https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

StageIV-IRC – A High-resolution Dataset of Extreme Orographic Quantitative Precipitation Estimates (QPE) Constrained to Water Budget Closure for Historical Floods in the Appalachian Mountains

Mochi Liao[1] and Ana P. Barros[1]

1. Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, IL

Corresponding Author:

Dr. Ana Barros

E-mail: barros@illinois.edu

Phone: +1 217-333-8038

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Abstract

Quantitative Flood Estimation (QFE) in complex terrain remains a grand challenge in operational hydrology due to the lack of accurate high-resolution Quantitative Precipitation Estimates (QPE) at spatial and temporal resolutions needed to capture the variability of orographic precipitation and where radar-based QPE are available there are significant biases due to the geometry and constraints of radar operations. Here, we present a high-resolution (i.e. 250m, 5minute-hourly) QPE dataset for 215 extreme (flood-producing) events from 2008 to 2024 for 26 gauged basins in the Appalachian mountains constrained to meet basin-scale water budget closure through inverse rainfall-runoff modeling to correct the Next Generation Weather Radar (NEXRAD) Stage IV analysis (4 km resolution, hourly) using a fully-distributed uncalibrated hydrological model that leverages recent advances in hydrologic modeling in mountainous regions (e.g. improved river routing and initial soil moisture estimation). The corrected Stage IV analysis QPE is referred to as StageIV-IRC (Inverse Rainfall Correction). Previously, a subset of this dataset informed the construction of a generalized QPE error prediction model and providing physics insights into orographic QPE uncertainties for various radar-based QPE products in complex terrain. The unique advantage of the StageIV-IRC QPE is it is in agreement with ground-based rainfall measurements and achieves water budget closure at the storm-flood event scale within observational uncertainty of streamflow observations when it is used to drive hydrological simulations of historical floods, that is the golden standard in hydrological modeling. The QPE dataset is publicly available at: https://doi.org/10.5281/zenodo.14028867 (Liao and Barros, 2024).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

1. Introduction

Over the past few decades, extreme precipitation has become an increasingly important research topic due to its social, economic, and environmental impacts (Alimonti et al., 2022; Wernberg et al., 2013). Studies show that both total annual precipitation and extreme precipitation events have increased in the US and in other parts of the world during the last century (e.g. Milly et al., 2002), often resulting in floods (e.g. Pielke and Downtown, 2002), and flash floods in the context of complex terrain due to steep slopes (e.g. Schumacher, 2017; Czigány et al., 2010). Flash floods are characterized by fast rainfall runoff responses on the scale of a few hours (< 6 hours) after extreme precipitation events, with watershed areas often ranging from a few tens to hundreds of square kilometers (Borga et al., 2010; Lumbroso and Gaume, 2012). As one of the deadliest natural hazards, flash floods often are associated with landslide events (e.g. Gupta et al., 2016; Deijns et al., 2022) and cause severe life loss and property damage (Saharia et al., 2017; Špitalar et al., 2014) such as recently in the USA and in Spain. An estimated 94 million people are affected worldwide every year since 2000 (Guha-Sapir et al., 2018; Wu et al., 2020) and the average annual economic loss in the U.S. subjected to freshwater flooding exceeds USD $8 billion (Wing et al., 2018). A recent noteworthy example is Hurricane Helene in late September 2024, with a death toll over 200 in the Southeast U.S. Despite extensive studies on improving flash flood simulations in small headwater basins, hydrological skill scores (e.g. Kling-Gupta Efficiency or KGE) remain poor at event scales largely due to significant difficulties involved in estimating highly localized orographic precipitation in complex terrain (e.g. Mtibaa and Asano, 2022; Ghimire et al., 2022;

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Arulraj and Barros, 2021; Barros and Arulraj, 2020; Zhang et al., 2012; Huffman et al., 2007; Andrieu et al. 1997).

With global warming and projected heavier precipitation and higher probability of floods in climate change hotspots and significantly modification of the hydrologic cycles in complex terrain (Nijssen et al., 2001), substantial efforts have been made understand and quantify precipitation uncertainties in the mountains and nearby lowlands (Pepin et al., 2022). Increased warming causes reduction in snowpack during winter/spring impacting seasonal streamflow patterns (Moraga et al., 2021; Saunders et al., 2008; Arnell 2003). Increasing probabilities of severe summer thunderstorms (e.g., Brooks 2013) are already one of the largest contributors to global losses (> $10 billion USD a year, Allen 2018). The urgent need to improve precipitation estimation and forecasting, particularly for warm-season flood-producing precipitation events, is unprecedented.

Current approaches involved in precipitation measurement and Quantitative Precipitation Estimation (QPE) broadly include in-situ point-scale observations using rain gauges and disdrometers, and remote spatial observations using ground-based radar and space-based sensors. In complex terrain, there is often a scarcity of in-situ measurements due to difficult access. Radar-based QPE products are plagued by uncertainties from various sources (e.g. ground clutter effects, retrieval uncertainties, radar viewing geometry, (Villarini and Krajewski, 2010; Arulraj and Barros, 2021; Barros and Arulraj, 2020; Zhang et al., 2012; Kreklow et al., 2020; Huffman et al., 2007; Andrieu et al. 1997; Creutin et al., 1997; Durden et al., 1998)). Numerical weather prediction (NWP) is an alternative to measurement. However, QPE produced by NWP models are characterized by large uncertainties when evaluated against rain gauges (e.g. Zhang and Anagnostou, 2019) leading to large runoff deviations from streamflow measurements when used

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

as inputs to hydrological models (e.g., Tao et al., 2016; Weiland et al., 2015; Diomede et al., 2008; Kobold and Suselj 2005). Due to these uncertainties and errors involved, significant efforts have been devoted to improving QPE via various approaches in the past few decades such as radar-raingauge data fusion (e.g. McKee and Binns, 2016; Goudenhoofdt and Delobbe, 2009; Delrieu et al., 2014; Nanding et al., 2015; Sideris et al. 2013; Schiemann et al. 2011; Berndt et al. 2014), radar reflectivity and retrieval corrections (e.g. Vignal et al., 2000; Shao et al., 2021; Dinku et al., 2002) and data assimilation of various radar products (e.g. Rafieeinasab et al., 2015; Wehbe et al., 2020). Rain gauge and disdrometer measurements often provide the ground truth for these QPE optimization approaches (e.g. Harrison et al., 2000; Shao et al., 2021; Fulton et al., 1998). The 'ground truth', however, has its own error (e.g., wind effects around the gauge orifice, Kochendorfer et al., 2017), and fails to capture highly localized orographic enhancement (e.g. Prat and Barros, 2010; Gentilucci et al., 2021; Buytaert et al., 2006). Gauge-radar fusion often relies on geostatistical assumptions that are primarily distance-based (e.g. Areerachakul et al., 2022; Cassiraga et al., 2021; Wang et al., 2020; Maggioni and Massari, 2018), lacking the full picture of basin topography which has a regulating role in orographic QPE. Although there is no definite consensus on guidance for the placement of rain gauges in mountainous basins (e.g. Suri and Azad, 2024), Barros et al. (2004), Prat and Barros (2010), Barros (2013), Barros et al. (2014), and Duan et al. (2015) provide consistence guidance regarding the importance to install precipitation networks on the topographic envelope and at regular intervals along ridges and adjacent valleys using examples from the Central Himalayas, the Central Andes and the Southern Appalachians.

To address this long standing QPE challenge in complex terrain mainly, a general QPE error quantification framework was developed leveraging widely available United States Geological Survey (USGS) streamflow observations at the outlet of headwater basins in complex

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

terrain, consisting of 2 distinct pathways: 1) rain gauge bias correction, and 2) grid-level QPE correction constrained to watershed-scale water budget closure. The first pathway includes rain gauge bias corrections at gauge locations both at diurnal and climatological scale, and geostatistical distribution of raingauge biases across a basin. The second pathway includes an innovative inverse QPE correction method by backward propagating runoff uncertainty using a hydrological model via streamlines to precipitation at storm-event scale (i.e. Inverse Rainfall Correction or IRC, Liao and Barros, 2022 or LB22). It is also worth noting that runoff uncertainty in hydrological modeling stems from various sources, generally including forcing uncertainties, land surface condition uncertainties and model specific uncertainties (e.g. Clark, et al., 2008; Beven and Binley, 1992).

LB22 found that initial soil moisture uncertainty can prevent the IRC framework from achieving water budget closure because large initial condition errors cause significant travel time uncertainties. Soil moisture is considered a particularly important factor among soil properties due to its significant role in regulating runoff generation, hence dramatically altering flood timing and magnitudes (e.g. Marchi et al., 2010; Penna et al., 2011; Yin et al., 2022; Vivoni et al., 2007), and soil moisture can change very fast at sub-daily time scales changing saturation to nearly wilting point levels depending on soil texture and land-use (Grillakis et al., 2016). Initial soil moisture conditions can therefore determine whether a rainstorm produces a major flash flood or not (Zehe and Blöschl, 2004, Komma et al., 2007). However, due to the lack of in situ soil moisture sensors and reliable high resolution soil moisture products, only a limited number of previous studies in the peer-review literature focused on soil moisture uncertainty and the impact on flood simulation (e.g. Laiolo et al., 2016; Zappa et al., 2011; Tao et all. 2016; Silvestro et al., 2019; Silvestro and Rebora, 2014; Uber et al., 2018). Liao and Barros (2024b) developed an Initial Condition

Correction (ICC), based upon travel time theory, which is consistent with the general IRC framework, demonstrating large improvements in initial soil moisture estimation. Note that when implementing the IRC and ICC, we are using a fully distributed physics-based uncalibrated hydrological model (i.e. Duke Coupled Hydrological Model, DCHM) that has been used for over 25 years with great success in the Southern Appalachians (e.g., Tao and Barros, 2013, 2014, 2016), and consequently uncertainty from the model and model parameters is assumed to be negligible. Hydrological model parameters have impact on rainfall-runoff response, but they are generally only of secondary importance compared to the precipitation proper and initial soil moisture conditions particularly in small basins (e.g. Mockler et al., 2016; Dobler et al., 2012; Zappa et al., 2011). Liao and Barros (2024b) also demonstrate that QPE uncertainty dominates runoff uncertainty for extreme precipitation events compared to initial condition uncertainty, and initial conditions only begin to play an important role for less extreme events particularly early in the event prior to the rapid rise of the hydrograph, which is consistent with previous studies where initial soil moisture uncertainty usually has a decreasing impact on runoff uncertainty as precipitation continues (e.g., Figure 8 in Iwasaki et al., 2020).

In this work, IRC and ICC are coupled into one framework (referred to as the coupled IRC-ICC framework) to construct a high resolution QPE dataset aiming to close the water budget at the scale of storm-flood events along the latitudinal range of the Appalachian Mountains across diverse hydroclimatic and geomorphic regions. The coupled IRC-ICC framework is applied to 26 headwater basins and 215 flood-producing events from 2008 to 2024 using Next Generation Weather Radar (NEXRAD) StageIV dataset as original inputs, at a spatial and temporal resolution of 250 m and 5 minutes respectively, and the improved post IRC-ICC QPE data (i.e. StageIV-IRC) are made available in this study.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

The manuscript is organized as follows. The data sources, and QPE error quantification framework which consists of raingauge bias correction and the coupled IRC-ICC framework, are detailed in Section 2. Section 3 presents this new dataset (StageIV-IRC) along with data assessment from various aspects. Section 4 discuss the potential application of this new dataset and future work. Section 5 provides the access to the dataset and summary of the work.

2. Data and Methodology

2.1 Radar QPE StageIV

The NCEP/EMC (Environmental Modeling Center) StageIV is a QPE product from the National Weather Service (NWS) derived from the regional hourly and 6-hourly multisensor (radar + NWS raingauges) precipitation analyses (MPEs), which is further improved with new analyses from River Forecast Centers (RFCs) over the conterminous United States (CONUS) (Lin and Mitchell, 2005). In mountainous regions, StageIV datasets suffer from the blockage of complex terrain, resulting in ground clutter, overshooting and retrieval uncertainties, demonstrating significant biases and errors in rainfall detection. In support of the NASA's Precipitation Measurement Missions (PMM) program ground validation (GV) activities (Prat and Barros, 2010a), a dense network of raingauges was installed in the Southern Appalachians in 2007 and has been well-maintained since 2007 (Barros at al., 2014). In this study, the raingauge observations from the Southern Appalachians are used to correct StageIV.

2.2 GV Raingauge Observations

A high resolution raingauge network consisting of 34 tipping bucket raingauges has been maintained in the Pigeon River basin, over the ten-year reference period 2007-2018, during and

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

immediately after IPHEx (Integrated Precipitation and Hydrology Experiment or IPHEx, Barros et al. 2014). A map of the raingauge network is shown in Figure 1 with each raingauge labelled with a number, and the detailed locations of these gauges are documented in Table 1. Besides spatial representativeness errors related to the setup of the raingauge network as stated earlier, common errors include funnel wetting, pipe clogging and turbulent winds near gauges (e.g. Wang and Wolff, 2010). The raingauge network is regularly visited and maintained for at least three times a year including on-site cleaning and calibration. In this work, we use these rainfall measurements to adjust hourly StageIV QPE. In-situ rain-gauge data are publicly and available and can be found at http://dx.doi.org/10.5067/GPMGV/IPHEX/GAUGES/DATA301.(Barros et al., 2017) In addition to raingauges, a network of Parsivel disdrometers was installed during the IPHEx EOP (Extended Observing Period, 2013-2014), with each disdrometer location denoted by the letter P in Figure 1. Due to the limited duration of the disdrometer measurements, the disdrometer data were used only for the purpose of independent evaluation. Note that raingauges are placed mostly on ridges while disdrometers are generally located on the hillslopes and lowlands.

<Figure 1 here please>

2.3 Methodology

The methodology of this work includes 3 major components: 1) raingauge bias correction, 2) grid-scale QPE correction by closing the water budget using streamgauge measurements, and methods involved in 3) basin and event selection procedures, and model setup.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

2.3.1 Raingauge Bias Correction

A schematic drawing of the raingauge bias correction framework to derive gauge-improved QPE (named $\text{StageIV}_{\text{DBKC}}$) is summarized in Figure 2.

<Figure 2 here please>

First, to make meaningful comparison between StageIV and raingauges in space, a fractal downscaling algorithm is used to generate high spatial resolution $\text{StageIV}_{\text{D}}$ at 1km from the original StageIV product (4 km resolution). After downscaling, bias correction at event scale and ordinary kriging are applied consecutively to modify the $\text{StageIV}_{\text{D}}$ to $\text{StageIV}_{\text{DB}}$ at hourly time-scale. Subsequently, $\text{StageIV}_{\text{DB}}$ data are evaluated against the raingauge climatology from 2008 to 2017 to reduce any remaining biases conditional on weather regime, and climatological biases are geostatistically interpolated using the ordinary Kriging method. The resulting dataset is named $\text{StageIV}_{\text{DBKC}}$ (abbreviated as $\text{STIV}_{\text{DBKC}}$).

2.3.2 Fractal downscaling

Aiming to derive high resolution QPE datasets in complex terrain, the assumption of self-similarity is imposed. The Hurst coefficient *H*, fractal dimension *D*, and the spectral exponent β are described as the following:

$$D = \frac{7-\beta}{2} \quad (1)$$

$$H = \frac{\beta-1}{2} \quad (2)$$

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

The parameter β describes rainfall statistics at different scales, and it is estimated as the slope of the power spectral density curve in 2D Fourier domain of the rainfall field (log-log plot). The power spectral density Z(u,v) in the 2D Fourier domain is :

$$Z(u,v) = \left(\frac{L}{N}\right)^2 \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} z(x,y)\, exp\left[-\frac{2\pi i}{N}(ux+vy)\right] \tag{3}$$

where u and v represent the transform of x and y in Fourier domain, N is the total number of grid points in each direction, and z(x,y) is the rainfall field. Additionally, the spectral density at wavenumber k = 1 is defined as the roughness factor, that is the variance of the field. The Hurst coefficient describes the auto correlation strength (range from 0 to 1) with higher values of H implying higher auto-correlation, that is persistence. The mean power spectral density in 2-D Fourier domain is given:

$$S_j = \frac{1}{L^2 N_j} \sum_{1}^{N_j} |Z(u,v)|^2 \tag{4}$$

where N_j is the number of coefficients that satisfy the condition $j < \sqrt{u^2+v^2} < j+1$. The mean power spectral density has a power-law relationship with wave number k, and k is defined as below:

$$k = \frac{2\pi}{\sqrt{u^2+v^2}} \tag{5}$$

$$S \sim k^{-\beta-1} \tag{6}$$

where β is the spectral exponent calculated as the slope of power density spectrum. Assuming the rainfall fields are self-similar, then the information at fine resolutions can be derived from the information at coarser resolution. This is accomplished using a Brownian surface (Z_b, H=0.5) at the desired fine resolution as spatial support for the interpolation, which is modified in the Fourier domain (Z_D) to replicate the distribution of energy slope of the spectral slope and roughness factor according to Bindlish and Barros (1996):

$$Z_D(u,v) = \frac{Z_b(u,v)}{k_r^{(\beta-\beta_b)/2}} exp\left[\frac{1}{2}\left(S_{r,1} - \frac{\beta+1}{\beta_b+1} S_{r,2}\right)\right] \tag{7}$$

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

where β, β_b, $Z_D(u,v)$ and $Z_b(u,v)$ are respectively the original rainfall fields spectral exponent, Brownian surface spectral exponent, interpolation surface in Fourier domain and Brownian surface, respectively; k_r is the wavenumber and $S_{r,1}$ and $S_{r,2}$ are the roughness factor of the original rainfall fields and Brownian surface. Due to the non-uniqueness of Brownian surfaces, multiple replicates of interpolation surfaces Z_D can be obtained. In this study, an ensemble of ND interpolation surfaces is derived, thus ND rainfall fields at finer resolution preserving the same rainfall statistics at coarse resolution are generated. Following Nogueira and Barros (2015), here ND=50 and the correction steps described in Figure 2 are applied to the ensemble mean of the downscaled rainfall fields.

