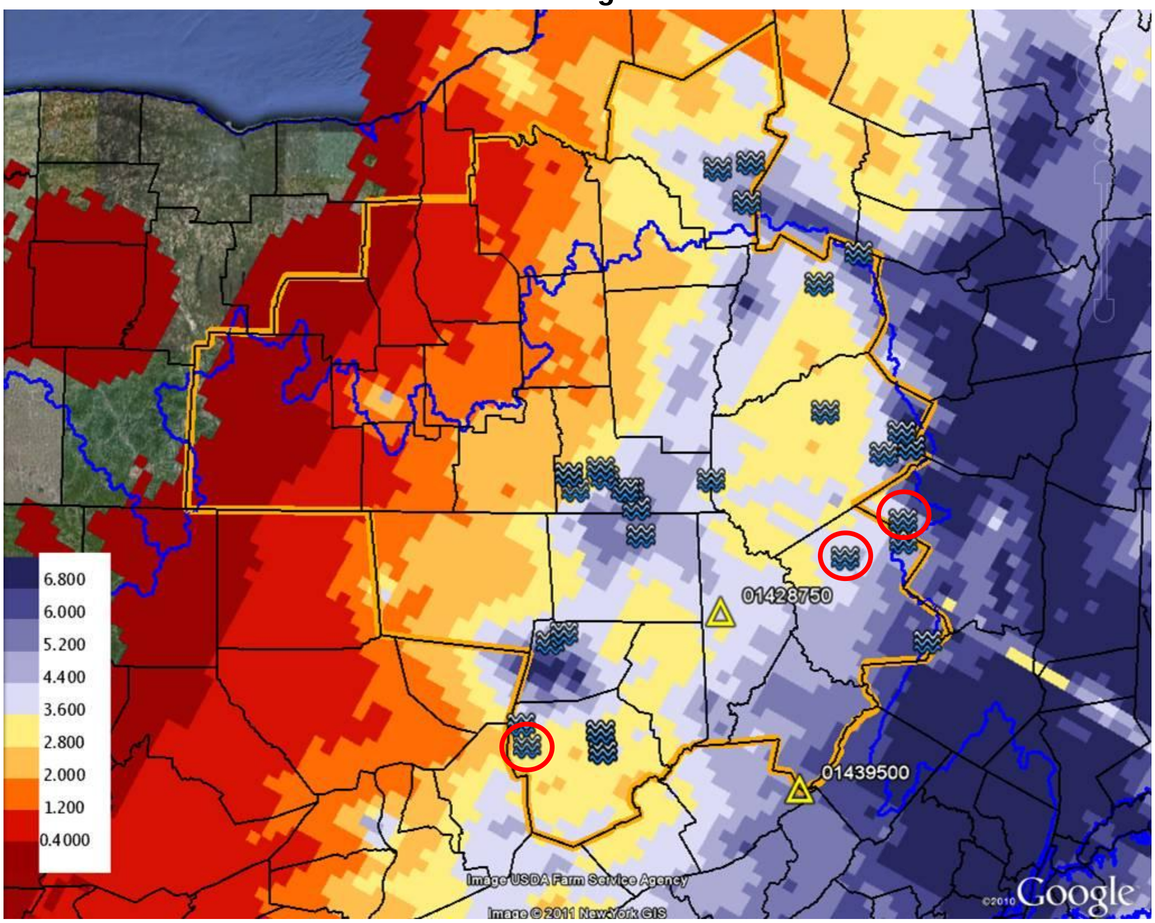
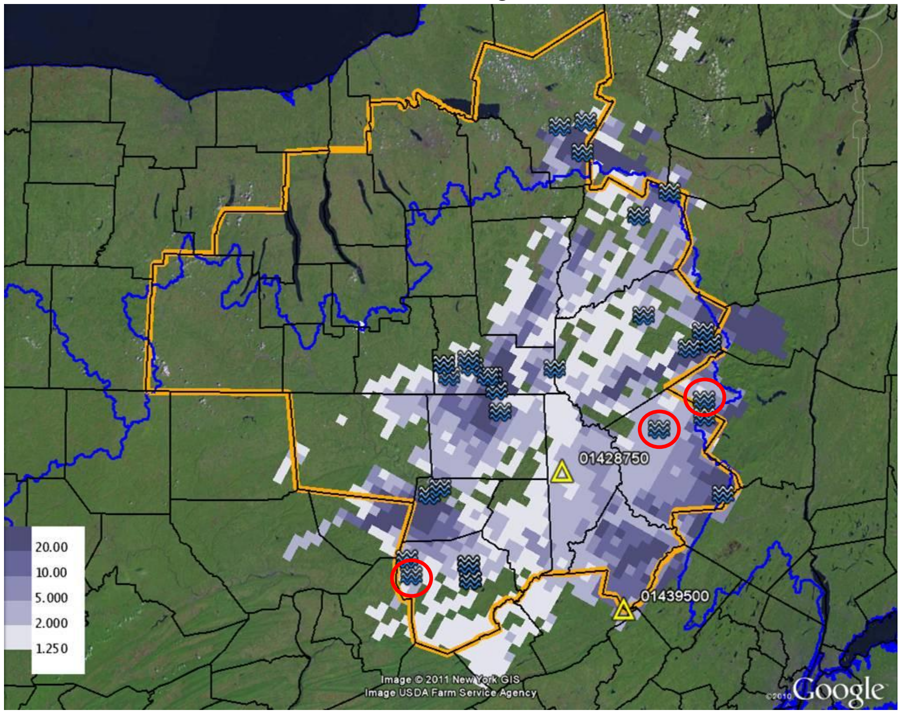


Figure 16. Accumulated precipitation from 11Z August 27th to 12Z August 29th, 2011 associated with flash flooding in the WFO BGM county warning area. The blue line represents the RFC boundary, the orange line represents the WFO boundary, the blue wave symbols represent observed locations of flash flooding, the yellow triangles represent USGS gauges, and the red circles represent unwarned storm reports.

