Precipitation Estimation Methods in Continuous, Distributed Urban Hydrologic Modeling

David Woodson

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Civil Engineering

Randel L. Dymond, Chair
Clayton C. Hodges
Kevin D. Young

May 14, 2019
Blacksburg, VA

Keywords: Urban hydrology, Quantitative Precipitation Estimation, Modeling

Precipitation Estimation Methods in Continuous, Distributed Urban Hydrologic Modeling

David Woodson

ACADEMIC ABSTRACT

Quantitative precipitation estimation (QPE) remains a key area of uncertainty in hydrological modeling, particularly in small, urban watersheds which respond rapidly to precipitation and can experience significant spatial variability in rainfall fields. Few studies have compared QPE methods in small, urban watersheds, and studies which have examined this topic only compared model results on an event basis using a small number of storms. This study sought to compare the efficacy of multiple QPE methods when simulating discharge in a small, urban watershed on a continuous basis using an operational hydrologic model and QPE forcings. The Research Distributed Hydrologic Model (RDHM) was used to model a basin in Roanoke, Virginia, USA forced with QPEs from four methods: mean field bias (MFB) correction of radar data, kriging of rain gauge data, uncorrected radar data, and a basin-uniform estimate from a single gauge inside the watershed. Based on comparisons between simulated and observed discharge at the basin outlet for a 6-month period in 2018, simulations forced with the uncorrected radar QPE had the highest accuracy, as measured by root mean square error (RMSE) and peak flow relative error, despite systematic underprediction of mean areal precipitation (MAP). Simulations forced with MFB corrected radar data consistently and significantly overpredicted discharge but had the highest accuracy in predicting the timing of peak flows.

Precipitation Estimation Methods in Continuous, Distributed Urban Hydrologic Modeling

David Woodson

GENERAL AUDIENCE ABSTRACT

Estimating the amount of rain that fell during a precipitation event remains a key source of error when predicting how much stormwater runoff will be produced, particularly in small, urban watersheds which respond rapidly to precipitation and can experience significant spatial variability in rainfall distribution. Rainfall estimation in small, urban watersheds has received relatively little attention, and studies which have examined this topic have generally only examined a small number of discrete storm events. This study sought to compare the efficacy of multiple precipitation estimation methods when simulating discharge in a small, urban watershed on a continuous basis using an operational hydrologic model and precipitation inputs. The Research Distributed Hydrologic Model (RDHM), commonly used by the National Weather Service, was used to model a basin in Roanoke, Virginia, USA forced with rainfall estimates from four methods: mean field bias (MFB) correction of radar data, kriging of rain gauge data, uncorrected radar data, and a basin-uniform estimate from a single gauge inside the watershed. Based on comparisons between simulated and observed discharge at the basin outlet for a 6-month period in 2018, simulations forced with the uncorrected radar QPE had the highest accuracy, as measured by several performance statistics, despite systematic underprediction of actual precipitation. Simulations forced with MFB corrected radar data consistently and significantly overpredicted discharge but had the highest accuracy in predicting the timing of peak flows.

Acknowledgements

I would like to extend a big thank you to Dr. Randy Dymond for giving me a number of great opportunities and learning experiences during my time as an undergraduate and graduate student at Virginia Tech, and for serving as an excellent mentor, role model, advisor, and boss. I would also like to sincerely thank Dr. Tom Adams, who tirelessly helped me with this thesis, and without whom this would not have been possible. Tom has also been a great mentor and role model over the last year. Thanks to Dr. Clay Hodges and Mr. Kevin Young, for serving on my committee and for feedback on drafts. Thanks to Dr. Marcus Aguilar for his mentorship and advice over the years, I have learned much from him.

Thanks also to my colleagues in Patton 310 for good times inside and outside of the office, the future seems bright with folks like these in it. And finally, thank you to my many teachers over the years, friends, and family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 – INTRODUCTION

1.1 Background

Precipitation is a key driver in the hydrologic cycle and associated modeling efforts. There have been significant advances in simulating runoff flow rate and volume due to higher resolution digital elevation models and land cover rasters (e.g. Fonstad et al., 2013; Mayer, 1999; Tokarczyk et al., 2015), better mapping and incorporation of storm sewer networks (e,g, Gironás et al., 2010; Smith et al., 2013), improved computational efficiency, and a wider variety of hydrology and hydraulic (H&H) models from which to choose. However, quantitative precipitation estimation (QPE) remains a key component of model uncertainty, regardless of the resolution of the remaining model components. QPE uncertainty is exacerbated in areas with orographic or convective precipitation due to heterogeneity in rainfall spatiotemporal distribution (Cristiano et al., 2017; Li Pen Wang et al., 2015). Similarly, some studies have shown that even small, fragmented urbanized areas can cause significant increases and/or decreases in precipitation due to impacts on temperature and wind (Daniels et al., 2016; Freitag et al., 2018). With growing urbanization and climatic changes increasing the frequency and magnitude of hydrologic extremes (National Academies of Sciences, Engineering, and Medicine, 2019), the ability of QPEs to accurately simulate and predict hydrologic response at high spatiotemporal resolutions is becoming increasingly important.

Accurate, high resolution rainfall data is required to reduce the error in hydrologic models, particularly those in small to mid-sized, urban watersheds. Several factors drive the need for high resolution data: first is the small size, high variability in land cover, and rapid response of urban catchments, and second is the potential for rainfall to vary significantly in space and time at small scales (Cristiano et al., 2017; Yoon and Lee, 2017). Relatively small variability in spatiotemporal rainfall distributions have resulted in large errors in predicted streamflow such that estimated rainfall was identified as the biggest contributor to error in modeling, rather than land-based parameters (Ogden et al., 2000).

In an effort to improve the accuracy of QPEs and associated hydrologic simulations, many studies have compared multiple methods of precipitation estimation (e.g. gauge only, radar only, gauge-radar hybrid) for streamflow prediction and all have found that a hybrid approach yields the most accurate streamflow predictions (James et al., 1993; Kim et al., 2008; Looper and Vieux, 2012; Pessoa et al., 1993; Seo et al., 2018; Sun et al., 2000). However, these studies have primarily occurred under one or more of the following conditions: large (>1000 km^2) or rural watersheds, varying degrees of gauge coverage (sometimes sparse), or using spatiotemporal resolutions too coarse for urban hydrology (e.g. 1 hour, 4 km). Hence, the results may not be applicable to small, urban watersheds.

Few studies have compared the differences in hydrologic simulations of small, urban watersheds forced with various QPE products (e.g. gauge network, uncorrected radar, gauge-radar hybrid). Such studies (Ochoa-Rodriguez et al., 2015; Wang, Ochoa-Rodríguez, et al., 2015a; Wang, Ochoa-Rodriguez, et al., 2015b; Wang et al., 2013; Yoon and Lee, 2017) have generally performed advanced geostatistical merging techniques on radar and gauge data (e.g. conditional merging, Bayesian merging, error variance minimization). These studies have found marginal to moderate improvement over less complex correction techniques such as kriging and/or mean field bias (MFB) correction, but significant improvement over uncorrected radar.

Previous studies have relied on hydrologic and hydraulic (H&H) models calibrated with rain gauge data, run on an event basis. Further, only a small number of storms were examined in each study, typically between 4 and 10 events. In comparing various QPE forcings, use of an uncalibrated model, run on a continuous basis may be preferable since this may prevent calibration bias (e.g. towards a gauge-centric QPE if gauges are used in calibration) and would evaluate QPEs under a greater variety of hydrometeorological conditions. Further, use of an operational model and forcings has the potential to be later adapted for flash flood forecasting, warnings, etc. in urban areas at a high spatiotemporal resolution. Thus far, no study has evaluated hydrologic models of small, urban watersheds forced with Next Generation Radar (NEXRAD) dual-pol Level III data.

1.2 **Problem statement**

QPE accuracy and spatiotemporal resolution are perhaps the most important factors in determining H&H model accuracy, particularly in urban and small to medium sized watersheds. Precipitation inputs derived from merged gauge-radar methods have been shown to yield the most accurate streamflow simulations for small urban watersheds under certain conditions (Ochoa-Rodriguez et al., 2015; Wang, Ochoa-Rodríguez, et al., 2015a; Wang, Ochoa-Rodriguez, et al., 2015b; Wang et al., 2013; Yoon and Lee, 2017). However, such studies have been conducted on an event basis, typically only examining a small number of storms, and have calibrated to rain gauge forcings. Testing the various QPE products in a continuous, distributed model with 6 months of forcing data, rather than an event-based, lumped model run with a small number of events is a novel approach. Using a separate precipitation dataset for model initialization and initializing the model over a long period of time (e.g. one year), should reduce model bias that may occur when initializing or calibrating with one of the QPE datasets used in analysis or with a shorter calibration time period. Assuming significant improvement is gained through bias correction of the radar data, NEXRAD's ~20-year archive of radar data could become a valuable resource for urban H&H modeling at a high resolution.

Previously, the City of Roanoke had been using a single gauge for hydrologic analyses in the city's various watersheds. The improvement in accuracy gained by usage of the previously described, spatially heterogenous QPE methods instead of a single gauge with a uniform rainfall assumption should be quantified.

1.3 Purpose and Objectives

Purpose

The goal of this master's research is to investigate discharge simulation skill of various precipitation forcings with a distributed, operational H&H model run on a continuous basis. Three fundamental questions are investigated by this thesis as outlined below:

- Which QPE method (single gauge, kriging, uncorrected radar, or mean field bias corrected radar) yields the most accurate streamflow predictions for a small, urban watershed in a mountainous area?
- How well does the Research Distributed Hydrologic Model (RDHM) perform in a small, urban watershed with significant storm sewer infrastructure, and which precipitation forcing achieves the highest discharge skill when paired with RDHM?
- How do spatially varied QPEs compare to a single gauge assumption in simulating discharge?

Objectives

The objectives of this thesis are to:

1. Determine which QPE method (single gauge, kriging, uncorrected radar, or mean field bias corrected radar) yields the most accurate streamflow predictions from a fully distributed H&H model of an urban watershed subject to heterogenous rainfall fields (e.g. convective,

orographic rainfall) and potential beam blockage. In order to differentiate from past studies, continuous, monthly simulations will be run for a six-month period.

2. Evaluate the efficacy of RDHM for H&H modeling of a small urban watershed.
3. Determine the difference in streamflow simulation accuracy between a uniform watershed estimate based on a single gauge and estimates from the proposed spatially heterogenous QPE methods.

CHAPTER 2 – LITERATURE REVIEW

2.1 QPEs in Urban Hydrology

Few studies have compared the differences in hydrologic simulations of small, urban watersheds forced with various QPE products (e.g. gauge network, radar, gauge-radar hybrid). Yoon and Lee (2017) used three precipitation datasets (kriged gauge data, radar, and conditional merging of gauge and radar data) to simulate flow from several flood events in a 7.4 km^2, highly urbanized catchment in Seoul, South Korea. The gauge network in the greater Seoul area is the densest in the world, with a density of 1 gauge per 3 km^2, or about 1 km between gauges. Kriged gauge data was assumed, and later proved via cross validation, to be the most accurate QPE method, and was accordingly used to calibrate their SWMM model. The authors found that while the kriged gauge data had marginally better QPE estimates compared to conditional merging, conditional merging resulted in marginally more accurate streamflow predictions compared to estimates from kriged data; QPE and flow depth estimates from uncorrected radar were significantly worse than estimates from the two other QPE methods. The marginal differences in hydrologic output between the kriged- and merged-data is likely due to the high and relatively uncommon gauge density.

Yoon and Lee (2017) found that storm total QPEs from radar alone suffered from significant underprediction and had mean average errors (MAE) and root mean square errors (RMSE) 2-3 times higher than the MAE and RMSE of either kriged gauge or combined gauge-radar QPEs. Flow simulations forced with radar data had an average relative error in peak depth (REPD) and RMSE of 57% and 0.33 mm, respectively, compared to measured data, whereas the average REPD and RMSE for flow simulations from kriged gauge and combined gauge-radar QPEs were 9.2%, 0.13 mm and 7.7%, 0.13 mm respectively. The marginal differences in hydrologic output between the kriged gauge and combined gauge-radar data may be due to the extraordinarily high gauge density of the study area as well as the bias correction method used. Conditional merging involves overlaying deviations in the radar rainfall field onto the kriged gauge rainfall field: since the gauge density is so high, the bias correction process appears to minimally influence the gauge-based rainfall field.

Wang et al. (2013) examined multiple QPE methods for modeling an 8.65 km^2 urban basin in London, UK. The authors forced an Infoworks CS model of the watershed with four different QPEs: gauge data, uncorrected radar data, MFB corrected radar data, and radar corrected by an error variance minimization technique known as Bayesian combination (BC). Based on predicted discharge depth from four events, the authors found that both the MFB and BC techniques performed well, but that BC outperformed MFB in terms of reducing bias and also better captured the spatiotemporal variability of rainfall fields (Li Pen Wang et al., 2013).

