A Real-time Multi-source Flash Flood Verification Database in Support of WPC Research and Operations

Brian A. Cosgrove (NOAA/NWS/OHD)
Faye Barthold, Thomas Workoff, Mark Klein, and David Novak (NOAA/NWS/WPC)
Wallace Hogsett (formerly NOAA/NWS/WPC)
Jonathan J. Gourley and Zac Flamig (NOAA/NSSL)

Obtaining a complete and accurate assessment of flash flood occurrences is critical for verifying and improving the operational excessive rainfall and experimental flash flood forecasts produced by the National Weather Service's Weather Prediction Center (NWS/WPC). Unfortunately, a single authoritative, comprehensive source of flash flood verification data does not currently exist. As such, WPC, with assistance from NOAA's National Severe Storm Laboratory (NSSL), leveraged three real-time CONUS-wide hydrologic data sources to create a new experimental, merged, real-time verification dataset. These data sources include NWS flash flood Local Storm Reports (LSRs), NSSL Meteorological Phenomena Identification Near the Ground (mPING) reports, and United Stage Geological Survey (USGS) stream gage measurements. While each of these datasets is not without weaknesses, they also feature complementary strengths.

National Weather Service LSRs are an official NWS product, provide relatively dense coverage and, in many cases, include rich descriptive language of the event. However, they can be subjective in nature and are dependent on people actually witnessing and reporting an event; darkness, low population density, and poor weather itself are factors that can limit the number of events observed. Event categorization, location and timestamp errors can also occur, as can long time lags in the submission of reports. Like LSRs, mPING reports are dependent on submission by end users, though in this case via a mobile app or a website. mPINGs suffer from similar categorization, coverage and quality control issues. Unlike LSRs, they do not differentiate between floods and flash floods, though NSSL examination of the mPING reports indicates that they are mainly flash floods. Even with these weaknesses, as with many crowd-sourced social media-type applications, there is strong potential for this data source to quickly grow over time as more people become mPING reporters.

The third and final component of the multi-source flash flood database centers on USGS stream gage reports. The only objective and automated source of the three, stream gage reports are underutilized for flash flood verification, and, to the best knowledge of the authors, this research effort represents the first CONUS-wide attempt at leveraging them for real-time verification of flash flooding. The dataset is comprised of stage and discharge data collected at approximately 10,000 automated USGS stream gages across the CONUS every 5-60 minutes. Gage coverage is dense in many areas of the country, and as an added strength, gages provide fully automated operation in all weather conditions. However, the reports are necessarily limited to stream locations and are sparse in some sections of the interior western US. Natural stream flow signals can also be contaminated by regulation (i.e., dams and diversions). Differing from the LSRs and mPINGs, these reports are not event-based and are not specifically aimed at isolating flash floods or floods. Rather, they are simply ongoing reports of the stage and discharge of the river at a particular location, whether during drought-, average-, or flood-type conditions. To extract natural flash flood event signals from these observations, real-time data from each USGS basin smaller than 2000 km^2 are passed through a series of sequential filters, which include checks for exceedance of minor flood stage or two year return period flow, rate of rise, and total stage change.

Stream gage observations are downloaded and processed via automated scripts alongside mPING and LSR data. Upon retrieval, reports from each data source are inserted into a searchable Postgres database. Latitude, longitude and timestamp values are stored for all three data sources, with additional attributes (e.g., descriptive event remarks, stream rate-of-rise) archived as available. Since direct comparison of point-type verification observations to areal-type flash flood forecasts is challenging, the combined data are plotted in both point-type and areal-type fashions. Underpinning this is the Practically Perfect (PP) analysis technique, which converts point observations to contoured areas and is used by the NWS Storm Prediction Center (SPC) for verifying severe weather forecasts (Hitchens et al., 2013; personal communication Israel Jirak, SPC). The goal of this approach is to produce a flash flood forecast map that resembles that which would be produced by someone with perfect knowledge of future flash flood events. While the PP approach is a relatively simplistic spatial approach and ignores basin boundaries, it is suitable for broad CONUS-scale verification applications and has proved valuable in verifying both WPC Flash Flood and Intense Rainfall Experiment (FFAIR) predictions and, in an experimental fashion, WPC excessive rainfall forecasts. Many other government, academic, and private agencies focus on flash floods, and it is expected that this database will prove useful to a wide variety of applications within those groups.

This presentation will explore the details of the database and will cover several verification case studies that leverage the flash flood information contained therein.