Maximum DHM-TF Return Period (Calibrated, Analysis w/Cell-to-Cell Routing)
12Z 8/27/11 through 12Z 8/29/11



Maximum DHM-TF Return Period (Calibrated, Analysis w/o Cell-to-Cell Routing)
12Z 8/27/11 through 12Z 8/29/11

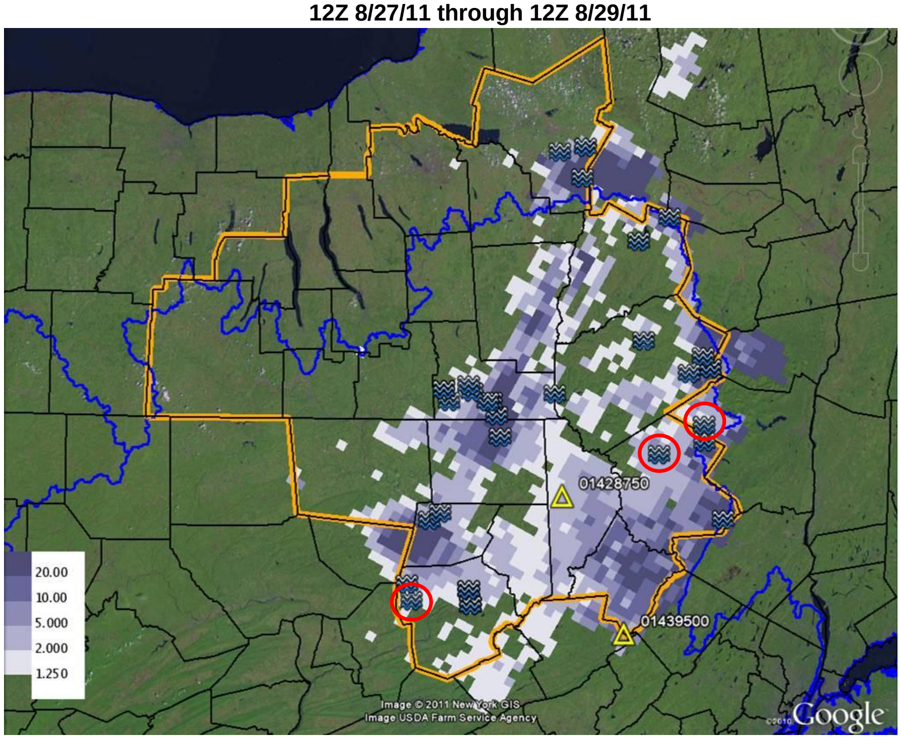


Figure 17 . Plots of maximum return period values during 8/27-8/29 2011 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangles indicate locations of USGS gauges used in verification, the red circle indicates reports of flooding that were not warned for, and blue wave symbols represent observed locations of flash flooding.

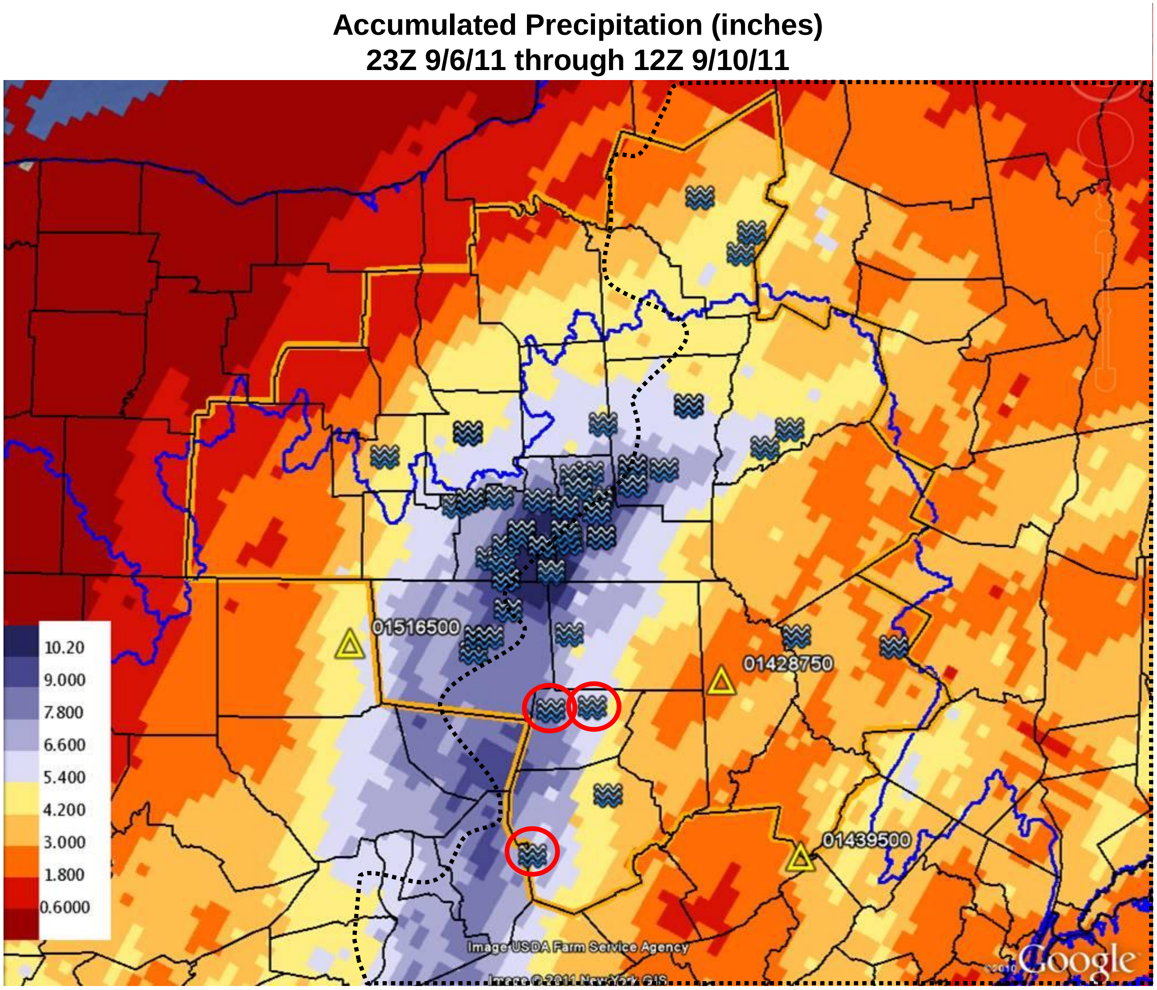
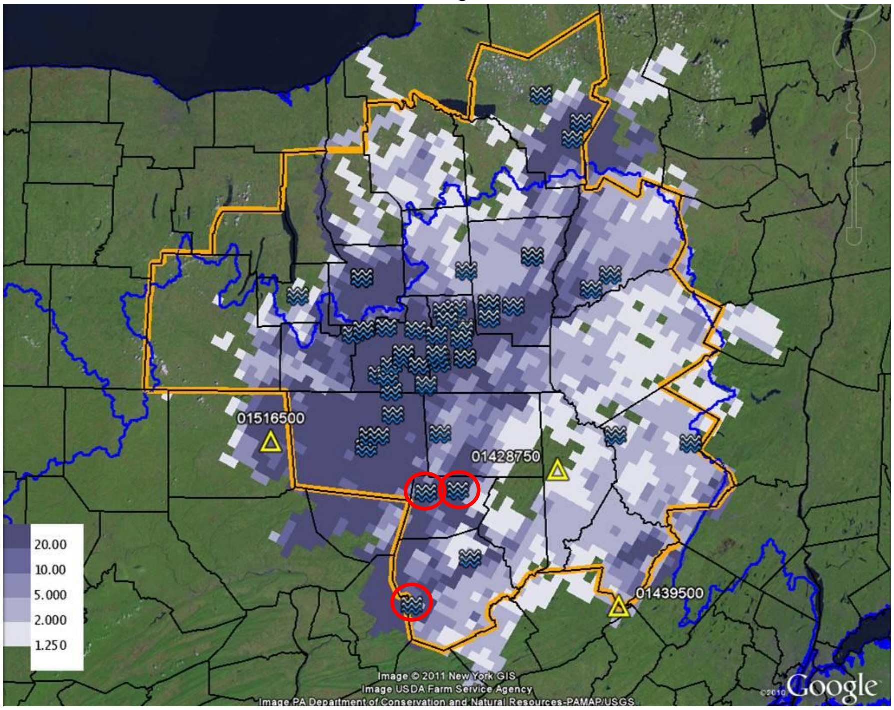