Wang et al. (2015a) forced a gauge-calibrated InfoWorks CS model of a 53 km^2 urban catchment in Edinburgh, UK with multiple QPEs, including: rain gauges, uncorrected radar, block kriging, MFB corrected radar, a Bayesian combination technique, and a Bayesian combination technique which preserves high intensity singularities detected by radar in the rainfall field. Based on four events, the authors found that all techniques worked relatively well, but that the rain gauge forcing performed the best, MFB correction led to overprediction of discharge, and the Bayesian combination techniques improved predictions compared to uncorrected radar (L. P. Wang et al., 2015a).

Löwe et al., (2014) used combined gauge-radar QPEs in urban watershed models of two separate watersheds (13 km^2 and 30 km^2) in Copenhagen, Denmark (Löwe et al., 2014). The authors found that gauge corrected radar QPEs improved probabilistic flood forecasts (100-minute horizon) from a lumped model compared to forecasts fed with gauge data alone.

Several studies have shown that, for H&H modeling, the required spatiotemporal resolution of precipitation data increases with decreasing catchment size, and that there is an optimal relationship between spatial and temporal resolution depending on catchment size (Ochoa-Rodriguez et al., 2015). Spatial resolution is defined in a variety of ways, including: the grid size of radar data (or gauge-corrected radar data), the areal coverage assumed for a single gauge, or the distance between adjacent gauges. Temporal resolution is defined as the time between consecutive measurements. For urban hydrology applications, Schilling (1991) recommended a minimum spatial and temporal resolution of 1 gauge per 1 km^2 (or 1 km^2 radar grids) and 5 minutes (ideally 1 minute), respectively (cited in Yoon and Lee 2017). More recent research has determined spatiotemporal resolution requirements as a function of watershed size. Using both gauge and radar data, Berne et al. (2004) suggested that urban watersheds around 1000 ha in size require a spatiotemporal resolution of 3 km and 5 minutes, while ~100 ha watersheds require a resolution of 2 km and 3 minutes. Other research based on radar data only suggests a minimum spatial and temporal resolution of 100 m and 1 minute, respectively, for urban catchments smaller than 1 ha.

However, for catchments between 1 and 100 ha, while the recommended resolution is 500 m/1 min, 1 km/1min resolutions lead to similar results (Ochoa-Rodriguez et al., 2015).

2.2 **QPE Datasets**

Multiple agencies operate platforms with both real-time access to- and long term archives of- rain gauge and other hydrometeorlogical data. Two of the most commonly used networks in the US are the National Water Information Service (NWIS) and the Automated Surface Observation System (ASOS). NWIS is operated by the United States Geological Survey (USGS) while ASOS is a joint project of the National Weather Service (NWS), Federal Aviation Administration (FAA), and Department of Defense (DOD). In Roanoke, Virginia the USGS recently installed a network of nine rain gauges that provide fairly good coverage of the metropolitan area. These gauges became operational at the beginning of March 2018. Similarly, an ASOS gauge has been operating at the Roanoke airport since 1948.

Radar derived QPEs are a useful tool for urban hydrology because of their capacity to capture the spatial variability of rainfall fields, but without gauge correction are subject to systematic underestimation as well as random error (Thorndahl et al., 2017). Unlike the point measurements of gauges, radar provides spatially distributed QPEs at varying resolutions. In the US, the National Weather Service (NWS) Next Generation Radar (NEXRAD) network offers a variety of S-band, dual-pol radar products, from base reflectivity to gauge corrected national mosaics. However, as the degree of processing and bias correction increases, the spatiotemporal resolution of the radar QPE decreases. The 1-hour and 4-km resolution of products such as NEXRAD Stage III/IV or Multisensor Precipitation Estimator (MPE) is too coarse for urban H&H modeling, which requires a resolution of at least 5 minutes and 1 km (Berne et al., 2004; Ochoa-Rodriguez et al., 2015; Schilling, 1991). Consequentially, a higher resolution but non-gauge corrected radar QPE is one of the few choices for modeling small, urban watersheds. Within the NEXRAD framework, the Level 3 dataset fills this role. The Level 3 dataset contains a variety of products at varying resolutions (e.g. instantaneous rainfall intensity, one hour accumulations, storm velocity, direction, hydrometeor type, etc.). A Level 3 dataset exists for each NWS radar site and a publicly available archive goes back to the late 1990s. Late in the first decade of the twenty-first century, dual-polarization upgrades to all NWS radars was completed, which significantly improved radar accuracy over single-polarization. For this study, the instantaneous precipitation rates, also known as digital precipitation rates (DPR) were used from the KFCX radar site located approximately 40 km southwest of Roanoke. DPR is generated through an empirical formula (i.e. ZR relationships) that converts reflectivity to a precipitation rate and has not been corrected with gauge data. However, DPR is available at a very high resolution: 600 m and ~3-5 minutes when in precipitation mode (~20 minutes when in clear air mode). Given the irregular time interval of the DPR dataset, temporal standardization (e.g. uniform 5-minute time steps) is required before bias correction or model input. Standardization can be achieved through temporal interpolation followed by aggregation.

Generally, improvements in radar measurements and processing over the last few decades have led to significant improvements in radar QPE accuracy but only marginal improvements in radar QPF accuracy (Adams and Dymond, 2018 – in review).

2.3 **QPE Accuracy**

Each precipitation measurement method has its own sources of error. Rain gauge error caused by wind effects ranges between 2-10% for rain and potentially 50% for snowfall. However, including other factors - particularly issues present in urban areas such as splashing and wind field changes due to structures - can further increase rainfall error to up to 30% (Cristiano et al., 2017). Tipping buckets, a common gauge type, are susceptible to significant underestimation, typically 10-15%, during high intensity events (over 200 mm/hour) due to rainfall losses during tipping, however gauge error can be reduced to 1% with appropriate calibration and modification (Lanza and Stagi, 2008; Molini et al., 2005). Error may also be greatest at high temporal resolutions but has been shown to become negligible with time intervals greater than 15 minutes (Habib et al., 2001). However, a 15-minute rainfall time step can lead to significant underestimation of simulated flows compared to a 5-minute or smaller interval (Guo, 2007). A gauge measurement represents precipitation at a single point, and while various methods are used to interpolate between gauge measurements, the error of these interpolation methods depends on the degree of rainfall heterogeneity (Cristiano et al., 2017). Accordingly, gauge density and rainfall type are important factors in gauge network accuracy.

Underprediction and other errors in radar QPEs stem from beam blockage, rainfall fields which are smaller than the sampling resolution, errors in ZR empirical relationships, beam widening and overshoot, earth curvature, changes in drop size distribution (DSD), calibration bias, etc. (Cunha et al., 2015; Niemi et al., 2017; J. Zhang et al., 2012).

2.4 **Research Distributed Hydrologic Model**

The National Weather Service (NWS) Hydrology Laboratory previously developed a distributed H&H model called the Research Distributed Hydrologic Model (RDHM), which is commonly used for flood forecasting in various NWS River Forecasting Centers (RFCs). Key model mechanisms include Sacramento-Soil Moisture Accounting (SAC-SMA), RUTPIX overland and channel routing (kinematic wave), and SNOW-17 for snow operations. RDHM comes with a set of a-priori SAC-SMA parameters that have been developed for the entire conterminous United States (CONUS) and were derived using soil and land cover data from the Soil Survey Geographic Database (SSURGO) and National Land Cover Dataset (NLCD), respectively ("Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) User Manual V. 3.5.0," 2012). Similarly, a CONUS-wide set of RUTPIX parameters is provided based on elevation data. However, the parameters, forcings, and resolution are typically derived and operated at a coarse scale, most commonly 1 HRAP grid (nominally ~4 km). Model initialization is generally recommended, and typically involves a 1-year or more warm-up simulation to determine initial model states for the study period. Calibration is also recommended but is not possible in ungauged watersheds. The a-priori parameters are therefore a useful tool for modeling in ungauged watersheds and aid in the calibration process for gauged watersheds (Z. Zhang et al., 2012).

2.5 **QPE Processing and Bias Correction**

2.5.1 Rain gauge kriging

Among gauge-based estimates there are a variety of methods for spatial interpolation between gauges. Two such methods - ordinary kriging and inverse distance weighting - have generally been shown to yield the most accurate rainfall estimates (Chen et al., 2010; Shope and Maharjan, 2015). Kriging is a sophisticated method of interpolation and generally performs well in areas where orographic effects on rainfall are present. Kriging involves creation of a variogram, or relationship between distance and variance between points in a topological dataset, which is then used to estimate values between points.

2.5.2 *Mean Field Bias (MFB) Correction*

Mean field bias (MFB) correction is perhaps the most common and easiest method for bias correction of radar rainfall data. MFB correction involves calculation of a ratio of average rainfall depth measured by selected gauges to the average depth determined by radar at the gauge locations (Equation 1).

$$B = \frac{\frac{1}{n}\sum_{i=1}^{n} P_{g-i}}{\frac{1}{n}\sum_{i=1}^{n} P_{ur-i}} = \frac{\sum_{i=1}^{n} P_{g-i}}{\sum_{i=1}^{n} P_{ur-i}} \qquad \text{Equation 1}$$

Where:
B = mean field bias
P_{g-i} = Precipitation depth at gauge i measured by rain gauge (in, mm)
P_{ur-i} = Precipitation depth at gauge i measured by uncorrected radar (in, mm)
n = number of gauges

The MFB ratio can be calculated on any time scale (e.g. 1 day, 1 hour, 5-minutes), using incremental depth measurements of the selected time scale. Once calculated, the ratios are applied to the entire uncorrected radar rainfall field at each time step (Equation 2).

$$P_{cr-t} = P_{ur-t} * B_t \qquad \text{Equation 2}$$

Where:

P_{cr-t} = Corrected radar rainfall depth field at time t (in, mm)

P_{ur-t} = Uncorrected radar rainfall depth field at time t (in, mm)

B_t = MFB ratio at time t

CHAPTER 3. PRECIPITATION ESTIMATION METHODS IN CONTINUOUS, DISTRIBUTED URBAN HYDROLOGIC MODELING

3.1 Introduction

Precipitation is a key driver in the hydrologic cycle and associated modeling efforts. There have been significant advances in simulating runoff flow rate and volume due to higher resolution digital elevation models and land cover rasters (e.g. (Fonstad et al., 2013; Mayer, 1999; Tokarczyk et al., 2015)), better mapping and incorporation of storm sewer networks (e.g. (Gironás et al., 2010; Smith et al., 2013)), improved computational efficiency, and a wider variety of hydrology and hydraulic (H&H) models from which to choose. However, quantitative precipitation estimation (QPE) remains a key component of model uncertainty, regardless of the resolution of the remaining model components. QPE uncertainty is exacerbated in areas with orographic or convective precipitation due to heterogeneity in rainfall spatiotemporal distribution (Cristiano et al., 2017; Li Pen Wang et al., 2015). Some studies have shown that even small, fragmented urbanized areas can cause significant increases and/or decreases in precipitation due to impacts on temperature and wind (Daniels et al., 2016; Freitag et al., 2018). With growing urbanization and climatic changes increasing the frequency and magnitude of hydrologic extremes (National Academies of Sciences, Engineering, and Medicine, 2019), the ability of QPEs to accurately simulate and predict hydrologic response at high spatiotemporal resolutions is becoming increasingly important.

Accurate, high resolution rainfall data is required to reduce the error in hydrologic models, particularly those in small to mid-sized, urban watersheds. Several factors drive the need for high resolution data: first is the small size, high variability in land cover, and rapid runoff response of urban catchments, and second is the potential for rainfall to vary significantly in space and time at small scales (Cristiano et al., 2017; Krajewski et al., 2003; Yoon and Lee, 2017). Relatively small variability in spatiotemporal rainfall distributions have resulted in large errors in predicted streamflow, with estimated rainfall identified as the biggest contributor to error in modeling, rather than land-based parameters (Ogden et al., 2000).

In an effort to improve the accuracy of QPEs and associated hydrologic simulations, many studies have compared multiple methods of precipitation estimation (e.g. gauge only, radar only, gauge-radar hybrid) for streamflow prediction and all have found that a hybrid approach yields the most accurate streamflow predictions (James et al., 1993; Kim et al., 2008; Looper and Vieux, 2012; Pessoa et al., 1993; Seo et al., 2018; Sun et al., 2000). However, these studies have primarily occurred under one or more of the following conditions: large (>1000 km^2) or rural watersheds, varying degrees of gauge coverage (sometimes sparse), or using spatiotemporal resolutions too coarse for urban hydrology (e.g. 1 hour, 4 km). Hence, the results may not be applicable to small, urban watersheds.