2.3.3 Bias Correction

The *first* phase of bias correction is carried out at event scale: a linear regression is established between raingauge measurements and collocated downscaled radar pixel estimates using the following formula:

$$R_g^t(i_g, j_g) = \kappa R_r^t(i_g, j_g) + \varepsilon \quad (8)$$

where R_r represent radar measurements, R_g represent raingauge observations, κ and e are respectively the slope (conditional bias correction) and the intercept (systematic bias correction). For each hour, collocated pairs of StageIV_D estimates and raingauge observations within a radius of 5 km centered on the StageIV_D pixel were identified as long as more than two raingauges measure non-zero rainfall. Regional least-square linear regression was applied subsequently to all StageIV_D pixels within ±1-σ deviation of the regional regression line at hourly time-scale by assuming homogeneity of variances or homoscedasticity.

The *second* phase of bias correction is done at climatological scale: aiming to reduce systematic radar errors caused by retrieval uncertainties and viewing geometry in complex terrain, demonstrating strong diurnal (time of day) and seasonal (weather regime) error dependencies when

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

comparing against 10-year raingauge observations due to miss detection of shallow rainfall related to radar overshooting in the Southern Appalachian (e.g. Wilson and Barros, 2014; Duan and Barros, 2017; Arulraj and Barros, 2017). For this purpose, the following corrections were added for rainfall below and above a threshold X, where X=2mm/hr in the Pigeon River Basin. When raingauge measurements are less than 2mm/hr and Stage IV_D estimates are 0, the $StageIV_D$ value was replaced by the raingauge observations, here termed Light Rainfall Correction (LRC). Furthermore, for each hour, nil $StageIV_D$ estimates where raingauge measurements are greater than X=2mm/hr were identified and replaced by the mean of the corresponding collocated raingauge measurements, hereafter Mean Rainfall Correction (MRC). Finally, for localized precipitation (i.e. only two or fewer rainguages detect nonzero rainfall) normally associated with isolated convective activity, the anomalies calculated as the differences between the $StageIV_D$ and the local raingauge measurements were linearly distributed among the surrounding 25 pixels (5 pixel window centered at the $StageIV_D$ pixel)– Convective Rainfall Correction (CRC).When more than 2 raingauges measured rainfall, then the anomalies for each pixel were spatially distributed using ordinary Kriging as described below – Global Rainfall Correction (GRC).

2.3.4 Ordinary Kriging

The Ordinary Kriging (OK) is a linear weighted geostatistical estimator that interpolates values of a variable at a specific location using weights derived from spatial covariances aiming to minimize prediction variance. In our case, the rainfall differences among raingauge measurements and $StageIV_{DB}$ at all raingauge locations were calculated and denoted as $G(x_i)$ at gauge location i. To produce estimates at any location within the study domain, a continuous model describing spatial covariance structure of the data is necessary. A commonly used semi-variogram model is the spherical model, exhibiting linear behavior at the origin. A review and comparison of different types of semivariogram models can be found (e.g. Li and Heap 2008; Oliver and Webster, 2014; Zimmerman and Zimmerman, 2012). Bohling (2005)

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

analyzed the differences of several commonly used semivariogram models and pointed out that, given the same variogram parameters (nugget, sill and range), spherical models reach to the maximum for comparatively shorter spatial lags, and thus are suitable to capture strong spatial dependencies over short distances as in the case of orographic precipitation (see also McBratney and Webster, 1986, for detailed description of spherical model):

$$\gamma(h) = C_0 + (C - C_0)\left(\frac{3h}{2d} - \frac{1}{2}\left(\frac{h}{d}\right)^3\right) \quad \text{if} \quad 0 \leq h \leq d \tag{9.1}$$

$$= C \qquad \text{if} \quad h>d \tag{9.2}$$

$$\gamma_{0i} = \frac{1}{N_A}\sum_{k=1}^{N_A} \gamma_{ki} \tag{9.3}$$

$$\gamma_{00} = \frac{1}{N_A}\sum_{k=1}^{N_A}\sum_{l=1}^{N_A} \gamma_{kl} \tag{9.4}$$

where d is the range, h is the lag, N_A is the number of available gauge locations, C and C_0 are the sill and nugget values. Neglecting local variability and measurement error at the spatial scale of the downscaled radar and raingauge (point) measurements, the nugget is constant and equal to zero (Diggle & Ribeiro, 2007). The rainfall difference at a target point x_0 $Z_{ok}^*(x_0)$ is calculated using a weighted linear combination of all available differences with constraints of unbiased estimator

$$Z_{ok}^*(x_0) = \sum_{i=1}^{n} \lambda_i^{ok} G(x_i) \tag{10.1}$$

$$\sum_{i=1}^{n} \lambda_i^{ok} = 1 \tag{10.2}$$

Optimal weights can be obtained by solving following equation by employing Lagrange multiplier μ:

$$\begin{pmatrix} \gamma_{11} & \cdots & \gamma_{n1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{1n} & \cdots & \gamma_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1^{\mathrm{OK}} \\ \vdots \\ \lambda_n^{OK} \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_{01} \\ \vdots \\ \gamma_{0n} \\ 1 \end{pmatrix} \tag{11}$$

In this study, OK distributes spatially the differences between available raingauge measurements and radar data, resulting in the generation of $\mathrm{STIV}_{\mathrm{DBKC}}$ dataset.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

2.3.5 Ordinary Kriging

Standard performance metrics (McBride and Ebert 2000; Wang, 2014) including false alarm rate (FR), probability of detection (PD), threat score (TS), and Heidke skill score (HSS), as well as bias, and the root-mean-square error (RMSE) are used to evaluate the corrected downscaled hourly rainfall. An instance when both radar QPE and raingauge observation exceed a specified rain rate threshold is a hit (H); when observation matches the criterion and radar QPE does not, it is classified as a miss (M), if the opposite happens, then it is a false alarm (FA). The performance metrics are determined by combination of Hs, Ms and FAs:

$$Bias = \frac{1}{N}\sum_{n=1}^{N}(O_n - R_n) \tag{12}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(O_n - R_n)^2} \tag{13}$$

$$FR = \frac{FA}{H+FA}, 0 \leq FR \leq 1 \tag{14}$$

$$PD = \frac{H}{H+M}, 0 \leq PD \leq 1 \tag{15}$$

$$TS = \frac{H}{H+FA+M}, 0 \leq TS \leq 1 \tag{16}$$

$$HSS = 2 * \frac{Z*H - FA*M}{((H+FA)*(Z+FA)) + ((M+H)*(M+Z))}, -1 \leq HSS \leq 1 \tag{17}$$

where O is raingauge observation and R is radar QPE, and N is the number of points. Z is the overall number of zeros (when neither radar data nor raingauge measurements match the threshold criterion). A TS of 0.5 implies that the criterion is satisfied at least 50% of the time, and a higher value is indicative of superior performance. A TS=0.33 is indicative of performance similar to persistence, meaning predicted values in the next hour are the same values in the previous hour. HSS describes the fractional improvement of the corrected $STIV_{DBKC}$ over original StageIV. An HSS of 0 means that the performance is not better than random chance.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

2.3.6 Hydrologic Correction

At flash flood timescale in headwater basins, streamflow uncertainty and precipitation uncertainty are strongly connected in a nonlinear way through rainfall runoff processes. Liao and Barros (2022) developed a Lagrangian-based framework named Inverse Rainfall Correction (IRC) allowing backpropagating streamflow uncertainty to precipitation inputs in space and time through an uncalibrated distributed hydrological model (i.e. DCHM), achieving water budget closure at event scale in small headwater basins. As stated earlier, the uncertainties associated with parameters and the hydrological model DCHM are neglected since the model configurations have been used and improved over the past two decades for this region accounting for various soil, vegetation, and river processes (e.g. Tao and Barros, 2013, 2014, and 2018; Lowman and Barros, 2016), and the IRC framework has been tested in multiple headwater basins extensively in this region with consistent success.

It is worth noting that IRC is a general framework to improve QPE at watershed scale that can be incorporated into any distributed hydrological models. Liao and Barros (2024a, 2024b) investigated the impact of model structure uncertainty and initial condition uncertainty on IRC and then the downstream product: the resulting IRC-improved QPE. The results suggest with improved watershed physics at finer resolution (e.g. river bank storage, Liao and Barros, 2024a), river routing algorithms (e.g. XY routing, Liao and Barros, 2024a) and improved antecedent soil moisture distributions (Liao and Barros, 2024b), post-IRC QPE demonstrate much more realistic precipitation features at high resolution that are aligned with basin topography with ridges associated with higher precipitation than valleys in general, showing a significant improvement from the original StageIV dataset which is characterized by unnatural boxy precipitation patterns in complex terrain due to resolution issue.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

As briefly mentioned before, LB22 reviewed various sources of uncertainty that can prevent post-IRC QPE from achieving water budget closure, among which initial condition uncertainty in soil moisture is a noteworthy source. Improved initial condition estimation results in significantly improved post-IRC precipitation features in complex terrain by better capturing transient travel time distributions (Liao and Barros, 2024b). They found that the uncertainty tied to initial conditions is relatively more important for less extreme events. Nevertheless, the initial condition correction (i.e. ICC) method is coupled with the IRC framework and the complete framework is named the coupled IRC-ICC framework since Liao and Barros (2024b) to reflect the importance of Initial Condition Correction (or ICC). The specifics of IRC, ICC and the coupled IRC-ICC framework are schematically drawn in Figure 3.

<Figure 3 here please>

Based on the characteristic timing definitions in panels d) and c), different temporal windows are identified corresponding to different flow regimes. In principle, many more windows can be identified if a rather complex hydrograph with more peaks and inflection points is presented. ICC is only applied to window 2 and 5 (i.e. rising point mismatch segment, and slow recession, respectively), assuming precipitation uncertainty is dominating streamflow differences for window 3 and 4 (i.e. the neighborhood of flow peak). W_{nm} represents precipitation after window m procedure at iteration n. DCHM stands for Duke Coupled Hydrology Model, which is an uncalibrated physics-based distributed model, and the spatial and temporal resolution are 250m and 5 min.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

2.3.7 Lagrangian Tracking

A flood event is simulated by the DCHM to simulate streamflow at the basin outlet with grid-based time-varying velocity fields for different soil layers. When the precipitation starts (i.e. basin-averaged precipitation > 0.1mm/hr), new particles (passive tracers) are launched at non-zero precipitation grids at *every* model time step (i.e. 5 minutes) in all soil layers following the velocity fields calculated by the DCHM, and the tracking resolution is 10 seconds, amounting to a release of approximately 600,000 particles for basin with an area of 120km^2 over a 24-hour period. During the tracking phase, each particle is saved along with information regarding its source location (grid-point where it originates), time of release ti, and travel time tT (tT is defined as the difference between current time t and the time of release ti, i.e., tT = t – ti). Multiple particles from different source locations can have the same travel time, which is the basis for identifying the number of trajectories contributing to the hydrograph at the outlet as a function of time.

2.3.8 QPE Correction

At time t, the water difference *wd(t)* between the observed and simulated streamflow over the time *Δt* between two consecutive discharge observations represents the fraction of runoff that eventually leave the basin as streamflow. Errors in precipitation forcing propagate to the runoff, under the assumption of negligible model and parameter uncertainties, *wd(t)* can be entirely attributed to precipitation error, which is the focus of this work.

$$wd(t) = [Q_{obs}(t) - Q_{simu}(t)] \times \Delta t \tag{18}$$

The subscripts obs and $simu$ refer to observed and simulated discharge, respectively. The strategy for the inverse rainfall correction (IRC) using hydrograph analysis is to follow the trajectories available from the Lagrangian tracking backward from the basin outlet to the source

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

locations at time t_i and apply a correction at the source locations proportional to original QPE magnitude to reduce *wd* at time *t*. The embedded assumption is that larger QPE values have larger uncertainties. Note that QPE corrections happened earlier in time will have an impact on runoff simulation at future times, and this is the reason why IRC framework is a recursive framework. The detailed rainfall correction steps can be found in (Liao and Barros, 2022).

2.3.9 Methods for Reducing Uncertainties from Other Sources

As briefly mentioned before, uncertainties from other sources (e.g. river routing, model physics, antecedent soil moisture, etc) have certain impacts on travel time distributions and simulated streamflow. Previous studies demonstrate that, for flood producing events in small headwater basins, streamflow response is largely controlled by precipitation inputs (e.g. Iwasaki et al., 2020). In this section, we briefly describe the methods used to minimize the impacts from other sources to facilitate the IRC framework to achieve water budge closure.

First and foremost, the DCHM is an uncalibrated model with parameters strongly tied to this region of study, demonstrating great success over the past 25 years. DCHM has been used extensively without significant biases, therefore parameter uncertainty and model structure uncertainty are ignored. The impact of routing algorithm on peak flood timing is investigated in Liao and Barros (2024a) and they pointed out that variable parameter Muskingum-Cunge routing leads to artificially fast rising limb of flash flood hydrographs in headwater basins due to the existence of mild slopes over the valleys. They developed a general routing framework which is more suitable for stream routing particularly for estimating flood timing in headwater basins (see Liao and Barros, 2024a for details). Their results also suggest meandering effects, riverbank storage, and initial soil moisture distributions can impact the early rising period of the hydrographs. Later, significant and consistent improvements are made when introducing an initial condition correction (ICC) module to reduce initial condition uncertainty (Liao and Barros, 2024b). In fact, numerous studies also point out that precipitation and initial condition are the 2 most important factors in hydrologic forecasting and simulation. Therefore, this innovative ICC module is coupled with the IRC framework since then. The red arrows in Figure 3e indicate where ICC are executed in the general architecture of the IRC framework and the specifics of the ICC module are stated below.

Particles launched during the IRC process that reached the outlet at time t are traced back directly to the IC timing or time 0, and their locations at the IC timing are shown in the bottom maps in Figure 3d as the control points of time t. The downstream area of the control points has shorter transportation time to arrive at the outlet (e.g. water difference ΔS_1), and upstream area of the control points takes longer to get to the basin outlet (e.g. water difference ΔS_2). Similarly, soil moisture in the impacted area can greatly impact the size of ΔS_2 and flow conditions after the timing t_2. Assuming initial condition are only impactful during early period and late recession of hydrograph, which is supported by the fact that these events are flood-producing events with large QPE uncertainties dominating the vicinity of peak flow, ICC is used for hydrological windows near the peak flow. Following the same notation in the IRC framework (Eq. 18), using backward-in-time numerical notation, $wd(t)$ represents the flow volume difference between simulated and observed flow between time t and $t - \Delta t$. A 'band' of region can therefore be identified, that is the region downstream to control points of time t and upstream to control points of time $t - \Delta t$. This 'band' is then referred to as the impacted area for time t, and the initial soil moisture in the impacted area significantly influence basin discharge between time $t - \Delta t$ and time t. Finally, $wd(t)$ is then applied to initial soil moisture within the impacted area (i.e. the 'band') and the details can be found in Liao and Barros (2024b).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

2.3.10 Hydrological Skill Metrics

The hydrological skill metrics used in this study include the Kling-Gupta Efficiency (KGE) of the streamflow calculated at time-interval τ (here 15 minutes) determined by the frequency of observations (i.e. USGS gauge records) over the event duration (here 24 hours):

$$KGE_{\tau} = 1 - \sqrt{(r-1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2} \quad (19)$$

where r is the correlation, σ_{obs} is the standard deviation in discharge observations, σ_{sim} the standard deviation in discharge simulations, μ_{sim} and μ_{obs} are the mean streamflow of the simulations and the observations respectively. The subscript τ denotes the time-scale dependence of the KGE. KGE ranges from negative infinity to 1.

The Nash Sutcliffe Efficiency (NSE) of the streamflow is also calculated at time-interval τ (here 15 min):

$$NSE_{\tau} = 1 - \frac{\sum_{t=1}^{T}(Q_O^t - Q_S^t)^2}{\sum_{t=1}^{T}(Q_O^t - \overline{Q_O})^2} \quad (20)$$

where Q_O^t and Q_S^t are the streamflow observation and simulation at time t, and t is ranging from the first time step to the last time step T. $\overline{Q_O}$ is the mean of observed streamflow. The subscript τ denotes the time-scale dependence of the NSE. NSE ranges from negative infinity to 1.

The relative volume error (EV) is the difference between the time integral of the simulated and observed hydrographs over the event discharge volume:

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

$$\mathrm{EV} = \frac{\text{Simulated hydrograph volume} - \text{Observed hydrograph volume}}{\text{Observed hydrograph volume}} \tag{21}$$

An overestimation case has EV>0 and an underestimation event has EV<0.

EPT is the error in the timing of peak discharge on the rising limb of the hydrograph. When calculating EPT, if multiple peaks are present, only the highest peak timing is considered. To better capture the difference between rising limbs of observations and simulations, EPT is calculated using both the rising point and the highest peak point; thus, EPT compares the difference between the mid points of the two rising limbs.

EPV or error in peak volume is a relative error between simulated and observed flood peak, and the equation is below:

$$\mathrm{EPV} = \frac{\text{Simulated peak flow} - \text{Observed peak flow}}{\text{Observed peak flow}} \tag{22}$$

2.3.11 Study Domain and Model Setup

28 headwater basins are selected in the Appalachians as illustrated in Figure 1 with basin drainage area ranging from 50 km^2 to 500 km^2. These headwater basins cover a wide range of geographic regions (e.g. Basin01 and Basin30 are over 2,000 km apart) with diverse weather and climate regimes, associated with large differences in geomorphology and hydrogeology.