Figure 18. Accumulated precipitation from 23Z September 6th to 12Z September 10th, 2011 associated with flash flooding in the WFO BGM county warning area. The blue line represents the RFC boundary, the orange line represents the WFO boundary, the blue wave symbols represent observed locations of flash flooding, the yellow triangles represent USGS gauges, the red circles represent unwarned storm reports, and areas within the black dashed line received over 3 inches of rain from Hurricane Irene 8/27-8/29.

Maximum DHM-TF Return Period (Calibrated, Analysis w/Cell-to-Cell Routing)
00Z 9/7/11 through 12Z 9/10/11



Maximum DHM-TF Return Period (Calibrated, Analysis w/o Cell-to-Cell Routing)
00Z 9/7/11 through 12Z 9/10/11

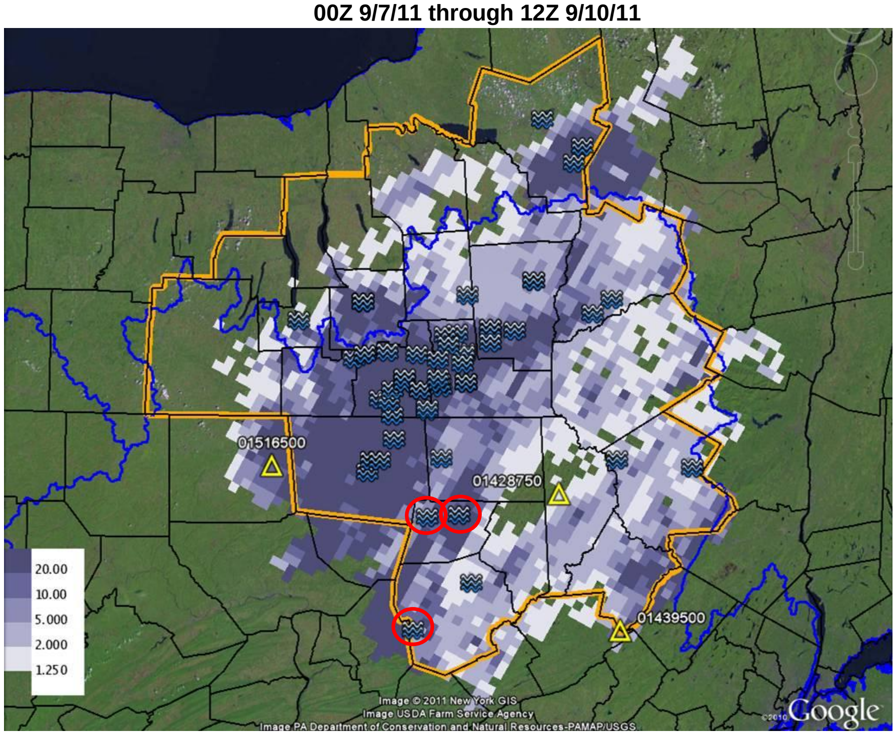


Figure 19. Plots of maximum return period values during 8/27-8/29 2011 event generated by DHM-TF with cell-to-cell routing (top) and without cell-to-cell routing (bottom). Triangles indicate locations of USGS gauges used in verification, the red circle indicates reports of flooding that were not warned for, and blue wave symbols represent observed locations of flash flooding.