Radar derived QPEs are a useful tool for urban hydrology because of their capacity to capture the spatial variability of rainfall fields (Thorndahl et al., 2017). Unlike the point measurements of gauges, radar provides spatially distributed QPEs at varying resolutions. In the US, the National Weather Service (NWS) Next Generation Radar (NEXRAD) network offers a variety of S-band, dual-pol radar products, from base reflectivity to gauge corrected national mosaics. However, as the degree of processing and bias correction increases, the spatiotemporal resolution of the radar QPE decreases. The 1-hour and 4-km resolution of products such as NEXRAD Stage III/IV or Multisensor Precipitation Estimator (MPE) is too coarse for urban H&H modeling, which requires a resolution of at least 5 minutes and 1 km (Berne et al., 2004; Ochoa-Rodriguez et al., 2015; Schilling, 1991). Consequentially, a higher resolution, but non-gauge corrected radar QPE is one of the few choices for modeling small, urban watersheds. Within the NEXRAD framework, the Level 3 dataset fills this role. The Level 3 dataset contains a variety of products at varying resolutions (e.g. instantaneous rainfall intensity, one hour accumulations, storm velocity, direction, hydrometeor type, etc.). A Level 3 dataset exists for each NWS radar site and a publicly available archive goes back to the late 1990s. Late in the first decade of the twenty-first century, dual-polarization upgrades to all NWS radars was completed, which significantly improved radar accuracy over single-polarization. For this study, the instantaneous precipitation rates, also known as digital precipitation rates (DPR) were used from the KFCX radar site located approximately 40 km southwest of Roanoke. DPR is generated through an empirical formula (i.e. ZR relationships) that converts reflectivity to a precipitation rate and has not been corrected with gauge data. However, DPR is available at a very high resolution: 600 m and ~3-5 minutes when in precipitation mode (~20 minutes when in clear air mode). Given the irregular time interval of the DPR dataset, temporal standardization (e.g. uniform 5-minute time steps) is required before bias correction or model input.

Few studies have compared the differences in hydrologic simulations of small, urban watersheds forced with various QPE products (e.g. gauge network, uncorrected radar, gauge-radar hybrid). Such studies (Ochoa-Rodriguez et al., 2015; L. P. Wang et al., 2015; Li Pen Wang et al., 2015, 2013; Yoon and Lee, 2017) have generally performed advanced geostatistical merging techniques on radar and gauge data (e.g. conditional merging, Bayesian merging, error variance minimization). These studies have found marginal to moderate improvement over less complex correction techniques such as kriging and/or mean field bias (MFB) correction, but significant improvement over uncorrected radar.

Previous studies have relied on hydrologic and hydraulic (H&H) models calibrated with rain gauge data, run on an event basis. Further, only a small number of storms were examined in each study, typically between 4 and 10 events. In comparing various QPE forcings, use of an uncalibrated model, run on a continuous basis may be preferable since this may prevent calibration bias (e.g. towards a gauge-centric QPE if gauges are used in calibration) and would evaluate QPEs under a greater variety of hydrometeorological conditions. Additionally, use of an operational model and forcings has the potential to be later adapted for flash flood forecasting, warnings, etc. in urban areas at a high spatiotemporal resolution. Thus far, no study has evaluated hydrologic models of small, urban watersheds forced with Next Generation Radar (NEXRAD) dual-pol Level III data.

The goal of this paper was to compare various gauge and radar QPE methods in a small urban watershed in Roanoke, Virginia on a continuous basis using an operational, uncalibrated H&H model. Specifically of interest was the evaluation of high-resolution dual-pol Level III radar data due to its wide availability and long-term archive. A continuous record of 5-minute, 300-m rainfall fields using four different QPE forcings was created for a 6-month period during 2018. The QPE methods used include kriging of rain gauge data, MFB correction of NWS radar data, uncorrected radar data, and a basin-uniform rainfall depth based on a single gauge inside the watershed. An ancillary goal of the study was to evaluate the efficacy of the Research Distributed Hydrology Model (RDHM) ("Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) User Manual V. 3.5.0," 2012), an operational model originally designed for large watersheds, for use in a small, urban watershed. The study region, QPE methods, and hydrologic model are described in Section 2, QPE and model results are presented in section 3, discussion occurs in Section 4, and conclusions are made in Section 5.

3.2 **Materials and Methods**

3.2.1 Study Area

This study examines the Lick Run watershed (~19.4 km^2), a tributary of the Roanoke River located in Roanoke, Virginia, US (Figure 1). The watershed is approximately 35% impervious and subject to a variety of land uses (Dymond et al., 2017). Located in a mountain valley, the Roanoke-metropolitan area has a history of frequent flooding, resulting in property damage and sometimes death. Much of the basin is served by the city's municipal separate storm sewer system (MS4). In the study area, the average distance between rain gauges is ~6.5 km and the gauge density in the Lick Run basin is 1 gauge per 9.7 km^2.

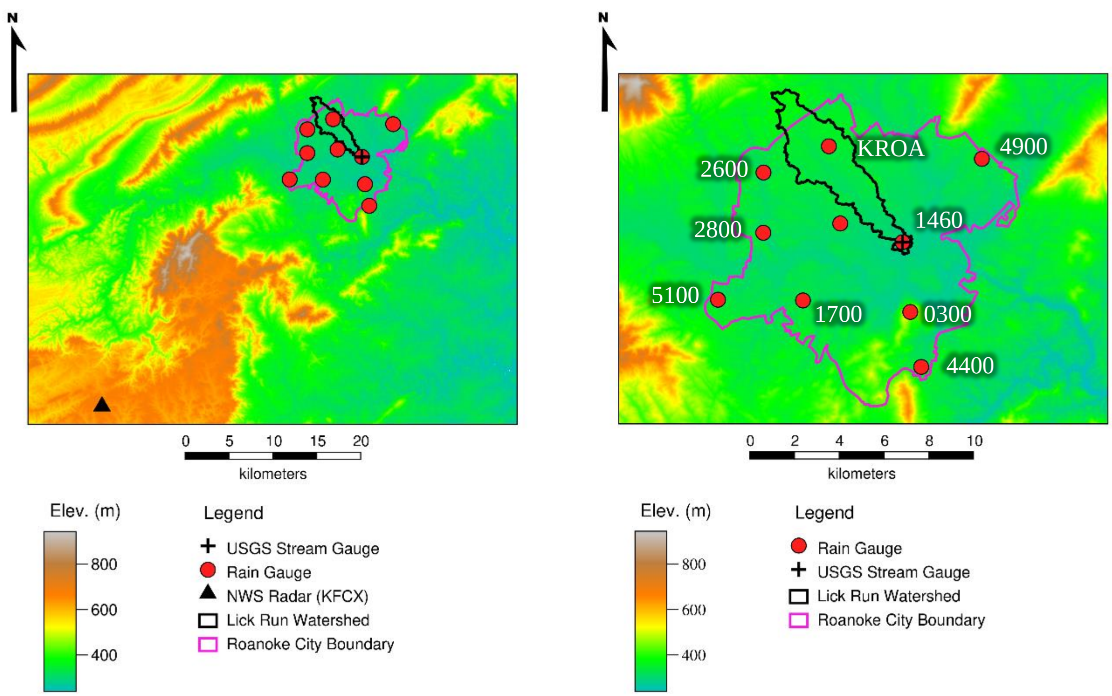


Figure 1. (a) Regional view of study area, including DEM, Lick Run watershed, KFCX radar site, rain gauges, and stream gauge (co-located with one rain gauge at basin outlet). (b) Large scale view of study area. Last four digits of gauge ID are shown for NWIS gauges. NWS gauge marked by ASOS ID.

3.2.2 QPE Datasets

Multiple agencies operate platforms with both real-time access to- and long term archives of- rain gauge and other hydrometeorlogical data. Two of the most commonly used networks in the US are the National Water Information Service (NWIS) operated by the United States Geological Survey (USGS) and the Automated Surface Observation System (ASOS) operated jointly by the

National Weather Service (NWS), the Department of Defense (DOD), and the Federal Aviation Administration (FAA). In Roanoke, Virginia the USGS recently installed a network of nine rain gauges that provide fairly good coverage of the metropolitan area. These gauges became operational at the beginning of March 2018. The ASOS gauge (KROA) located at the Roanoke airport has an archive going back to 1948. Full gauge identifiers and corresponding abbreviations are shown in Table A1. Gauge data was used to create three of the four QPE datasets: rainfall fields derived by kriging 5-minute incremental depths measured at all 10 rain gauges, MFB corrected 5-minute incremental depths, and a basin-uniform estimate based on the measured 5-minute depth at the KROA gauge. Level 3 instantaneous precipitation rates from the KFCX radar near Roanoke were used to create two of the four QPE datasets: uncorrected 5-minute incremental depths and MFB corrected 5-minute depths.

3.2.3 QPE Processing

Among gauge-based estimates there are a variety of methods for spatial interpolation between gauges. Two such methods - ordinary kriging and inverse distance weighting - have generally been shown to yield the most accurate rainfall estimates (Chen et al., 2010; Shope and Maharjan, 2015). Kriging is a sophisticated method of spatial interpolation and generally performs well in areas where orographic effects on rainfall are present. Kriging involves creation of a variogram, or relationship between distance and variance between points in a topological dataset, which is then used to estimate values between points.

Mean field bias (MFB) correction is perhaps the most common and easiest method for bias correction of radar rainfall data. MFB correction involves calculation of a ratio of average rainfall depth measured by selected gauges to the average depth determined by radar at the gauge location(s) (Equation 1).

$$\beta = \frac{\frac{1}{n}\sum_{i=1}^{n} P_{g-i}}{\frac{1}{n}\sum_{i=1}^{n} P_{ur-i}} = \frac{\sum_{i=1}^{n} P_{g-i}}{\sum_{i=1}^{n} P_{ur-i}} \qquad \text{Equation 1}$$

Where: β = mean field bias

P_{g-i} = Precipitation depth at gauge i measured by rain gauge (mm)

P_{ur-i} = Precipitation depth at gauge i measured by uncorrected radar (mm)

$n =$ number of gauges

The MFB ratio can be calculated on any time scale (e.g. 1 day, 1 hour, 5-minutes), using incremental depth measurements of the selected time scale. Once calculated, the ratios are applied to the entire uncorrected radar rainfall field at each time step (Equation 2).

$$P_{cr-t} = P_{ur-t} * \beta_t \qquad \text{Equation 2}$$

Where: $P_{cr-t} =$ Corrected radar rainfall depth field at time t (mm)

$P_{ur-t} =$ Uncorrected radar rainfall depth field at time t (mm)

$B_t =$ MFB ratio at time t

For this study, the mean field bias correction process involved correction of the Level 3 NEXRAD data using the 10 rain gauges in the study area. First, using Linux bash scripting and GRASS GIS 7.4 (GRASS Development Team, 2018), the irregular time series of Level 3 instantaneous intensity raster grids were interpolated to a 1-minute resolution. These 1-minute intensity rasters were then converted to 1-minute depths and aggregated to 5-minute depths, resulting in the uncorrected radar QPE (Figure 2).

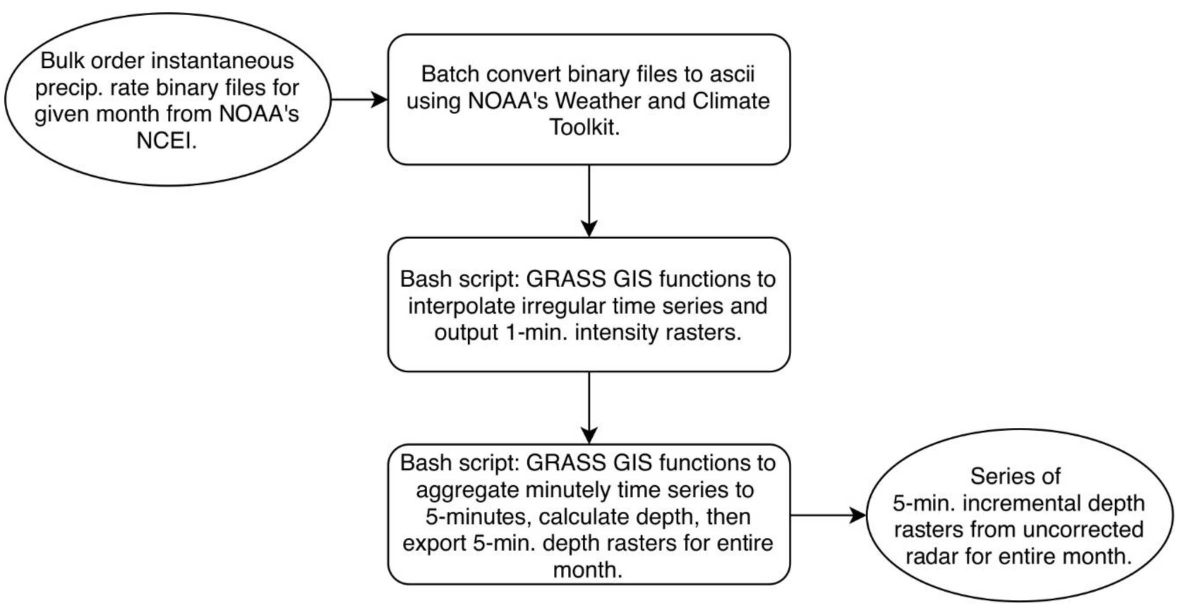


Figure 2. Algorithm for creation of uncorrected, 5-minute incremental radar depth QPE

Gauge precipitation data from NWIS and ASOS was retrieved and processed in R (R Core Team, 2018) to create a corresponding time series of 5-minute incremental depths. At each gauge location, a sample of the radar estimated precipitation depth was taken every five minutes for the study period. MFB ratios were calculated every five minutes from the sampled radar depths and gauge depths. MFB ratios of infinity, NaN, or 0 (resulting from #/0, 0/0, or 0/# respectively) were replaced with 1. The MFB ratios were then applied to the archive of 5-minute radar rainfall fields, creating a new set of corrected rainfall fields (Figure 3). A final bash script was used to spatially re-project the rasters to the Hydrologic Rainfall Analysis Project (HRAP) projection (Fulton, 1998) and convert the ascii radar grids to xmrg files prior to model input. xmrg is a binary file format commonly used by NWS and required for RDHM input (National Weather Service, 2011).