<Figure 4 here please>

Soil-related parameters are downloaded from a global high resolution (1 km) soil data repository (Zhang et al., 2018). For each basin, the vertical hydraulic conductivity remains the same for the entire soil column. The lateral hydraulic conductivity in the unsaturated zone was assumed to be two-three orders of magnitude larger than the vertical conductivity in the shallow soil layers, with higher values where the stone fraction in the soils is higher (Carlson, 2010, Freeze and Cherry, 1979). The final scaling factors were obtained through simple sensitivity analysis to match the curvature and slope of the observed subsurface runoff recession curves (Yildiz and Barros, 2007; Chen and Kumar, 2001; Linsley et al., 1982), and the final scaling factors are: 1500, 150, 15 and 1.5 for layer 1 (0-10 cm), layer 2 (10-75 cm), layer 3 (75-200 cm) and layer 4 (2-20 m), respectively. No parameter calibration is done in this work as the primary focus of this work is to develop a QPE dataset that can consistently close the water budget while controlling uncertainties from other sources, largely advancing the understanding of QPE uncertainties across climate, weather, and geomorphological regimes.

Flood-producing events are selected for the headwater basins identified in the Appalachians for recent years (i.e. the study period is from January 2021 to April 2024). The selection criteria are threshold-based, specifically the peak flow must be greater than 95% of the flow records in the study period. The choice of 95% is a compromise because 99% would yield too few events while 90% would be too close to the annual flood. Additionally, rainfall runoff response time must be shorter or equal to 6 hours to be qualified as a flash flood event. Only warm season liquid precipitation events 2021-2024 are finally selected during this systematic event selection process. Here, the warm season is specifically defined as from April 1st to September

30th. Note, data quality control is enforced and events with missing streamflow records are discarded.

For the Cataloochee Creek Basin (Basin05), located in the SAM known to have experienced multiple flash floods in the past (Tao and Barros, 2013), Liao and Barros (2024a) created a Historical Flood Record database (HFR) for this basin, which includes numerous flood-producing events from 2008 to 2017. The event selection criteria when developing HFR are also using the same 95% flow threshold method. The difference is that the HFR also includes multiple winter-time liquid precipitation events that result in flash floods. In total, there are 54 events for Basin 05 in the HFR and these events are also used to expand the study sample size in this work.

The hydrological spin-up process starts with iterative DCHM runs from April 30th to September 30th, 2021, including a total of 5 iterations (i.e. reaching a stable/equilibrium model state). Then DCHM runs continuously from October 1st 2021, to April 1st, 2024 to derive initial conditions for events happened after September 30th 2021. During this spin-up process, no parameter calibration is involved. The initial conditions used for the events in this study are from the 5th iteration from April 30th to September 30th, 2021, and from the subsequent continuous run from October 1st, 2021 to April 1st, 2024.

2.4 Caveat

In the entire study domain, rain gauges are only installed in the Southern Appalachians specifically in the vicinity of the Cataloochee Creek Basin (Basin 05) in the core area of the IPHEx rain gauge mesonet. The remainder of the studied basins are not monitored by raingauges, and therefore no raingauge bias correction is done for those basins and the downscaled original dataset StageIV (i.e. $STIV_D$) is used as inputs for the IRC method and hydrological simulations in this study.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

As an important component of the IRC framework, the Lagrangian tracking algorithm is only implemented when transitioning from one hydrological window to the next window, instead of being implemented every model time step (i.e. 5 minutes), and this is because of computational constraints. Additionally, we do not differentiate peak flow points and recession inflection points between simulations and observations when classifying hydrological flow regimes/windows, and consistently use observations as the reference basis for hydrological window identification because 1) precise locations of particles become much more uncertain later in the hydrograph due to numerical rounding errors and grid-based abruptly-changing velocity fields used in the Lagrangian tracking algorithm, and 2) the computational costs associated with excessive running of the tracking algorithm. Very short travel times (i.e. <15 minutes) are ignored because of temporal resolution restrictions from streamflow observations. A systematic use of 24 hours for event total duration is imposed in this work to reduce excessive tracking workload, which might be problematic for events with very long and heavy tails though not common for flash flood events in headwater basins.

Nevertheless, the coupled IRC-ICC recursive framework allows us to quantify QPE uncertainties more realistically by improving initial soil moisture estimation, and this framework proves to be numerically efficient in achieving good and stable hydrological state after only a few iterations. In this work, the stable state of IRC-ICC is reached when the Kling-Gupta Efficiency (KGE) oscillations are within 0.05 from iteration to iteration, calculated at 15-minutes intervals.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

3. Results and Discussion

3.1 Raingauge Bias Correction

Raingauge bias correction includes linear bias correction for radar-gauge pairs (see Eq. 8) and a series of biases corrections listed in Section 2.3.1: LRC, MRC, CRC and GRC. Analysis of the diurnal cycle on a seasonal basis reveals bias patterns linked to radar operations, and in particular terrain blockage, radar beam overshooting, and excessive attenuation that may vary from hour to hour but when taken over a long period of time indicate localized errors in space and time that reflect the site hydrometeorology. Light and shallow rainfall is a particular challenge in the region of study (e.g. Duan et al. 2015; Duan and Barros, 2017; Arulraj and Barros, 2017). The peak number of missed rainfall corresponds to about 10-15% of the total number of hours for each season in the late afternoon. The missed events correspond to both light and moderate rainfall, and occasionally to isolated heavy rainfall likely associatd with isolated thunderstorms.

The climatologically corrected $STIV_{DBKC}$ fields have significantly accurate diurnal cycle comparing to only event-scale bias corrected $STIV_{DBK}$. This processes is illustrated in Figure 5 for one raingauge in eastern ridges (left panel) and another in the western ridges (right panel).

<Figure 5 here please>

Biases in original $StageIV_{D}$ are more significant over the western ridges (e.g. right panel) at all times of day reflecting the impact of cloud immersion and seeder-feeder enhacement of low level precipitation over the ridges (Duan and Barros, 2017), with mid-day bias being a problem across the region (e.g., Barros and Arulraj, 2019). Overall, analysis of the $StageIV_{DBKC}$ fields

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

demonstrates that the climatology corrections work well in terms of mean rainfall, as well as reducing miss detection errors. Figure 6 shows the diurnal cycle of missed precipitation at two selected gauge locations (top row) in the winter (Januray-February and March – JFM) in StageIV that are preserved in StageIV$_{D}$ (black) and StageIV$_{DBK}$ (cyan). These missed precipitation events correspond to instances of very light rainfal (bottom row) at the raingauge locations (< 1.5 mm/hr). After applying the LRC and MRC climatology corrections, the missed detection problems (cyan) in StageIV$_{DBK}$ are largely eliminated for the StageIV$_{DBKC}$ fields (green).

<Figure 6 here please>

When integrated over the ten-year period, the averaged seasonal HSS, TS, and RMSE statistics of STIV$_{DBKC}$ demonstrate significantly better performance comparing to STIV$_{D}$ for all hours of the day (Figure 7a). Moreover, note that there is no decreasing trend in TS with increasing precipitation rate threshold (Figure 7b) which indicates that climatology correction is working for the heavy rainfall amounts linked to localized thunderstorm activity. Figure 7c shows the diurnal cycle and seasonality distribution of RMSE conditional on rain rate. The RMSE generally stays below 0.1 mm/hr except in the early morning and in the late afternoon in the cold season. In part this error could be related to snowfall which is not properly accounted for as the raingauges are not heated.

<Figure 7 here please>

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

3.2 Hydrologic Correction

The coupled IRC-ICC framework is originally developed and applied in Basin 05, the Cataloochee Creek Basin, and an example showing the results from iterations is demonstrated in Figure 8. The notation follows the definition in Figure 3. Note $STIV_{DBKC}$ data derived in Section 3.1 are further downscaled to 250m and used for hydrological simulations in this section. For all other basins (except Basin05), raingauges are not available and $STIV_{D}$ data are used instead.

<Figure 8 here please>

It is demonstrated that for this extreme flood-producing event, IRC-ICC produces stable results after about 3 to 4 iterations without significant oscillations. In general, for less significant events, IRC-ICC often reaches an equilibrium state faster (merely 3 iterations), providing fast and convergent corrections. As explained in the Section 3, the equilibrium state is considered when oscillations in the KGE values are within 0.05, and then IRC-ICC is stopped immediately. This study suggests that for most events 3 iterations is a good rule of thumb.

3.2.1 Systematic Application of IRC-ICC

Systematic application of the coupled IRC-ICC framework is conducted in the 28 basins selected in the Appalachians for 225 events, and examples are displayed in Figure 9.

<Figure 9 here please>

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

The performance of IRC-ICC is in general slightly better in the Southern and Northern Appalachian Mountains (SAM, NAM) than the Central Appalachian Mountains (CAM). In CAM, particularly along the border of the state of West Virginia and the state of Virginia, residing expanded karst terrain, and numerous caverns are identified (see the documented caverns: http://www.wvgs.wvnet.edu/www/geology/docs/WV_Tax_Districts_Containing_Karst_Terrain.pdf). The current version of DCHM hydrological model does not solve physics involved in Karst terrain. Here, the advantage of not calibrating parameters becomes obvious because these Karst terrain related physics can easily be ignored by parameter calibration when domain knowledge is not sufficient. Being in Karst terrain, Basin 13 and 14 (not shown) demonstrate noticeably poor simulations with severely underestimated baseflow contribution and artificial peaks due to the lack of subterranean river representation. This is apparently beyond the resolved scale using the current DCHM model with current spatiotemporal resolutions (250m, 5minutes). Here, resolved scale refers to a reasonable scale range where physical processes are represented in the hydrological model. At coarse scales, physical processes are substantially averaged, and information is potentially lost during averaging. At fine scales, some physical processes are not yet known or not parameterized in the model. The 10 events in Basin 13 and 14 are therefore discarded.

Event 2021-06-10 in Basin 19 (Figure 9) is an example when more hydrological windows (see Figure 3) are required to capture the subtle changes in the hydrograph for a relatively more complex hydrograph (e.g. multiple mild peaks around the major peak). These subtle changes could be a shifting of dominant river branch in the basin due to the movement of rainfall. Again, this requires much finer resolution both for the hydrological model and for the tracking algorithm to represent this detailed level of physics for summer thunderstorms. With limited computational power, this study systematically uses a 4-window IRC-ICC framework, including pre-rising-point

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

segment, rising limb, early recession, and late recession (separated by the recession inflection point).

3.2.2 IRC and IRC-ICC Precipitation Corrections

Event total precipitation fields are calculated after IRC-only and IRC-ICC frameworks reaching an equilibrium state, and these fields are compared with product $STIV_{DBKC}$ used as inputs for these frameworks. Examples are shown in Figure 10 categorized by seasons in the Cataloochee Creek Basin (Basin05). Again, warm season is defined as April 1st to September 30th, and the rest is defined as the cold season, with only liquid precipitation events are studied in this work.

<Figure 10 here please>

The original QPE (**a1** and **b1**) shows boxy patterns and abrupt spatial changes, which is a common issue of radar observations when used at high spatial resolution. By contrast the IRC-corrected precipitation maps (from both the IRC-only framework and the coupled IRC-ICC framework) are aligned better with terrain gradients, showing strong spatial patterns with higher precipitation along ridges and lower precipitation in adjacent valleys. IRC-ICC precipitation fields have similar patterns to IRC-only precipitation. Note the dark blue colors corresponding to very low precipitation near the basin outlet are an artifact of the IRC tied to very short travel times that cannot be fully resolved even at the fine scales of 250m and 5minutes. However, with proper IC uncertainty reduction, these artifacts are dramatically reduced as shown for the 2009-10-14, 2009-04-20, and 2013-04-12 events because of overall basin-wide travel time improvements attributed to improved IC. These three events are relatively mild events, indicating larger importance of IC for events of lower magnitude because of the critical role of IC in runoff generation mechanisms and travel times distributions.

3.2.3 Precipitation and Hydrologic Statistics

Event-total precipitation maps are calculated for each basin and event, and basin scale precipitation statistics (e.g. mean and standard deviation) are derived for each event-total precipitation map. These statistics are plotted in Figure 11, and subregions are separated by vertical black lines. Basins 01 to 11 are located in the SAM, Basins 12 to 20 are located in the CAM, and Basins 21 to 30 are located in the NAM. Given the dramatic impact of the Karst terrain on the hydrological performance related to Basin 13 and 14, the results from these two basins are not included in the statistics.

<Figure 11 here please>

It is clearly demonstrated that the change in the mean (i.e. basin-averaged event total QPE) is relatively small (from 36.10mm to 38.07mm) compared to the change in the standard deviation (from 6.63mm to 14.08mm) after the application of the coupled IRC-ICC. The small standard deviation of the original QPE suggests that original QPE data are spatially tightly clustered with low variability (see Figure 10a for boxy rainfall features), while larger standard deviation post-IRC-ICC indicates spatial variability is enhanced, which is highlighted by the terrain-aligning precipitation features in Figure 10c. The relatively small change in the mean indicates that the original input precipitation (i.e. $\mathrm{StageIV}_{\mathrm{DBKC}}$ for Basin 05, and $\mathrm{StageIV}_{\mathrm{D}}$ for the remainder basins) does not contain significant unconditional systematic biases across basins and events, which would lead to consistent positive or negative flood volume errors. This argument is supported by the fact that only small changes in the mean are introduced by the IRC-ICC framework. As an

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

exception, it is worth noting that the standard deviation of Basin 05 events does not change significantly after the IRC-ICC compared to other basins and events because rain gauge corrections are employed in Basin 05 but not anywhere else. It can never be overly emphasized that even after rain gauge bias correction, essentially as a point-scale correction method, the resulting QPE is still subjected to large water budget closure errors (see Figure 12 for more discussion) on account of the highly heterogeneous nature of QPE in complex terrain.

The hydrologic statistics described in Table 1 using all studied events are plotted in Figure 12.

<Figure 12 here please>

Figure 12 shows that the median KGE in each sub-region across the basins and events is improved from 0.36, 0.39, 0.27 to 0.89, 0.74, 0.84 for SAM, CAM and NAM, respectively. It should be pointed out that QPE changes for Basin 05 events (event number 55 to 108) are important for improving water budget closure, albeit small in magnitude compared to other events in other basins as shown in Figure 11 and 12, and yet critical to capture the complex precipitation heterogeneity in complex terrain to close the water budget. Basin 05 is a good example illustrating not only the contributions but also the limitations of rain gauge bias corrections in complex terrain in general. The relatively mild improvement in the CAM is explained by lacking physics representation of subterranean rivers in the Karst terrain in the DCHM model, causing large baseflow errors during hydrograph recession and thus low KGE values. Nevertheless, for applications in flash floods research, peak flood discharge, flood peak timing, and flood volume are the most important factors (see the second, third and fourth horizontal subplots in Figure 12. Flood volume error (the second panel) is controlled within ±10% for over 90% of the flood-producing events in the Appalachians, with the median EV error being less than 5% for post IRC-

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

ICC products in SAM and NAM. Flood peak volume (the third panel) is controlled within 20% for most of the events, which is significant because these events are extreme events. This is demonstrated by Tropical Storm Fred on 2021-08-17 and event that caused floods in multiple SAM basins, caused 5 deaths and an estimated economic loss of over 1 billion dollars: the KGE improves to 0.9 and peak timing errors are less than 30 minutes using IRC-ICC. For most of the studied events, timing errors (shown in the fourth panel) of the post IRC-ICC product are bounded by ±60 minutes, though some outliers are observed in the CAM and NAM potentially due to complex surface conditions such as antecedent snow on the ground for April events.

Events with relatively large timing errors (±90 minutes) are investigated in detail. These include the 2023-07-08 event in Basin 27 in New Hampshire (event number 185, which is 2.5 hours too early). This was a localized summer thunderstorm event, only taking half an hour to reach its peak flow, posing a challenge in separating flow regimes using hydrological windows defined in the IRC-ICC framework at the current model resolution. The event on 2022-05-27 in Basin16 (a relatively big basin >400 km^2) is a relatively slow rising event (event number 118, 2 hours too early) in West Virginia with rain-on-snow conditions and potentially snowmelt effects involved at high elevations. Finally, event 2021-09-22 in Basin19 (event number 133, 2 hours too late) is a relatively more complex event with multiple rain cells moving over the basin close to each other in time, therefore requiring many more windows to capture highly transient hydrological regimes than the 4-window default structure (i.e. pre-rising limb, rising limb, early recession, late recession) used in the IRC-ICC.

Overall, there are significant improvements in QPE corresponding to improvements in flood volume, flood peak and flood timing as a result of IRC-ICC. Because the IRC-ICC is a framework built upon runoff travel time, it cannot be used when precipitation is missing or there

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

are severe timing errors due to the lack of water travel time trajectories to distribute corrections. From a practical point of view, the QPE IRC-ICC corrections amount to space-time bias correction. The improved QPE data can be used to build general QPE error prediction models such as Liao and Barros (2023) and therefore correct remote-sensing products to improve orographic QPE data to support hydroclimatic studies and model calibration under reduced forcing uncertainty.

4. Discussion and Future Work

Limitations in this study stem mainly from computational constraints rather than the methodology. A systematic definition of 24-hour flood duration is imposed, implying that for floods with longer high-flow tails slow contributions from deeper soil layers are not considered. The current IRC-ICC framework was built to support flash flood studies and only utilizes shallow subsurface moisture transport information, consistent with the governing role of shallow soil moisture dynamics in steep topography. It is expected that for long duration precipitation events or basins with large mild-slope areas, deeper interflows would play a much more important role in improving flood timing, volume estimation and resulting QPE via IRC-ICC.

