Assessing the Impact of Radar-Rainfall Uncertainty on Streamflow Prediction

Nicolás Velásquez[1], Witold F. Krajewski[1], Bong-Chul Seo[2]

[1]IIHR-Hydrosciences & Engineering, Department of Civil and Environmental Engineering, The University of Iowa, Iowa City, IA, USA

[2]Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO, USA

Corresponding author: Nicolás Velásquez (nicolas-giron@uiowa.edu)

Key Points:

- Radar-rainfall biases strongly influence streamflow prediction.
- Radar locations and their interplay with the watersheds create strong spatial bias patterns.
- Radar-rainfall biases vary with the magnitude of the rainfall events.

Abstract

Hydrological models and quantitative precipitation estimation (QPE) are critical elements of flood forecasting systems. Both are subject to considerable uncertainties. Quantifying their relative contribution to the forecasted streamflow and flood uncertainty has remained challenging. Past work documented in the literature focused on one of these elements separately from the other. With this in mind, we present a systematic approach to assess the impact of QPE uncertainty in streamflow forecasting. Our approach explores the operational Iowa Flood Center (IFC) hydrological model performance after altering two radar-based QPE products. We ran the Hillslope Link Model (HLM) for Iowa between 2015 and 2020, altering the Multi-Radar/Multi-Sensor System (MRMS) and the specific attenuation-based (IFCA) IFC radar-derived product with a multiplicative error term. We assessed the forecasting system performance at 112 USGS streamflow gauges using the altered QPE products. Our results suggest that addressing rainfall uncertainty has the potential for much-improved flood forecasting spatially and seasonally. We identified spatial patterns linking prediction improvements to the radar's location and the magnitude of rainfall. Also, we observed seasonal trends suggesting underestimations during the cold season (October to April). The patterns for different radar products are generally similar but also show some differences, implying that the QPE algorithm plays a role. This study's results are a step towards separating modeling and QPE uncertainties. Future work involving larger areas and different hydrological and error models is essential to improve our understanding of the impact of QPE uncertainty.

Plain Language Summary

This study investigates the impact of radar-rainfall on flood forecasting uncertainty. Previous research focused on rainfall-runoff models, ignoring the errors in rainfall estimation. We used a systematic approach to adjust two radar-rainfall products, forcing a simple hydrological model. Results show the potential improvement in streamflow prediction by correcting basin-wide bias in rainfall. The optimal correction varies with basin size, location, season, and rainfall amount.

1 Introduction

Streamflow prediction errors depend on multiple factors, the rainfall-runoff model and quantitative precipitation estimation (QPE) uncertainty being arguably among the most significant. Literature has attributed most streamflow predictive uncertainties to the formulation of the models (e.g. Addor & Melsen, 2019; Fenicia et al., 2008; Gharari et al., 2021), their parameters (Beven & Binley, 2014; Schoups & Vrugt, 2010; Stedinger et al., 2008), and the calibration procedures (Beven, 2012; Duan et al., 1994; Fenicia et al., 2007; Shen et al., 2022). Uncertainty in the QPE has received considerably less attention with few notable exceptions (Ehlers et al., 2019; Kavetski et al., 2006a, 2006b; Bardossy et al., 2022; Liao and Barros, 2022).

In this study, we focus our considerations on radar-based QPE and use comparisons of simulated and observed discharge values of streamflow to quantify uncertainty (Arsenault et al., 2018; Klemeš, 1986; Legates & McCabe, 1999). While our study is motivated by the needs of operational streamflow forecasting (e.g. Krajewski et al., 2017), strictly speaking, we only evaluate model simulation skills. This is because to talk about forecasting, the rainfall forcing would also have to be forecasted. The effect of quantitative precipitation forecasts (QPF) on streamflow forecasting is outside of the scope of this work. For relevant insights, see Ghimire et al. (2022) and Seo et al., (2018).

Another important note is that our operational goal is to forecast streamflow "everywhere," i.e. at small streams and large rivers, and "all the time," i.e., with frequent (e.g. hourly or shorter) forecast updates in nearly real-time. This goal implies that we need to use fully distributed hydrologic models with predictive capabilities across scales ranging from 0.1- to 100,000 km^2. The same goal makes performance evaluation more difficult as the stream gauge network is rather sparse and favors monitoring larger basins.

From a scientific point of view, as opposed to operational, our study is a step towards disentangling the predictive uncertainty into that due to the input versus that due to the model. This is an unsolved problem in hydrology. As rainfall is the key but not the only agent of basin response, errors in the input estimates will affect model calibration and streamflow prediction. Hydrologists have studied this problem mostly in idealized simulation-based experiments or by comparing QPE errors only with gauge observations (He et al., 2013). One simple reason why this problem has remained unsolved is that a comprehensive error structure of radar-based QPE is still unknown despite considerable effort over the past 30 years. See Berne & Krajewski (2013); Krajewski & Smith (2002); Villarini & Krajewski, 2010; and Krajewski & Smith, (2023) for summaries.

However, several recent studies revealed some crucial insights into the aspects of the QPE uncertainty that are particularly important for streamflow prediction. Mantilla et al., (2023) show that small-scale variability and uncertainty in various aspects of runoff production, including rainfall, are effectively filtered out by the river network structure, which aggregates flow. Also, Ghimire et al., (2022) demonstrated that arguably the most important aspect of skillful prediction is using accurate rainfall volume over a given basin. These authors show that while the small-scale (hillslope) errors are not important, the overall space-time distribution is.

If the total rainfall (QPE) volume over an event is the most important, a natural question arises: how much off are the current most used products? Since this question cannot be easily answered over a regional scale with the existing (sparse) rain gauge network density, we have designed a data-based simulation experiment to address it. We modified two existing radar-based QPE products in a simple way, just by scaling them up or down with a multiplicative factor. This does not affect the space-time distribution of storms' key features, such as intensive cells, their velocity, and direction.

To justify our focus on the total rainfall volume, consider the results of an experiment similar to those described in Ghimire et al., (2022) and Krajewski & Smith, (2023). In Figure 1, we show the Kling-Gupta Efficiency index (KGE) calculated to describe the performance of the Hillslope-Link Model (e.g. Krajewski et al., 2017; Mantilla et al., 2022) using two different radar-based QPE products, called here MRMS and IFCA. Each dot in panel (a) denotes model performance at a river basin monitored by a USGS stream gauge. The index is calculated over 112 watersheds over a period of seven years. Significant scatter is evident, with the input being the only difference. In fact, both input products use data from the same radars. Thus, the difference is due to the algorithms converting radar observables to rainfall quantities. When evaluated against rain gauge data, the two products show similar performance (Seo & Krajewski, 2020). In panel (b), we show the change in the index after one of the products (e.g. IFCA) was equalized in the mean over the basin to the other product, i.e. MRMS. Remarkably, the scatter is much reduced. We obtain similar results (not shown) with the equalization in the other direction, i.e. MRMS to IFCA. Note that the overall range of performance has not changed much. We still do not know whether the basin-wide rainfall volume is correct or not or which product is clearly better (according to the KGE measure).

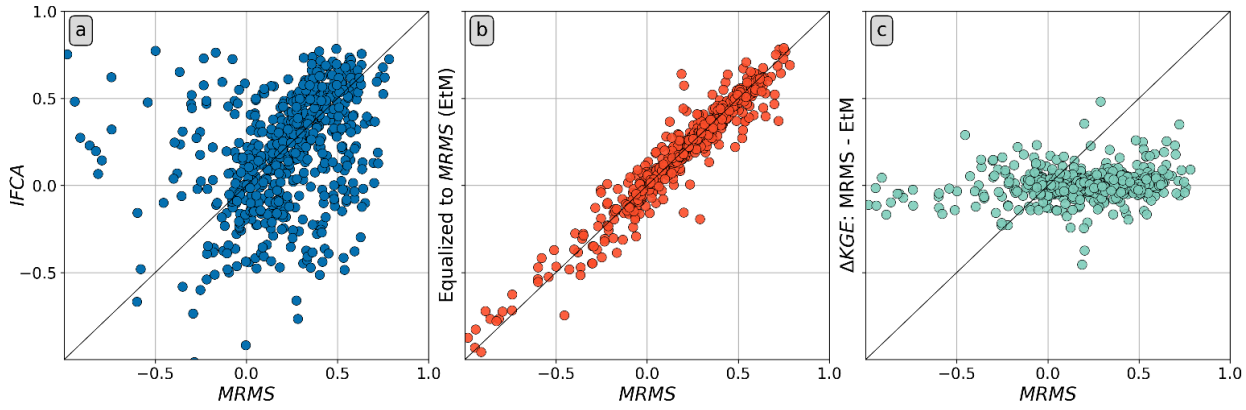


Figure 1. HLM KGE performance forced two QPE products before and after equalization. a) MRMS vs. IFCA KGE before equalization, b) MRMS vs. IFCA equalized to MRMS (EtM), and c) MRMS vs. MRMS - EtM.

Additionally, evident is a spatial component, as shown in Figure 2, where we colored the gauged watersheds with the equalization factors for both products. Looking at the factors, we found spatial patterns and divergences between both products (Figures 2a and b). For example, MRMS estimates more rainfall in Iowa than IFCA. This trend becomes stronger over the western part of the state (draining to the Missouri River). In some regions, the spatial trends seem linked to the locations of the NWS weather radars and their domain coverage. Note how the factors change in

some nested watersheds, suggesting discrepancies between their QPE errors even at relatively local scales. Several features in the QPE products may explain the described differences. For example, MRMS estimates are corrected using rain gauge data (Zhang et al., 2016, 2020), while IFCA are not. Each watershed has a varying distance to the respective radars, which, mixed with the storm tracks, creates variability in the QPE quality. Regardless of the factors affecting the rainfall products, radar bias affects streamflow prediction (**¡Error! No se encuentra el origen de la referencia.**). The observed variability increases the challenge of assessing QPE errors in streamflow predictions and the need to incorporate the spatial component into its analysis.