Model Calibration and Snowpack Modeling

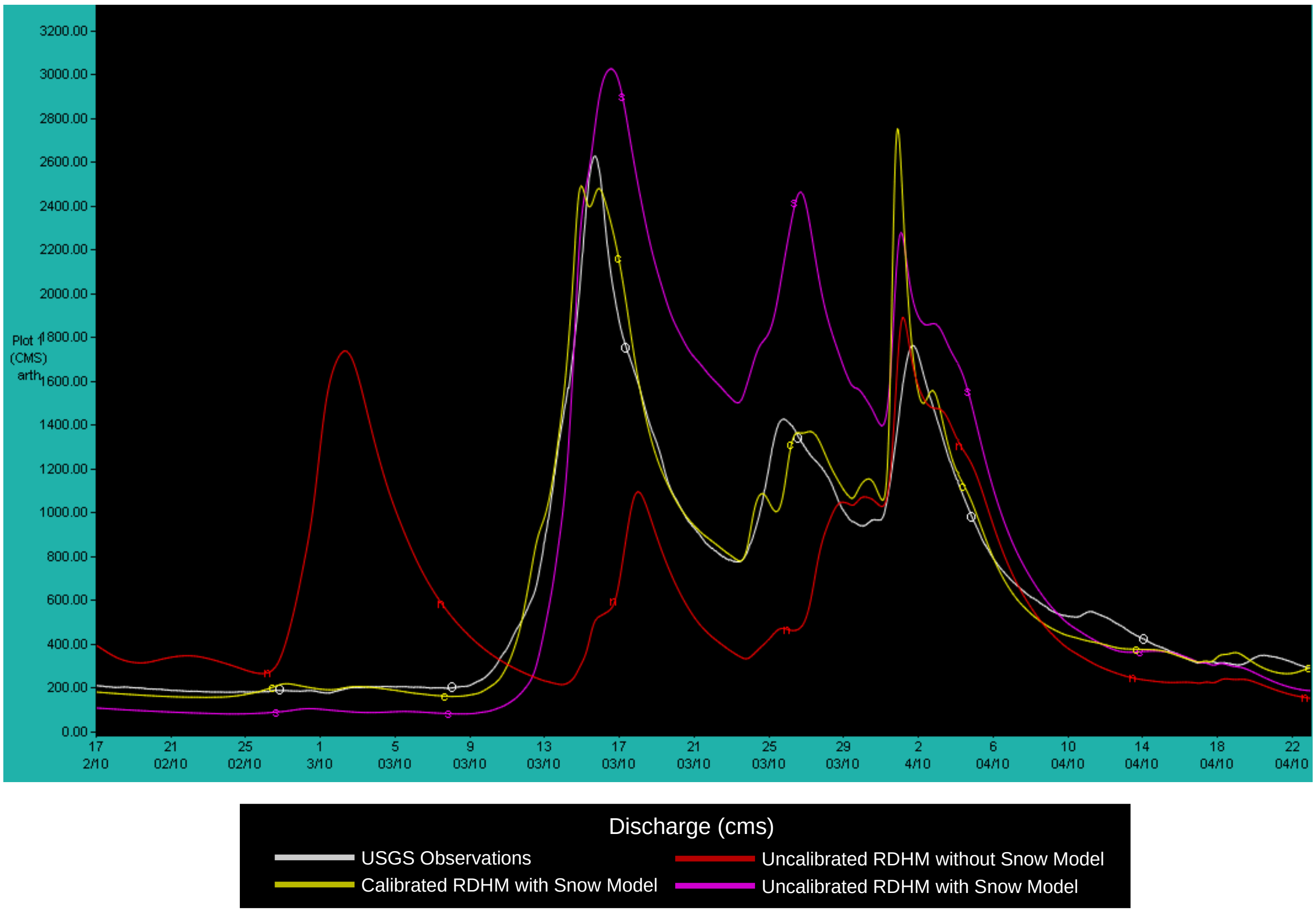


Figure 20. Observed USGS (white) and simulated HL-RDHM discharge using uncalibrated parameters and no snow model (Red), uncalibrated parameters and the Snow17 model (Purple), and calibrated parameters and the Snow17 model (Yellow) for Februrary 2010 through April 2010.

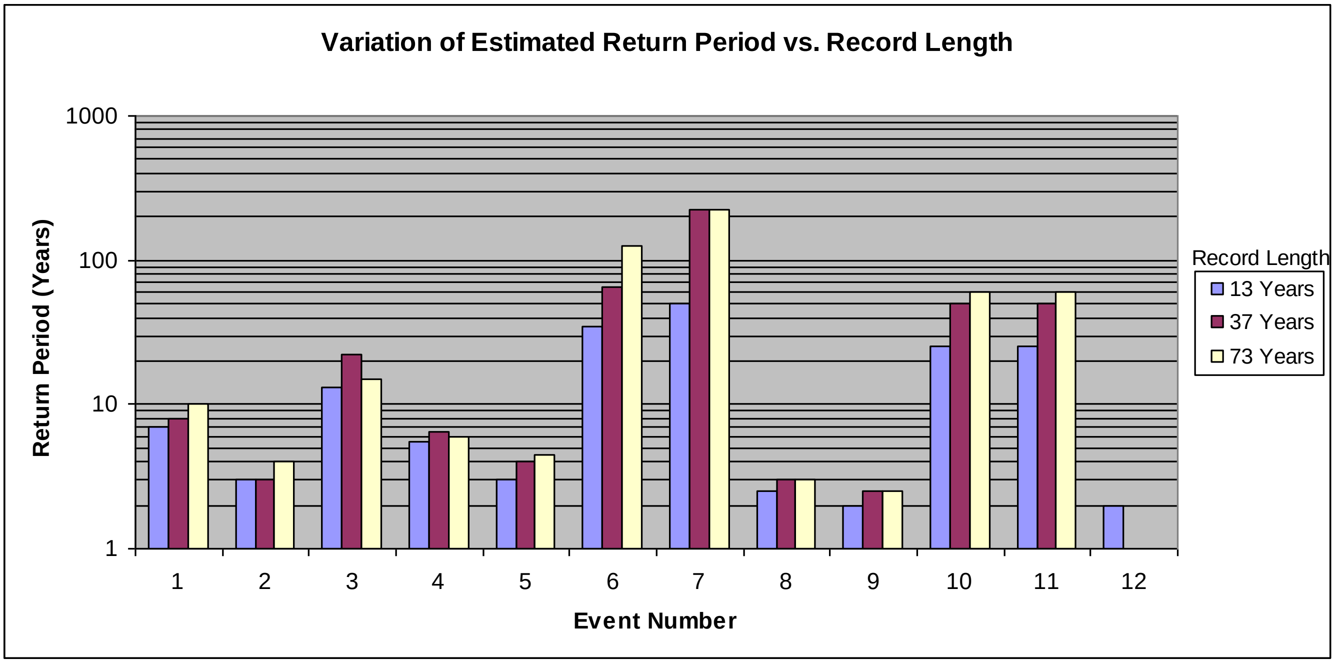


Figure 21. Return periods computed by applying Log Pearson Type III approach to 13, 37, and 73 years of USGS stream gauge data for 12 flood events.