We plan to improve the StageIV-IRC product by further improving the IRC-ICC framework through improved model physics and resolution and utilizing 3D velocity fields to capture the full travel time distributions. When computational resources allow, the IRC can be carried out at the same frequency as the model resolution, therefore eliminating any artifacts produced due to inadequate sampling and updating of travel time distributions. This dataset can even be used in near real time in operational hydrology to improve Quantitative Precipitation Forecasts (QPF), advancing flood forecasting and emergency management.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

5. Data Availability Statement

The StageIV-IRC dataset at 250 m 5 minute resolution for 26 basins and 215 events is available at: https://doi.org/10.5281/zenodo.14028867.(Liao and Barros, 2024). Associated geographic documentation of the selected basins is also provided via the same link.

6. Conclusion

QPE has been an enduring challenge particularly in complex terrain. Radar QPE are plagued with uncertainties from multiple sources while rain gauge networks are scarce and suffer from the lack of representativeness in the mountains. To address this grand challenge, we develop a series of corrections from point-scale to watershed-scale encompassing event, climatology, and water budget closure corrections for radar QPE: the IRC-ICC framework. To our knowledge, this is the first QPE dataset aiming to close the water budget at high resolution for flood events, consistent with fundamental physics at watershed scale, and achieving superior hydrological performance at sub-hourly scale in headwater basins.

The coupled IRC-ICC framework is applied to 26 headwater basins in the Appalachians for 215 events with robust success yielding significant improvements in streamflow simulation, particularly on flood timing and volume. The tracking algorithm in the IRC-ICC framework is only updated when shifting from one hydrological window to another but not every time step. With enough computational resources, post IRC-ICC QPE data should further improve by capturing transient travel time distributions between model time steps.

Over 90% of the events have flood timing errors within one hour using the StageIV-IRC product compared to fewer than 20% of the events without the use of IRC-ICC, while the median KGE improved from 0.34 to 0.86 across the events. Results show that initial conditions are more

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

important for less severe precipitation events, especially during the slow rising period of hydrograph, which influence subsequent streamflow simulations. It is also worth noting that physical parameters used in this work are not calibrated for any precipitation event in any basin. This physics-based IRC-ICC framework can capture the fundamental physics involved in flash flood events, that is the fast hydrological response in surface and shallow subsurface soil layers due to steep slopes and gravity, therefore skillful hydrologic prediction is achieved without model calibration by instead focusing on getting the forcing right.

The IRC-ICC is a general framework that can be incorporated into any distributed hydrological models. Thus, the StageIV-IRC dataset also enables meaningful intercomparison among different radar QPE dataset, providing physics insights into QPE error structure from water budget closure perspective, toward improving radar retrievals and to characterize radar specific errors related to radar operations at high spatial resolution in the mountains. The demonstrated success of StageIV-IRC in ungauged basins strongly supports the use of IRC-ICC in the vast area of remote mountains worldwide where raingauges are generally not available. This dataset can be utilized as a reference for building machine learning models (or even deep-learning models when the number of studied precipitation events is expanded) that can learn the QPE uncertainties conditional on time of day, weather, climate and geomorphological regimes for both radar QPE analysis and forecasts, advancing the understanding of orographic precipitation uncertainties at high resolution across global mountains.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

CRediT author statement

M. Liao: Methodology, Data curation, Writing- first draft preparation, Analysis, Investigation. A. P. Barros: Conceptualization, Methodology, Analysis, Writing- revisions and editing, Supervision, Project administration, Funding acquisition.

Competing interests

The authors declare there are no competing interests.

Acknowledgments

The work was supported in part by NASA Precipitation Measurement Mission Program under NASA grant 80NSSC19K0685 with the second author, a NASA Earth System Science Fellowship with the first author, and grant from the IBM Accelerator program with the second author.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

REFERENCES

Abbott, M. B. (1976). Computational hydraulics: A short pathology. *Journal of Hydraulic Research*, *14*(4), 271-285. https://doi.org/10.1080/00221687609499661

AghaKouchak, A., Mehran, A., Norouzi, H., & Behrangi, A. (2012). Systematic and random error components in satellite precipitation data sets. Geophysical Research Letters, 39(9). https://doi.org/10.1029/2012GL051592

Ahmed, S. I., Rudra, R., Goel, P., Khan, A., Gharabaghi, B., & Sharma, R. (2022). A comparative evaluation of using rain gauge and NEXRAD radar-estimated rainfall data for simulating streamflow. Hydrology, 9(8), 133. https://doi.org/10.3390/hydrology9080133

Ajami, H., McCabe, M. F., Evans, J. P., & Stisen, S. (2014). Assessing the impact of model spin-up on surface water-groundwater interactions using an integrated hydrologic model. Water Resources Research, 50(3), 2636-2656. https://doi.org/10.1002/2013WR014258

Alimonti, G., Mariani, L., Prodi, F., & Ricci, R. A. (2022). A critical assessment of extreme events trends in times of global warming. *European Physical Journal - Plus*, 137, 112. https://doi.org/10.1140/epjp/s13360-021-02243-9

Allen, J. T. (2018). Climate change and severe thunderstorms. In Oxford research encyclopedia of climate science. https://doi.org/10.1093/acrefore/9780190228620.013.62

Areerachakul, N., Prongnuch, S., Longsomboon, P., & Kandasamy, J. (2022). Quantitative precipitation estimation (QPE) rainfall from meteorology radar over Chi Basin. *Hydrology*, *9*(10), 178. https://doi.org/10.3390/hydrology9100178

Arnell, N. W. (2003). Effects of IPCC SRES* emissions scenarios on river runoff: a global perspective. *Hydrology and Earth System Sciences*, *7*(5), 619-641. https://doi.org/10.5194/hess-7-619-2003

Arulraj, M., & Barros, A. P. (2017). Shallow precipitation detection and classification using multifrequency radar observations and model simulations. *Journal of Atmospheric and Oceanic Technology*, *34*(9), 1963-1983. https://doi.org/10.1175/JTECH-D-17-0060.1

Arulraj, M., & Barros, A. P. (2021). Automatic detection and classification of low-level orographic precipitation processes from space-borne radars using machine learning. *Remote Sensing of Environment*, *257*, 112355. https://doi.org/10.1016/j.rse.2021.112355

Barros, A. P. (2013). Orographic precipitation, freshwater resources, and climate vulnerabilities in mountainous regions. In *Climate vulnerability: Understanding and addressing threats to essential resources* (pp. 57-78). Elsevier Inc.. 10.1016/B978-0-12-384703-4.00504-9

Barros, A., & Arulraj, M. (2020). Remote sensing of orographic precipitation. In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk (Eds.), Satellite precipitation measurement. Springer. https://doi.org/10.1007/978-3-030-35798-6_6

Barros, A., Miller, D., Wilson, A., Cutrell, G., Arulraj, M., Super, P., & Petersen, W. (2017). GPM Ground Validation Southern Appalachian Rain Gauge IPHEx [indicate subset used]. Dataset available online from the NASA Global Hydrometeorology Resource Center DAAC,

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

Huntsville, Alabama, U.S.A. DOI: http://dx.doi.org/10.5067/GPMGV/IPHEX/GAUGES/DATA301

Barros, A. P., Kim, G., Williams, E., & Nesbitt, S. W. (2004). Probing orographic controls in the Himalayas during the monsoon using satellite imagery. *Natural Hazards and Earth System Sciences*, *4*(1), 29-51. https://doi.org/10.5194/nhess-4-29-2004

Barros, A. P., & Lettenmaier, D. P. (1993). Dynamic modeling of the spatial distribution of precipitation in remote mountainous areas. *Monthly weather review*, *121*(4), 1195-1214. https://doi.org/10.1175/1520-0493(1993)121<1195:DMOTSD>2.0.CO;2

Barros, A.P., Petersen, W., Schwaller, M., Cifelli, R., Mahoney, K., Peters-Liddard, C., Shepherd, M., Nesbitt, S., Wolff, D., Heymsfield, G. & Starr, D. (2014) NASA GPM-Ground Validation: Integrated Precipitation and Hydrology Experiment 2014 Science Plan. EPL/Duke University: Durham, N.C. https://doi.org/10.7924/G8CC0XMR.

Berndt, C., Rabiei, E., & Haberlandt, U. (2014). Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *Journal of Hydrology*, *508*, 88-101. https://doi.org/10.1016/j.jhydrol.2013.10.028

Beven, K. J. (1996). A discussion of distributed hydrological modelling. In *Distributed hydrological modelling* (pp. 255-278). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-0257-2_13

Beven, K., & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, *6*(3), 279-298. https://doi.org/10.1002/hyp.3360060305

Bindlish, R., & Barros, A. P. (2000). Disaggregation of rainfall for one-way coupling of atmospheric and hydrological models in regions of complex terrain. *Global and Planetary Change*, *25*(1-2), 111-132. https://doi.org/10.1016/S0921-8181(00)00024-2

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: a review. *Hydrological processes*, *9*(3-4), 251-290. https://doi.org/10.1002/hyp.3360090305

Bohling, G. (2005). Introduction to geostatistics and variogram analysis. Kansas geological survey, 1(10), 1-20.

Brooks, H. E. (2013). Severe thunderstorms and climate change. *Atmospheric research*, *123*, 129-138.https://doi.org/10.1016/j.atmosres.2012.04.002

Brunner, G. 1989. Muskingum-Cunge Channel Routing. Lecture Notes. HEC, U.S. Army Corps of Engineer, Davis, CA

Burcea, S., Cheval, S., Dumitrescu, A., Antonescu, B., Bell, A., & Breza, T. (2012). Comparison between radar estimated and rain gauge measured precipitation in the Moldavian Plateau. Environmental Engineering & Management Journal (EEMJ), 11(4).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

Buytaert, W., Celleri, R., Willems, P., De Bievre, B., & Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. Journal of hydrology, 329(3-4), 413-421. https://doi.org/10.1016/j.jhydrol.2006.02.031

Cai, M., Yang, S., Zeng, H., Zhao, C., & Wang, S. (2014). A distributed hydrological model driven by multi-source spatial data and its application in the Ili River Basin of Central Asia. Water resources management, 28, 2851-2866. https://doi.org/10.1007/s11269-014-0641-z

Cassiraga, E., Gómez-Hernández, J. J., Berenguer, M., Sempere-Torres, D., & Rodrigo-Ilarri, J. (2021). Spatiotemporal precipitation estimation from rain gauges and meteorological radar using geostatistics. *Mathematical Geosciences*, *53*, 499-516. https://doi.org/10.1007/s11004-020-09882-1

Castillo, V. M., Gomez-Plaza, A., & Martınez-Mena, M. (2003). The role of antecedent soil water content in the runoff response of semiarid catchments: a simulation approach. Journal of Hydrology, 284(1-4), 114-130. https://doi.org/10.1016/S0022-1694(03)00264-6

Catto, J. L., & Pfahl, S. (2013). The importance of fronts for extreme precipitation. Journal of Geophysical Research: Atmospheres, 118(19), 10-791. https://doi.org/10.1002/jgrd.50852

Chavez, S. P., Silva, Y., & Barros, A. P. (2020). High‐elevation monsoon precipitation processes in the Central Andes of Peru. *Journal of Geophysical Research: Atmospheres*, *125*(24), https://doi.org/10.1029/2020JD032947

Chen, T. H., Henderson-Sellers, A., Milly, P. C. D., Pitman, A. J., Beljaars, A. C. M., Polcher, J., ... & Zeng, Q. (1997). Cabauw experimental results from the project for intercomparison of land-surface parameterization schemes. Journal of Climate, 10(6), 1194-1215. https://doi.org/10.1175/1520-0442(1997)010<1194:CERFTP>2.0.CO;2

Chen, F., & Mitchell, K. (1999). Using the GEWEX/ISLSCP forcing data to simulate global soil moisture fields and hydrological cycle for 1987-1988. *Journal of the Meteorological Society of Japan. Ser. II*, *77*(1B), 167-182. https://doi.org/10.2151/jmsj1965.77.1B_167

Chen, Y., Li, J., Wang, H., Qin, J., & Dong, L. (2017). Large-watershed flood forecasting with high-resolution distributed hydrological model. Hydrology and Earth System Sciences, 21(2), 735-749. https://doi.org/10.5194/hess-21-735-2017

Chow, V. T. 1959. Open-channel Hydraulics, Mc-Graw Hill, New York, p. 607.

Clark MP, et al. 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Adv. Water Resour. 31(10):1309–1324. https://doi.org/10.1016/j.advwatres.2008.06.005

Cosgrove, B. A., Lohmann, D., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., ... & Tarpley, J. D. (2003). Land surface model spin‐up behavior in the North American Land Data Assimilation System (NLDAS). Journal of Geophysical Research: Atmospheres, 108(D22). https://doi.org/10.1029/2002JD003316

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Cunge, J. A. (1969). On the subject of a flood propagation computation method (Musklngum method). *Journal of Hydraulic Research*, *7*(2), 205-230. https://doi.org/10.1080/00221686909500264

Czigány, S., Pirkhoffer, E., & Geresdi, I. (2010). Impact of extreme rainfall and soil moisture on flash flood generation. *Quarterly Journal of the Hungarian Meteorological Service*, *114*(1-2), 79-100.

David, C. H., Maidment, D. R., Niu, G. Y., Yang, Z. L., Habets, F., & Eijkhout, V. (2011). River network routing on the NHDPlus dataset. *Journal of Hydrometeorology*, *12*(5), 913-934. https://doi.org/10.1175/2011JHM1345.1

de Goncalves, L. G. G., Shuttleworth, W. J., Burke, E. J., Houser, P., Toll, D. L., Rodell, M., & Arsenault, K. (2006). Toward a South America land data assimilation system: Aspects of land surface model spin-up using the simplified simple biosphere. *Journal of Geophysical Research: Atmospheres*, *111*(D17). https://doi.org/10.1029/2005JD006297

Deijns, A. A., Dewitte, O., Thiery, W., d'Oreye, N., Malet, J. P., & Kervyn, F. (2022). Timing landslide and flash flood events from SAR satellite: a regionally applicable methodology illustrated in African cloud-covered tropical environments. *Natural Hazards and Earth System Sciences*, *22*(11), 3679-3700. https://doi.org/10.5194/nhess-22-3679-2022

Delrieu, G., Wijbrans, A., Boudevillain, B., Faure, D., Bonnifait, L., & Kirstetter, P. E. (2014). Geostatistical radar–raingauge merging: A novel method for the quantification of rain estimation accuracy. *Advances in Water Resources*, *71*, 110-124. https://doi.org/10.1016/j.advwatres.2014.06.005

Diggle, P. and Ribeiro, P. J., 2007. Model-based geostatistics. Springer Series in Statistics, 230.

Dinku, T., Anagnostou, E. N., & Borga, M. (2002). Improving radar-based estimation of rainfall over complex terrain. Journal of Applied Meteorology and Climatology, 41(12), 1163-1178. https://doi.org/10.1175/1520-0450(2002)041<1163:IRBEOR>2.0.CO;2

Diomede, T., Davolio, S., Marsigli, C., Miglietta, M. M., Moscatello, A., Papetti, P., ... & Malguzzi, P. (2008). Discharge prediction based on multi-model precipitation forecasts. *Meteorology and atmospheric physics*, *101*, 245-265. https://doi.org/10.1007/s00703-007-0285-0

Dobler, C., Hagemann, S., Wilby, R. L., & Stötter, J. (2012). Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. *Hydrology and Earth System Sciences*, *16*(11), 4343-4360. https://doi.org/10.5194/hess-16-4343-2012

Duan, Y., & Barros, A. P. (2017). Understanding how low-level clouds and fog modify the diurnal cycle of orographic precipitation using in situ and satellite observations. *Remote Sensing*, *9*(9), 920. https://doi.org/10.3390/rs9090920

Duan, Y., Wilson, A. M., & Barros, A. P. (2015). Scoping a field experiment: Error diagnostics of TRMM precipitation radar estimates in complex terrain as a basis for IPHEx2014. Hydrology and Earth System Sciences, 19(3), 1501-1520. https://doi.org/10.5194/hess-19-1501-2015

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Durand, M., Rodriguez, E., Alsdorf, D. E., & Trigg, M. (2009). Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 3(1), 20-31. https://doi.org/10.1109/JSTARS.2009.2033453

Durden, S. L., Haddad, Z. S., Kitiyakara, A., & Li, F. K. (1998). Effects of nonuniform beam filling on rainfall retrieval for the TRMM precipitation radar. Journal of Atmospheric and Oceanic Technology, 15(3), 635-646. https://doi.org/10.1175/1520-0426(1998)015<0635:EONBFO>2.0.CO;2

Durkee, J. D., Campbell, L., Berry, K., Jordan, D., Goodrich, G., Mahmood, R., & Foster, S. (2012). A synoptic perspective of the record 1-2 May 2010 Mid-South heavy precipitation event. Bulletin of the American Meteorological Society, 93(5), 611-620. https://doi.org/10.1175/BAMS-D-11-00076.1

Ekmekcioğlu, Ö., Demirel, M. C., & Booij, M. J. (2022). Effect of data length, spin-up period and spatial model resolution on fully distributed hydrological model calibration in the Moselle basin. *Hydrological Sciences Journal*, *67*(5), 759-772. https://doi.org/10.1080/02626667.2022.2046754

Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. *Information & Management*, *24*(3), 159-167. https://doi.org/10.1016/0378-7206(93)90064-Z

Fulton, J., & Ostrowski, J. (2008). Measuring real-time streamflow using emerging technologies: Radar, hydroacoustics, and the probability concept. *Journal of Hydrology*, *357*(1-2), 1-10. https://doi.org/10.1016/j.jhydrol.2008.03.028

Gentilucci, M., Bufalini, M., D'Aprile, F., Materazzi, M., & Pambianchi, G. (2021). Comparison of data from rain gauges and the IMERG product to analyse precipitation in mountain areas of central Italy. ISPRS International Journal of Geo-Information, 10(12), 795. https://doi.org/10.3390/ijgi10120795

Grayson, R. B., Moore, I. D., & McMahon, T. A. (1992). Physically based hydrologic modeling: 2. Is the concept realistic?. *Water resources research*, *28*(10), 2659-2666. https://doi.org/10.1029/92WR01259

Goodrich, D. C., Schmugge, T. J., Jackson, T. J., Unkrich, C. L., Keefer, T. O., Parry, R., ... & Amer, S. A. (1994). Runoff simulation sensitivity to remotely sensed initial soil water content. *Water Resources Research*, *30*(5), 1393-1405. https://doi.org/10.1029/93WR03083

Goudenhoofdt, E., & Delobbe, L. (2009). Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, *13*(2), 195-203. https://doi.org/10.5194/hess-13-195-2009

Guha-Sapir, D., Below, R., and Hoyois, P. H.: EM-DAT: International Disaster Database, Université Catholique de Louvain, Brussels, Belgium, available at: http://www.emdat.be, last access: 19 February 2018.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

Guilloteau, C., & Foufoula-Georgiou, E. (2020). Multiscale evaluation of satellite precipitation products: Effective resolution of IMERG. Satellite Precipitation Measurement, V. Levizzani et al., Ed., Springer, 533–558, https://doi.org/10.1007/978-3-030-35798-6_5

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, *377*(1-2), 80-91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Gupta, V., Nautiyal, H., Kumar, V., Jamir, I., & Tandon, R. S. (2016). Landslide hazards around Uttarkashi township, Garhwal Himalaya, after the tragic flash flood in June 2013. Natural Hazards, 80, 1689-1707. https://doi.org/10.1007/s11069-015-2048-4

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. (2009) Multivariate Data Analysis. 7th Edition. Upper Saddle River, NJ: Prentice Hall.