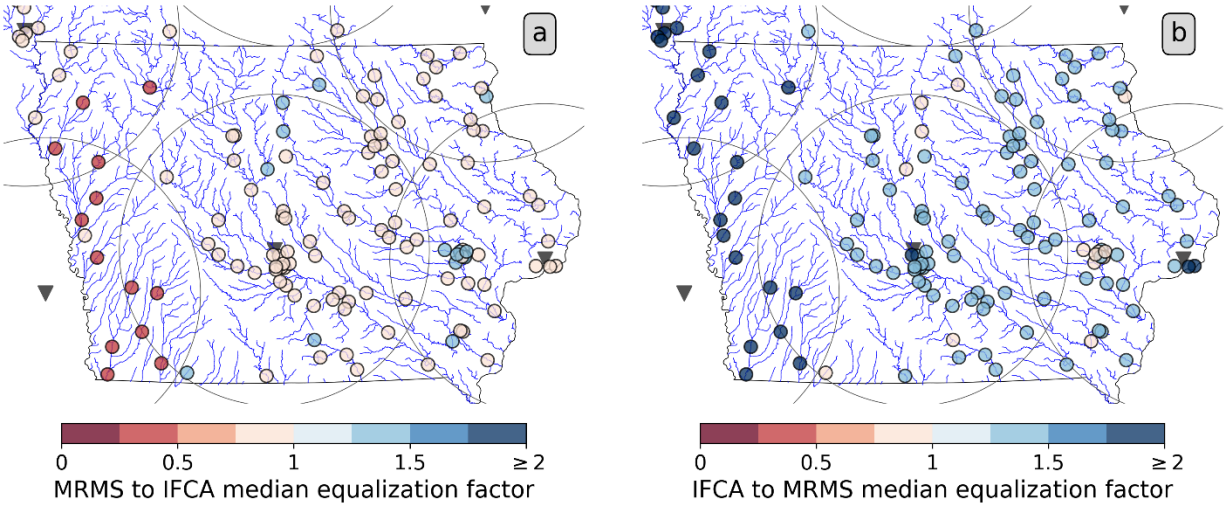


Figure 2. The median value of the multiplicative factors (f_c) used to equalize the QPE products at each USGS gauge (colored dots). Panel (a) presents the f_c values to equalize MRMS to IFCA and panel (b) the f_c valules for the IFCA to MRMS case. The blue lines represent the river network, the black triangles represent the meteorological radars, and the circles have a coverage radius of 150km.

If the rainfall volume dominates the streamflow simulation, the question arises of whether we can improve it further. Note that the MRMS product has already been rain gauge data corrected, but not necessarily in the mean-basin sense. Since the "potential" of using rain gauge data to adjust the rainfall input has been exhausted, we resort to a simulation experiment. In the remainder of this paper, we describe our methodology in Section 2, the key results in Section 3, and we close with a discussion in Section 4.

2 Materials and Methods

To examine the implications of QPE errors on flood forecasting uncertainty, we conducted a simple simulation experiment over multiple watersheds and over several years of data. We used two radar-based QPE products, forcing the Hillslope Link Model (HLM). The two QPE products are the Multi-Radar/Multi-Sensor (MRMS) (Zhang et al., 2016, 2020), which has national distribution, and an in-house product developed by IFC (IFCA) based on the specific attenuation algorithm proposed by (Ryzhkov et al., 2014) and implemented by Seo et al., (2020). Both products are derived as a mosaic of data from the following seven WSR-88DP weather radars: KDVN in Davenport, Iowa, KDMX in Des Moines, Iowa, KEAX in Kansas City,

Missouri, KOAX in Omaha, Nebraska, KFSD in Sioux Falls, South Dakota, KMPX in Minneapolis, Minnesota, and KRAX in La Cross, Wisconsin. We clarify that MRMS QPE used for our study period was generated using reflectivity-based algorithms (Zhang et al., 2016), and the synthetic QPE algorithm (Zhang et al., 2020) mainly based on specific attenuation was not operationally implemented for the period. In our approach, we altered the QPE products with multiplicative factors (f_c) that took values between 0.1 and 5. We applied the proposed framework to 112 watersheds defined by the USGS streamflow gauges in Iowa between 2015 and 2020.

We also used a prior, reflectivity-based IFC product (e.g. Seo et al., 2019), but we do not show the results as they are qualitatively similar. For each product, we assumed a multiplicative error represented by f_c. We ran the HLM configuration described in Velásquez et al., (2023) for each product after altering it and evaluated the performance metrics at each observed hydrograph. Finally, we assessed the impact of rainfall bias by analyzing the results corresponding to the highest performance factors. In Figure 3, we present a schematic of the experiment setup and describe it in detail.

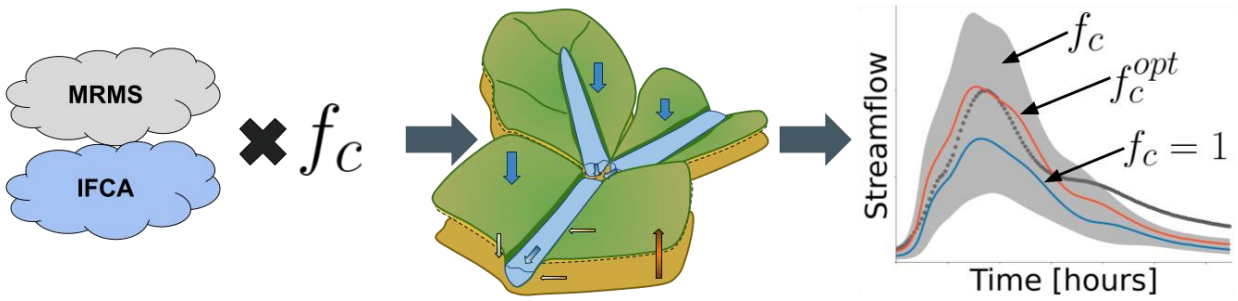


Figure 3. Experiment setup from left to right: QPE alteration using the multiplicative factor (f_c), open-loop streamflow simulation using the Hillslope Link Model, and simulations evaluation at the event scale by comparing $f_c = 1$ (blue line) with the factor that provides the optimal performance f_c^{opt} (red line).

2.1 Study area

For our analysis, we used 112 USGS gauges (colored dots in Figure 2) within Iowa, USA. The gauges monitor watersheds with areas ranging between 10 and 36,000 km^2. With a landscape dominated by gently rolling plains, Iowa's land use is predominantly agricultural with two main crops being corn and soybeans. Nevertheless, the region also has prairie potholes (in the Des Moines Lobe area), deep loess deposits (around 30 meters deep) at the Loess Hills, and two extensive alluvial plains next to the Missouri and Mississippi rivers. Due to its landscape and weather, Iowa has been affected by several historical floods, usually in spring and summer.

2.2 The Hillslope Link Model

As described in detail by Mantilla et al., (2022), the Hillslope Link Model (HLM) is a distributed hydrological modeling framework that solves a set of ordinary differential equations (ODEs) representing the hillslope processes and channel routing while preserving the river network topological connectivity. As a framework, HLM allows the formulation of runoff mechanisms with different complexity and parameter requirements. For this experiment, we used the HLM

formulation that includes a nonlinear representation of the subsurface fluxes (Fonley et al., 2021; Velasquez et al., 2021) and a snowmelt parametrization (Koya et al., 2023; Velásquez et al., 2023). The nonlinear sub-surface fluxes better represent the hydrograph recession and baseflow. At the same time, the snow component captures the winter-to-spring transition events. We set up HLM using a network derived from a 10 m digital elevation model (DEM) (USGS, 2022) that closely follows the high-resolution version of the NHDPlus network (USGS, 2017), obtaining around one million links over Iowa.

2.3 Rainfall Uncertainty Assessment

We opted to perform our study at the event scale to avoid the effects of prolonged no-rain periods in continuous simulation. We identified around 7,800 basin response events in the 112 discharge gauges between 2015 and 2022. Following, we provide a detail of the event identification and the relative rainfall (R_r) estimation procedures.

2.4 Events Identification

We evaluated all the model outputs using the standardized discharge (Z) by dividing the observed and simulated discharges (Q_o and Q_s, respectively) by the mean annual peak flow ($\bar{Q}_{peak}$) of each gauge, allowing us to compare results from watersheds with areas varying between 10 and 36,000 km^2. After the standardization, we identified the events at each watershed by separating the runoff from baseflow using the (Nathan and McMahon, 1990) filter:

$$Z_r(t) = a \cdot Z_r(t-1) + \frac{1+a}{2} \cdot (Z_o(t) - Z_o(t-1)) \quad (1)$$

where $Z_r(t)$ is the runoff component at time t, a a parameter set to 0.98, and Z_o the standardized observed discharge. Using Z_r, we create a binary time series (B_r) where $B_r = 1$ if $Z_r > Z_{crit}$ and 0 otherwise. Z_{crit} is the runoff threshold set equal to 0.01. Then, we take the first derivative of B_r (ΔB_r) to identify the beginning of each event (ΔB_r = 1). The end of the events corresponds to the beginning of the next one or the watershed response time (whichever happens first). After using the described procedure, we evaluated only the events with a relative peak flow greater than 20% of the mean annual peak flow.

2.5 Performance Metrics

For each event, we computed the performance metrics, the average relative rainfall (R_r) and identified the optimal factor (f_c^{opt}) for the performance metrics. The metrics include the peak flow bias (Q_p), the time-to-peak difference (Δt_p), and the volumetric bias (V). With the described metrics, our goal is to obtain a robust assessment of the QPE basin-wide bias while analyzing its impact on features of the simulated hydrographs related to flood forecast performance.

2.6 Relative Rainfall (R_r) Estimation

We used R_r to perform conditional analyses of our results. To identify R_r, we first obtained the accumulated rainfall (R $[mm]$) of each event between a temporal lag and the peak flow time. We estimated the lag for each watershed by looking at the correlation between the observed discharge and the MRMS average rainfall aggregated between t and $t - \Delta t$ with Δt varying

between 2 and 400 hours. The lag corresponds to the Δt value that maximizes the correlation. Finally, we obtained the total relative rainfall of the events (R_r) dividing R by the average of the sum of the observed events.

Additionally, we estimated event-based mean areal precipitation over 60 watersheds using rain gauge records from the National Weather Service (NWS) Cooperative (COOP) Hourly Precipitation Data (HPD) Network, Version 2.0 (Lawrimore et al., 2020). Mean areal precipitation was estimated by taking a simple average of rainfall observations from corresponding rain gauges within each watershed. We used these estimates to calculate QPE bias and compare the calculated bias with the optimal factor (f_c^{opt}) derived from our experiment.

3 Results and Discussion

As described in Section 2, we use a collection of multiplicative factors (f_c) to represent the QPE uncertainty of two products (MRMS and IFCA). This simple representation of QPE uncertainty does not affect small-scale (pixel) variability and errors and, thus, the spatial distribution of the radar-rainfall patterns. Instead, it focuses on the basin-wide bias of the radar-rainfall input. In the following, we present the assessment of the results obtained at 112 USGS gauges with discharge observations. For each watershed and each event, we evaluated the $QPE \cdot f_c$ combination that provided the best performance. As shown in Figure 4, changes in f_c provide significant differences in streamflow simulation performance compared with the original result for $f_c = 1$. In our analysis, we assumed that QPE bias explains the f_c variability allowing us to analyze possible structural biases in the radar-rainfall products.