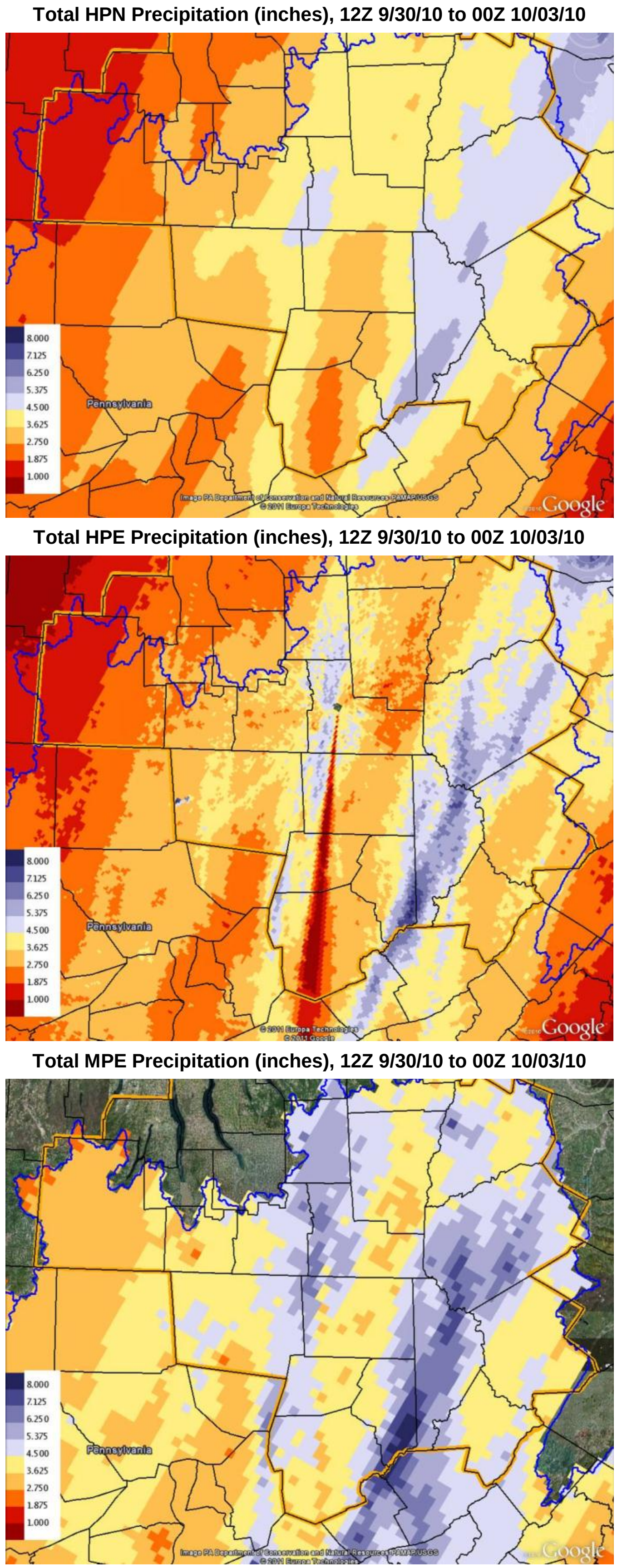


Figure 22. Total (top) HPN, (middle) HPE, and (bottom) MPE precipitation (in) over the MARFC portion of the BGM CWA for 12Z 9/30 to 00Z 10/03 2010.

APPENDIX A: Feedback from WFOs on performance and utility of DHM-TF with regard to selected events from this analysis report.

Subject: Baltimore/Washington WFO—August 6th-August 7th 2011 Rappahannock flood event, regarding uncalibrated output from DHM-TF
Source: Jason Elliott, LWX Service Hydrologist

"We were just looking at the DHM-TF output for this past weekend. As you may have seen in various accounts, there was a flash flood in Culpeper County, VA...followed the next day by people being rescued from the Rappahannock River in Fredericksburg and Falmouth. It looks like DHM-TF did a really nice job capturing both the initial Culpeper flash flood, and the ensuing water moving down the Rappahannock."

-8/9/2011

"Well...I'm impressed. The 15z file has a push of water of nearly 6000 cfs on the Rappahannock above the confluence. That alone (without any Rapidan water) equates to about 4.5 feet at the Fredericksburg gauge. So if we'd seen that in near-real-time, we would've had an idea that a multiple feet rise was coming. We might not have known the sharp rapidness of the rise, but we would've known there would be a decent water level increase and could have translated it downstream. The 17z value in the grid box at the Rapidan-Rappahannock confluence of 6721 cfs. At the gauge location, the model had 3157 cfs (observed was 5140 cfs, but it had been less than 3000 cfs just 30 minutes prior). The 21z file for the grid box at the gage location was 7220 cfs (actual gage value at 21z was 8850 cfs). At the confluence on the 21z file, it had 8224 cfs."

-8/19/2011

Subject: Binghamton WFO—August 27th-August 29th 2011 Hurricane Irene flood event, regarding calibrated output from DHM-TF
Source: Michael Evans, BGM Meteorologist

"Sorry for the long delay in getting back to you. I'm finally getting a chance to look over what you sent in some detail. My first thought on the Irene data is that there did not seem to be big differences in the routed vs. non-routed output for most areas. Both graphics through 12z look really good. One thing that I am still trying to understand is when these graphics would be available in real-time. For example, I see a slide that says "13z through 16z, zero precip input after 13z". Could this be available as early as 1330z? Or not until after 16z? Catastrophic flooding occurred in extreme eastern Delaware county by 1330z. Looking at the slide that says "zero input after 12z", I see 2 to 5 year returns in the no routing slide in that area, and some 5 to 10 year returns in the slide to the left with routing. By the next hour (zero input after 13z), I see a few pixels of 5 to 10 in the no routing slide to the right, and some 10 to 20 year returns in the routing slide to the left. So it looks like the routing may have been important for this event. I think that this is correct, as very heavy rain over the northern Catskills ran down into that area. Certainly I think that the data on slides 10 and 11 would have been useful to us in real time. There was also some big flooding in extreme eastern Sullivan county early this morning where we were late with a warning. I see some returns of greater than 10 years developing in both the routed and non-routed output in that area by slide 8 (no precip input after 10z), so I think that this would have been very useful to us in real time. Again, the key would be how timely the data would be. Is it correct to think that the graphic shown in slide 8 could be made available shortly after 10z?"

-11/7/2011

Subject: Binghamton WFO—September 7th-September 10th 2011 Tropical Storm Lee flood event, regarding calibrated output from DHM-TF
Source: Michael Evans, BGM Meteorologist

"Taking a look at the output from Lee. Amazing how fast the return periods increased from slide 15 (zero input after 12z) to slide 17 (zero input after 14z). This is around the time (12z-14z) that really significant flash flooding was developing in southwest Broome / southeast Tioga county, so looks really good. Also interesting to note that there were big differences between the return values on the routing vs. no routing slides along the Susquehanna river during the early afternoon. For example look at slide 19. To me this says that the flooding early in the afternoon in Owego (just west of Binghamton, right on the Susquehanna river), was a flash flood, with the routing component pretty insignificant through about 17z. After 17z the routing catches up, with greater than 10 year returns along the Susquehanna west of Binghamton by 18z. So, from a hydrologists point of view, it seems to me that these slides tell an interesting story on the impact of local run-off vs. routing for the area. Really neat! So my question is what do you want to do now. These graphics would clearly be useful to us, as long as they are timely. For example, I would like to see a graphic saying "zero precip input after 18z" available no later than 19z, preferably by 1830z."