Harrison, D. L., Driscoll, S. J., & Kitchen, M. (2000). Improving precipitation estimates from weather radar using quality control and correction techniques. Meteorological Applications, 7(2), 135-144. https://doi.org/10.1017/S1350482700001468

Hayami, S. (1951). On the propagation of flood waves. *Bulletins-Disaster Prevention Research Institute, Kyoto University*, *1*, 1-16.

He, X., Sonnenborg, T. O., Refsgaard, J. C., Vejen, F., & Jensen, K. H. (2013). Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling. Water Resources Research, 49(9), 5989-6005. https://doi.org/10.1002/wrcr.20471

Honnert, R., Efstathiou, G. A., Beare, R. J., Ito, J., Lock, A., Neggers, R., Plant, R. S., Shin, H. H., Tomassini, L. & Zhou, B. (2020). The atmospheric boundary layer and the "gray zone" of turbulence: A critical review. Journal of Geophysical Research: Atmospheres, 125(13), https://doi.org/10.1029/2019JD030317

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. J. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeorol., 8(1), 38–55. https://doi.org/10.1175/JHM560.1

Huffman, G.J. et al. (2020). Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). In: Levizzani, V., Kidd, C., Kirschbaum, D.B., Kummerow, C.D., Nakamura, K., Turk, F.J. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 67. Springer, Cham. https://doi.org/10.1007/978-3-030-24568-9_19

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2014), NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 4.4, PPS. NASA/GSFC, 30 pp.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Iwasaki, K., Katsuyama, M., & Tani, M. (2020). Factors affecting dominant peak-flow runoff-generation mechanisms among five neighbouring granitic headwater catchments. Hydrological Processes, 34(5), 1154-1166. https://doi.org/10.1002/hyp.13656

Li, H., Luo, L., Wood, E. F., & Schaake, J. (2009). The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *Journal of Geophysical Research: Atmospheres*, *114*(D4). https://doi.org/10.1029/2008JD010969

Kabir, T., Pokhrel, Y., & Felfelani, F. (2022). On the precipitation‐induced uncertainties in process‐based hydrological modeling in the Mekong River Basin. Water Resources Research, 58(2), e2021WR030828. https://doi.org/10.1029/2021WR030828

Kalinga, O. A., & Gan, T. Y. (2006). Semi-distributed modelling of basin hydrology with radar and gauged precipitation. *Hydrological Processes*, *20*(17), 3725-3746. https://doi.org/10.1002/hyp.6385

Khu, S. T., Madsen, H., & Di Pierro, F. (2008). Incorporating multiple observations for distributed hydrologic model calibration: An approach using a multi-objective evolutionary algorithm and clustering. *Advances in Water Resources*, *31*(10), 1387-1398. https://doi.org/10.1016/j.advwatres.2008.07.011

Kim, G., & Barros, A. P. (2001). Quantitative flood forecasting using multisensor data and neural networks. Journal of Hydrology, 246(1-4), 45-62. https://doi.org/10.1016/S0022-1694(01)00353-5

Kim, K. B., Kwon, H. H., & Han, D. (2018). Exploration of warm-up period in conceptual hydrological modelling. *Journal of Hydrology*, *556*, 194-210. https://doi.org/10.1016/j.jhydrol.2017.11.015

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water resources research*, *42*(3). https://doi.org/10.1029/2005WR004362

Kirstetter, P. E., Delrieu, G., Boudevillain, B., & Obled, C. (2010). Toward an error model for radar quantitative precipitation estimation in the Cévennes–Vivarais region, France. Journal of Hydrology, 394(1-2), 28-41. https://doi.org/10.1016/j.jhydrol.2010.01.009

Kirstetter, P. E., Hong, Y., Gourley, J. J., Chen, S., Flamig, Z., Zhang, J., ... & Amitai, E. (2012). Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar–based National Mosaic QPE. Journal of Hydrometeorology, 13(4), 1285-1300. https://doi.org/10.1175/JHM-D-11-0139.1

Knoben, W. J., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323-4331. https://doi.org/10.5194/hess-23-4323-2019

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Kobold, M., & Sušelj, K. (2005). Precipitation forecasts and their uncertainty as input into hydrological models. *Hydrology and Earth System Sciences*, *9*(4), 322-332. https://doi.org/10.5194/hess-9-322-2005

Kochendorfer, J., Rasmussen, R., Wolff, M., Baker, B., Hall, M. E., Meyers, T., ... & Leeper, R. (2017). The quantification and correction of wind-induced precipitation measurement errors. Hydrology and Earth System Sciences, 21(4), 1973-1989. https://doi.org/10.5194/hess-21-1973-2017

Kreklow, J., Tetzlaff, B., Burkhard, B., & Kuhnt, G. (2020). Radar-based precipitation climatology in Germany—developments, uncertainties and potentials. *Atmosphere*, *11*(2), 217. https://doi.org/10.3390/atmos11020217

Kuligowski, R. J., & Barros, A. P. (1998). Using artificial neural networks to estimate missing rainfall data 1. *JAWRA Journal of the American Water Resources Association*, *34*(6), 1437-1447. https://doi.org/10.1111/j.1752-1688.1998.tb05443.x

Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, *49*(1), 360-379. https://doi.org/10.1029/2012WR012195

Lagerquist, R., Allen, J. T., & McGovern, A. (2020). Climatology and variability of warm and cold fronts over North America from 1979 to 2018. Journal of Climate, 33(15), 6531-6554. https://doi.org/10.1175/JCLI-D-19-0680.1

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., ... & Puca, S. (2016). Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. *International Journal of Applied Earth Observation and Geoinformation*, *48*, 131-145. https://doi.org/10.1016/j.jag.2015.06.002

Lee, D. H., & Kang, D. S. (2016). The application of the artificial neural network ensemble model for simulating streamflow. *Procedia engineering*, *154*, 1217-1224. https://doi.org/10.1016/j.proeng.2016.07.434

Leendertse, J. J. 1967. Aspects of a computational model for long-period water wave propagation. Report RM-5294-PR, The Rand Corporation, Santa Monica, California, May.

Li, J., & Heap, A. D. 2008. A review of spatial interpolation methods for environmental scientists. 11-12.

Liao, M., & Barros, A. P. (2019). The Integrated Precipitation and Hydrology Experiment-Hydrologic Applications for the Southeast US (IPHEX-H4SE) Part IV: High-Resolution Enhanced Stage IV-Raingauge Combined Precipitation Product [Dataset]. Durham, NC: Duke Digital Repository. https://idn.duke.edu/ark:/87924/r4pc2zd75

Liao, M., & Barros, A. P. (2022). Toward optimal rainfall – Hydrologic QPE correction in headwater basins. Remote Sensing of Environment, 279, 113107. https://doi.org/10.1016/j.rse.2022.113107

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Liao, M., & Barros, A.P., 2024a. Model Celerity-Discharge Behavior in Complex Terrain to Improve Orographic Quantitative Precipitation Estimation and Hydrologic Prediction in Headwater Basins. (under review).

Liao, M., & Barros, A.P., 2024b. Toward optimal rainfall – Hydrologic QPE Correction in Headwater Basins — Closing the Water Budget within Observational Uncertainty Through Correcting Initial Soil Moisture Conditions. (under review).

Liao, M., & Barros, A. (2024). StageIV-IRC – A High-resolution Dataset of Extreme Orographic Quantitative Precipitation Estimates (QPE) Constrained to Water Budget Closure for Historical Floods in the Appalachian Mountains (Version v1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14028867

Lin, Y., & Mitchell, K. E. (2005), The NCEP stage II/IV hourly precipitation analyses: Development and applications, paper presented at 19th Conference on Hydrology [Dataset], Am. Meteorol. Soc., San Diego, Calif., 9–13 Jan.

Mälicke, M., Hassler, S. K., Weiler, M., Blume, T., & Zehe, E. (2018). Exploring hydrological similarity during soil moisture recession periods using time dependent variograms. Hydrology and Earth System Sciences Discussions, 2018, 1-25. https://doi.org/10.5194/hess-2018-396

McKee, J. L., & Binns, A. D. (2016). A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology. Canadian Water Resources Journal/Revue canadienne des ressources hydriques, 41(1-2), 186-203. https://doi.org/10.1080/07011784.2015.1064786

Miller, D. K., Miniat, C. F., Wooten, R. M., & Barros, A. P. (2019). An expanded investigation of atmospheric rivers in the southern Appalachian Mountains and their connection to landslides. Atmosphere, 10(2), 71. https://doi.org/10.3390/atmos10020071

Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, *415*(6871), 514-517. https://doi.org/10.1038/415514a

Minet, J., Laloy, E., Lambot, S., & Vanclooster, M. (2011). Effect of high-resolution spatial soil moisture variability on simulated runoff response using a distributed hydrologic model. Hydrology and Earth System Sciences, 15(4), 1323-1338. https://doi.org/10.5194/hess-15-1323-2011

Mockler, E. M., Chun, K. P., Sapriza-Azuri, G., Bruen, M., & Wheater, H. S. (2016). Assessing the relative importance of parameter and forcing uncertainty and their interactions in conceptual hydrological model simulations. *Advances in Water Resources*, *97*, 299-313. https://doi.org/10.1016/j.advwatres.2016.10.008

Moore, B. J., Neiman, P. J., Ralph, F. M., & Barthold, F. E. (2012). Physical processes associated with heavy flooding rainfall in Nashville, Tennessee, and vicinity during 1–2 May 2010: The role of an atmospheric river and mesoscale convective systems. *Monthly Weather Review*, *140*(2), 358-378. https://doi.org/10.1175/MWR-D-11-00126.1

Moradkhani, H., Hsu, K., Hong, Y., & Sorooshian, S. (2006). Investigating the impact of remotely sensed precipitation and hydrologic model uncertainties on the ensemble streamflow forecasting. Geophysical research letters, 33(12). https://doi.org/10.1029/2006GL026855

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Moraga, J. S., Peleg, N., Fatichi, S., Molnar, P., & Burlando, P. (2021). Revealing the impacts of climate change on mountainous catchments through high-resolution modelling. Journal of Hydrology, 603, 126806. https://doi.org/10.1016/j.jhydrol.2021.126806

Moulin, L., Gaume, E., & Obled, C. (2009). Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations. Hydrology and Earth System Sciences, 13(2), 99-114. https://doi.org/10.5194/hess-13-99-2009

Mtibaa, S., & Asano, S. (2022). Hydrological evaluation of radar and satellite gauge-merged precipitation datasets using the SWAT model: Case of the Terauchi catchment in Japan. *Journal of Hydrology: Regional Studies*, *42*, 101134. https://doi.org/10.1016/j.ejrh.2022.101134

Nanding, N., Rico-Ramirez, M. A., & Han, D. (2015). Comparison of different radar-raingauge rainfall merging techniques. Journal of Hydroinformatics, 17(3), 422-445. https://doi.org/10.2166/hydro.2015.001

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, *10*(3), 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

Nijssen, B., O'Donnell, G. M., Hamlet, A. F., & Lettenmaier, D. P. (2001). Hydrologic sensitivity of global rivers to climate change. *Climatic change*, *50*, 143-175. https://doi.org/10.1023/A:1010616428763

Nikolopoulos, E. I., Anagnostou, E. N., Borga, M., Vivoni, E. R., & Papadopoulos, A. (2011). Sensitivity of a mountain basin flash flood to initial wetness condition and rainfall variability. Journal of Hydrology, 402(3-4), 165-178. https://doi.org/10.1016/j.jhydrol.2010.12.020

Pechlivanidis, I. G., Jackson, B. M., Mcintyre, N. R., & Wheater, H. S. (2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST journal*, *13*(3), 193-214. https://doi.org/10.30955/gnj.000778

Pepin, N. C., Arnone, E., Gobiet, A., Haslinger, K., Kotlarski, S., Notarnicola, C., ... & Adler, C. (2022). Climate changes and their elevational patterns in the mountains of the world. Reviews of Geophysics, 60(1), e2020RG000730. https://doi.org/10.1029/2020RG000730

Penna, D., Tromp-van Meerveld, H. J., Gobbi, A., Borga, M., & Dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. Hydrology and Earth System Sciences, 15(3), 689-702. https://doi.org/10.5194/hess-15-689-2011

Pielke, R. A., & Downton, M. W. (2000). Precipitation and damaging floods: Trends in the United States, 1932–97. *Journal of climate*, *13*(20), 3625-3637. https://doi.org/10.1175/1520-0442(2000)013<3625:PADFTI>2.0.CO;2

Ponce, V. M., & Yevjevich, V. (1978). Muskingum-Cunge method with variable parameters. *Journal of the Hydraulics Division*, *104*(12), 1663-1667. https://doi.org/10.1061/JYCEAJ.0005119

Ponce, V. M. (1979). Simplified Muskingum routing equation. *Journal of the Hydraulics Division*, *105*(1), 85-91. https://doi.org/10.1061/JYCEAJ.000514

Ponce, V. M., & Changanti, P. V. (1994). Variable-parameter Muskingum-Cunge method revisited. *Journal of Hydrology*, *162*(3-4), 433-439. https://doi.org/10.1016/0022-1694(94)90241-0

Ponce, V. M., & Vuppalapati, B. (2016). Muskingum-Cunge amplitude and phase portraits with online computation. Online article.

Prat, O. P., & Barros, A. P. (2007). Exploring the use of a column model for the characterization of microphysical processes in warm rain: Results from a homogeneous rainshaft model. *Advances in Geosciences*, *10*, 145-152. https://doi.org/10.5194/adgeo-10-145-2007

Prat, O. P., & Barros, A. P. (2010). Assessing satellite-based precipitation estimates in the Southern Appalachian mountains using rain gauges and TRMM PR. *Advances in Geosciences*, *25*, 143-153. https://doi.org/10.5194/adgeo-25-143-2010

Prat, O. P., & Barros, A. P. (2010). Ground observations to characterize the spatial gradients and vertical structure of orographic precipitation–Experiments in the inner region of the Great Smoky Mountains. *Journal of Hydrology*, *391*(1-2), 141-156. https://doi.org/10.1016/j.jhydrol.2010.07.013

Pritchard, M. S., & Somerville, R. C. (2009). Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi‐scale climate model. Geophysical Research Letters, 36(5). https://doi.org/10.1029/2008GL036964

Rafieeinasab, A., Norouzi, A., Seo, D. J., & Nelson, B. (2015). Improving high-resolution quantitative precipitation estimation via fusion of multiple radar-based precipitation products. Journal of Hydrology, 531, 320-336. https://doi.org/10.1016/j.jhydrol.2015.04.066

Rodell, M., Houser, P. R., Berg, A. A., & Famiglietti, J. S. (2005). Evaluation of 10 methods for initializing a land surface model. *Journal of Hydrometeorology*, *6*(2), 146-155. https://doi.org/10.1175/JHM414.1

Rozalis, S., Morin, E., Yair, Y., & Price, C. (2010). Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions. Journal of hydrology, 394(1-2), 245-255. https://doi.org/10.1016/j.jhydrol.2010.03.021

Sangati, M., Borga, M., Rabuffetti, D., & Bechini, R. (2009). Influence of rainfall and soil properties spatial aggregation on extreme flash flood response modelling: an evaluation based on the Sesia river basin, North Western Italy. *Advances in Water Resources*, *32*(7), 1090-1106. https://doi.org/10.1016/j.advwatres.2008.12.007

Saunders, S., Montgomery, C., Easley, T., & Spencer, T. (2008). Hotter and drier: the West's changed climate. Rocky Mountain Climate Organization and the Natural Resources Defense Council. *Online: http://www. rockymountainclimate. org [2012, Nov 1]*.