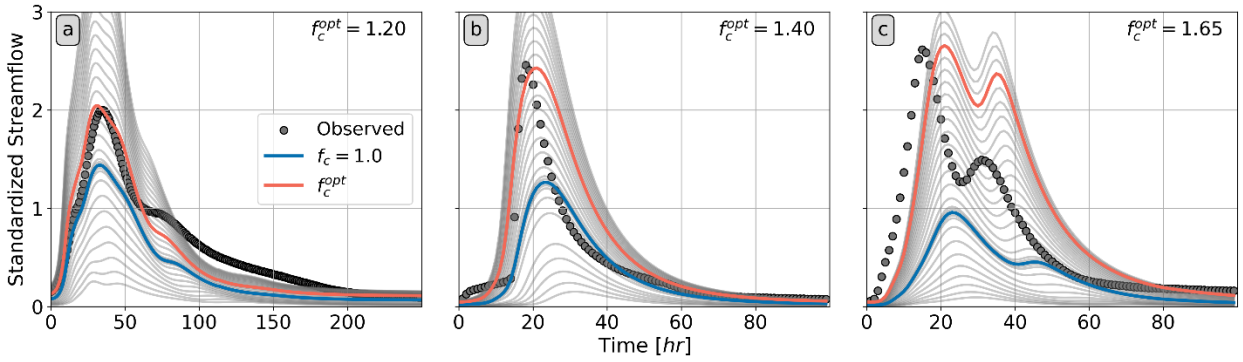


Figure 4. Comparison of simulated hydrographs using $f_c = 1$ (blue) and f_c^{opt}(red) according to the peak flow bias (Q_p bias). The gray hydrographs correspond to simulations using f_c values between 0.2 and 1.6.

3.1 Multiplicative Factors and Performance Frequencies

As shown in Figure 4, QPE magnitude changes induced by f_c can bring significant predictive improvements. To achieve these improvements, f_c^{opt} exhibited notable variability (between 0.7 and 2) representing QPE over- and under-estimations, respectively (see Figure 5). Despite the f_c^{opt} range, its distribution and magnitude are comparable to the gauge/QPE biases obtained at the watersheds (black line in Figure 5), where the range and median values (~1.2 for f_c^{opt}, 1.14 for MRMS, and 1.12 for IFCA) provide a validation of our results. On the other hand, the f_c^{opt}

distributions exhibited similar histograms for the selected performance metrics with some differences between our two QPE products.

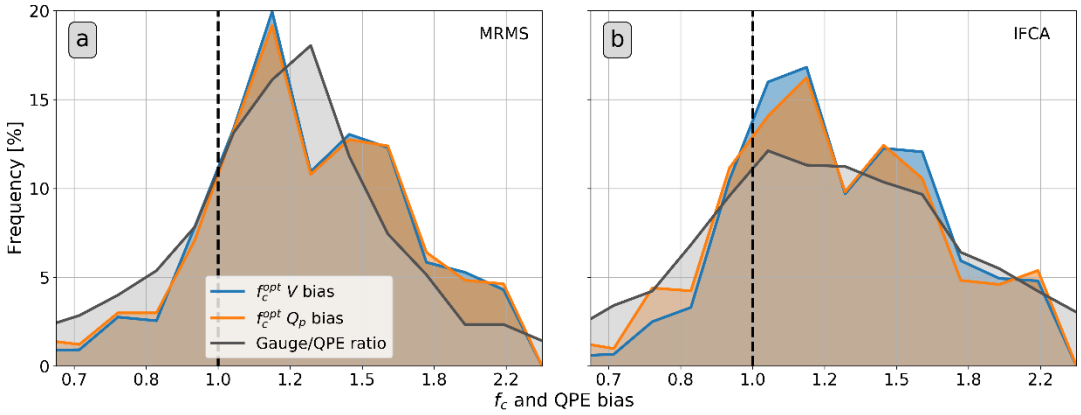


Figure 5. Multiplicative factors frequency distribution for the V bias (blue), Q_p bias (orange) metrics, and for the rainfall Gauge/QPE ratio (black). Panels a and b correspond to MRMS and IFCA results, respectively.

We further compare f_c (using V bias) with the gauge/QPE bias by analyzing their joint distribution (Figure 6). In contrast with Figure 5, this approach allows us to present in more detail the differences between f_c and the bias of both products. In the MRMS case, both f_c and the gauge/QPE ratio exhibit some converge for values between 0.9 and 1.3 with significant discrepancies for f_c values above 1.3. The IFCA case, exhibits a stronger agreement between f_c and the gauge/QPE ratio with a higher frequency around 1 and more disagreement on the QPE for values below 0.9. We attribute these differences to discrepancies in the spatial organization of both QPEs and the low density of gauges used in the gauge/QPE bias estimation. These differences between the two QPE products may come as a surprise. Recall that the MRMS is gauge-corrected.

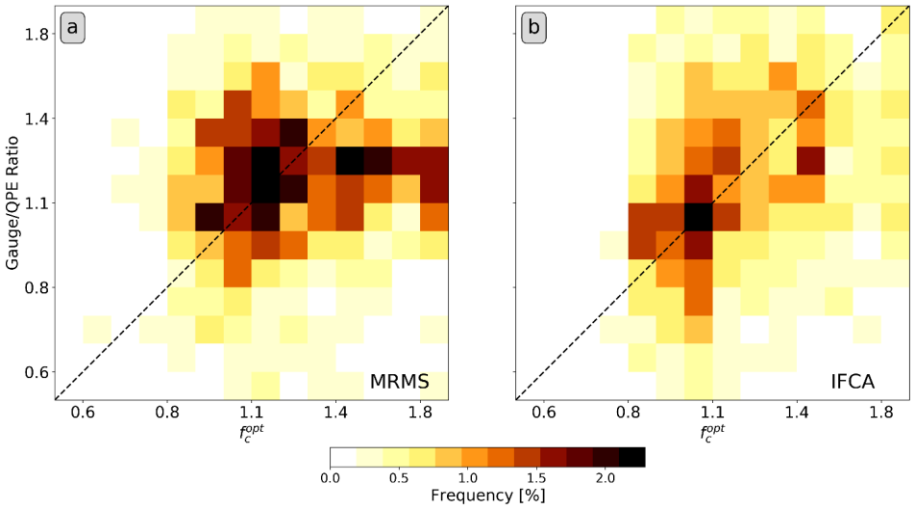


Figure 6. The joint distribution between the V bias f_c^{opt} values and the Gauge/QPE ratio for MRMS (a) and IFCA (b).

We attribute the described discrepancies to differences in the way we obtained f_c^{opt} and the gauge/QPE ratio. The f_c^{opt} represents the bias at a watershed scale looked through the lens of a hydrological model. On the other hand, the gauge/QPE ratio is the bias of a direct comparison between gauged rainfall at a given point and the collection of pixels that correspond to a given watershed. Due to differences on both approaches, drawing strong conclusions about their correctness is difficult. Nevertheless, their magnitude and distribution similarities point to an existing issue on the QPE.

Additionally, we evaluated the potential bias induced by the QPE uncertainty by comparing the performance frequency of using a multiplicative factor of 1 and f_c^{opt} value (Figure 7). According to our results, f_c^{opt} induces a large improvement in both QPE products. The improvement is impressive for the three metrics, with differences bounded within 50% and centered around 0. As discussed before, model uncertainties likely explain a significant portion of the errors in the $f_c = 1$ case. Improvements in the model parameters (Francés et al., 2007; Refsgaard, 1997) or a better representation of the runoff and routing processes (Velásquez et al., 2021; Velasquez et al., 2022) can also increase model performance.

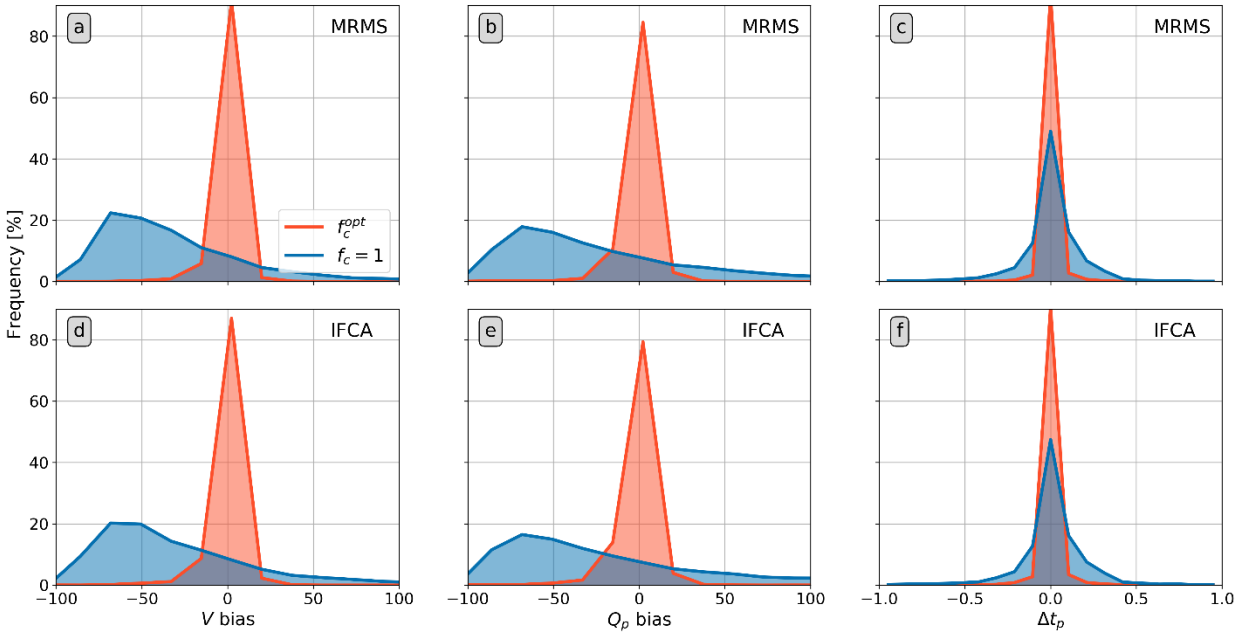


Figure 7. Model performance frequencies for the case of $f_c = 1$ (blue) and for f_c^{opt} (red) during each event. The rows correspond to the MRMS and IFCA results, respectively. The columns correspond to V bias, Q_p bias, and Δt_p histograms.

In the Q_p bias case, improvements represented changes from around 50% to 2% for MRMS and IFCA (see Figure 8). In the MRMS case (Figure 8a), we observe more dispersion for the f_c^{opt} Q_p bias case with values oscillating between 1 and 7%. On the other hand, the IFCA case (Figure 8b), most of the change happens around 4%. Despite the differences, both QPE products exhibit a similar improvement, indicating that the basin-wide magnitude is a key feature controlling the hydrograph. The results presented here and by Ghimire et al., (2022) highlight the essential role

of the QPE and how critical it is to understand its uncertainty at the watershed level to improve our forecasting capabilities.