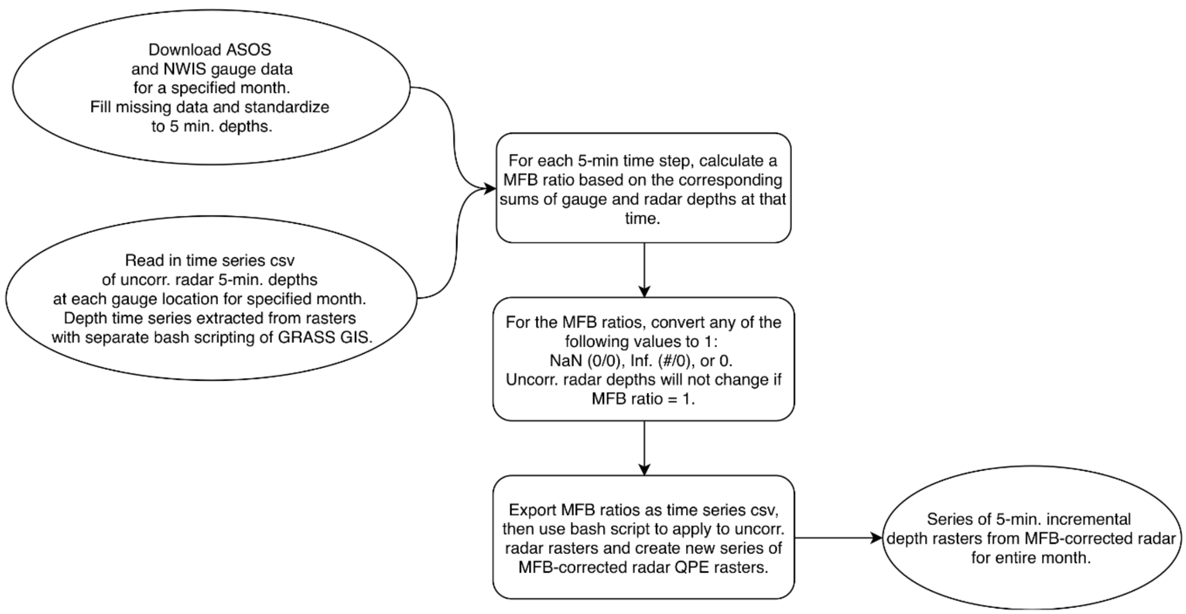


Figure 3. Algorithm for creation of MFB-corrected radar QPE

The 5-minute incremental gauge data was spatially interpolated by kriging using the autoKrige function in the automap R library (Hiemstra et al., 2009). autoKrige automatically determines which model to use for variogram fitting (e.g. spherical, gaussian, exponential) based on the given point data, then uses the resulting variogram to perform kriging. The resulting rainfall fields were then reprojected to HRAP and converted to xmrg files (Figure 4).

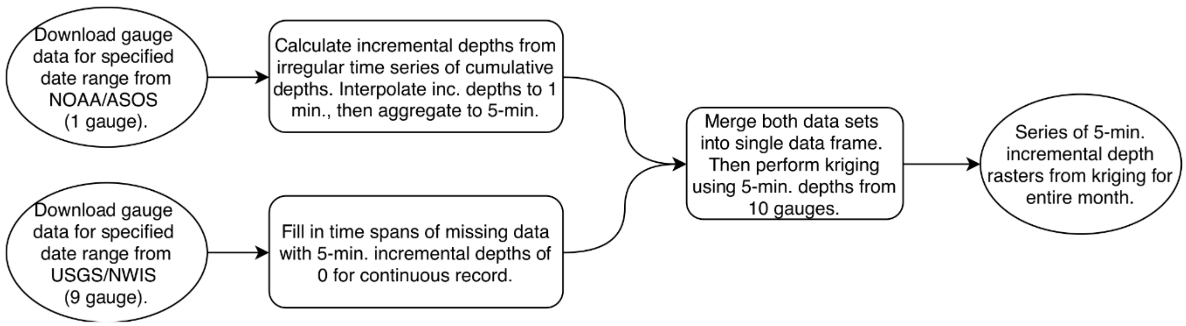


Figure 4. Algorithm for creation of kriged QPE

Creation of the single-gauge, uniform-basin QPE simply involved using the 5-minute incremental depths from the KROA gauge time series to create a series of rasters populated with the gauge depths (Figure 5).

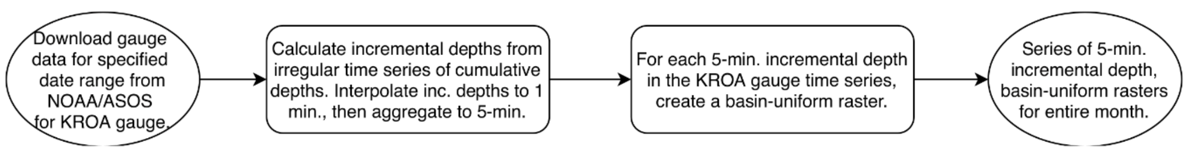


Figure 5. Algorithm for creation of single-gauge, uniform-basin QPE

All of the QPE products were created at a resolution of 600-m to match the native resolution of the original Level 3 radar data but were later re-sampled to 300-m before RDHM input to match the model resolution. In total, an archive of high-resolution data was created for the four QPE products (kriging, MFB corrected radar, uncorrected radar, and single gauge) spanning from 1 May 2018 to 31 October 2018, a period during which approximately two-thirds of the total annual precipitation occurred.

3.2.4 Research Distributed Hydrologic Model (RDHM)

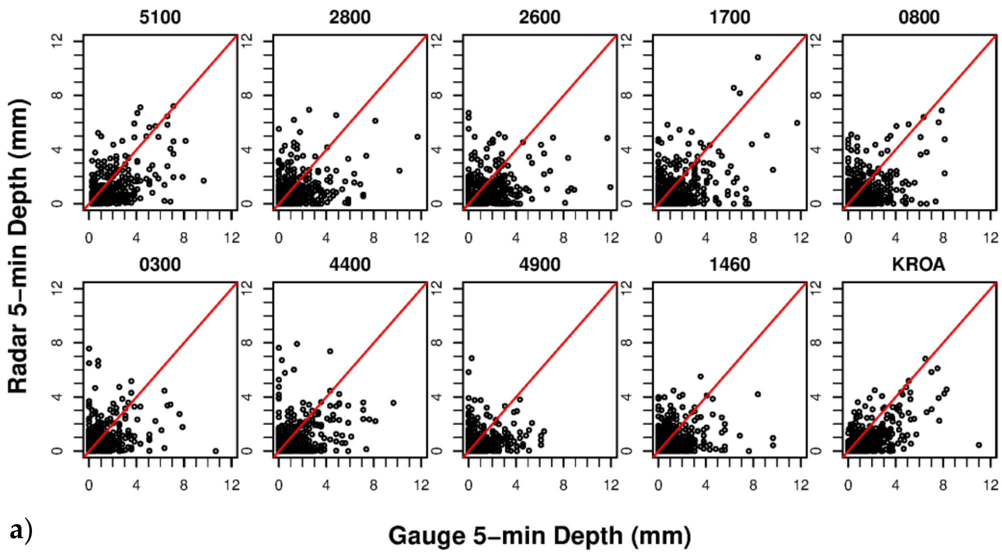
The National Weather Service (NWS) Hydrology Laboratory previously developed a distributed H&H model called the Research Distributed Hydrologic Model (RDHM), which is one of the models used for hydrologic forecasting in various NWS River Forecasting Centers (RFCs). Principal model components include Sacramento-Soil Moisture Accounting (SAC-SMA), RUTPIX overland and channel routing (kinematic wave), and SNOW-17 for snow operations. RDHM comes with a set of a-priori SAC-SMA parameters that have been developed for the entire conterminous United States (CONUS) and were derived using soil and land cover data from the Soil Survey Geographic Database (SSURGO) and National Land Cover Dataset (NLCD), respectively. Similarly, a CONUS-wide set of RUTPIX parameters is provided based on elevation data originally at a 400-m resolution ("Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) User Manual V. 3.5.0," 2012). However, the parameters, forcings, and resolution are typically derived and operated at a coarse scale, most commonly 1 HRAP grid (nominally ~4 km). Model initialization is generally recommended, and typically involves a 1-year or more warm-up simulation to determine initial model states for the study period. Calibration is also recommended but is not possible in ungauged watersheds. The a-priori parameters are therefore a useful tool for modeling in ungauged watersheds and aid in the calibration process for gauged watersheds (Z. Zhang et al., 2012).

For this study, a warm up simulation was run from 1 January 2017 through 31 December 2018 using precipitation and temperature data from NASA's Land Data Assimilation Systems (NLDAS) (Xia et al., 2012). From this historical simulation, the model states (e.g. channel depth, flow rate, soil water content, etc.) at the beginning of each month were used as the initial states for each monthly simulation (May-November 2018) forced with the high-resolution QPE datasets. RDHM was operated at a 1/16 HRAP grid resolution (approximately 300 m) due to the small size of the watershed and was run uncalibrated to avoid introducing bias towards one of the QPE methods.

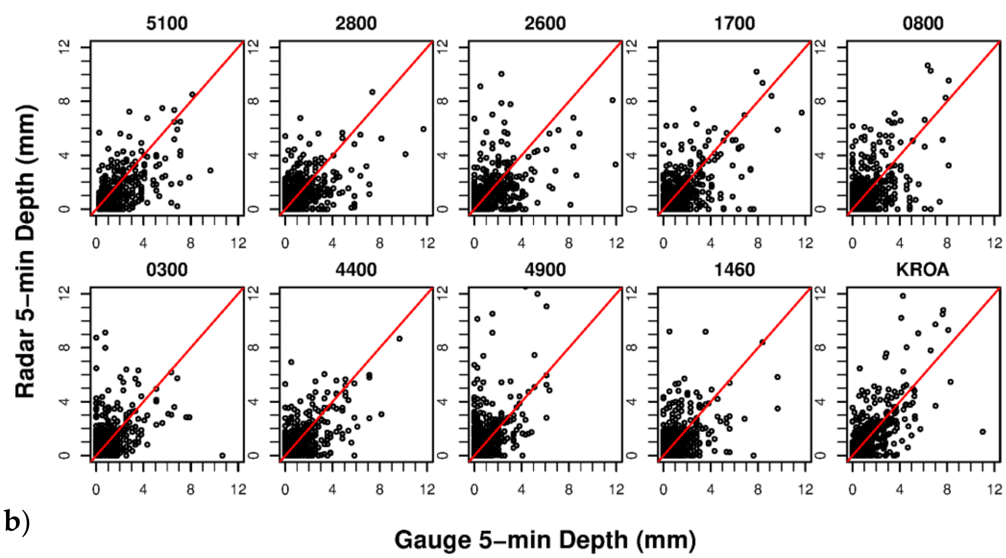
3.3 **Results**

3.3.1 QPE Products

Multiple QPE products were generated at a 5-minute, 600-m resolution for model input, including: uncorrected Level 3 radar data, MFB corrected Level 3 radar data, kriged rain gauge data (10 gauges), and a watershed uniform depth based on measurements at the KROA rain gauge. Figure 6 shows a comparison of gauge and radar 5-minute incremental depths before (a) and after (b) MFB correction.



a)



b)

Figure 6. Five-minute incremental depths at each gauge location for 6-month study period, as indicated by gauge ID (a) Uncorrected radar vs. gauge (b) Corrected radar vs. gauge (several outliers not shown).

From the 5-minute incremental depths, cumulative depths were calculated over the study period for gauge-, uncorrected radar-, and MFB corrected radar-QPEs at each gauge location (Figure 7).