Schiemann, R., Erdin, R., Willi, M., Frei, C., Berenguer, M., & Sempere-Torres, D. (2011). Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. *Hydrology and Earth System Sciences*, *15*(5), 1515-1536. https://doi.org/10.5194/hess-15-1515-2011

Schumacher, R. S. (2017). Heavy rainfall and flash flooding. In *Oxford research encyclopedia of natural hazard science*. https://doi.org/10.1093/acrefore/9780199389407.013.132

Seck, A., Welty, C., & Maxwell, R. M. (2015). Spin-up behavior and effects of initial conditions for an integrated hydrologic model. *Water Resources Research*, *51*(4), 2188-2210. https://doi.org/10.1002/2014WR016371

Senarath, S. U., Ogden, F. L., Downer, C. W., & Sharif, H. O. (2000). On the calibration and verification of two-dimensional, distributed, Hortonian, continuous watershed models. Water Resources Research, 36(6), 1495-1510. https://doi.org/10.1029/2000WR900039

Shamseldin, A. Y., Nasr, A. E., & O'connor, K. M. (2002). Comparison of different forms of the multi-layer feed-forward neural network method used for river flow forecasting. hydrology and earth system sciences, 6(4), 671-684. https://doi.org/10.5194/hess-6-671-2002

Shao, Y., Fu, A., Zhao, J., Xu, J., & Wu, J. (2021). Improving quantitative precipitation estimates by radar-rain gauge merging and an integration algorithm in the Yishu River catchment, China. Theoretical and Applied Climatology, 144, 611-623. https://doi.org/10.1007/s00704-021-03526-y

Shrestha, P., Barros, A. P., & Khlystov, A. (2010). Chemical composition and aerosol size distribution of the middle mountain range in the Nepal Himalayas during the 2009 pre-monsoon season. *Atmospheric Chemistry and Physics*, *10*(23), 11605-11621. https://doi.org/10.5194/acp-10-11605-2010

Shrestha, R., & Houser, P. (2010). A heterogeneous land surface model initialization study. Journal of Geophysical Research: Atmospheres, 115(D19). https://doi.org/10.1029/2009JD013252

Sideris, I. V., Gabella, M., Erdin, R., & Germann, U. (2014). Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 1097-1111. https://doi.org/10.1002/qj.2188

Silvestro, F., & Rebora, N. (2014). Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain. *Journal of Hydrology*, *519*, 1052-1067. https://doi.org/10.1016/j.jhydrol.2014.07.042

Silvestro, F., Rossi, L., Campo, L., Parodi, A., Fiori, E., Rudari, R., & Ferraris, L. (2019). Impact-based flash-flood forecasting system: Sensitivity to high resolution numerical weather prediction systems and soil moisture. *Journal of Hydrology*, *572*, 388-402. https://doi.org/10.1016/j.jhydrol.2019.02.055

Skofronick-Jackson, G., Kirschbaum, D., Petersen, W., Huffman, G., Kidd, C., Stocker, E., & Kakar, R. (2018). The Global Precipitation Measurement (GPM) mission's scientific achievements and societal contributions: Reviewing four years of advanced rain and snow observations. Quarterly Journal of the Royal Meteorological Society, 144, 27-48. https://doi.org/10.1002/qj.3313

Sungmin, O., & Kirstetter, P. E. (2018). Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. Quarterly Journal of the Royal Meteorological Society, 144, 270-281. https://doi.org/10.1002/qj.3218

Tao, J., & Barros, A. P. (2013). Prospects for flash flood forecasting in mountainous regions–An investigation of Tropical Storm Fay in the Southern Appalachians. *Journal of Hydrology*, *506*, 69-89. https://doi.org/10.1016/j.jhydrol.2013.02.052

Tao, J., & Barros, A. P. (2014). Coupled prediction of flood response and debris flow initiation during warm-and cold-season events in the Southern Appalachians, USA. *Hydrology and Earth System Sciences*, *18*(1), 367-388. https://doi.org/10.5194/hess-18-367-2014

Tao, J., & Barros, A. P. (2018). Multi-year atmospheric forcing datasets for hydrologic modeling in regions of complex terrain–Methodology and evaluation over the Integrated Precipitation and Hydrology Experiment 2014 domain. *Journal of Hydrology*, *567*, 824-842. https://doi.org/10.1016/j.jhydrol.2016.12.058

Tao, J., Wu, D., Gourley, J., Zhang, S. Q., Crow, W., Peters-Lidard, C., & Barros, A. P. (2016). Operational hydrological forecasting during the IPHEx-IOP campaign–Meet the challenge. *Journal of hydrology*, *541*, 434-456. https://doi.org/10.1016/j.jhydrol.2016.02.019

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., & Samaniego, L. (2019). The multiscale routing model mRM v1. 0: Simple river routing at resolutions from 1 to 50 km. Geoscientific Model Development, 12(6), 2501-2521. https://doi.org/10.5194/gmd-12-2501-2019

Tian, Y., Peters‐Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., ... & Zeng, J. (2009). Component analysis of errors in satellite‐based precipitation estimates. Journal of Geophysical Research: Atmospheres, 114(D24). https://doi.org/10.1029/2009JD011949

Todini, E. (1991). Hydraulic and hydrologic flood routing schemes. In *Recent advances in the modeling of hydrologic systems* (pp. 389-405). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-3480-4_18

Todini, E. (2007). A mass conservative and water storage consistent variable parameter Muskingum-Cunge approach. *Hydrology and Earth System Sciences*, *11*(5), 1645-1659. https://doi.org/10.5194/hess-11-1645-2007

Uber, M., Vandervaere, J. P., Zin, I., Braud, I., Heistermann, M., Legoût, C., ... & Nord, G. (2018). How does initial soil moisture influence the hydrological response? A case study from southern France. Hydrology and Earth System Sciences, 22(12), 6127-6146. https://doi.org/10.5194/hess-22-6127-2018

Vignal, B., Galli, G., Joss, J., & Germann, U. (2000). Three methods to determine profiles of reflectivity from volumetric radar data to correct precipitation estimates. *Journal of Applied Meteorology and Climatology*, *39*(10), 1715-1726. https://doi.org/10.1175/1520-0450-39.10.1715

Villarini, G., & Krajewski, W. F. (2010). Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surveys in geophysics*, *31*, 107-129. https://doi.org/10.1007/s10712-009-9079-x

Vivoni, E. R., Entekhabi, D., Bras, R. L., & Ivanov, V. Y. (2007). Controls on runoff generation and scale-dependence in a distributed hydrologic model. Hydrology and Earth System Sciences, 11(5), 1683-1701. https://doi.org/10.5194/hess-11-1683-2007

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, *19*, 378-387. https://doi.org/10.1007/s00477-005-0006-5

Wang, K. H., Chu, T., Yang, M. D., & Chen, M. C. (2020). Geostatistical based models for the spatial adjustment of radar rainfall data in typhoon events at a high-elevation river watershed. *Remote Sensing*, *12*(9), 1427. https://doi.org/10.3390/rs12091427

Weiland, F. C. S., Vrugt, J. A., Weerts, A. H., & Bierkens, M. F. (2015). Significant uncertainty in global scale hydrological modeling from precipitation data errors. *Journal of Hydrology*, *529*, 1095-1115. https://doi.org/10.1016/j.jhydrol.2015.08.061

Wehbe, Y., Temimi, M., & Adler, R. F. (2020). Enhancing precipitation estimates through the fusion of weather radar, satellite retrievals, and surface parameters. *Remote Sensing*, *12*(8), 1342. https://doi.org/10.3390/rs12081342

Wernberg, T., Smale, D., Tuya, F., Thomsen, M. S., Langlois, T. J., de Bettignies, T., Bennett, S., & Rousseaux, C. S. (2013). An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change*, 3, 78–82. https://doi.org/10.1038/nclimate1627

Wilson, A. M., & Barros, A. P. (2014). An investigation of warm rainfall microphysics in the southern Appalachians: Orographic enhancement via low-level seeder–feeder interactions. *Journal of the Atmospheric Sciences*, *71*(5), 1783-1805. https://doi.org/10.1175/JAS-D-13-0228.1

Wing, O. E. J., Bates, P. D., Smith, A. M., Sampson, C. C., Johnson, K. A., Fargione, J., & Morefield, P. (2018). Estimates of present and future flood risk in the conterminous United States. Environmental Research Letters, 13(3), 1–7. https://doi.org/10.1088/1748-9326/aaac65

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

Wu, W., Emerton, R., Duan, Q., Wood, A. W., Wetterhall, F., & Robertson, D. E. (2020). Ensemble flood forecasting: Current status and future opportunities. WIREs Water, 7(3), 1–32. https://doi.org/10.1002/wat2.1432

Yildiz, O., & Barros, A. P. (2005). Climate variability, water resources, and hydrologic extremes—Modeling the water and energy budgets. *Climate and Hydrology, in Mountain Areas*. https://doi.org/10.1002/0470858249.ch20

Yildiz, O., & Barros, A. P. (2007). Elucidating vegetation controls on the hydroclimatology of a mid-latitude basin. *Journal of Hydrology*, *333*(2-4), 431-448. https://doi.org/10.1016/j.jhydrol.2006.09.010

Yıldız, O., & Barros, A. P. (2009). Evaluating spatial variability and scale effects on hydrologic processes in a midsize river basin. *Sci. Res. Essays*, *4*, 217-225. https://doi.org/10.5897/SRE.9000465

Yin, D., Xue, Z. G., Bao, D., RafieeiNasab, A., Huang, Y., Morales, M., & Warner, J. C. (2022). Understanding the role of initial soil moisture and precipitation magnitude in flood forecast using a hydrometeorological modelling system. Hydrological Processes, 36(10), e14710. https://doi.org/10.1002/hyp.14710

Yu, D., Yang, J., Shi, L., Zhang, Q., Huang, K., Fang, Y., & Zha, Y. (2019). On the uncertainty of initial condition and initialization approaches in variably saturated flow modeling. Hydrology and Earth System Sciences, 23(7), 2897-2914. https://doi.org/10.5194/hess-23-2897-2019

Zappa, M., Jaun, S., Germann, U., Walser, A., & Fundel, F. (2011). Superposition of three sources of uncertainties in operational flood forecasting chains. *Atmospheric Research*, *100*(2-3), 246-262. https://doi.org/10.1016/j.atmosres.2010.12.005

Zhang, J., Qi, Y., Kingsmill, D., & Howard, K. (2012). Radar-based quantitative precipitation estimation for the cool season in complex terrain: Case studies from the NOAA Hydrometeorology Testbed. *Journal of Hydrometeorology*, *13*(6), 1836-1854. https://doi.org/10.1175/JHM-D-11-0145.1

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., & Kitzmiller, D. (2016). Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities [Dataset]. Bulletin of the American Meteorological Society, 97(4), 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, X., & Anagnostou, E. N. (2019). Evaluation of numerical weather model–based satellite precipitation adjustment in tropical mountainous regions. *Journal of Hydrometeorology*, *20*(3), 431-445. https://doi.org/10.1175/JHM-D-18-0008.1

Zhang, Y., Schaap, M. G., & Zha, Y. (2018). A high-resolution global map of soil hydraulic properties produced by a hierarchical parameterization of a physically based water retention model. *Water Resources Research*, *54*(12), 9774-9790. https://doi.org/10.1029/2018WR023539

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Zimmerman, D. L., & Zimmerman, M. B. (1991). A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. Technometrics, 33(1), 77-91. https://doi.org/10.1080/00401706.1991.10484771

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

LIST OF TABLES

Table 1 – Index, and coordinates for the raingauge stations marked in Figure1. The index is used to identify specific gauges in some of the graphs. Two raingauges at Purchase Knob, a supersite in the inner mountain region, are highlighted in bold font. Shaded rows indicate stations with collocated raingauges that have different temporal resolution (e.g. tip size).

Table 2: Hydrologic skill metrics used in this study.

Table 3 - Basin information including basin index used in this work for simplicity, USGS streamflow gauge ID, the corresponding drainage area, highest elevation in the basin and basin relief.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

Table 1 – Index, and coordinates for the raingauge stations marked in Figure1. The index is used to identify specific gauges in some of the graphs. Two raingauges at Purchase Knob, a supersite in the inner mountain region, are highlighted in bold font. Shaded rows indicate stations with collocated raingauges that have different temporal resolution (e.g. tip size).

NO.	Site ID.	Lat.	Lon.	Elev. (m)
1	RG001	35.39830	-82.91300	1156
2	RG002	35.41750	-82.97140	1731
3	RG003	35.38460	-82.91610	1609
4	RG004	35.36830	-82.99020	1922
5	RG005	35.40890	-82.96460	1520
6	RG008	35.38210	-82.97360	1737
7	RG010	35.45640	-82.94680	1478
8	RG100	35.58610	-83.07250	1495
9	RG100T	35.58767	-83.06468	1485
10	RG101	35.57500	-83.08820	1520
11	RG102	35.56370	-83.10360	1635
12	RG103	35.55340	-83.11790	1688
13	RG104	35.55490	-83.08800	1584
14	RG106	35.43210	-83.02910	1210
15	RG109	35.49560	-83.04040	1500
16	RG110	35.54810	-83.14820	1563
17	RG300	35.72653	-83.21692	1558
18	RG301	35.70552	-83.25595	2003
19	RG302	35.72135	-83.24675	1860
20	**RG303PK**	**35.58610**	**-83.07253**	**1495**
21	**RG303S**	**35.76295**	**-83.16222**	**1490**
22	RG304	35.67010	-83.18287	1820
23	RG305	35.69150	-83.13190	1630
24	RG306	35.74597	-83.17148	1536
25	RG307	35.65163	-83.19952	1624
26	RG308	35.73027	-83.18237	1471
27	RG309	35.68297	-83.15003	1604
28	RG310	35.70273	-83.12263	1756
29	RG311	35.76507	-83.14042	1036
30	RG400	35.70273	-83.12263	1756
31	RG401	35.65163	-83.19952	1624
32	RG402	35.72135	-83.24675	1860
33	RG403	35.51777	-83.10113	925
34	RG407	35.51777	-83.10113	925

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

1426 **Table 2**: Hydrologic skill metrics used in this study.

Metric	Description/Unit	Formula/Reference
KGE	Kling-Gupta efficiency	Eq. (19) /Gupta et al. (2009)
NSE	Nash-Sutcliffe efficiency	Eq. (20) /Nash and Sutcliffe. (1970)
EV	Error in area under the hydrograph	Eq. (21)
EPT	Error in Time to Peak (minutes)	Time diff. between mid-points of rising limbs
EPV	Relative Error in Peak Volume	Eq. (22)

1427

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

1428 **Table 3** - Basin information including basin index used in this work for simplicity, USGS
1429 streamflow gauge ID, the corresponding drainage area, highest elevation in the basin and basin
1430 relief.

Basin index	USGS Gauge ID	Drainage area (km^2)	Basin highest elevation (m)	Basin relief (m)	Location
1	3544970	118.7	1442	847	GA
2	2178400	176.1	1629	1051	GA
3	3504000	149.9	1667	1032	NC
4	3497300	317.6	1999	1651	TN
5	3460000	148.1	1879	1174	NC
6	3456500	152.8	1873	1157	NC
8	344894205	41.3	1995	1221	NC
9	3463300	134.3	1989	1425	NC
10	3400500	234.7	1257	1257	KY
11	3479000	283.3	1772	1216	NC
13	3182700	447.3	1111	717	WV
14	2011460	194.4	1388	763	VA
15	1620500	54.5	1321	712	VA
16	3180500	426.8	1416	621	WV
17	3068800	437.1	1471	908	WV
18	1595000	234.8	1230	560	MD
19	1595300	130.3	1069	712	WV
20	1544500	445.9	765	457	PA
21	1422747	81.4	766	394	NY
22	1415000	106.8	1019	636	NY
23	1413398	152.8	1094	754	NY
24	13621955	41.7	1074	717	NY
25	1421610	51.3	970	497	NY
26	1074520	389.4	1582	1582	NH
27	10642505	294.9	1895	1693	NH
28	1137500	300.3	1894	1546	NH
29	1133000	183.2	975	719	VT
30	1055000	334.1	1143	975	MAINE

1431

1432

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

LIST OF FIGURES

Figure 1 - Map of ground-based observations. Locations marked by numbers-only are raingauges; locations marked by numbers preceded by P are disdrometers. See Table 1 for list of stations and geographical coordinates.

Figure 2 – Workflow to generate the product $STIV_{DBKC}$.

Figure 3 – An illustration of the structure of IRC, ICC and the coupled IRC-ICC framework including **a)** the residual hydrograph between the observed and simulated discharge, with the discharge water difference *wd(t)* being distributed across the time window T; **b)** Example of travel time distribution TT(t) and map (inset) illustrating a hypothetical distribution of runoff source areas (in red, ns=3) with travel time x_2 contributing to streamflow at time t, meaning that at time t-x_2 there are three pixels (ns=3) generating runoff that reaches the outlet at time t. T is the time window over which runoff source areas with TT < T are mapped and the inverse rainfall correction (IRC) are applied; **c)** Example of IRC windows guided by timescales of dominant hydrological processes. The first window solely covers the initial streamflow conditions before the target event. The second window depicts the early rising limb of the hydrograph. The third window captures the steep rising limb of the hydrograph until it reaches the peak flow. The fourth and fifth windows correspond to interflow-dominant and baseflow-dominant stages of the recession curve respectively, separated by the recession inflection point; **d)** A schematic drawing that illustrates the Initial Condition Correction (ICC). Characteristic timing of the hydrographs: T_{r*} and T_r are rising points of simulated and observed hydrograph respectively. T_p is peak timing of observed flow. T_I is the inflection point of observed flow. Flow differences between simulated and observed flow at t_1 and t_2 are denoted as ΔS_1 and ΔS_2 respectively as examples for discussion in the text. IC, P, and Q stand for initial condition, precipitation and discharge, respectively; **e)** A recursive framework consists of Initial Condition Correction and Inverse Rainfall Correction (i.e. the coupled IRC-ICC framework) for illustration purposes.