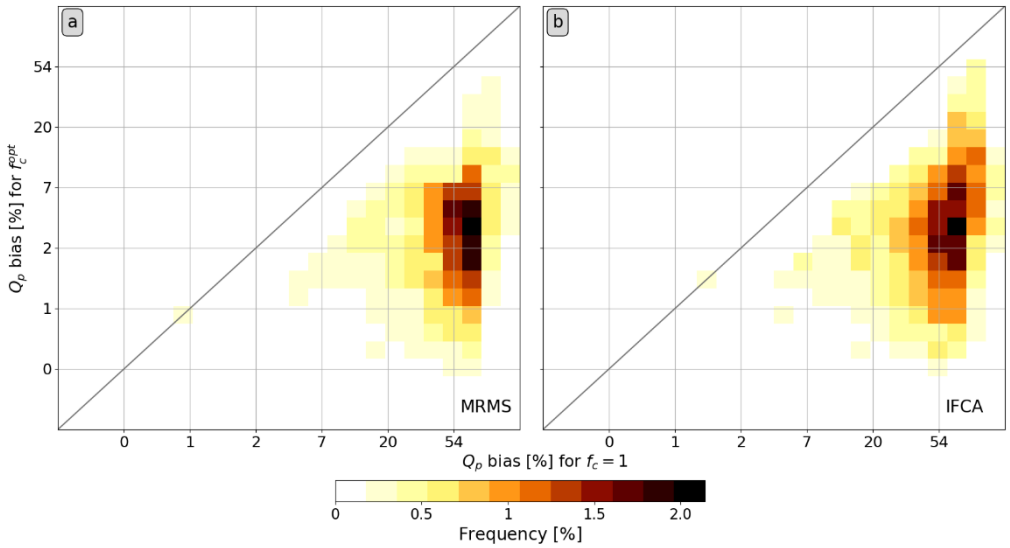


Figure 8. Q_p bias performance (in log space) for f_c equals to 1 vs f_c^{opt} for MRMS (a) and IFCA (b).

Up to this point we analyzed the corrective factors as a distribution over all stream gauges and events. Therefore, a fair question could be whether these factors are random which would make them less insightful. To address this question we analyzed their temporal variations at the seasonal scale and the spatial variations as basin scale and simple maps.

3.2 Seasonal trends

It is well-established that QPE uncertainty follows seasonal patterns (Bytheway et al., 2020; Ciach et al., 2007; Derin et al., 2016; Gupta et al., 2010) due to atmospheric changes (Lu et al., 2010) that induce error sources such as melting layer height changes (Cocks et al., 2017). Considering this, we explored the performance improvements and f_c^{opt} variability for the MRMS case (Figure 9). The three metrics exhibited performance increases in the transition periods of winter to spring and summer to fall. V bias (Figure 9b) have the most pronounced improvements in the winter to spring transition while Q_p bias (Figure 9a) has it during the summer. f_c^{opt} also follows the similar pattern for all the metrics except Δt_p (Figure 9d and e). During the winter to spring transition f_c^{opt} tends towards values above 1.5 and in the summer, it oscillates around one. The described seasonal patterns suggest a link between the QPE-modeling uncertainties and the atmospheric changes worth of exploring.

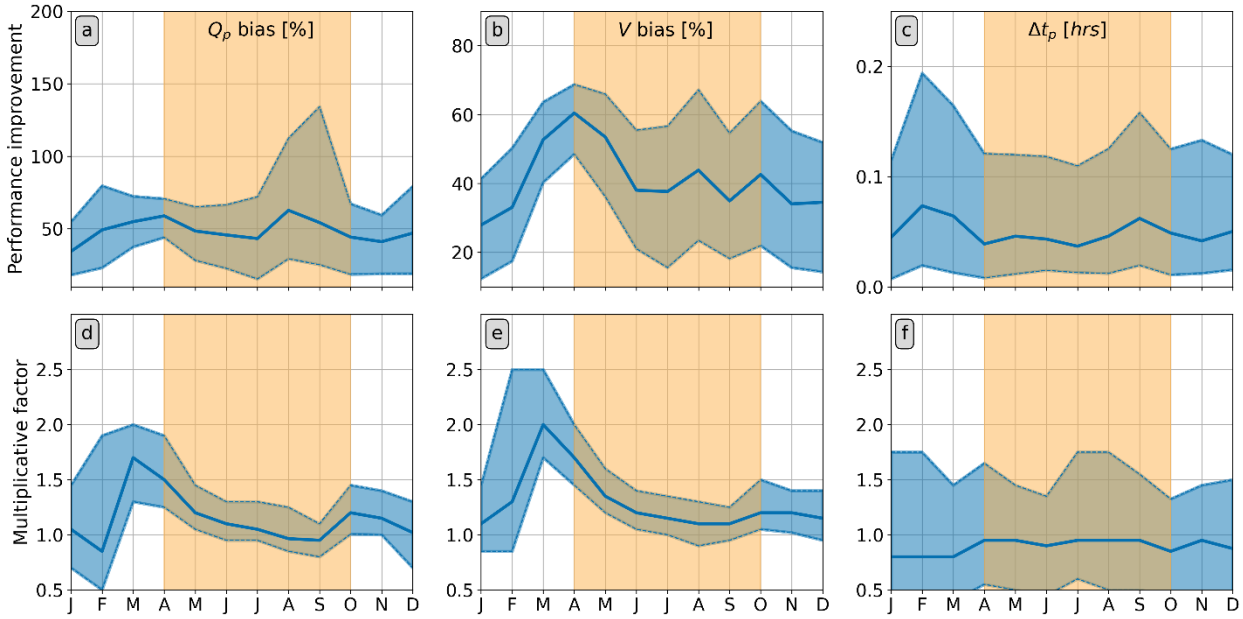


Figure 9. Performance improvement (first row) and f_c^{opt} (second row) seasonal variability for HLM runs with MRMS. The yellow vertical band indicates the warm season, the blue band corresponds to the interquartile range, and the tick blue line corresponds to the median value.

3.3 Scale analysis

Scaling of flood peaks at the event scale has been a well-researched problem (e.g. Ayalew et al. 2014a, b; Gupta et al., 2010). Therefore, in addition to the seasonal uncertainties, we analyzed the role of the watershed scale. We compared the performance improvement and the multiplicative factor variations with the upstream area of the gauged watersheds (Figure 10). According to the figure, the performance improvement has a weak link with the upstream area where only Δt_p shows an improvement for areas ranging between 10 and 100 km^2. Conversely, Q_p bias, V bias, and f_c^{opt} (second row) do not present a clear relationship with the scale. For most of the performance indexes, f_c^{opt} oscillates between 1 and 1.5. For Q_p bias, f_c^{opt} slightly decreases with the scale, and in the V bias case, its variability (blue bands) seems to increase. Nevertheless, a strong connection with the watershed area is not evident.

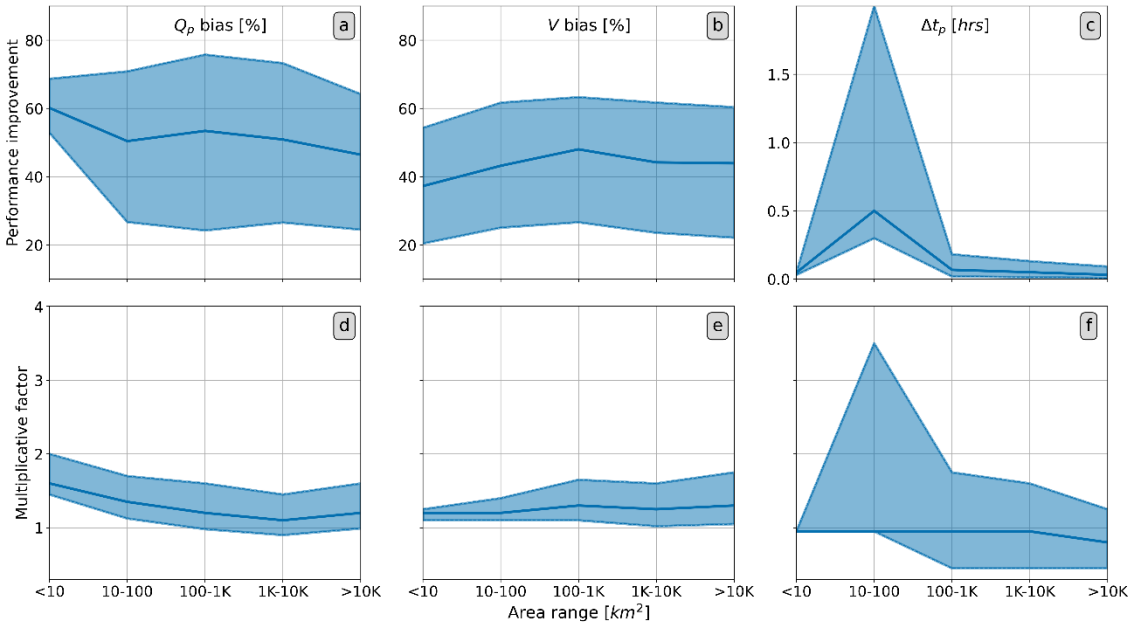


Figure 10. Performance improvement (first row) and f_c^{opt} (second row) versus the watershed upstream area. a) and d) correspond to Q_p bias, b) and e) to V bias, and c) and f) to Δt_p.

3.4 Spatial patterns

Here, we expand the assessment to the spatial domain, considering the median value of the performance at each gauge, its improvement, and f_c^{opt} value. Figure 11 presents the median Q_p bias performance at each USGS gauge for the two QPE products. According to the figure, the products have similar spatial patterns but with some differences. The most significant difference corresponds to performance differences between MRMS and IFCA over the south-west and north-east areas. We attribute the similarity between the results to the model parameterization and the shared origin of the radar data used to develop the QPEs. The performance shown in Figure 11 is a reference for the following analysis, where we compared the forecast improvements for the optimal f_c^{opt} at each gauge.