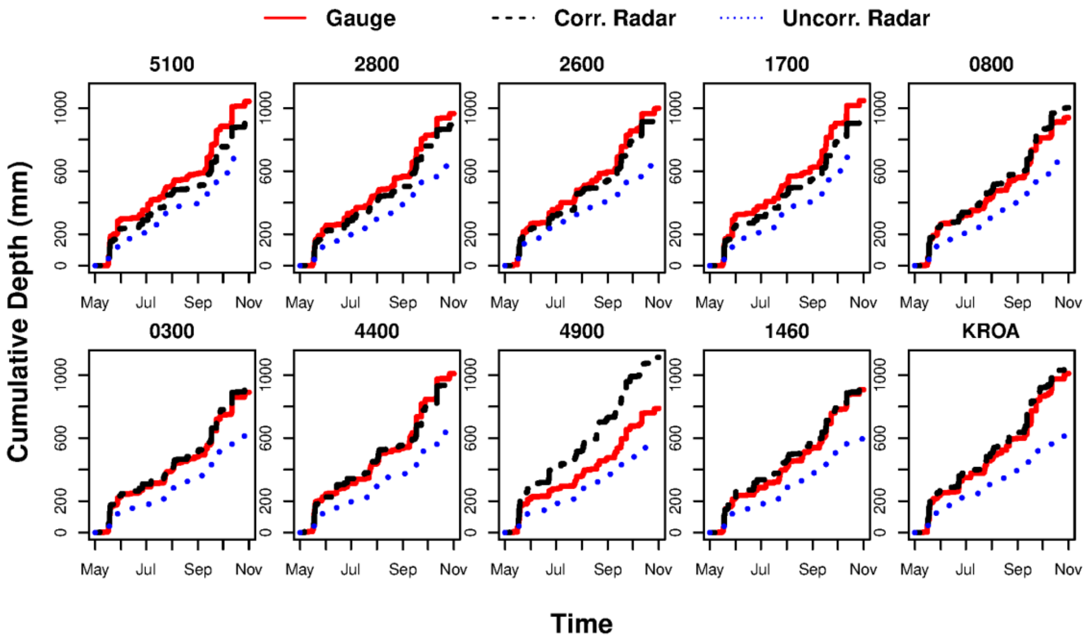


Figure 7. Cumulative precipitation depths over 6-months for gauge-, uncorrected radar, and MFB corrected radar at each rain gauge location.

The 5-minute mean areal precipitation (MAP) over the Lick Run basin was calculated by RDHM during each simulation along with the discharge time series. From the incremental depths, cumulative MAP was calculated for each month (Figure 8).

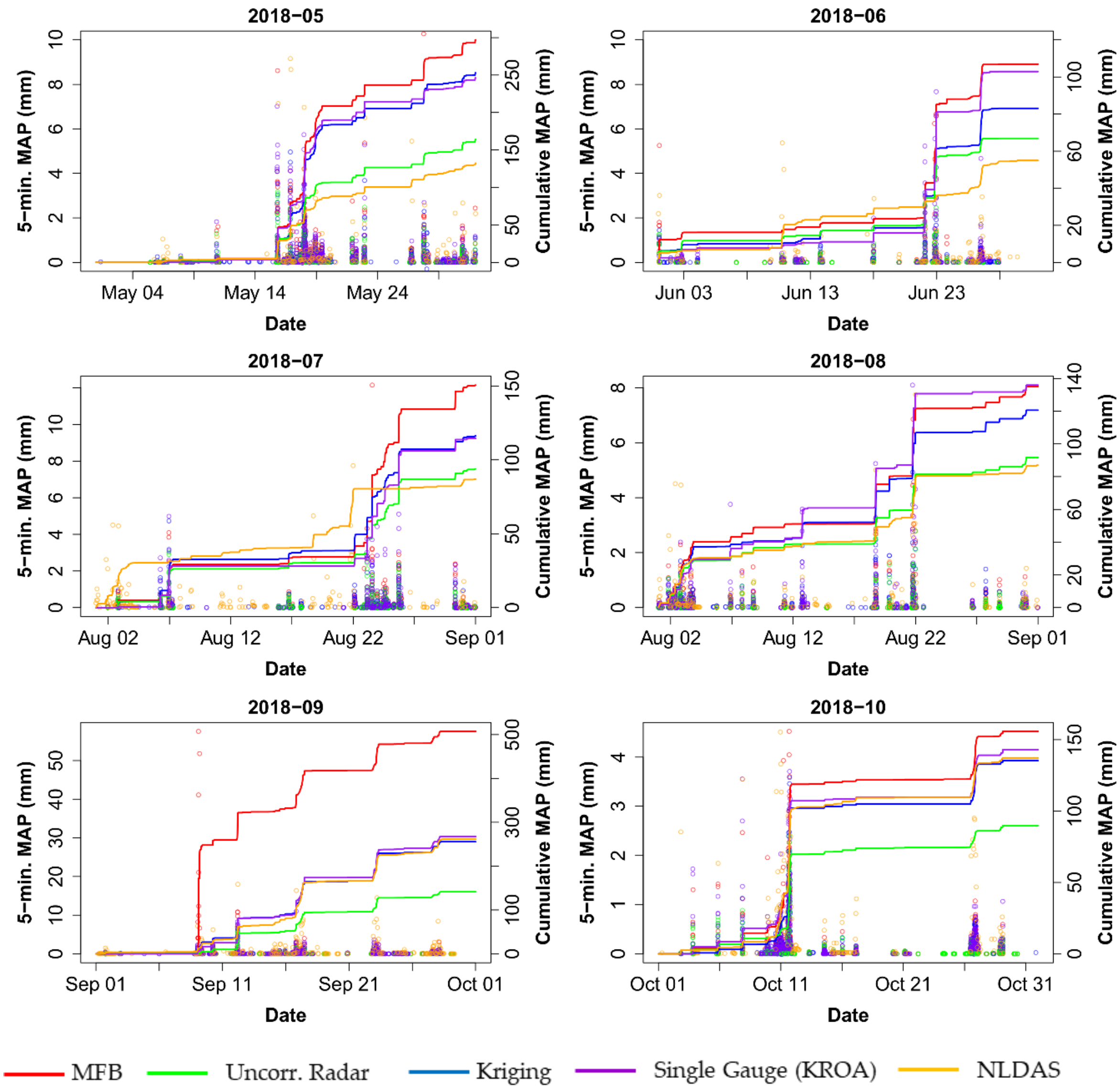


Figure 8. Incremental and cumulative mean areal precipitation (MAP) over the Lick Run basin by month for each of the high resolution QPE methods.

In order to evaluate over-prediction by certain high-resolution QPE methods (e.g. MFB correction and kriging), cumulative MAP from NLDAS and the KROA gauge between the beginning of 2017 to the end of 2018 was calculated and compared (Figure 9). Cumulative MAP from KROA and NLDAS was 2612 mm and 2447 mm, respectively; a difference of 165 mm.

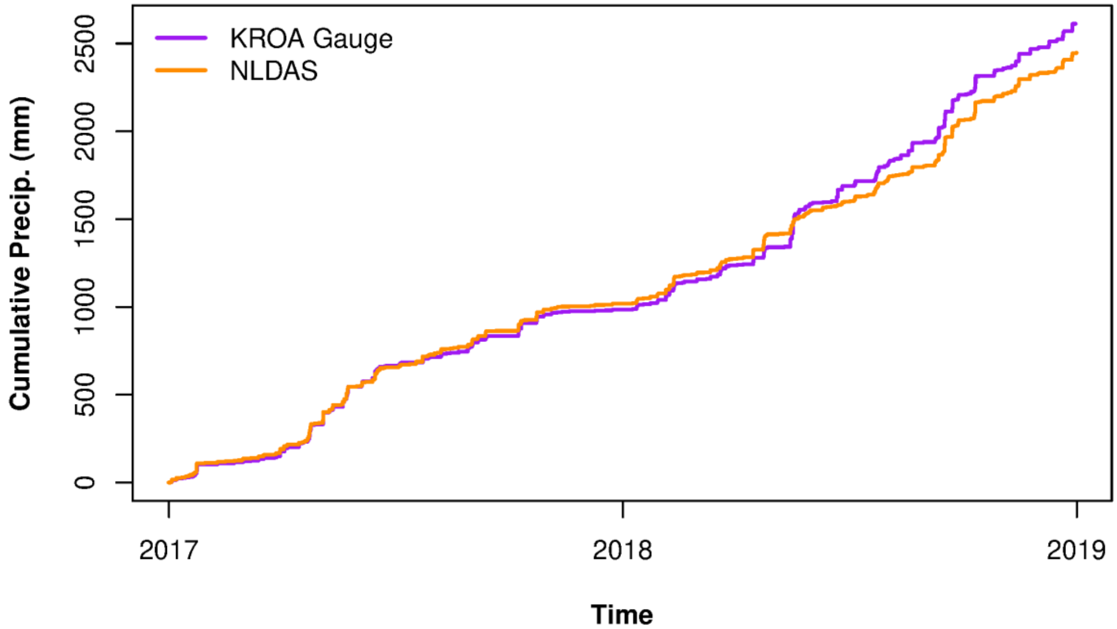


Figure 9. Cumulative mean areal precipitation (MAP) over the Lick Run basin for 2017 and 2018 as measured by NLDAS and KROA gauge.

For the 6-month study period in 2018, cumulative precipitation between NLDAS and each gauge in the network was calculated and compared (Figure 10). NLDAS had the lowest cumulative depths over the study period and appears to underpredict for larger storm events.

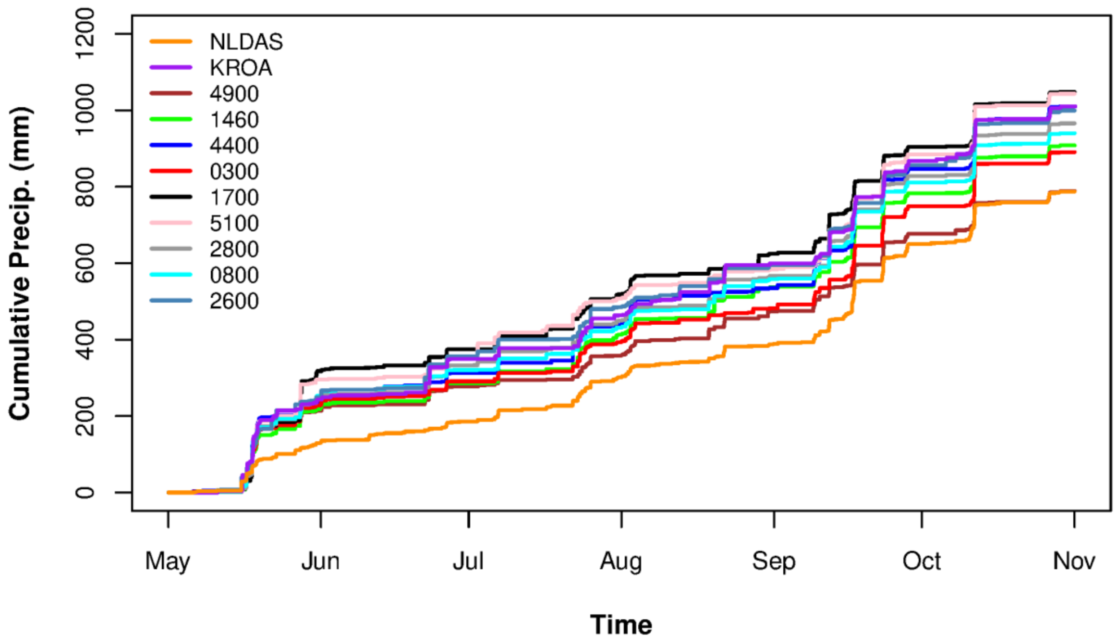


Figure 10. Cumulative precipitation comparison between NLDAS and gauge network for 6-month study period in 2018.

3.3.2 Model Results

From the uncalibrated historical simulation, a time series of predicted discharge at the basin outlet was generated by RDHM, then compared with USGS stream gauge measurements at the same location (Figure 11).

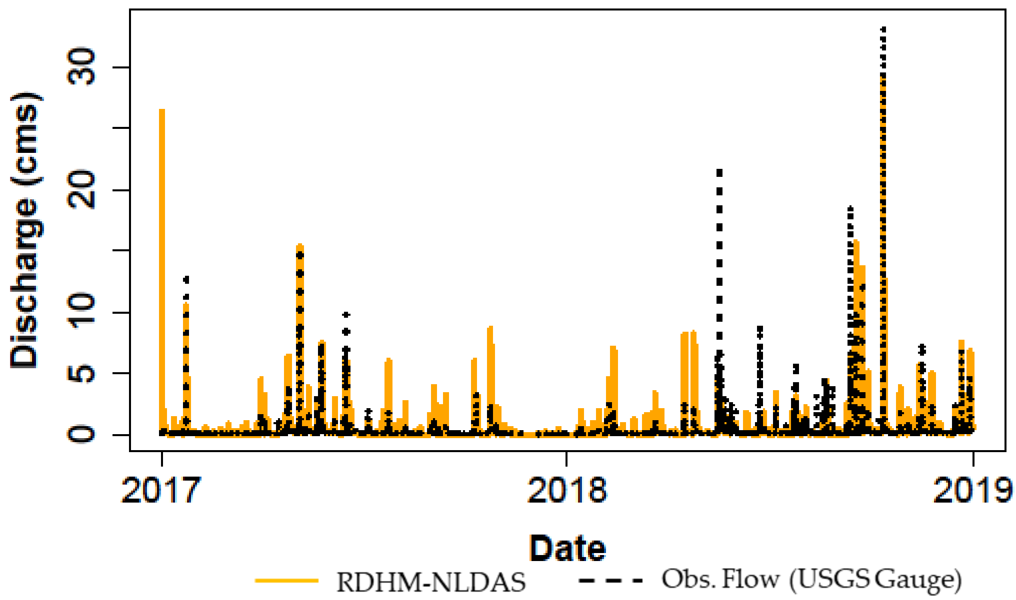


Figure 11. Historical simulation predicted flow compared to measured discharge (January 2017 through 31 December 2018).