-11/7/2011

Subject: Binghamton WFO— August 27th-August 29th 2011 Hurricane Irene flood event and September 7th-September 10th 2011 Tropical Storm Lee flood event, regarding calibrated output from DHM-TF
Source: Jim Brewster, BGM Senior Service Hydrologist

I concur on all of Mike's assessments to the model performance and the clarification on the date/time stamping. Thanks Brian, it helped a lot. I don't really have any further comments, but do have a suggestion as far as output goes. I REALLY like the KMZ output for display in GE. This format is crucial in assisting forecasters to drill down into the geographic areas of concern, whereas the PNG images are a bit more broad based. They are good for testing this model, but operationally, a focus toward the XML/KML type output, I think, is the way to go. It's also likely to be a more friendly format for AWIPS 2, as well. That said...GE in operations has been limping along here at our office for a variety of reasons. I wonder if it could be explored to drop the KMZ into a Google Maps API (or Open Layers API), which we could then easily link to our intranet pages and still have the basin zoom capability of GE. I know our ITO Ron has done this with other geocentric applications (e.g. special event weather notifications, daily observations, etc.). Due to varying API keys, this may have to be done more during the local WFO implementation, but it may be worth investigating on your end for inclusion into any set up instruction material. Keep up the good work.

-11/7/2011

APPENDIX B: Excerpt from NSSL evaluation of DHM-TF, GFFG, and FFG techniques.

Text and images directly taken from:

Gourley, J., Z. Flamiq, Y. Hong, and K. Howard, "Evaluation of past, present, and future tools for radar-based flash flood prediction", submitted, Hydrological Sciences Journal.

5. Summary and recommendations

This study provides a comprehensive evaluation and inter-comparison of tools that have been proposed or are presently used in operational agencies to predict flash flooding. The FFG approach has been in operation for decades in the USA NWS and is the primary tool for decision making in flash flood forecasting. This method, along with a newer gridded approach called GFFG, both derive rainfall thresholds that will result in flooding by running hydrologic models under multiple rainfall scenarios until bankfull conditions are reached at the basin outlet. Forecasters then monitor observed or forecast rainfall and consider issuing flash flood watches or warnings when theses rainfall thresholds are reached or exceeded.

The advent of high-resolution, radar-based rainfall estimates, computational resources to run distributed hydrologic models in real time, and GIS datasets to describe spatially variable land surface and soil characteristics has led to a new forward modelling approach to flash flood prediction called DHM-TF. This method is similar to FFG and GFFG in that it relies on static runoff thresholds to identify rare, flood producing discharges, but it uses observed or forecast rainfall to directly force the hydrologic model in real time. The runoff thresholds are computed by running the distributed hydrologic model over a sufficiently long period of time in which an archive of gridded rainfall observations are available. The first objective of this study was to evaluate each approach using a common observational database and the results are summarised as follows:

- The best overall skill as measured by a CSI of 0.39 occurred with DHM-TF at a return period of 2.2 years on gauged basins < 260 km^2
- DHM-TF was more skillful than FFG and GFFG over a range of return period flows from 1.9 to 4.1 years
- FFG was slightly more skillful than GFFG
- Independent reports of flash flooding from trained spotters confirm DHM-TF was more skillful than FFG and GFFG

These encouraging results regarding DHM-TF skill relative to operational tools motivated us to expand the observational database by increasing the number of study basins from 15 to 70 and lengthening the time period of evaluation from 24 to 81 months. The following points summarise the sensitivity of DHM-TF results on basins with different scales, hydroclimatological conditions, urbanization effects, and flood controls:

- DHM-TF had no skill on basins that have a significant contribution to runoff from snowmelt
- Regulation to natural flows including dams and modifications to channels reduced the skill of DHM-TF flash flood forecasts
- The skill of DHM-TF forecasts were found to decrease at an approximate linear rate with decreasing basin area
- The mean CSI for DHM-TF with optimised warning thresholds was 0.47 representing a 38% improvement over its calibrated FFG benchmark
- DHM-TF forecasts based on exceedance of simulated two-yr return period flows had a mean CSI of 0.32, which was 68% better than the uncalibrated FFG benchmark

Although DHM-TF generally outperformed the FFG and GFFG methods, there were several shortcomings found in this study that should guide future research. Specifically, surface precipitation phase should be taken into account using, for example, the USA National Severe Storms Laboratory's National Mosaic and Quantitative Precipitation Estimation (NMQ) product (http://nmq.ou.edu). There is a snowmelt module called Snow-17 available within the HL-RDHM modeling framework that can readily incorporate NMQ's surface precipitation phase to model the contribution to surface runoff from snowmelt. The performance of DHM-TF in small, urban basins needs improvement. There are channel routing parameters within HL-RDHM that can be modified to appropriately describe Manning's coefficients that have been significantly altered due to channelization of natural streams. In addition, all studied urban drainages had catchments < 46 km^2, which meant basin-scale processes were modelled with less than four grid cells. Future work should evaluate the potential improvements to DHM-TF skill by refining the model grid cell resolution and precipitation forcing to one km and five min, which is more commensurate with flash flooding impacts.

Additional design features of flash flood forecasting systems based on distributed hydrologic models should consider improving the estimation of state variables by assimilating soil moisture and streamflow observations. Forcing to the model can be improved by incorporating satellite data in multisensor rainfall products where radar coverage is inadequate, and forecast lead time will improve in conjuction with rapid-update stormscale modelling efforts that include radar data assimilation. Uncertainty estimation of each modelling component (i.e., forcing, states, structure, observations) is paramount and ensemble methods and derived products are recommended. Lastly, forecasts should consider a continuum of severity thresholds beyond the two-year return period flow and, more importantly, should be targeted on the specificity of the anticipated flash flood impacts.

The following figures from "Evaluation of past, present, and future tools for radar-based flash flood prediction" are not explicitly referenced in the Appendix text above, but rather in the main body of the journal article's text which is not included in this Appendix. They provide an informative graphical depiction of DHM-TF's performance alongside that of FFG and GFFG. Descriptions of the content of each figure are given in captions.