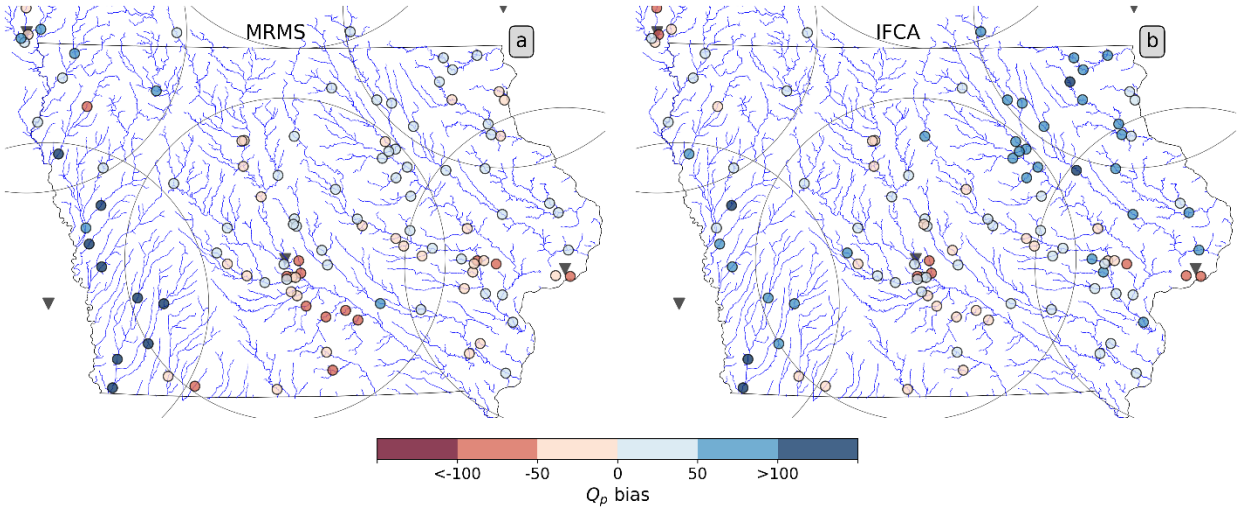


Figure 11. Event-based Q_p bias performance for HLM forced with MRMS (a) and IFCA (b). The colored dots represent the median Q_p bias, the black triangles represent the radars and the dark circles, the radar coverage at 150 km.

In Figure 12, we present the improvements for Q_p and V bias performance metrics for the MRMS and IFCA products. Accordingly, gauges of relatively low improvement coincide with high performance at $f_c^{opt} = 1$ (Figure 11). Moreover, we also found similarities and differences between the QPE products. In the Q_p bias case (Figure 12a and b), we found more differences with higher improvements over the north for IFCA and around the south-west for MRMS. With larger improvements in IFCA, the V bias pattern (Figure 12c and d) is similar for both QPEs presenting more significant differences over the North-West and Center. The described patterns suggest the existence of spatial QPE biases, probably due to the radars' characteristics (and limitations), the algorithms used to merge them, and differences between the atmospheric conditions at the radar domains.

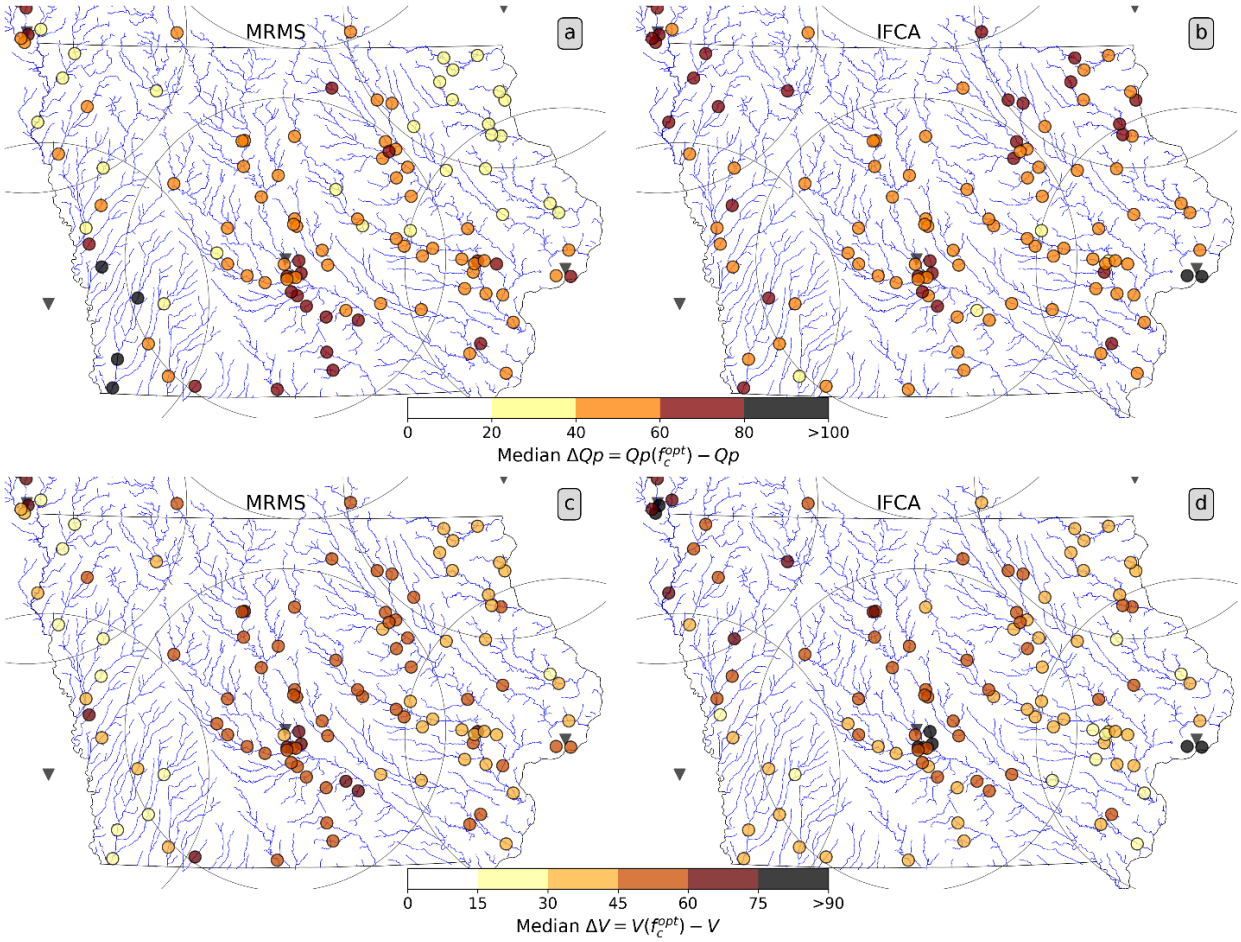


Figure 12. The median value of the Q_p bias (a and b) and V bias (c and d) performance improvements after comparing events using $f_c = 1$ and f_c^{opt} for MRMS (a and c) and IFCA (b and d).

Following the performance improvement biases, the median f_c^{opt} value at each gauge (Figure 13) also arises a spatial pattern. The f_c^{opt} variability coincides at some extent with the performance improvement results. Nevertheless, Figure 13 exhibits stronger spatial trends and some differences between the Q_p and V bias cases (first and second rows, respectively). In contrast with Figure 12, high and low f_c^{opt} values tend to be more influenced by the radar's coverage radius. For instance, watersheds close to the South-West radar tend to have f_c^{opt} values around 0.8 while the watersheds around the central radar have f_c^{opt} values around 1.3. We observe a similar behavior with the other radars, on the East, f_c^{opt} values oscillate around 1.1 and, in the center, it takes values between 1.0 and 1.5. However, the described f_c^{opt} values are in overall larger for the V bias performance metric.

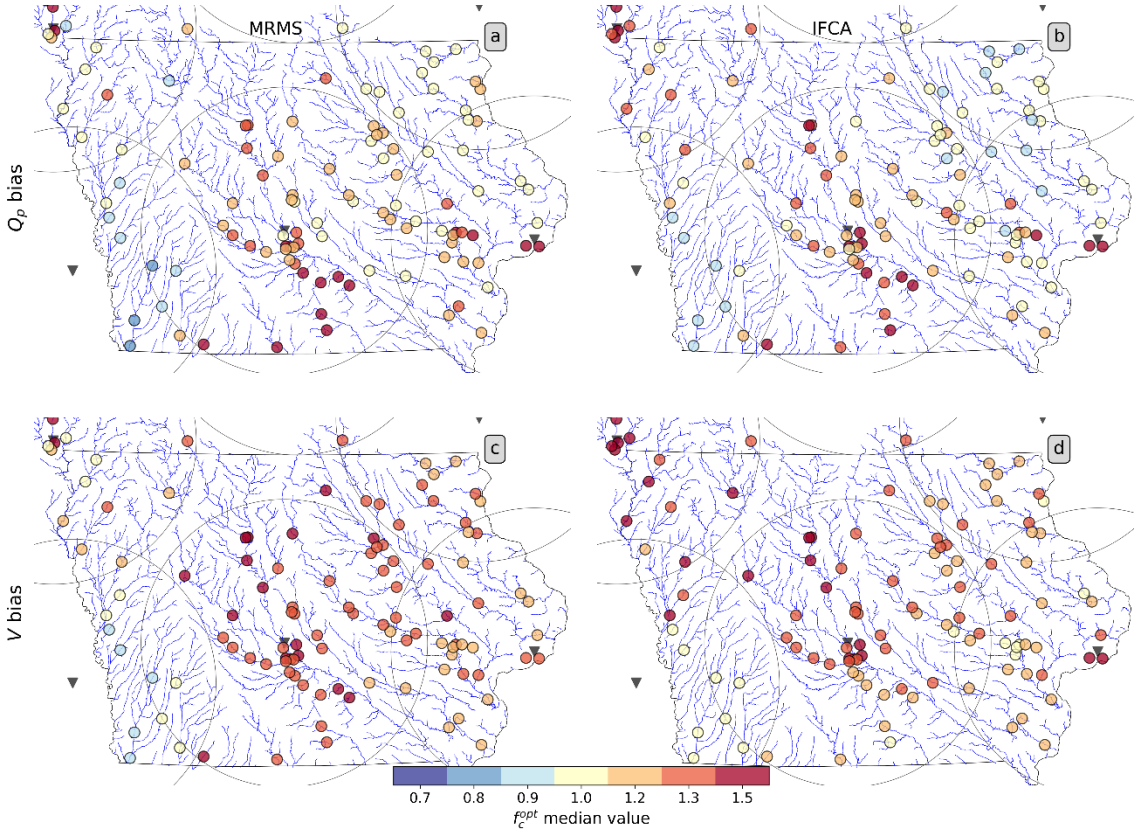


Figure 13. Median f_c^{opt} values for each product (columns) and for the Q_p bias and V bias metrics (rows). Yellow colors indicate f_c^{opt} values close to 1.0, blue values below 1.0, and red values over 1.0.

The spatial variability of f_c^{opt} exhibits some differences among the QPE products. In the IFCA case (Figure 13 b and d), f_c^{opt} values around the North-West radar tend towards higher values, while in the MRMS case (Figure 13 a and d), South-West values are lower. We attribute the differences to the algorithms used to develop each product. Nevertheless, both products exhibit similar patterns, which are also present when we compare f_c^{opt} with the performance increase (Figure 14). Here, we find similar patterns with increased performance for f_c^{opt} values different

from one. On the other hand, the MRMS case (blue dots) exhibit lower f_c^{opt} values for the ΔQ_p and ΔV bias cases corresponding to larger improvements indicating QPE overestimations.