Figure 4 – Headwater basins selected in the Appalachian Mountains, USA. The USGS gauge ID for each basin and basic information are listed in Table 2. For simplicity, sub-regions are identified and conveniently named as: Southern Appalachian Mountains (SAM, including Basin 01-11), Central Appalachian Mountains (CAM, including Basin 13-20), and Northern Appalachian Mountains (NAM, including Basin 21-30).

Figure 5 - Diurnal cycle of rainfall (mean and ±standard deviation) for different seasons and gauge locations. Left panel - Summer (JAS: July-August-September) at RG008 in the eastern ridges. Right panel – Spring (AMJ; April-May-June) at RG302 in the western ridges. Raingauge measurements (blue); $StageIV_{DBK}$ (black); $StageIV_{DBKC}$ (green).

Figure 6 –Top row - Wintertime (January-February-March, JFM) diurnal cycle of missing precipitation in the eastern ridges (RG003) and in the inner region (RG103) for each of the RR products: . Bottom row- same as top row for the raingauge climatology of hourly rainfall (blue). $StageIV_D$ (black); $StageIV_{DBK}$ (cyan); $StageIV_{DBKC}$ (green).

Figure 7 – Statistics summary: a) Diurnal cycle of mean HSS and TS statistics including all raingauges over the 10-year reference period (2008-2017): $STIV_D$ (black) and $STIV_{DBKC}$ (green); b) Seasonal mean HSS and TS statistics conditional on different rainfall thresholds over the 10-

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

year reference period; c) Diurnal cycle of RMSE at hourly time-scale and seasonal-scale RMSE conditional on observed rainfall rate. Note only the winter season (JFM, January February, March) is demonstrated in this Figure.

Figure 8 – The application of the coupled IRC-ICC framework to Basin05 (Cataloochee Creek Basin, NC) for 2017-10-23 event. This extreme hydrological response is caused by the remnants of Hurricane Nate in 2017. This figure includes **a)** the original simulation results using $\mathrm{STIV_{DBKC}}$ as inputs; **b)** the dashed rectangular plot consisting of intermediate results including each iteration from the coupled IRC-ICC framework as explained in Figure 3; **c)** the equilibrium state reached by the coupled framework after 5 iterations, and **d)** KGE value compilation graph calculated using 15 minutes intervals for each iteration in the coupled framework.

Figure 9 – The systematic application of the coupled IRC-ICC framework to the 28 basins selected in the Appalachians. The results include **a)** 5 events from the Southern Appalachians; **b)** 5 events from the Central Appalachians; and **c)** 5 events from the Northern Appalachians. The IRC-ICC KGE evolution plots from iterations are included below the hydrographs. The black dash line uses the original $\mathrm{STIV_D}$ and the pink line is the IRC-ICC equilibrium state ($STIV_{\mathrm{D}}^{\mathrm{IRC}*}$), and the corresponding colored numbers are KGE values calculated at 15 minutes interval.

Figure 10 – Event total precipitation maps for **a)** cold season events and **b)** warm season events. Each category includes 5 columns representing different events and 3 rows with the first row (**a1**, and **b1**) representing original precipitation input $\mathrm{STIV_{DBKC}}$, and the second row (**a2**, and **b2**) representing $\mathrm{STIV_{DBKC}^{IRC*}}$ from IRC-only framework, and the third row (**a3**, and **b3**) representing $\mathrm{STIV_{DBKC}^{IRC*}}$ from the coupled IRC-ICC framework.

Figure 11 – Summary charts of precipitation statistics for all event-total precipitation maps. Basin mean and standard deviation for each event are represented by circles and triangles in the top and bottom panel, respectively. Each panel is separated into 3 sub-regions by vertical black lines: the Southern Appalachian Mountains, Central Appalachian Mountains, and Northern Appalachian Mountains (SAM, CAM and NAM). The list of events in Basin 05 (with event number ranging from 55 to 108) in the SAM is highlighted by a green rectangle for further discussion in the text. The average values of all events for both the mean and the standard deviation are calculated and shown in the top right corner. Black color and pink color represent pre and post IRC-ICC QPE statistics, respectively.

Figure 12 – Summary charts of hydrologic skill metrics for all events. Horizontal green dash lines (i.e. the perfect situation) and green envelopes are for reference purposes. Hydrologic statistics are explained: EPV: Error in Peak Volume (Unit: %), EPT: Error in Peak Timing (Unit: minutes), EV: Error in flow Volume (Unit: %), note KGE is calculated using **15-minute** intervals over a 24 hour period. Pink dots and black dots represent post IRC-ICC results, and original inputs results, respectively (each dot represents one event). Each panel is separated into 3 sub-regions: the SAM, CAM and NAM. Histograms graphs are attached right next to the scatter plot.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

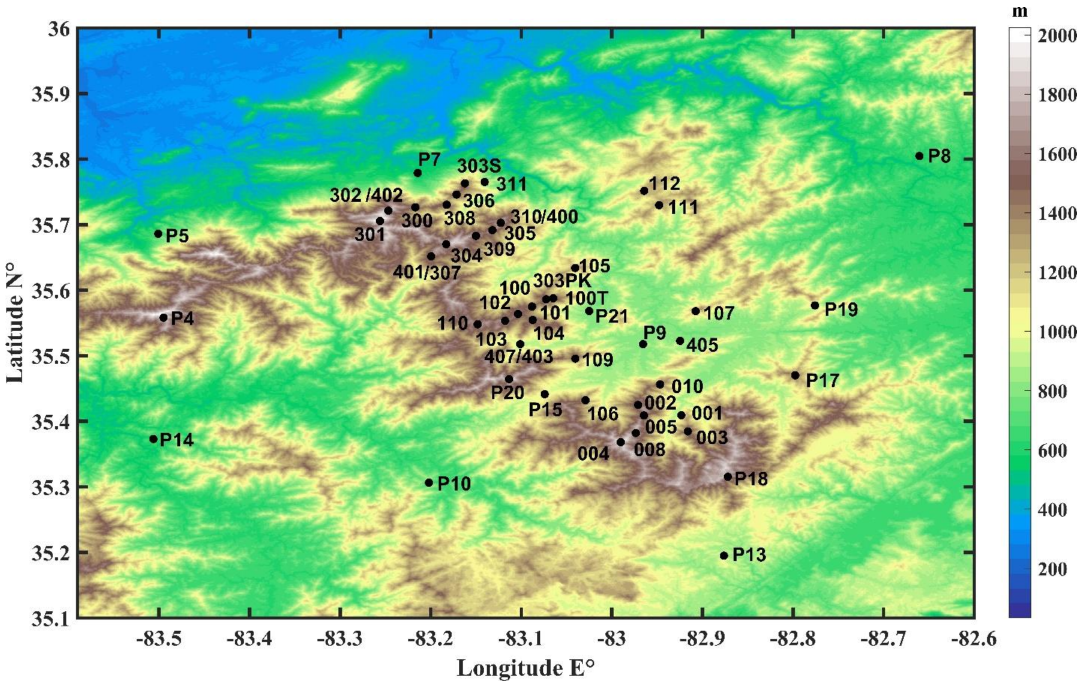

Figure 1 - Map of ground-based observations. Locations marked by numbers-only are raingauges; locations marked by numbers preceded by P are disdrometers. See Table 1 for list of stations and geographical coordinates.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

Open Access Earth System Science Data Discussions

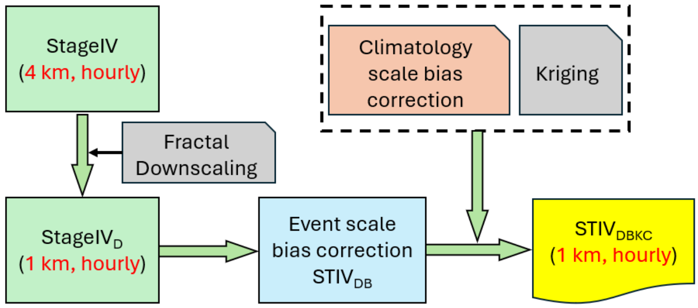

Figure 2 – Workflow to generate the product STIV$_{DBKC}$.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

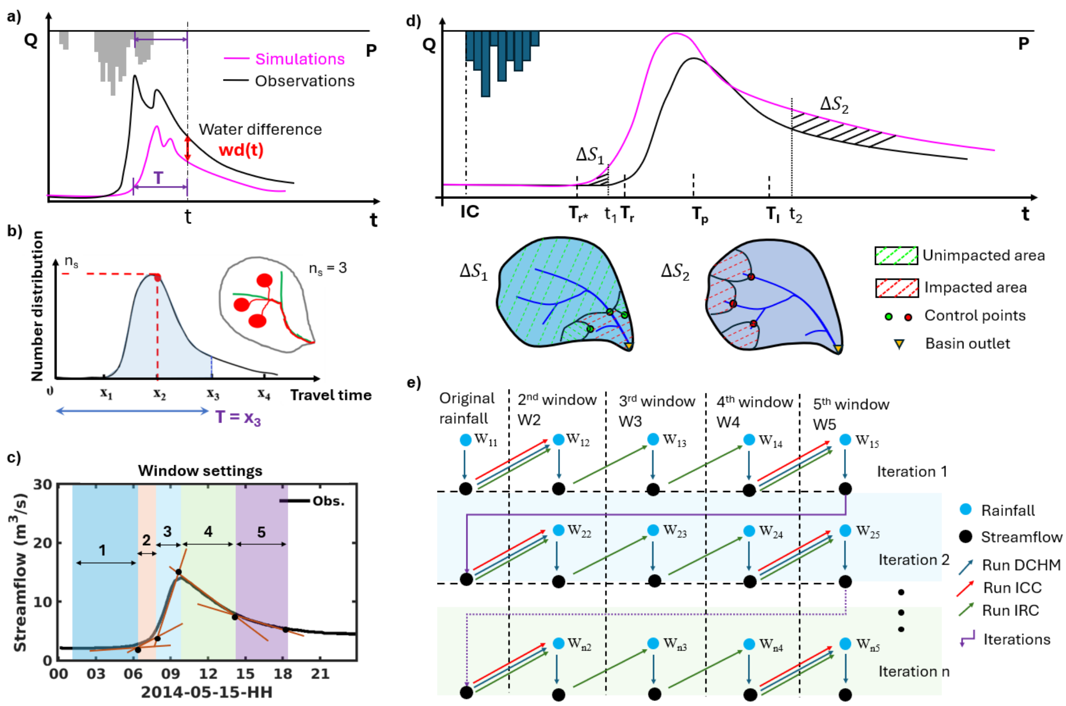

Figure 3 – An illustration of the structure of IRC, ICC and the coupled IRC-ICC framework including **a)** the residual hydrograph between the observed and simulated discharge, with the discharge water difference *wd(t)* being distributed across the time window T; **b)** Example of travel time distribution TT(t) and map (inset) illustrating a hypothetical distribution of runoff source areas (in red, ns=3) with travel time x_2 contributing to streamflow at time t, meaning that at time t-x_2 there are three pixels (ns=3) generating runoff that reaches the outlet at time t. T is the time window over which runoff source areas with TT < T are mapped and the inverse rainfall correction (IRC) are applied; **c)** Example of IRC windows guided by timescales of dominant hydrological processes. The first window solely covers the initial streamflow conditions before the target event. The second window depicts the early rising limb of the hydrograph. The third window captures the steep rising limb of the hydrograph until it reaches the peak flow. The fourth and fifth windows correspond to interflow-dominant and baseflow-dominant stages of the recession curve respectively, separated by the recession inflection point; **d)** A schematic drawing that illustrates the Initial Condition Correction (ICC). Characteristic timing of the hydrographs: T_{r*} and T_r are rising points of simulated and observed hydrograph respectively. T_p is peak timing of observed flow. T_I is the inflection point of observed flow. Flow differences between simulated and observed flow at t_1 and t_2 are denoted as ΔS_1 and ΔS_2 respectively as examples for discussion in the text. IC, P, and Q stand for initial condition, precipitation and discharge, respectively; **e)** A recursive framework consists of Initial Condition Correction and Inverse Rainfall Correction (i.e. the coupled IRC-ICC framework) for illustration purposes.

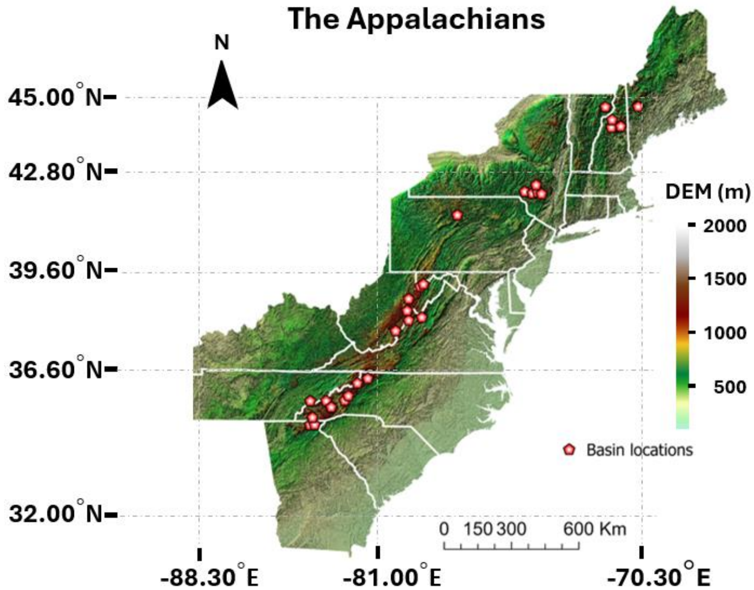

Figure 4 – Headwater basins selected in the Appalachian Mountains, USA. The USGS gauge ID for each basin and basic information are listed in Table 2. For simplicity, sub-regions are identified and conveniently named as: Southern Appalachian Mountains (SAM, including Basin 01-11), Central Appalachian Mountains (CAM, including Basin 13-20), and Northern Appalachian Mountains (NAM, including Basin 21-30).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

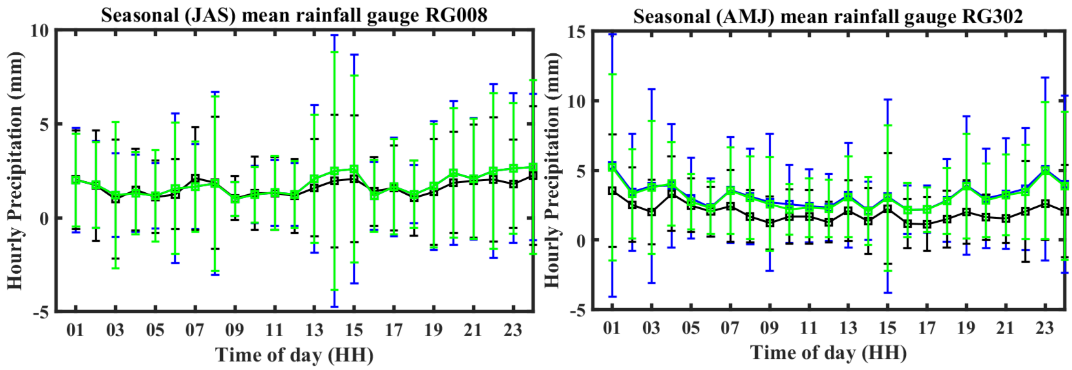

Figure 5 - Diurnal cycle of rainfall (mean and ±standard deviation) for different seasons and gauge locations. Left panel - Summer (JAS: July-August-September) at RG008 in the eastern ridges. Right panel – Spring (AMJ; April-May-June) at RG302 in the western ridges. Raingauge measurements (blue); StageIV$_{DBK}$ (black); StageIV$_{DBKC}$ (green).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

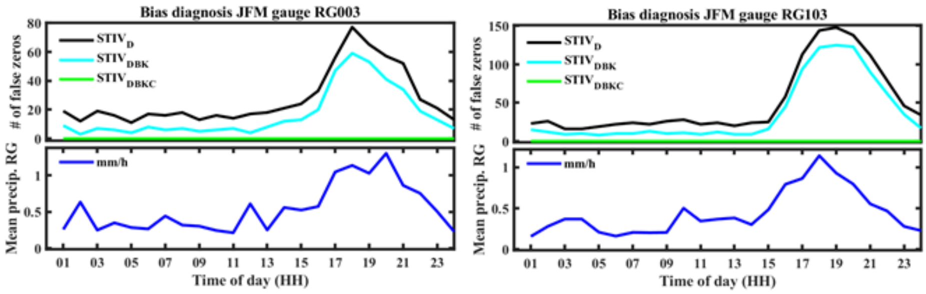

Figure 6 –Top row - Wintertime (January-February-March, JFM) diurnal cycle of missing precipitation in the eastern ridges (RG003) and in the inner region (RG103) for each of the RR products: . Bottom row- same as top row for the raingauge climatology of hourly rainfall (blue). StageIV$_D$ (black); StageIV$_{DBK}$ (cyan); StageIV$_{DBKC}$ (green).