The saved model states from this historical simulation were used as warm start values for the monthly simulations forced with each of the four high-resolution QPEs (Figure 12).

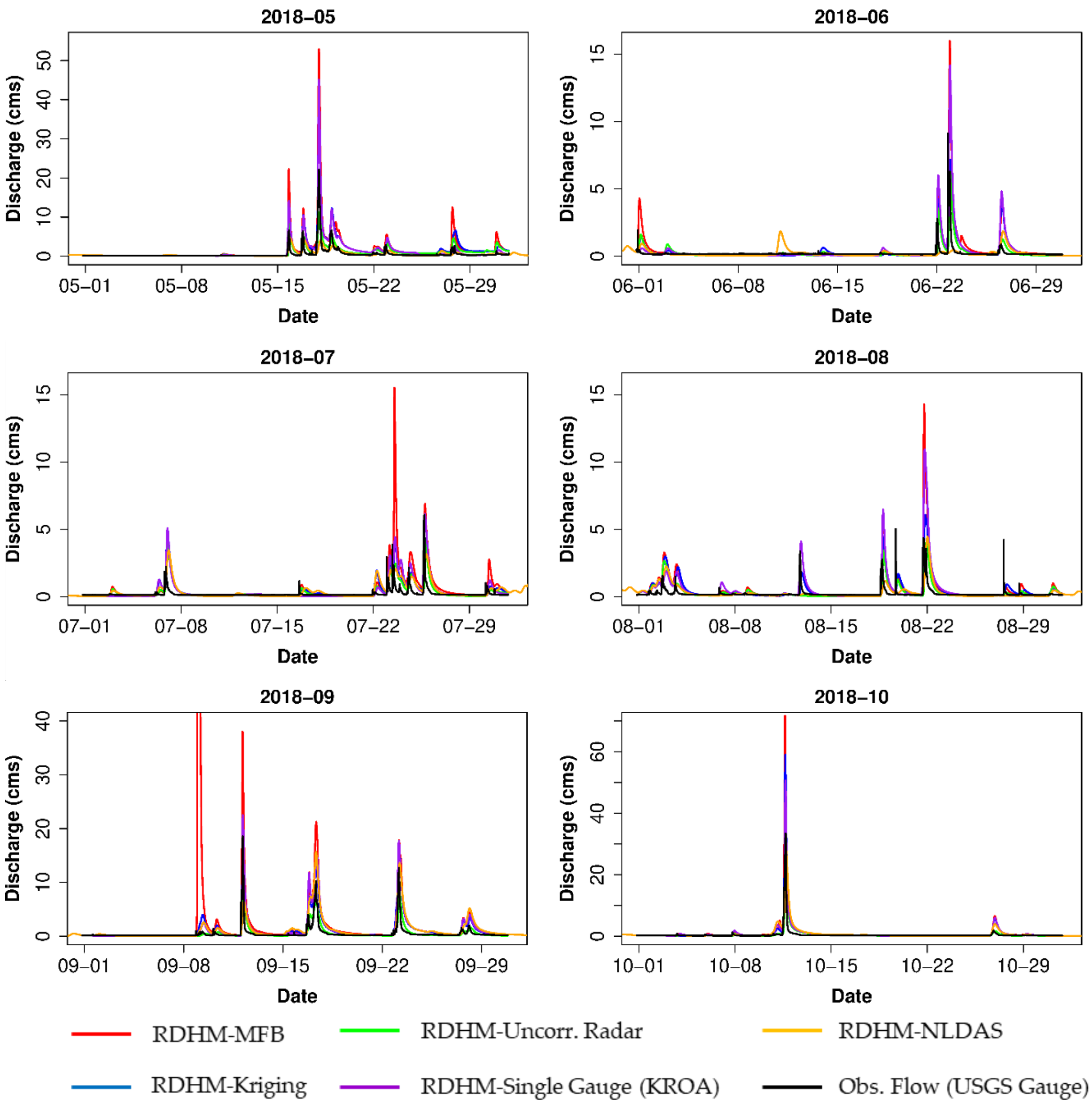


Figure 12. High resolution (5-min, ~300 m) simulation predicted flow compared to measured discharge by month.

Figure 13 shows the largest flow events from each month of the high-resolution simulations.

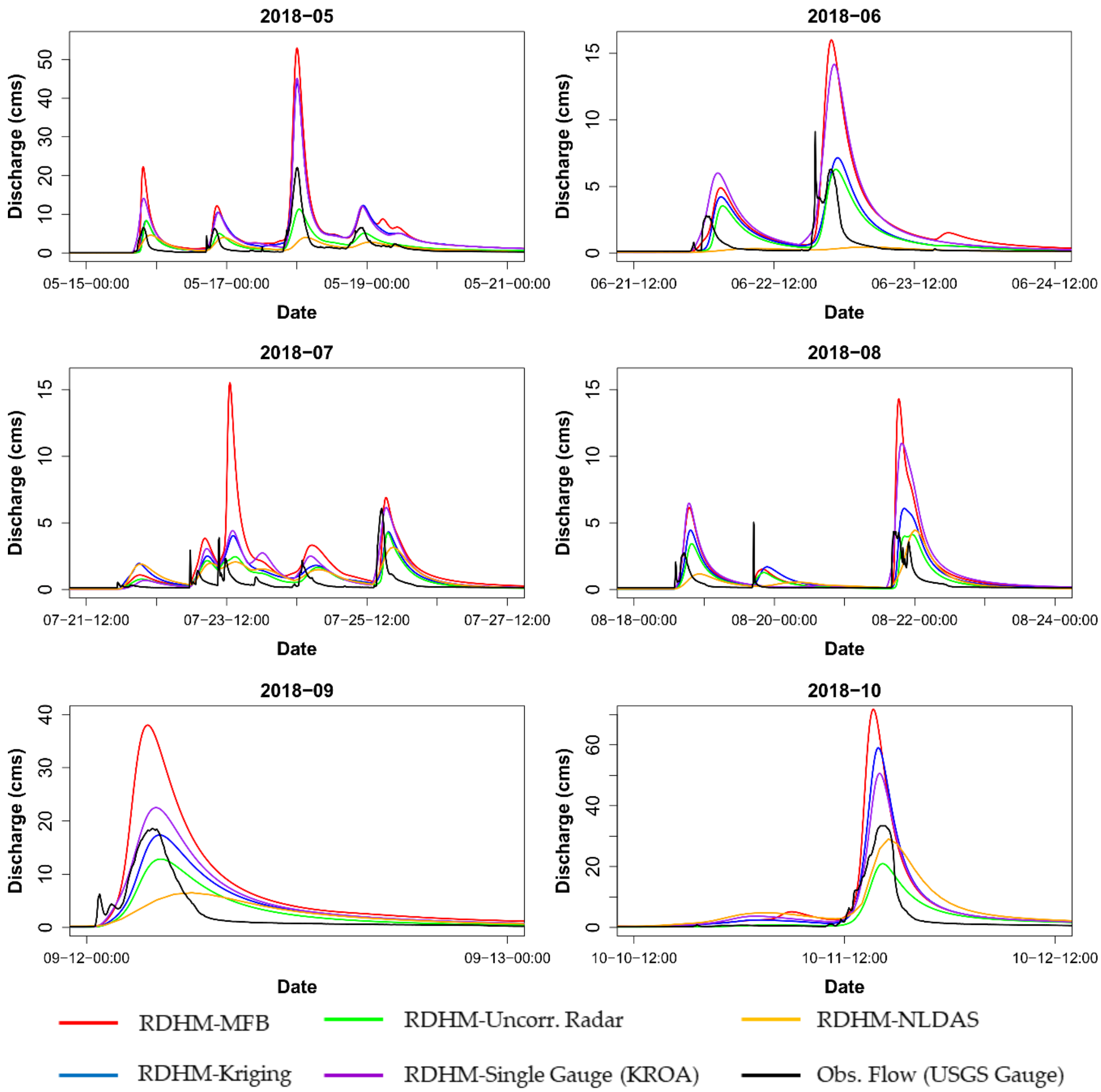


Figure 13. High resolution (5-min, ~300 m) simulation predicted flow compared to measured discharge for the biggest flow events each month.

Since the monthly simulations were conducted on a continuous basis, the root mean square error (RMSE) was calculated hourly rather than on an event basis. Similarly, RMSE values calculated during baseflow periods were removed. Here, baseflow was defined as anything less than 0.3 cms and any hourly RMSE values less than 0.3 cms were removed. Figure 14 shows the hourly RMSE values for each QPE forcing by month.

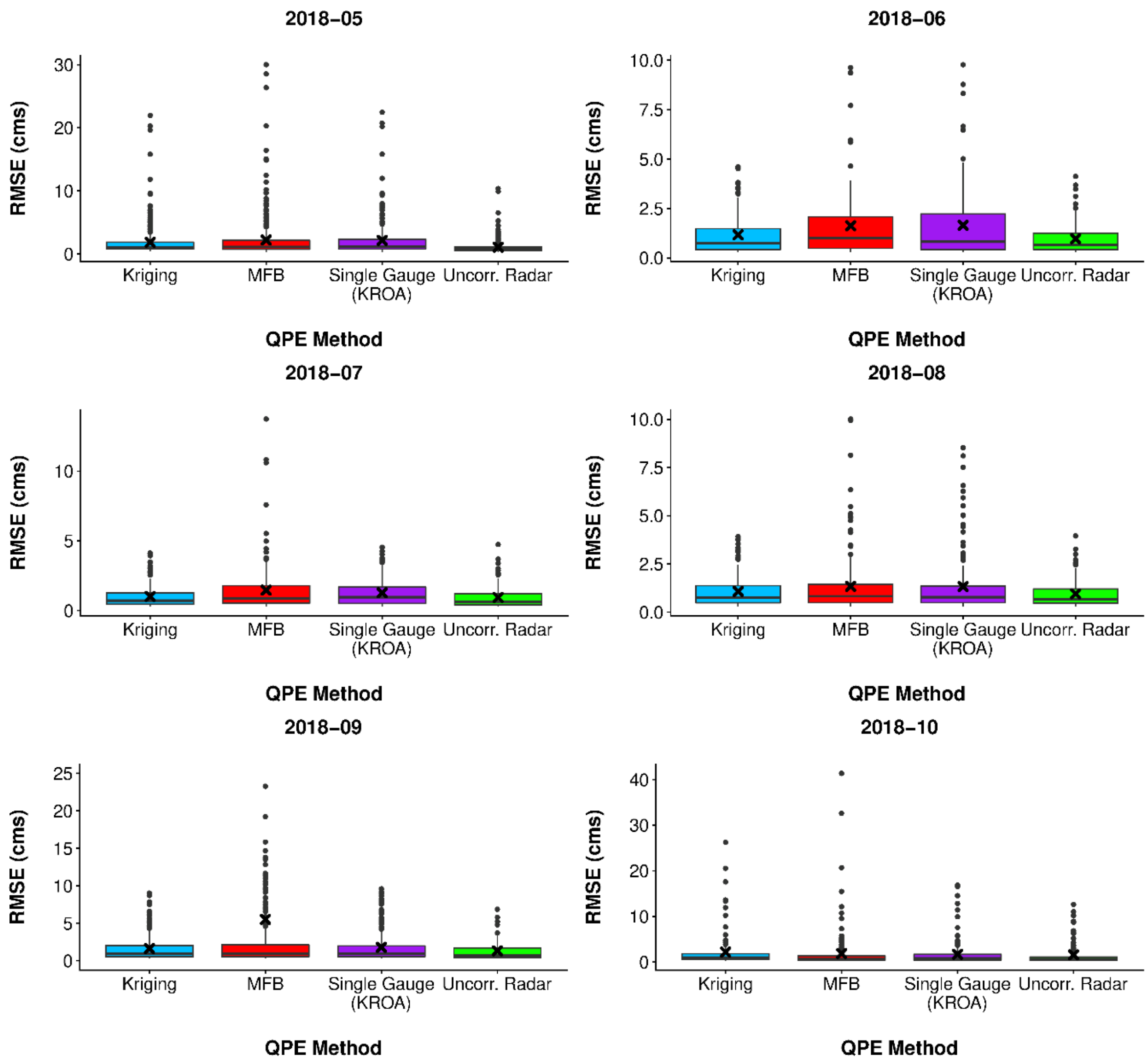


Figure 14. Hourly RMSE for each QPE method by month. RMSE values less than baseflow (0.3 cms) were removed.

Relative error in peak flow (REPQ) was analyzed by QPE forcing for each month by extracting the peaks for any flow event which surpassed 1 cms (Figure 15). The peak time error (PTE) for these events was also summarized (Figure 16). The number of flow events analyzed each month (n) ranged from 4 to 9.