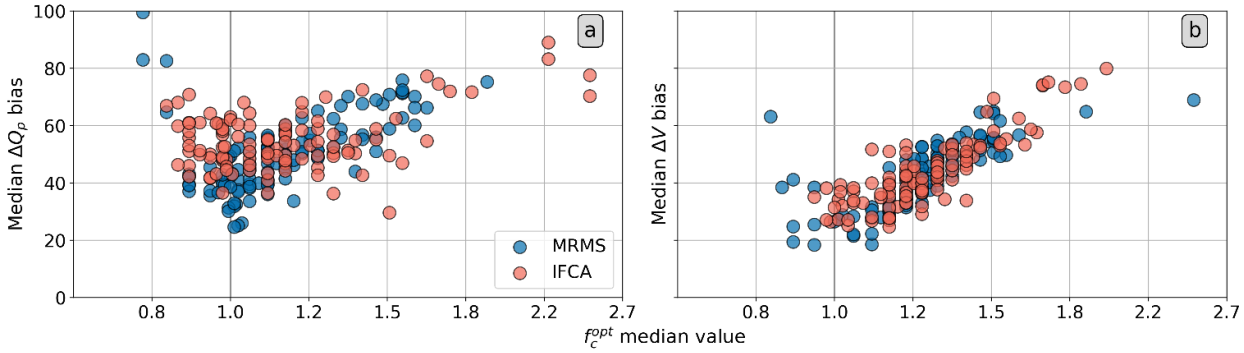


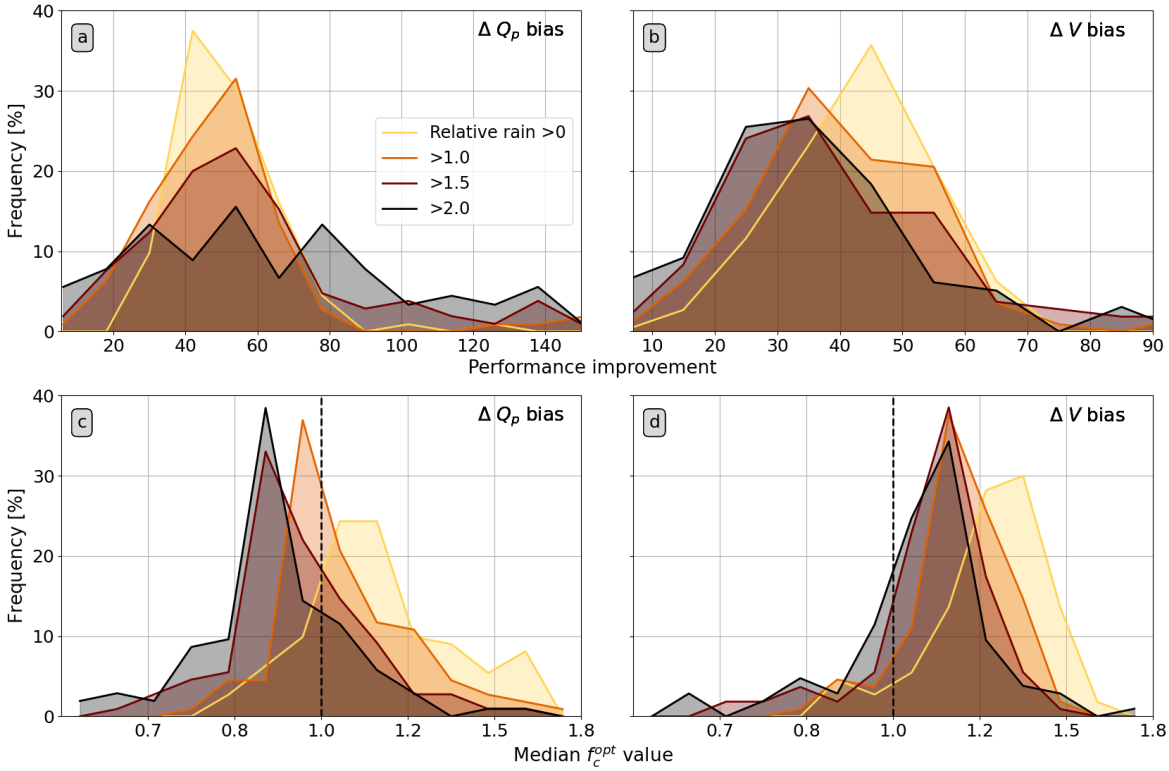
Figure 14. Comparison of the median f_c with the median of the performance improvement for $logKGE$ (a), Q_p bias (b), and V bias (c) for MRMS (blue) and IFCA (red).

The similarities between the QPE products suggest consistent biases affecting the simulations at the event scale. However, we also notice differences that highlight the uncertainties of each QPE. Previous work by Quintero et al., (2020) also presented performance differences when comparing MRMS, Stage IV (Lin & Mitchell, 2005), and IFC rainfall. In this case, we explored how the QPE uncertainty impacts the forecasting skill of a distributed hydrological model. According to our spatial analysis, there is variability in the function of the radars and their unique characteristics, as previously suggested by (Post & Krajewski, 2023). However, the relationship seems complex and requires a more in-depth study considering QPE from independent radars and the distance to the watersheds.

3.5 Rainfall Conditioned Assessment

The nature and magnitude of rainfall events also condition QPE uncertainty (Derin et al., 2016; Lu et al., 2010), and therefore it may also change the streamflow forecast biases. We used the relative total rainfall (R_r) of each event and watershed to explore this link. We repeated the previously described analysis, conditioning the events to R_r values greater than 0, 1, 1.5, and 2. Figure 15 presents a summary of this analysis for the performance improvement (first row) and f_c^{opt} (second row).

According to the performance improvement histograms, ΔV bias (Figure 15b and d) exhibit some reduction for R_r thresholds above 1.0. Conversely, the ΔQ_p bias increase with R_r (Figure 15a and c). Nevertheless, in both cases, the connection between R_r and the performance improvement seems weak. We attribute this weakness to additional uncertainty factors other than the QPE. On the other hand, the relationship is stronger in the f_c^{opt} case. In both cases, f_c^{opt} tends to decrease with the R_r magnitude.



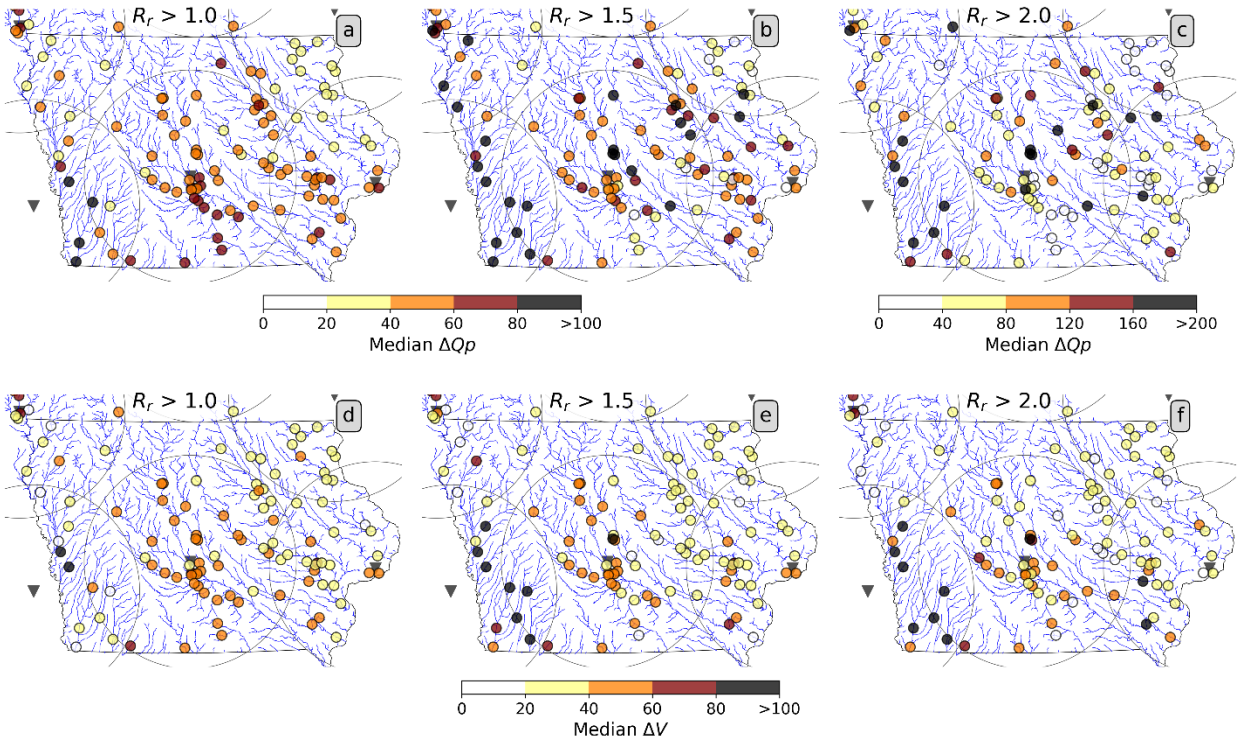
405

406 **Figure 15**. Summary of the performance improvement and f_c^{opt} change conditioned to MRMS
407 relative rainfall R_r. The first row presents the performance improvement histograms, and the
408 second row to the f_c^{opt} histogram. The colors from yellow to black correspond to increasing R_r
409 thresholds (0 to 2).

410 ### 3.6. Spatial analysis

411 The results we described in Figure 15 suggest a relationship between the magnitude of the
412 rainfall event and the QPE-modeling uncertainty. During large events, QPE overestimations may
413 lead to significant errors in estimating hydrological signatures, such as the peak flow, which in
414 this case are corrected through f_c^{opt}. However, the described relationships for the performance
415 and f_c^{opt} also have an spatial component as shown in (Figure 16 and Figure 17). In contrast with
416 Figure 12 a and c (corresponding to MRMS), the improvement exhibits an accentuated spatial
417 change for increasing R_r values.

418 The Q_p bias case (Figure 16 a to c), exhibits significant improvements at most of the watersheds
419 (> 100%) with radar-related solid patterns in the three cases. In the $R_r > 2.0$ case (Figure 16 c),
420 the ΔQ_pbias reached values larger than 200% indicating increased biases during events with a
421 significative rainfall accumulation. Also, there is a more evident difference between the
422 southwest and northeast radars. The ΔV case (Figure 16 a to c), also follows a spatial pattern that
423 accentuates with R_r. In contrast with ΔQ_p, ΔV has a stronger differentiation between radars for
424 increased R_r values.