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

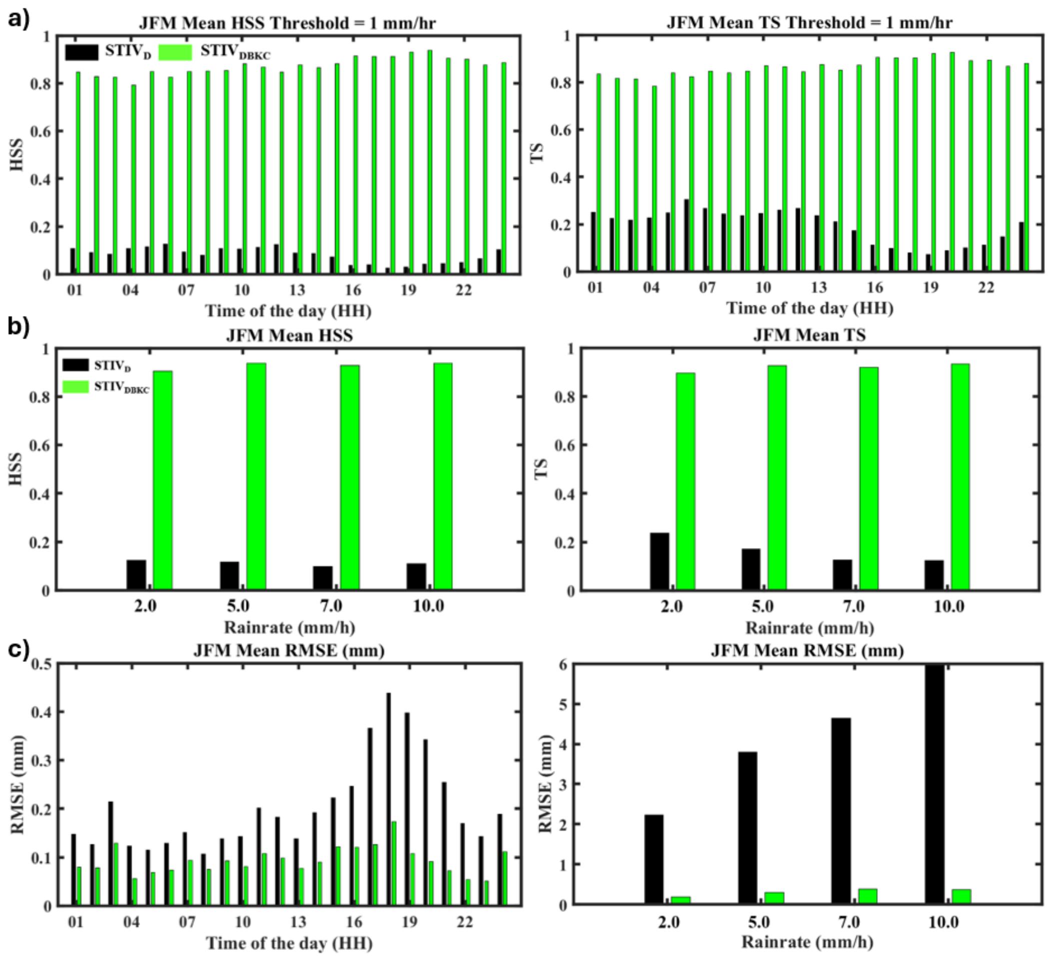

Figure 7 – Statistics summary: a) Diurnal cycle of mean HSS and TS statistics including all raingauges over the 10-year reference period (2008-2017): $STIV_D$ (black) and $STIV_{DBKC}$ (green); b) Seasonal mean HSS and TS statistics conditional on different rainfall thresholds over the 10-year reference period; c) Diurnal cycle of RMSE at hourly time-scale and seasonal-scale RMSE conditional on observed rainfall rate. Note only the winter season (JFM, January February, March) is demonstrated in this Figure.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

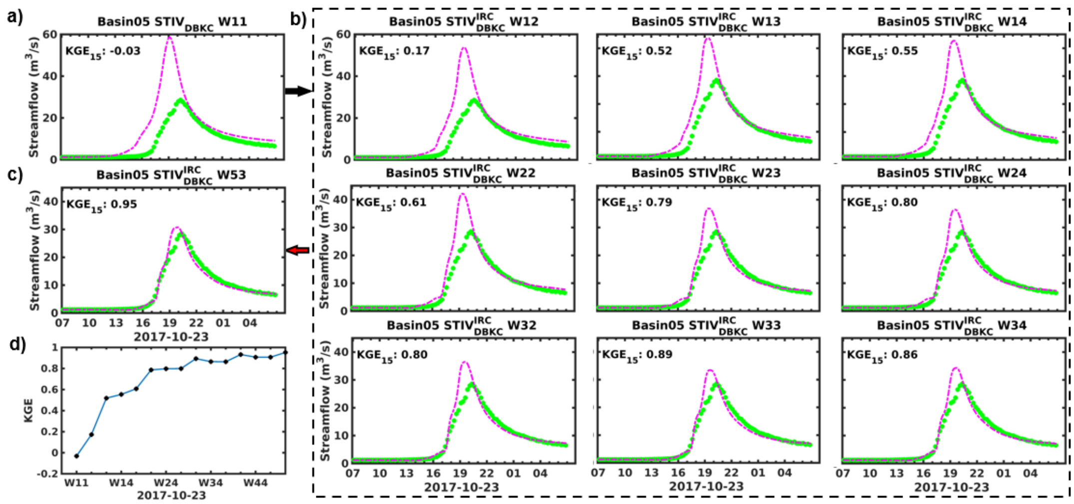

Figure 8 – The application of the coupled IRC-ICC framework to Basin05 (Cataloochee Creek Basin, NC) for 2017-10-23 event. This extreme hydrological response is caused by the remnants of Hurricane Nate in 2017. This figure includes **a)** the original simulation results using $STIV_{DBKC}$ as inputs; **b)** the dashed rectangular plot consisting of intermediate results including each iteration from the coupled IRC-ICC framework as explained in Figure 3; **c)** the equilibrium state reached by the coupled framework after 5 iterations, and **d)** KGE value compilation graph calculated using 15 minutes intervals for each iteration in the coupled framework.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

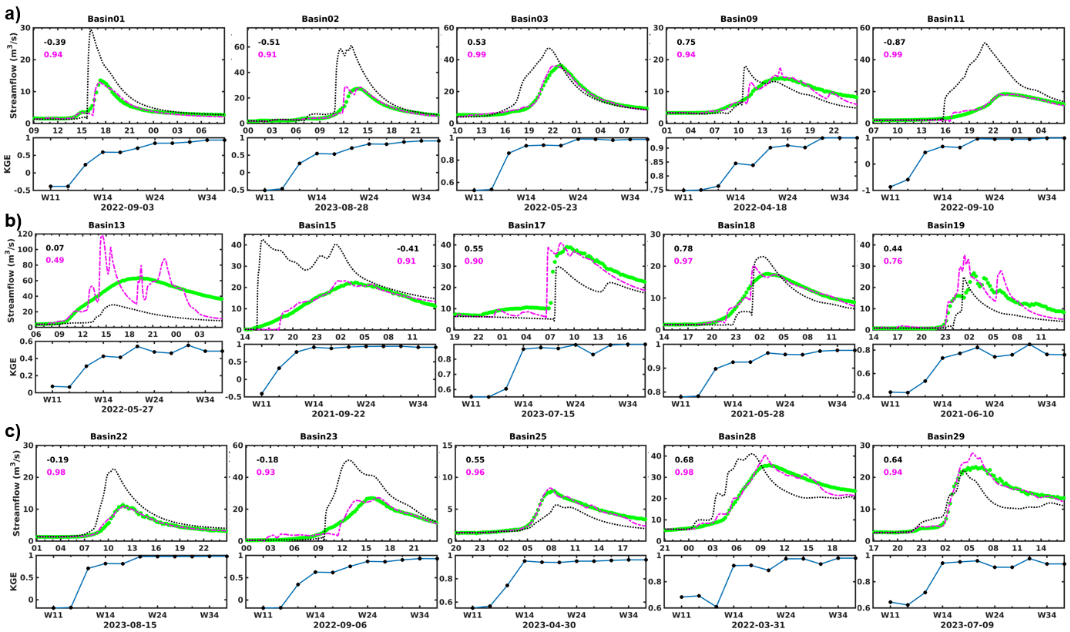

Figure 9 – The systematic application of the coupled IRC-ICC framework to the 28 basins selected in the Appalachians. The results include **a)** 5 events from the Southern Appalachians; **b)** 5 events from the Central Appalachians; and **c)** 5 events from the Northern Appalachians. The IRC-ICC KGE evolution plots from iterations are included below the hydrographs. The black dash line uses the original $STIV_D$ and the pink line is the IRC-ICC equilibrium state ($STIV_D^{IRC*}$), and the corresponding colored numbers are KGE values calculated at 15 minutes interval.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

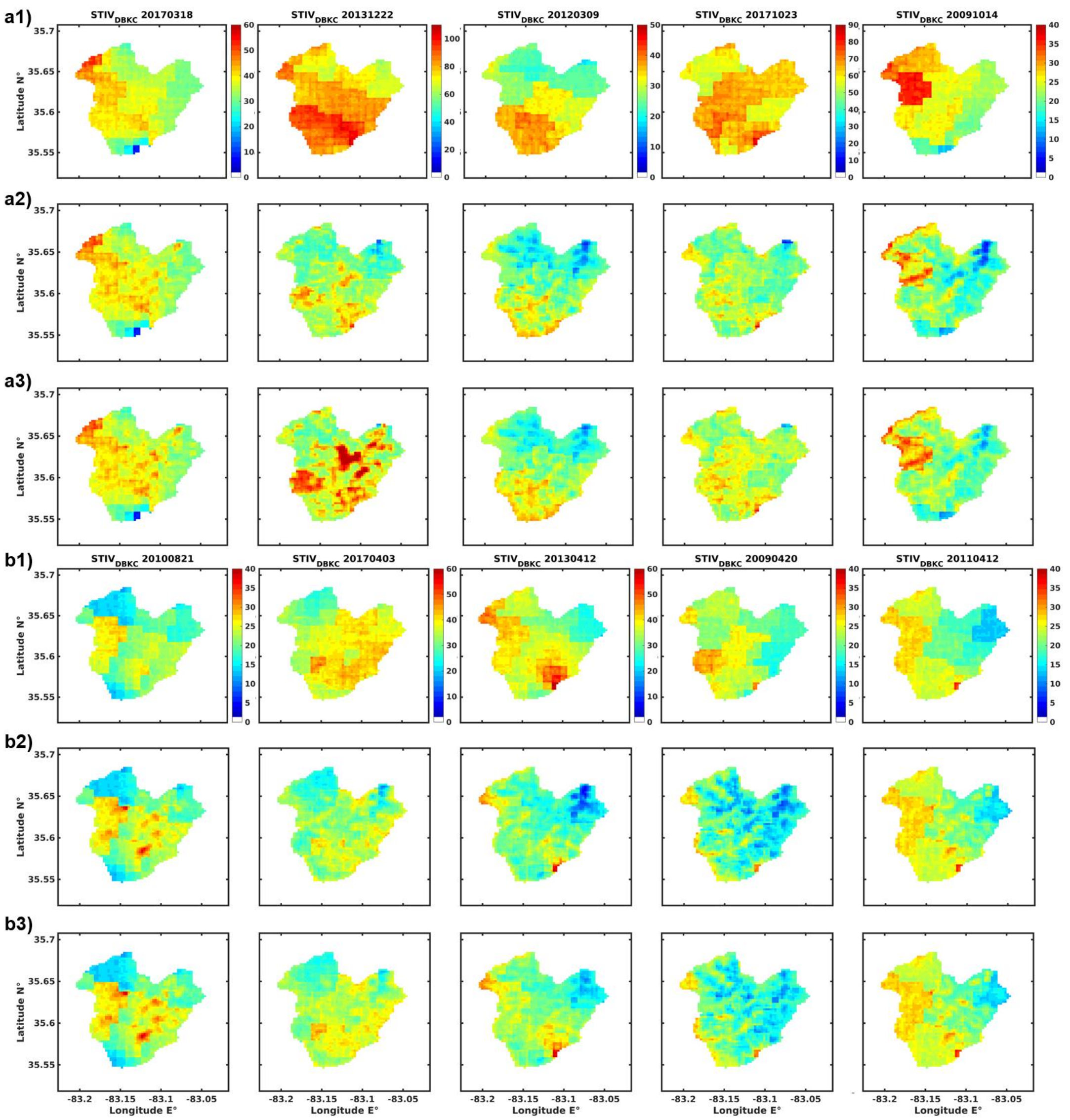

Figure 10 – Event total precipitation maps for **a)** cold season events and **b)** warm season events. Each category includes 5 columns representing different events and 3 rows with the first row (**a1**, and **b1**) representing original precipitation input $\mathrm{STIV_{DBKC}}$, and the second row (**a2**, and **b2**) representing $\mathrm{STIV_{DBKC}^{IRC*}}$ from IRC-only framework, and the third row (**a3**, and **b3**) representing $\mathrm{STIV_{DBKC}^{IRC*}}$ from the coupled IRC-ICC framework.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

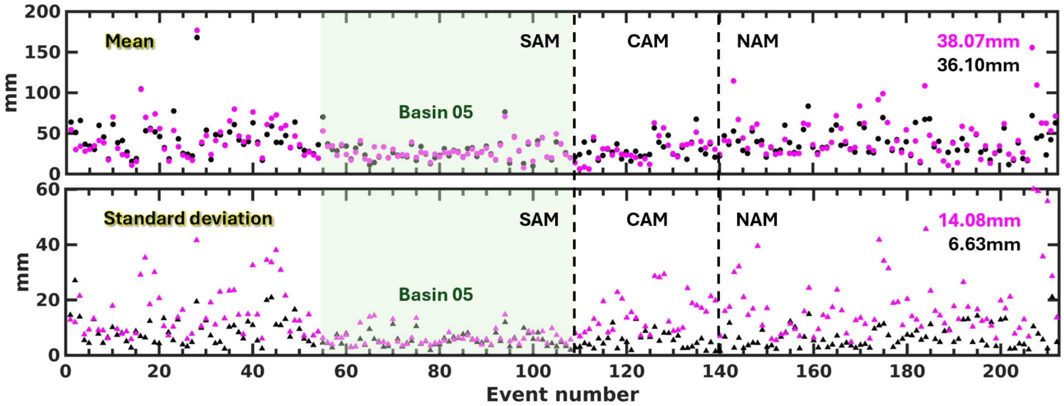

Figure 11 – Summary charts of precipitation statistics for all event-total precipitation maps. Basin mean and standard deviation for each event are represented by circles and triangles in the top and bottom panel, respectively. Each panel is separated into 3 sub-regions by vertical black lines: the Southern Appalachian Mountains, Central Appalachian Mountains, and Northern Appalachian Mountains (SAM, CAM and NAM). The list of events in Basin 05 (with event number ranging from 55 to 108) in the SAM is highlighted by a green rectangle for further discussion in the text. The average values of all events for both the mean and the standard deviation are calculated and shown in the top right corner. Black color and pink color represent pre and post IRC-ICC QPE statistics, respectively.

https://doi.org/10.5194/essd-2024-513
Preprint. Discussion started: 2 January 2025

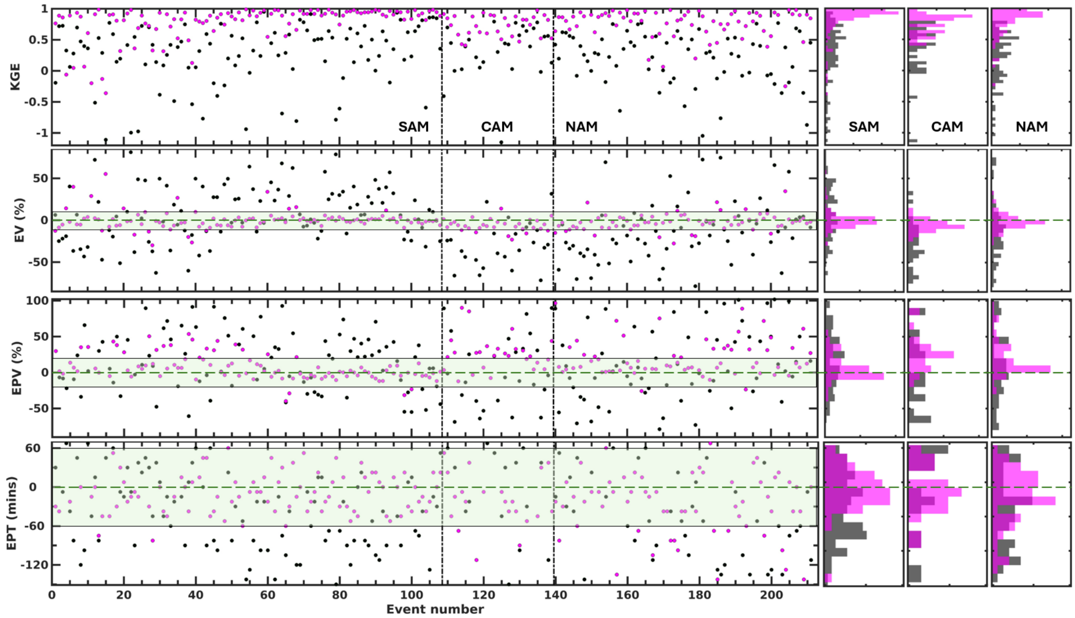

Figure 12 – Summary charts of hydrologic skill metrics for all events. Horizontal green dash lines (i.e. the perfect situation) and green envelopes are for reference purposes. Hydrologic statistics are explained: EPV: Error in Peak Volume (Unit: %), EPT: Error in Peak Timing (Unit: minutes), EV: Error in flow Volume (Unit: %), note KGE is calculated using **15-minute** intervals over a 24 hour period. Pink dots and black dots represent post IRC-ICC results, and original inputs results, respectively (each dot represents one event). Each panel is separated into 3 sub-regions: the SAM, CAM and NAM. Histograms graphs are attached right next to the scatter plot.