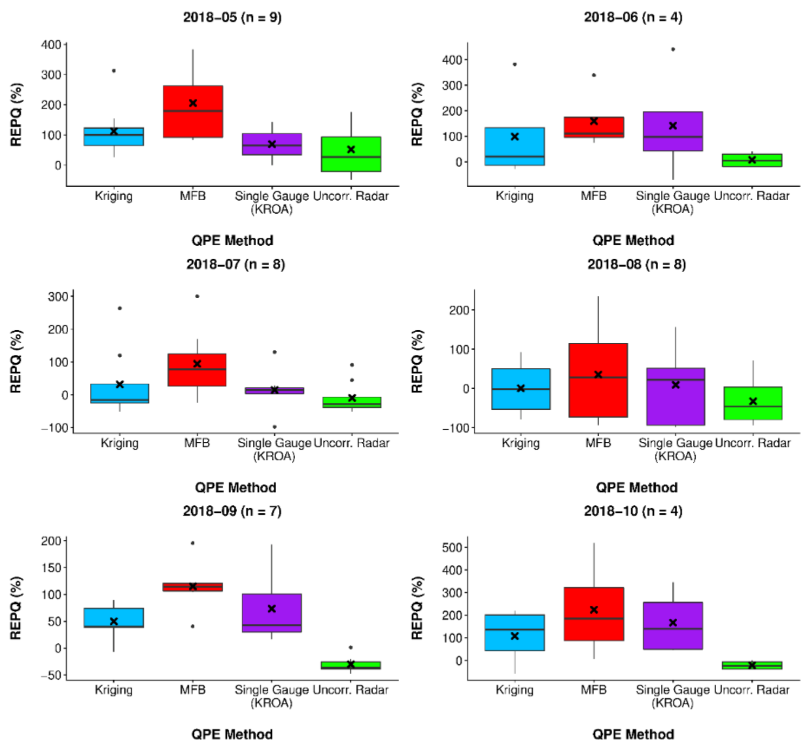


Figure 15. Relative error in peak flow (REPQ) by QPE method for *n* events each month.

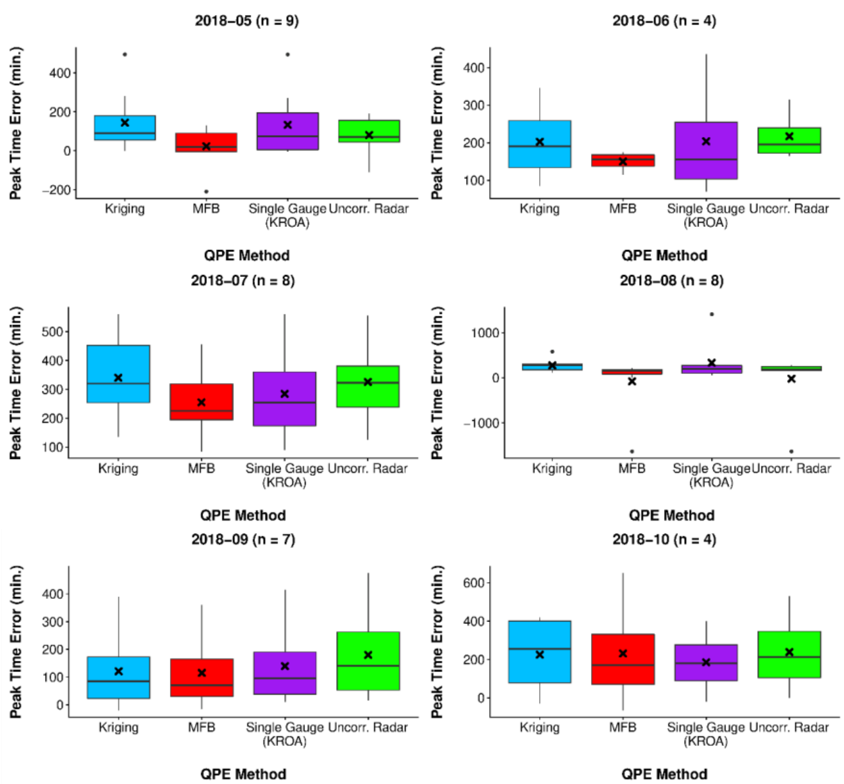


Figure 16. Error in peak flow timing by QPE method for *n* events each month.

Forty (40) flow events above 1 cms occurred between 1 May 2018 and 1 November 2018. A summary RMSE, REPQ, and PTE is shown in Table 1.

Table 1. Predicted discharge skill summary by QPE forcing.

Metric		**QPE**			
		MFB	Kriging	Single Gauge	Uncorr. Radar
Hourly RMSE (cms)	Mean	2.69	1.54	1.71	1.10
	Median	0.94	0.95	0.94	0.70
	Minimum	0.3	0.3	0.3	0.30
	Maximum	368	26.3	22.5	12.6
REPQ (%)	Mean	131	61.2	64.1	-3.4
	Median	110	45.5	40.9	-21.4
	Minimum	-92.7	-78.3	-97.6	-93.2
	Maximum	519	382	441	176
PTE (hours)	Mean	1.65	3.66	3.61	2.61
	Median	2.08	3.08	2.75	3
	Minimum	-27.3	-0.5	-0.33	-27.3
	Maximum	10.8	9.67	23.5	9.25

3.4 **Discussion**

The RDHM simulations yielded interesting and counterintuitive results: the uncorrected radar QPE forcing resulted in the most accurate discharge predictions, rather than the more sophisticated methods of kriging and MFB correction. Regarding predicted discharge magnitude, the simulations forced with uncorrected radar data had the lowest error as measured by hourly RMSE and peak flow relative error. The kriging and single gauge QPE forcings resulted in moderately worse and relatively similar performance statistics. While the MFB-corrected radar QPE led to significantly worse results, caused by severe and consistent overprediction. In terms of peak timing, the MFB forcing performed moderately better than the remaining QPE methods, all of which had similar aggregate results.

As expected, the uncorrected radar systematically underpredicted. Less expected was the overprediction of the three other QPE forcings (MFB, kriging, single gauge). There are several possible explanations for the over-estimation of the three QPE forcings: model parameterization, QPE methodology, and/or model initialization.

Here, model parameterization refers to the a-priori SAC-SMA and rutpix routing parameters based on soil, land cover, and elevation data. Potentially, the resolution of the datasets from which the parameters were derived, and/or the resolution of the resulting parameter grids (1 HRAP grid resolution), is too coarse for small, urban watersheds. Alternatively, mechanisms in the parameter derivation process lead to parameters which cause over-estimated discharge predictions. One or both of these rationales would explain the performance statistics of the uncorrected radar: even though radar data that has not been gauge-corrected is well known to unpredict precipitation, discharge overestimation attributed to RDHM parameterization may increase the predicted flow such that the uncorrected radar forcing had the highest skill. Re-derivation of model parameters using higher resolution data, and/or producing higher resolution parameter grids, may mitigate this over-estimation.

Besides model parameterization, another potential source of over-estimation is the QPE methodology used to create the various precipitation forcings. Potentially, even a gauge density of ~6.5 km is too coarse for small, urban watersheds. Similarly, single gauge coverage for a ~19.4 km^2 watershed appears to be inadequate. Both of these possibilities would support prior studies' recommendations of QPE resolution on the order of 1 km (Berne et al., 2004; Ochoa-Rodriguez et al., 2015; Schilling, 1991; Thorndahl et al., 2017).

There are several potential sources of error for the MFB corrected QPE. One possibility is random error associated with the MFB correction process. Much of the systematic bias was removed during the MFB process, but significant random error persisted – likely due to the small time scale (5-mintes) of the correction process (Figure 6). This supports prior findings (Krajewski W.F. and Smith J.A., 2002; Thorndahl et al., 2017) that a longer time scale is needed for MFB correction to remove random errors, but the time scale cannot be so long such that the drop size distribution (DSD) changes. Changes in DSD are a particular concern for short, intense convective storms common during summer months, where storm durations are likely to be sub-hourly and DSD can change over the course of several minutes. Further, while the MFB process provided relatively good correction of the radar data at each gauge location (Figure 7), precipitation over-prediction may have resulted from spatial variability in field bias between gauges. This suggests that even at small spatial scales using the mean of the field bias does not provide adequate correction, necessitating a spatially varied correction method. One example of this occurred in early September 2018, gauge and radar depths inside the watershed were 0 or negligible but moderate and trace depths for gauge and radar estimates, respectively, at gauge location(s) outside the watershed resulted in very high MFB ratios, such that the MAP inside the basin was erroneously and highly inflated. This phenomenon can be seen in the several high incremental MFB MAP values during September 2018 in Figure 8 and the resulting impact on predicted discharge as well as cumulative MAP. This also suggests that an outlier threshold is required for MFB correction.

A third possible source of overestimation may be the saved model states from the historical simulation that were used as warm start values for the monthly simulations. The coarse resolution of the NLDAS forcings (13.5-km, 1-hour) may not be adequate for estimating model states (e.g. soil moisture) in small, urban watersheds. Similarly, a finer model resolution (e.g. 30-m instead of 300-m) may improve performance at the cost of increased model run time.

Generally, the MAP from the MFB QPE was consistently higher than the remaining QPE methods, and except for one or two exceptions was relatively close to the MAP from the kriging and single gauge QPEs. Conversely, the MAP from the uncorrected radar was consistently and significantly the lowest of the four high-resolution QPEs. For three of six months, NLDAS had the lowest cumulative MAP, but for the other three months had cumulative MAP similar to the kriging and single gauge QPEs. The cumulative MAP from NLDAS and the KROA gauge between 1 January 2017 and 1 January 2019 tracked closely, with a final difference of 165 mm. As an operational dataset, NLDAS is a reliable QPE on hourly and daily time scales, which seems to confirm both the accuracy of the KROA gauge (Figure 9) and the overprediction of the MFB corrected QPE. However, it does appear that NLDAS underpredicts as storm total precipitation depth increases (e.g. first rainfall event in Figure 10).

Several other findings regarding QPE temporal resolution, gauge density, and model calibration can be posited from the study. Although overall cumulative MAP between NLDAS and several of the high-resolution QPEs was similar, consistent underprediction of discharge when forced with NLDAS suggests that hourly precipitation inputs are inadequate for modeling small, urban watersheds – a finding echoed by prior research (Berne et al., 2004; Ochoa-Rodriguez et al., 2015; Schilling, 1991; Thorndahl et al., 2017). Similar results from the kriging and single-gauge QPEs indicate that, for the gauge density and basin size of this study (1 gauge per 9.7 km^2 and 19.4 km^2, respectively), a single gauge QPE (using a gauge inside the watershed) can provide results on-par with estimates from kriging. While these two QPE products at the given density were not optimal for RDHM, better performance may be achieved with other models. Calibration would improve model skill, however calibration would necessarily bias model results towards the QPE forcing used in calibration, making comparisons between different QPE forcings tenuous.

The RDHM was not calibrated, so fast response hydrograph features are not captured well (e.g. Figure 13, second flow event in the 2018-06 series). Additionally, the model does not have the capability to model storm sewer infrastructure and does not account for detailed variations in channel cross-section geometry. However, with better parameterization, calibration, and possibly a higher model resolution, RDHM has the potential to skillfully predict peaks and volume, even in small, urban watersheds. The relatively good performance of the uncorrected radar QPE with RDHM makes this pairing an attractive option for high resolution discharge simulations in small, gauged- or ungauged-watersheds; especially considering that both the Level 3 dataset and the RDHM parameters have coverage for the entire CONUS.

3.5 **Conclusion**

Overall, this study examined various QPE methods for continuous, distributed hydrological modeling in a small urban watershed using an operational H&H model. Of the four high-resolution precipitation forcings used, the uncorrected Level 3 radar QPE resulted in the most accurate discharge predictions from RDHM. This suggests that when paired with RDHM, the Level 3 dataset has a potential for use in hydrologic modeling of small, urban watersheds – even without bias correction. Model calibration is recommended.

Future studies which use RDHM in small urban watersheds should compare various parameter and model resolutions as well as uncalibrated versus calibrated simulations. Similarly, the effect of outlier removal on MFB QPE simulation skill should be examined. Finally, more advanced bias correction algorithms, such as those used in the Multi-Radar Multi-Sensor (MRMS) system (J. Zhang et al., 2016), are needed at finer spatial and temporal resolutions.

CHAPTER 4. CONCLUSION

4.1 Implications

The ability to accurately and efficiently predict streamflow is becoming increasingly important, particularly due to two key reasons: accelerating urbanization throughout the world and climatic changes that increase the frequency and intensity of flood producing rainfall events in certain regions. Flash flooding in urban areas poses a risk to both people and property: improving hydrologic prediction skill can benefit design of infrastructure based on historical simulations as well as forecasting for flood warnings and evacuations.

While precipitation is a key piece of the hydrologic puzzle, the model and parameters used will also influence predictions. Particularly for resource constrained stakeholders or in areas where detailed infrastructure information is not available, simple models and pre-defined parameters that perform well are an important and potentially easy to implement tool for flood prevention and mitigation. These models must be paired with accurate rainfall forcings with a high spatiotemporal resolution – the NEXRAD system is one such example. Given the high capital and maintenance costs of a national radar network, a potential alternative to radar networks may lie with satellite derived QPEs due to their potential for a lower cost and wider coverage. Open source methods, such as those used in this study, should be used in such efforts to make these tools available to the widest possible audience.