425

426 **Figure 16**. Q_p bias and V bias performance increases conditioned on the MRMS relative rainfall
427 (R_r) of each event. The columns correspond to R_r values above 1, 1.5, and 2. The rows
428 correspond to performance differences in the Q_p bias (a to c) and V bias (d to f) performance
429 metrics.

430 Like the results described for the conditional performance increase (Figure 16), the
431 multiplicative factors also increased with R_r (Figure 17). In both, Q_p and V bias cases, f_c^{opt} took
432 more extreme values for larger R_r values. In the Q_p case (Figure 17a to c), f_c^{opt} bounds change
433 from 0.8 and 1.3 (Figure 17a) to values below 0.5 and above 1.8 (Figure 17c). The V bias case is
434 similar, with the difference that f_c^{opt} values tend to become lower than 1.0 for $R_r > 2$ (Figure
435 17f). Despite the differences between the f_c^{opt} values for both metrics, the spatial patterns remain
436 with the south-west region having lower values, the central Iowa values kept around the unity,
437 and the northeast values above 1.0 in the V bias case. Moreover, the accentuated f_c^{opt} values
438 coincide with the performance increase shown in Figure 16.

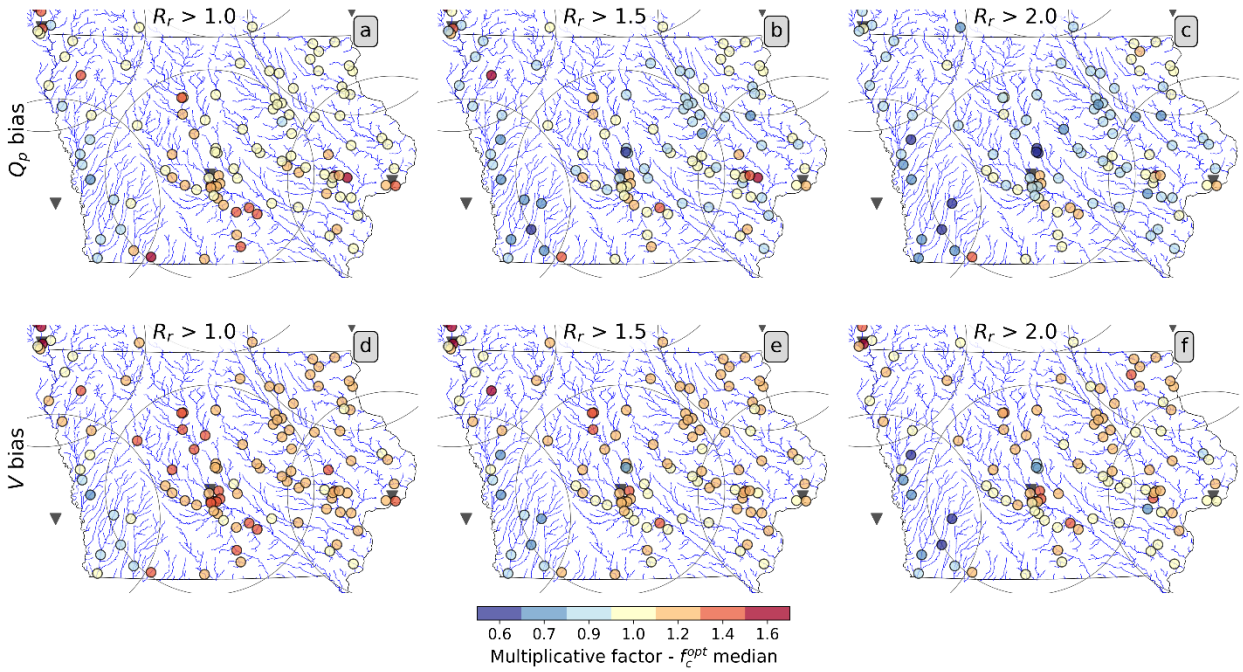


Figure 17. Median f_c^{opt} in function of the MRMS relative rainfall (R_r) of each event. The columns correspond to R_r values above 1, 1.5, and 2. The rows correspond to the f_c^{opt} values for the Q_p (a to c) and V bias (d to f) performance metrics.

4 Conclusions

Using a straightforward approach, we have comprehensively analyzed the connection between radar-based QPE uncertainties regarding basin-wide bias and streamflow forecasting errors. In our experiment, we assumed that QPE uncertainties represent a significant portion of the modeling errors. From this assumption, we described the QPE uncertainty using a multiplicative factor oscillating between 0.1 and 5. We ran the HLM distributed hydrological model in open-loop mode between 2015 and 2020 using two QPE products (MRMS and IFCA). We compared the results at the event scale using 112 USGS stream gauges in Iowa. For each event and gauge, we identified the best-performing multiplicative factor and the performance differences with respect to a factor of one. According to our results, we can draw the following conclusions:

- As Bárdossy et al., (2022) suggested, errors in precipitation are a significant contributor to streamflow forecast uncertainty. We obtained improvements over 50% in most cases and over 100% for some peak flow cases. Compared to previous calibration and data assimilation efforts, the magnitude of our improvements indicates that rainfall uncertainties induce a significant bias in hydrological models, probably also affecting our efforts to parametrize them and discriminate the proper rainfall-runoff mechanisms.
- As QPE quality has a seasonal component (Cocks et al., 2017), so does the QPE-modeling uncertainty structure. Our study obtained differentiable performance improvements between the warm and cold seasons (see Figure 9). We largely improved the V bias at the end of the cold season with factors around 0.4 and the Q_p bias near the end of the warm season with increments around 100%. The multiplicative factor f_c^{opt}

also reflects the seasonal trends with relatively large values during the cold season and lower values during the warm season. We did not explore the seasonal component further in this work. However, our results indicate that it is essential to understand this connection to improve our streamflow forecasts throughout the year.

- In our analysis, the spatial distribution of radar locations seems to determine streamflow forecasting performance and the QPE bias factors. Despite working with a mosaic product of radar data, we identified significant spatial differences that coincide with the distance between radars and the watersheds defined by the USGS stream gauges. We will explore this connectivity further by analyzing rainfall events relative to the watersheds and their relative distance to individual radars. This requires a specialized reprocessing of the radar data. More progress and results in this direction would help us create corrective algorithms for the QPE bias.
- We found that QPE-modeling uncertainty is not independent of the magnitude of the storm event. In our study, we found larger improvements during larger rainfall events. On the other hand, f_c^{opt} exhibited diverging results for the evaluated performance metrics (V and Q_p bias). The described results suggest that total rainfall determines the QPE uncertainty and the selection of the best-performing parameters. Similar results have been reported for the calibration case during high and low flows (Asadzadeh et al., 2014). This is an intricate issue as it shows significant variability in the QPE uncertainty between rainfall events.

We see the current study as a step towards separating the QPE and rainfall-runoff model uncertainty. At this point, we are still far from achieving the goal. Nevertheless, we consider our results relevant to the community as they indicate the existence of spatial, seasonal, and rainfall-magnitude conditional patterns. Our results set up the context for developing an effective algorithm for basin-wide correction of radar-rainfall. Moreover, we consider that there are multiple future work avenues in this topic. Further analysis may consider the connections between the observed spatial bias and the seasons. Also, we would like to analyze the relationships between nested watersheds, expand the analysis area, and evaluate our results using different models.

Acknowledgments

This work was completed with partial support from the Iowa Flood Center, Mid-America Transportation Center via a grant from the U.S. Department of Transportation's University Transportation Centers Program (USDOT UTC grant number for MATC: 69A3551747107), the Iowa Highway Research Board and Iowa Department of Transportation (Contract number:TR-699). Partial funding for this project was provided by the National Oceanic and Atmospheric Administration (NOAA), awarded to the Cooperative Institute for Research on Hydrology (CIROH) through the NOAA Cooperative Agreement with The University of Alabama, NA22NWS4320003.

References

Addor, N., & Melsen, L. A. (2019). Legacy, Rather Than Adequacy, Drives the Selection of

Hydrological Models. *Water Resources Research*, *55*(1), 378–390. https://doi.org/10.1029/2018WR022958

Arsenault, R., Brissette, F., & Martel, J. L. (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, *566*(September), 346–362. https://doi.org/10.1016/j.jhydrol.2018.09.027

Asadzadeh, M., Tolson, B. A., & Burn, D. H. (2014). A new selection metric for multiobjective hydrologic model calibration. *Water Resources Research*, *50*, 7082–7099. https://doi.org/10.1002/2013WR014979.Reply

Ayalew, T. B., Krajewski, W. F., & Mantilla, R. (2014). Connecting the power-law scaling structure of peak-discharges to spatially variable rainfall and catchment physical properties. *Advances in Water Resources*, *71*, 32–43. https://doi.org/10.1016/j.advwatres.2014.05.009

Ayalew, T. B., Krajewski, W. F., Mantilla, R., & Small, S. J. (2014). Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge. *Advances in Water Resources*, *64*, 9–20. https://doi.org/10.1016/j.advwatres.2013.11.010

Berne, A., & Krajewski, W. F. (2013). Radar for hydrology: Unfulfilled promise or unrecognized potential? *Advances in Water Resources*, *51*, 357–366. https://doi.org/10.1016/j.advwatres.2012.05.005

Beven, K., & Binley, A. (2014). GLUE: 20 years on. *Hydrological Processes*, *28*(24), 5897–5918. https://doi.org/10.1002/hyp.10082

Beven, K. J. (2012). Rainfall-Runoff Modelling. In *Rainfall-Runoff Modelling: The Primer: Second Edition*. https://doi.org/10.1002/9781119951001

Bytheway, J. L., Hughes, M., Mahoney, K., & Cifelli, A. R. (2020). On the uncertainty of high-resolution hourly quantitative precipitation estimates in California. *Journal of Hydrometeorology*, *21*(5), 865–879. https://doi.org/10.1175/JHM-D-19-0160.1

Ciach, G. J., Krajewski, W. F., & Villarini, G. (2007). Product-Error-Driven Uncertainty Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD Data. *Journal of Hydrometeorology*, *8*(6), 1325–1347. https://doi.org/10.1175/2007JHM814.1

Cocks, S. B., Zhang, J., Martinaitis, S. M., Qi, Y., Kaney, B., & Howard, K. (2017). MRMS QPE performance east of the Rockies during the 2014 warm season. *Journal of Hydrometeorology*, *18*(3), 761–775. https://doi.org/10.1175/JHM-D-16-0179.1

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C. H., Delrieu, G., Hong, Y., Hsu, Y. C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E. I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J. P., Vergara, H. J., & Yilmaz, K. K. (2016). Multiregional satellite precipitation products evaluation over complex terrain. *Journal of Hydrometeorology*, *17*(6), 1817–1836. https://doi.org/10.1175/JHM-D-15-0197.1