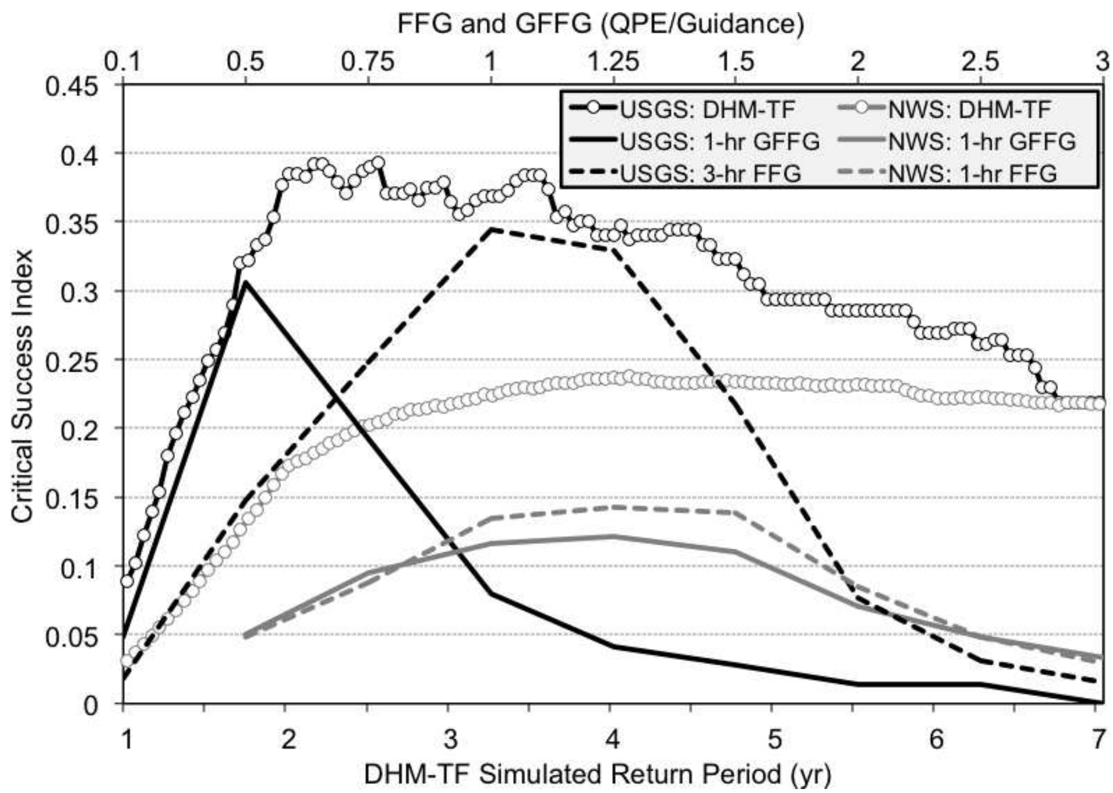


Fig. 1. Skill of FFG, GFFG, and DHM-TF method according to two-year return period flows from 15 USGS stream gauging stations and NWS trained spotter reports of flash flooding from 01 September 2006 to 31 August 2008 in south-central USA. Critical success index is computed for simulated return periods (primary abscissa) and for different ratios of exceedance of rainfall over FFG and GFFG (secondary abscissa).

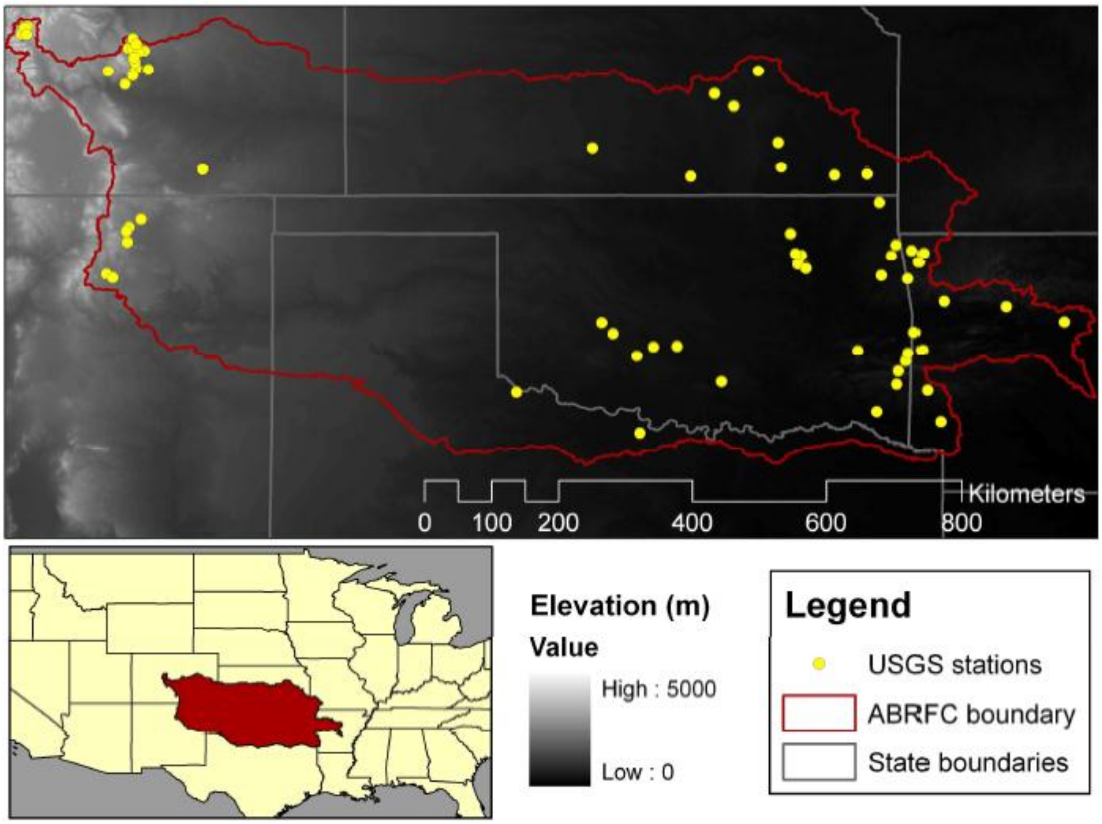


Fig. 2. Study domain showing the Arkansas-Red river basins outlined in red along with the 70 USGS stream gauge observations used in the study.