Overall, this study examined various QPE methods for continuous, distributed hydrological modeling in a small urban watershed using an operational H&H model. Of the four high-resolution precipitation forcings used, the uncorrected Level 3 radar QPE resulted in the most accurate discharge predictions from RDHM. This suggests that when paired with RDHM, the Level 3 dataset has a potential for use in hydrologic modeling of small, urban watersheds – even without bias correction. These results are encouraging for multiple reasons. First, RDHM is used for flood prediction and forecasting in multiple NWS RFCs, incorporation of a precipitation forcing with a high spatiotemporal resolution and CONUS-wide availability such as the Level III dataset into operational flood prediction can potentially be a benefit for flood warnings in small urban watersheds. Additionally, it was shown that a distributed operational model that did not incorporate storm sewer infrastructure or detailed channel geometry, and only relied on a water balance and simple kinematic routing at a coarse resolution, was able to perform relatively well in a small, urban watershed. Potentially, these results may serve as a rationale for greater focus on radar bias correction at high spatiotemporal resolutions for operational flood prediction in small urban basins. Finally, it would seem that simplistic methods such as MFB correction are not adequate for precipitation estimation in small urban watersheds.

4.2 **Future Work**

Future studies which use RDHM in small urban watersheds should compare various parameter and model resolutions as well as uncalibrated versus calibrated simulations. Whether, for example, a 30-m model and/or parameter resolution provides significant improvement over a 300-m resolution is an important question, as a higher resolution will significantly increase computational time and may not be warranted for only marginal improvements in skill. Similarly, the improvement gained by calibration should be quantified.

The source of over-prediction when using RDHM should also be determined. Does the overprediction observed in this study result from model parameterization and/or resolution, or does it originate from the QPE methods used? The distinction is important, as it could identify important shortcomings within the RDHM parameterization process or detect deficiencies in the QPE methods and sources used.

It seems that MFB correction is inadequate for urban hydrology, a finding supported by prior studies. However, it is a simple and easily implemented radar correction method. The effect of outlier removal and other quality control measures on MFB corrected radar simulation skill should be examined, as some operational datasets still use MFB correction and others may use it in the future.

Given the shortcomings of MFB, more sophisticated radar correction methods, such as conditional merging, which better maintains rainfall field spatial variability in the correction process, should be examined on a continuous basis under a variety of hydrometeorological conditions. Similarly, more work should be done to improve the spatiotemporal resolution and accuracy of satellite derived QPE methods, in order to provide a global dataset of high-quality precipitation measurements and hopefully negate the need for installation and maintenance of radar networks.

Comparisons between RDHM and other H&H models, particularly those designed for urban H&H modeling (e.g. SWMM, GSSHA, SewerGEMS), should be made. RDHM should be tested in a greater number of small urban watersheds, under diverse geographic and climatic conditions, in order to confirm its efficacy for such applications. Potentially, RDHM could be an effective operational model for flood prediction and forecasting at high spatiotemporal resolutions and an attractive alternative to traditional urban hydrology models.

REFERENCES

Berne, A., Delrieu, G., Creutin, J. D., and Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2004.08.002.

Chen, D., Ou, T., Gong, L., Xu, C. Y., Li, W., Ho, C. H., and Qian, W. (2010). Spatial interpolation of daily precipitation in China: 1951-2005. *Advances in Atmospheric Sciences*. doi:10.1007/s00376-010-9151-y.

Cristiano, E., ten Veldhius, M.-C., and van de Giesen, N. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - a review. *Hydrology and Earth System Sciences*. doi:10.5194/hess-21-3859-2017.

Cunha, L. K., Smith, J. A., Krajewski, W. F., Baeck, M. L., and Seo, B.-C. (2015). NEXRAD NWS Polarimetric Precipitation Product Evaluation for IFloodS. *Journal of Hydrometeorology*, *16*(4), 1676–1699. doi:10.1175/JHM-D-14-0148.1.

Daniels, E. E., Lenderink, G., Hutjes, R. W. A., and Holtslag, A. A. M. (2016). Observed urban effects on precipitation along the Dutch West coast. *International Journal of Climatology*. doi:10.1002/joc.4458.

Dymond, R. L., Aguilar, M. F., Bender, P., and Hodges, C. C. (2017). *Lick Run Watershed Master Plan*. Roanoke.

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L., and Carbonneau, P. E. (2013). Topographic structure from motion: A new development in photogrammetric measurement. *Earth Surface Processes and Landforms*. doi:10.1002/esp.3366.

Freitag, B. M., Nair, U. S., and Niyogi, D. (2018). Urban Modification of Convection and Rainfall in Complex Terrain. *Geophysical Research Letters*. doi:10.1002/2017GL076834.

Fulton, R. (1998). *WSR-88D Polar-to-HRAP Mapping*. Silver Spring. Retrieved from https://www.nws.noaa.gov/oh/hrl/papers/wsr88d/hrapmap.pdf.

Gironás, J., Niemann, J. D., Roesner, L. A., Rodriguez, F., and Andrieu, H. (2010). Evaluation of methods for representing urban terrain in stormwater modeling. *Journal of Hydrologic Engineering*. doi:10.1061/(ASCE)HE.1943-5584.0000142.

GRASS Development Team. (2018). Geographic Resources Analysis Support System (GRASS) Software. Open Source Geospatial Foundation. Retrieved from http://grass.osgeo.org/.

Guo, J. C. Y. (2007). Effect of Raingage Density on Runoff Simulation Modeling. *World*

Environmental and Water Resources Congress 2007. doi:10.1061/40927(243)2.

Habib, E., Krajewski, W. F., and Kruger, A. (2001). Sampling Errors of Tipping-Bucket Rain Gauge Measurements. *Journal of Hydrologic Engineering*. doi:10.1061/(ASCE)1084-0699(2001)6:2(159).

Hiemstra, P. H., Pebesma, E. J., Twenhöfel, C. J. W., and Heuvelink, G. B. M. (2009). Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. *Computers and Geosciences*. doi:10.1016/j.cageo.2008.10.011.

Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) User Manual V. 3.5.0. (2012). NOAA. Retrieved from https://www.cbrfc.noaa.gov/present/rdhm/RDHM_User_Manual.pdf.

James, W. P., Robinson, C. G., and Bell, J. F. (1993). Radar-Assisted Real-Time Flood Forecasting. *Journal of Water Resources Planning and Management*. doi:10.1061/(ASCE)0733-9496(1993)119:1(32).

Kim, B. S., Kim, B. K., and Kim, H. S. (2008). Flood simulation using the gauge-adjusted radar rainfall and physics-based distributed hydrologic model. *Hydrological Processes*. doi:10.1002/hyp.7043.

Krajewski, W. F., Ciach, G. J., and Habib, E. (2003). An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological Sciences Journal*. doi:10.1623/hysj.48.2.151.44694.

Krajewski W.F., and Smith J.A. (2002). Radar hydrology: rainfall estimation. *Advances in Water Resources*. doi:10.1016/S0309-1708(02)00062-3.

Lanza, L. G., and Stagi, L. (2008). Certified accuracy of rainfall data as a standard requirement in scientific investigations. *Advances in Geosciences*. doi:10.5194/adgeo-16-43-2008.

Looper, J. P., and Vieux, B. E. (2012). An assessment of distributed flash flood forecasting accuracy using radar and rain gauge input for a physics-based distributed hydrologic model. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2011.05.046.

Löwe, R., Thorndahl, S., Mikkelsen, P. S., Rasmussen, M. R., and Madsen, H. (2014). Probabilistic online runoff forecasting for urban catchments using inputs from rain gauges as well as statically and dynamically adjusted weather radar. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2014.03.027.

Mayer, H. (1999). Automatic object extraction from aerial imagery - a survey focusing on buildings. *Computer Vision and Image Understanding*. doi:10.1006/cviu.1999.0750.

Molini, A., Lanza, L. G., and La Barbera, P. (2005). The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications. *Hydrological Processes*. doi:10.1002/hyp.5646.

National Academies of Sciences, Engineering, and Medicine. (2019). *Framing the Challenge of Urban Flooding in the United States*. Washington D.C.

National Weather Service. (2011). The XMRG File Format and Sample Codes to Read XMRG Files. *Distributed Model Intercomparison Project Phase 2*. Retrieved from https://www.nws.noaa.gov/ohd/hrl/dmip/2/xmrgformat.html.

Niemi, T. J., Warsta, L., Taka, M., Hickman, B., Pulkkinen, S., Krebs, G., Moisseev, D. N., Koivusalo, H., and Kokkonen, T. (2017). Applicability of open rainfall data to event-scale urban rainfall-runoff modelling. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2017.01.056.

Ochoa-Rodriguez, S., Wang, L. P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., Van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., and Ten Veldhuis, M. C. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2015.05.035.

Ogden, F. L., Sharif, H. O., Senarath, S. U. S., Smith, J. A., Baeck, M. L., and Richardson, J. R. (2000). Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. *Journal of Hydrology*. doi:10.1016/S0022-1694(00)00146-3.

Pessoa, M. L., Bras, R. L., and Williams, E. R. (1993). Use of Weather Radar for Flood Forecasting in the Sieve River Basin: A Sensitivity Analysis. *Journal of Applied Meteorology*. doi:10.1175/1520-0450(1993)032<0462:uowrff>2.0.co;2.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/.

Schilling, W. (1991). Rainfall data for urban hydrology: what do we need? *Atmospheric Research*. doi:10.1016/0169-8095(91)90003-F.

Seo, B.-C., Krajewski, W. F., Quintero, F., ElSaadani, M., Goska, R., Cunha, L. K., Dolan, B., Wolff, D. B., Smith, J. A., Rutledge, S. A., and Petersen, W. A. (2018). Comprehensive Evaluation of the IFloodS Radar Rainfall Products for Hydrologic Applications. *Journal of*

Hydrometeorology, *19*(11), 1793–1813. doi:10.1175/jhm-d-18-0080.1.

Shope, C. L., and Maharjan, G. R. (2015). Modeling spatiotemporal precipitation: Effects of density, interpolation, and land use distribution. *Advances in Meteorology*. doi:10.1155/2015/174196.

Smith, B. K., Smith, J. A., Baeck, M. L., Villarini, G., and Wright, D. B. (2013). Spectrum of storm event hydrologic response in urban watersheds. *Water Resources Research*. doi:10.1002/wrcr.20223.

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F. (2000). Flood estimation using radar and raingauge data. *Journal of Hydrology*. doi:10.1016/S0022-1694(00)00350-4.

Thorndahl, S., Einfalt, T., Willems, P., Ellerbæk Nielsen, J., Ten Veldhuis, M. C., Arnbjerg-Nielsen, K., Rasmussen, M. R., and Molnar, P. (2017). Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences*. doi:10.5194/hess-21-1359-2017.

Tokarczyk, P., Leitao, J. P., Rieckermann, J., Schindler, K., and Blumensaat, F. (2015). High-quality observation of surface imperviousness for urban runoff modelling using UAV imagery. *Hydrology and Earth System Sciences*. doi:10.5194/hess-19-4215-2015.

Wang, L. P., Ochoa-Rodriguez, S., Onof, C., and Willems, P. (2015a). Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications. *Hydrology and Earth System Sciences*. doi:10.5194/hess-19-4001-2015.

Wang, L. P., Ochoa-Rodríguez, S., Simões, N. E., Onof, C., and Maksimović, Č. (2013). Radar-raingauge data combination techniques: A revision and analysis of their suitability for urban hydrology. *Water Science and Technology*. doi:10.2166/wst.2013.300.

Wang, L. P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R. D., Pessemier, M., Kroll, S., Willems, P., and Onof, C. (2015b). Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2015.05.049.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., Livneh, B., Lettenmaier, D., Koren, V., Duan, Q., Mo, K., Fan, Y., and Mocko, D. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research Atmospheres*. doi:10.1029/2011JD016048.

Yoon, S. S., and Lee, B. (2017). Effects of using high-density rain gauge networks and weather radar data on urban hydrological analyses. *Water (Switzerland)*. doi:10.3390/w9120931.

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., and Kitzmillller, D. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*. doi:10.1175/BAMS-D-14-00174.1.

Zhang, J., Qi, Y., Kingsmill, D., and 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*. doi:10.1175/JHM-D-11-0145.1.

Zhang, Z., Koren, V., Reed, S., Smith, M., Zhang, Y., Moreda, F., and Cosgrove, B. (2012). SAC-SMA a priori parameter differences and their impact on distributed hydrologic model simulations. *Journal of Hydrology*. doi:10.1016/j.jhydrol.2011.12.004.

APPENDIX A

Table A1 shows the full rain gauge identifier (ID) for each of the abbreviated IDs used in various graphics.

Table A1. Full and abbreviated rain gauge IDs

Full ID	Abbreviated ID
371840079534900	4900
205551460	1460
371339079554400	4400
371459079560300	0300
371518079591700	1700
371520080015100	5100
371657080002800	2800
371709079580800	0800
371824080002600	2600
KROA	KROA