Duan, Q., Sorooshian, S., & Gupta, V. K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, *158*, 265–284. https://doi.org/10.1201/9781351076586

Ehlers, L. B., Sonnenborg, T. O., & Refsgaard, J. C. (2019). Observational and predictive uncertainties for multiple variables in a spatially distributed hydrological model. *Hydrological Processes*, *33*(5), 833–848. https://doi.org/10.1002/hyp.13367

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2007). A comparison of alternative multiobjective calibration strategies for hydrological modeling. *Water Resources Research*, *43*(3), 1–16. https://doi.org/10.1029/2006WR005098

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2008). Understanding catchment behavior through stepwise model concept improvement. *Water Resources Research*, *44*(1), 1–13. https://doi.org/10.1029/2006WR005563

Fonley, M. R., Qiu, K., Velásquez, N., Haut, N. K., & Mantilla, R. (2021). Development and Evaluation of an ODE Representation of 3D Subsurface Tile Drainage Flow Using the HLM Flood Forecasting System. *Water Resources Research*, *57*(3). https://doi.org/10.1029/2020WR028177

Francés, F., Vélez, J. I., & Vélez, J. J. (2007). Split-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology*, *332*(1–2), 226–240. https://doi.org/10.1016/j.jhydrol.2006.06.032

Gharari, S., Gupta, H. V., Clark, M. P., Hrachowitz, M., Fenicia, F., Matgen, P., & Savenije, H. H. G. (2021). Understanding the Information Content in the Hierarchy of Model Development Decisions: Learning From Data. *Water Resources Research*, *57*(6). https://doi.org/10.1029/2020WR027948

Ghimire, G. R., Krajewski, W. F., Ayalew, T. B., & Goska, R. (2022). Hydrologic investigations of radar-rainfall error propagation to rainfall-runoff model hydrographs. *Advances in Water Resources*, *161*(September 2021), 104145. https://doi.org/10.1016/j.advwatres.2022.104145

Gupta, V. K., Mantilla, R., Troutman, B. M., Dawdy, D., & Krajewski, W. F. (2010). Generalizing a nonlinear geophysical flood theory to medium size river basins. *Geophysical Research Letters*, *37*(11), 1–6. https://doi.org/10.1029/2009GL041540

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

Kavetski, D., Kuczera, G., & Franks, S. W. (2006a). Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, *42*(3), n/a-n/a. https://doi.org/10.1029/2005WR004368

Kavetski, D., Kuczera, G., & Franks, S. W. (2006b). Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *Water Resources Research*, *42*(3), n/a-n/a.

https://doi.org/10.1029/2005WR004376

KlemeŠ, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, *31*(1), 13–24. https://doi.org/10.1080/02626668609491024

Koya, S. R., Velasquez, N., Mantilla, R. I., Rojas, M., Harvey, K., Ceynar, D., Krajewski, W. F., & Roy, T. (2023). A Prototype Flood Forecasting System for Nebraska Watersheds. *Environmental Modelling and Software*, *164*(April), 105693. https://doi.org/10.1016/j.envsoft.2023.105693

Krajewski, W. F., & Smith, J. A. (2002). Radar hydrology: Rainfall estimation. *Advances in Water Resources*, *25*(8–12), 1387–1394. https://doi.org/10.1016/S0309-1708(02)00062-3

Krajewski, W.F., & Smith, J. A. (2023). Radar Hydrology. In V. N. Mishra, K. V Mishra, & M. Thurai (Eds.), *Advances in Weather Radar*. IET Press.

Krajewski, Witoldld F., Ceynar, D., Demir, I., Goska, R., Kruger, A., Langel, C., Mantilllla, R., Niemeier, J., Quintero, F., Seo, B. C., Smallll, S. J., Weber, L. J., & Young, N. C. (2017). Real-time flood forecasting and information system for the state of Iowa. *Bulletin of the American Meteorological Society*, *98*(3), 539–554. https://doi.org/10.1175/BAMS-D-15-00243.1

Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233–241. https://doi.org/10.1029/1998WR900018

Lin, Y., & Mitchell, K. E. (2005). The NCEP Stage II/IV hourly precipitation analyses: development and applications. *Preprints, 19th Conf. on Hydrology, American Meteorological Society, San Diego, CA, 9-13 January 2005, Paper 1.2*, 2–5.

Lu, C., Yuan, H., Tollerud, E. I., & Wang, N. (2010). Scale-dependent uncertainties in global QPFs and QPEs from NWP model and satellite fields. *Journal of Hydrometeorology*, *11*(1), 139–155. https://doi.org/10.1175/2009JHM1164.1

Mantilla, R., Fonley, M., & Velasquez, N. (2023). Technical Note : Testing the Connection Between Hillslope Scale Runoff Fluctuations and Streamflow Hydrographs at the Outlet of Large River Basins. *Hydrology and Earth System Sciences*, *August*, 1–13.

Mantilla, R. I., Krajewski, W. F., Velasquez, N., Small, S. J., Ayalew, T. B., Quintero, F., Jadidoleslam, N., & Fonley, M. (2022). The Hydrological Hillslope-Link Model for Space-Time Prediction of Streamflow: Insights and Applications at the Iowa Flood Center. In *Extreme Weather Forecasting*. Elsevier.

Mantilla, R., Krajewski, W. F., Velasquez, N., Small, S. J., Ayalew, T. B., Quintero, F., Jadidoleslam, N., & Fonley, M. (2022). The hydrological hillslope-link model for space-time prediction of streamflow: insights and applications at the Iowa Flood Center. *Extreme Weather Forecasting*, 200.

615 Nathan, R.J., McMahon, T. A. (1990). Evaluation of Automated Techniques for Base Flow and
616 Recession Analyses. *Water Resources Research*, *26*(7), 1465–1473.
617 https://doi.org/10.1029/2006WR005467

618 Post, R., & Krajewski, W. F. (2023). Correction to: Examining the stage-IV radar-rainfall
619 product for Probabilistic rainfall estimation: case study over Iowa. *Stochastic*
620 *Environmental Research and Risk Assessment*, *37*(9), 3677. https://doi.org/10.1007/s00477-
621 023-02522-0

622 Quintero, F., Krajewski, W. F., Seo, B. C., & Mantilla, R. (2020). Improvement and evaluation
623 of the Iowa Flood Center Hillslope Link Model (HLM) by calibration-free approach.
624 *Journal of Hydrology*, *584*. https://doi.org/10.1016/j.jhydrol.2020.124686

625 Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological
626 models. *Journal of Hydrology*, *198*(1–4), 69–97. https://doi.org/10.1016/S0022-
627 1694(96)03329-X

628 Ryzhkov, A., Diederich, M., Zhang, P., & Simmer, C. (2014). Potential utilization of specific
629 attenuation for rainfall estimation, mitigation of partial beam blockage, and radar
630 networking. *Journal of Atmospheric and Oceanic Technology*, *31*(3), 599–619.
631 https://doi.org/10.1175/JTECH-D-13-00038.1

632 Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive
633 inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors.
634 *Water Resources Research*, *46*(10), 1–17. https://doi.org/10.1029/2009WR008933

635 Seo, B. C., Keem, M., Hammond, R., Demir, I., & Krajewski, W. F. (2019). A pilot
636 infrastructure for searching rainfall metadata and generating rainfall product using the big
637 data of NEXRAD. *Environmental Modelling and Software*, *117*(March), 69–75.
638 https://doi.org/10.1016/j.envsoft.2019.03.008

639 Seo, B. C., & Krajewski, W. F. (2020). Statewide real-time quantitative precipitation estimation
640 using weather radar and NWP model analysis: Algorithm description and product
641 evaluation. *Environmental Modelling and Software*, *132*, 104791.
642 https://doi.org/10.1016/j.envsoft.2020.104791

643 Seo, B. C., Krajewski, W. F., & Ryzhkov, A. (2020). Evaluation of the specific attenuation
644 method for radar-based quantitative precipitation estimation: Improvements and practical
645 challenges. *Journal of Hydrometeorology*, *21*(6), 1333–1347. https://doi.org/10.1175/JHM-
646 D-20-0030.1

647 Seo, B. C., Quintero, F., & Krajewski, W. F. (2018). High-resolution QPF uncertainty and its
648 implications for flood prediction: A case study for the eastern Iowa flood of 2016. *Journal*
649 *of Hydrometeorology*, *19*(8), 1289–1304. https://doi.org/10.1175/JHM-D-18-0046.1

650 Shen, H., A. Tolson, B., & Mai, J. (2022). *Shen-Tolson-Mai_2020_Time to Update the Split-*
651 *Sample Approach in Hydrological Model Calibration.pdf*.

Stedinger, J. R., Vogel, R. M., Lee, S. U., & Batchelder, R. (2008). Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. *Water Resources Research*, *44*(12), 1–17. https://doi.org/10.1029/2008wr006822

Survey, U. S. G. (2022). *USGS 1/3 Arc Second*.

USGS. (2017). *National Hydrography Dataset Plus High Resolution (NHDPlus HR) - USGS National Map Downloadable Data Collection*.

Velásquez, N., Mantilla, R., Krajewski, W., Fonley, M., & Quintero, F. (2021). Improving Hillslope Link Model Performance from Non-Linear Representation of Natural and Artificially Drained Subsurface Flows. *Hydrology*, *8*(4), 187.

Velásquez, N., Quintero, F., Koya, S. R., Roy, T., & Mantilla, R. (2023). Snow-detonated floods: Assessment of the U.S. midwest march 2019 event. *Journal of Hydrology: Regional Studies*, *47*. https://doi.org/10.1016/j.ejrh.2023.101387

Velasquez, N, Mantilla, R., Krajerwski, W. F., Quintero, F., & Zanchetta, A. (2022). Identification and Regionalization of Streamflow Routing Parameters Using Machine Learning for the HLM Hydrological Model in Iowa. *Journal of Advances in Modeling Earth Systems*, *14*(7).

Velasquez, Nicolas, Mantilla, R., Krajewski, W. F., Fonley, M., & Quintero, F. (2021). Improving Hillslope Link Model Performance fromNon-Linear Representation of Natural and Artificially Drained Subsurface Flows. *Hydrology*, *8*(187).

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*(1), 107–129. https://doi.org/10.1007/s10712-009-9079-x

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

Zhang, J., Tang, L., Cocks, S., Zhang, P., Ryzhkov, A., Howard, K., Langston, C., & Kaney, B. (2020). A dual-polarization radar synthetic QPE for operations. *Journal of Hydrometeorology*, *21*(11), 2507–2521. https://doi.org/10.1175/JHM-D-19-0194.1