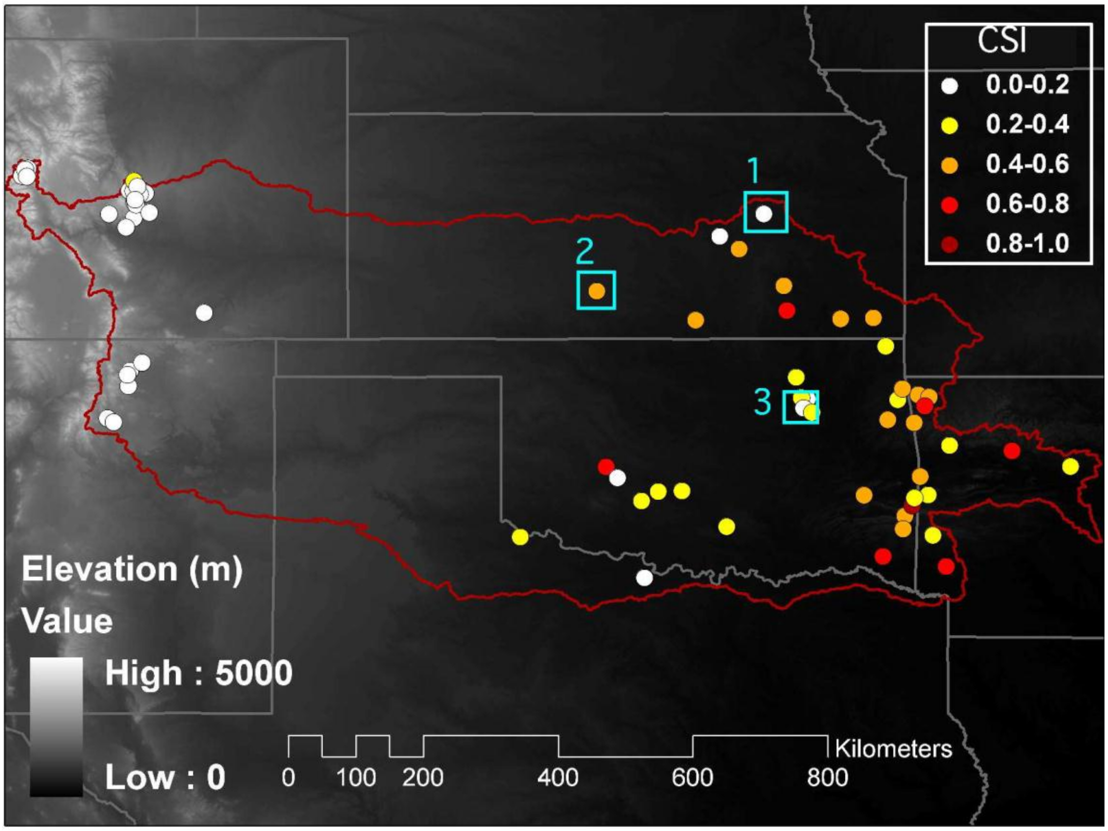


Fig. 3. Spatial distribution of skill from the DHM-TF method using streamflow observations at 70 USGS stations from 2003-2009. The three locations with cyan squares around them are shown in detail in Fig. 4.

Fig. 4. Satellite images obtained from Google MapsTM over USGS no. (a) 07179500, (b) 07144910, (c) 07164600, and (d) 07177650. The stream gauge locations are shown as yellow dots in each panel. Observations from the first two stations were deemed as regulated by the USGS whereas data for the third and fourth station were noted to be impacted by urban effects

.

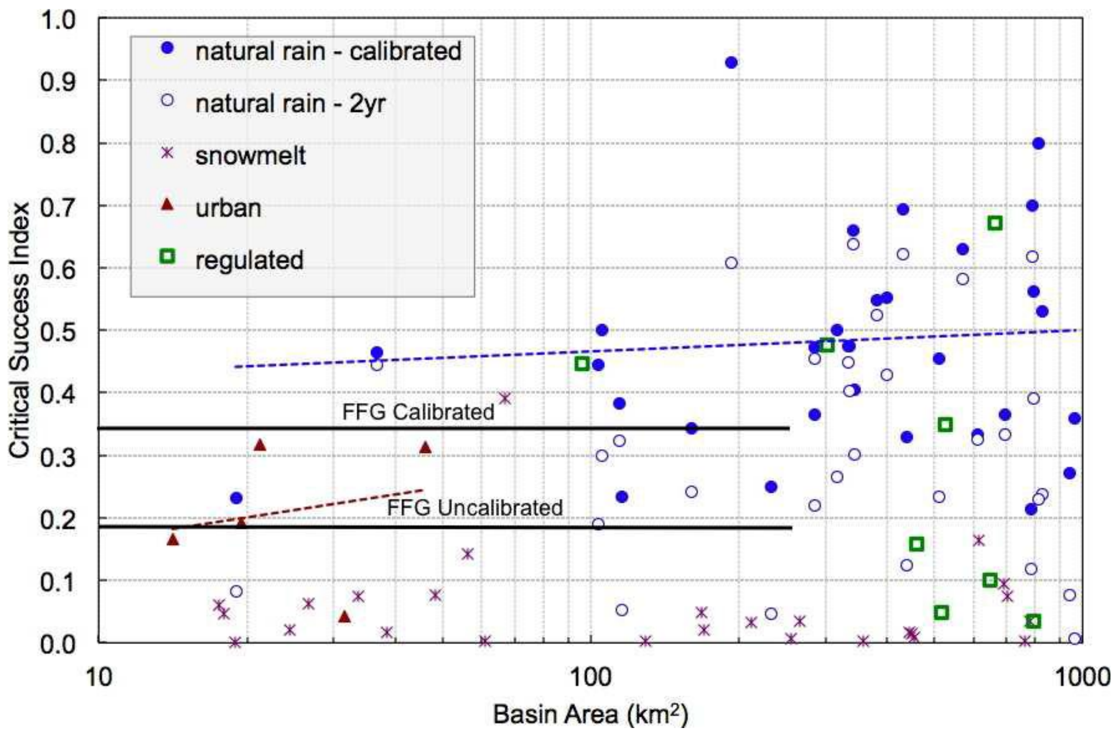


Fig. 5. Critical success index shown as a function of basin catchment area for USGS stations shown in Fig. 4. The USGS data records were interrogated and those having discharge influenced by snowmelt, urban effects, or man-made diversions or dams are distinguished in this analysis (refer to legend). The open circles are for CSI values for uncalibrated runoff simulations that exceeded their two-year return period flows whereas the filled circles are for simulations with calibrated runoff thresholds. The two black horizontal lines correspond to the calibrated and uncalibrated skill of FFG found in Gourley *et al.* (2012).

Gourley, J. J., Erlingis, J. M., Hong, Y. & Wells, E. B. (2012) Evaluation of tools used for monitoring and forecasting flash floods in the US. *Wea. Forecasting*, doi: 10.1175/WAF-D-10-05043.1.