Impact of radar-rainfall error structure on estimated flood magnitude across scales : an investigation based on a parsimonious distributed hydrological model

Cunha, Luciana K.; Mandapaka, Pradeep V.; Krajewski, Witold F.; Mantilla, Ricardo.; Bradley, Allen A.

2012

Cunha, L. K., Mandapaka, P. V., Krajewski, W. F., Mantilla, R., & Bradley, A. A. (2012). Impact of radar-rainfall error structure on estimated flood magnitude across scales: an investigation based on a parsimonious distributed hydrological model. *Water Resources Research*, 48.

<https://hdl.handle.net/10356/95165>

<https://doi.org/10.1029/2012WR012138>

© 2012 American Geophysical Union. This paper was published in *Water Resources Research* and is made available as an electronic reprint (preprint) with permission of American Geophysical Union. The paper can be found at the following official URL: [<http://dx.doi.org/10.1029/2012WR012138>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Impact of radar-rainfall error structure on estimated flood magnitude across scales: An investigation based on a parsimonious distributed hydrological model

Luciana K. Cunha,^{1,2} Pradeep V. Mandapaka,³ Witold F. Krajewski,¹ Ricardo Mantilla,¹ and Allen A. Bradley¹

Received 15 March 2012; revised 4 August 2012; accepted 8 August 2012; published 9 October 2012.

[1] The goal of this study is to diagnose the manner in which radar-rainfall input affects peak flow simulation uncertainties across scales. We used the distributed physically based hydrological model CUENCAS with parameters that are estimated from available data and without fitting the model output to discharge observations. We evaluated the model's performance using (1) observed streamflow at the outlet of nested basins ranging in scale from 20 to 16,000 km² and (2) streamflow simulated by a well-established and extensively calibrated hydrological model used by the US National Weather Service (SAC-SMA). To mimic radar-rainfall uncertainty, we applied a recently proposed statistical model of radar-rainfall error to produce rainfall ensembles based on different expected error scenarios. We used the generated ensembles as input for the hydrological model and summarized the effects on flow sensitivities using a relative measure of the ensemble peak flow dispersion for every link in the river network. Results show that peak flow simulation uncertainty is strongly dependent on the catchment scale. Uncertainty decreases with increasing catchment drainage area due to the aggregation effect of the river network that filters out small-scale uncertainties. The rate at which uncertainty changes depends on the error structure of the input rainfall fields. We found that random errors that are uncorrelated in space produce high peak flow variability for small scale basins, but uncertainties decrease rapidly as scale increases. In contrast, spatially correlated errors produce less scatter in peak flows for small scales, but uncertainty decreases slowly with increasing catchment size. This study demonstrates the large impact of scale on uncertainty in hydrological simulations and demonstrates the need for a more robust characterization of the uncertainty structure in radar-rainfall. Our results are diagnostic and illustrate the benefits of using the calibration-free, multiscale framework to investigate uncertainty propagation with hydrological models.

Citation: Cunha, L. K., P. V. Mandapaka, W. F. Krajewski, R. Mantilla, and A. A. Bradley (2012), Impact of radar-rainfall error structure on estimated flood magnitude across scales: An investigation based on a parsimonious distributed hydrological model, *Water Resour. Res.*, 48, W10515, doi:10.1029/2012WR012138.

1. Introduction

[2] In this paper we explore the effects that uncertainties in radar-estimated rainfall have on streamflow simulation by a distributed hydrologic model at a range of spatial scales. Rainfall is the main input to many hydrological models, and its uncertainties affect streamflow simulation [e.g.,

Arnaud *et al.*, 2011]. While a number of direct and indirect methods are available to estimate rainfall rate and accumulation, weather radars remain the only operational instruments capable of providing rainfall quantities over large domains with the space-time resolution required for flood prediction across a large range of scales (~ 0.1 –100,000 km²). Radar-rainfall estimates are subject to significant uncertainty (see Villarini and Krajewski [2010] for a review), and their use for flood prediction requires a better understanding of how the estimation errors propagate through hydrological models and affect streamflow prediction across scales.

[3] The main goal of this study is to evaluate the impact that radar-rainfall uncertainties, associated with rainfall maps (products) that were generated by the Next-Generation Radar (NEXRAD) system of WSR-88D weather radars [e.g., Fulton *et al.*, 1998], have on flood simulation. As our goal is to have the capability to predict potential flooding conditions anywhere within a large region (basin), we have been developing a distributed, data-intensive rainfall-runoff model. It is clear that for such a model to have useful

¹IIHR-Hydrosience & Engineering, University of Iowa, Iowa City, Iowa, USA.

²Now at Department of Civil and Environmental Engineering, School of Engineering and Applied Science, Princeton University, Princeton, New Jersey, USA.

³MeteoSwiss, Switzerland.

Corresponding author: L. K. Cunha, Department of Civil and Environmental Engineering, School of Engineering and Applied Science, Princeton University, E422 Engineering Quad, Princeton, NJ 08544, USA. (lcunha@princeton.edu)

predictive skill in many places within the basin, the model's parameterizations (physics-based) should depend on data that characterize details of the domain and the forcing rather than the parameter calibration (e.g., Gupta, 2004). However, since there is no hydrologic theory that could guide the development of such a model, and due to the data's limitations, it is difficult to know which processes should be represented in detail and which could be neglected or simplified. The importance of various processes depends on the location and the hydrologic "circumstances" such as seasons and antecedent conditions. On the other hand, there is plenty of evidence documented in the literature that rainfall variability, resolution, and accuracy play crucial roles in the model's development and performance. Given the above circumstances, we pose the following question: given a hydrologic model characterized by a certain skill level, what is the effect of radar-rainfall uncertainty on the model's prediction ability? We seek to answer this question across a range of spatial scales for continuous real-time simulation. To obtain meaningful answers, we need to avoid calibrating the model to discharge data because such fitting would compensate for the uncertainty due to the rainfall input and thereby compromise our objectives.

[4] Our methodology can be viewed as a data-based simulation. It consists of generating ensembles of "equally probable" rainfall fields that have the same error structure as the NEXRAD products. These fields are then propagated through a rainfall-runoff hydrological model that simulates streamflow for each stream link in a watershed. This approach requires two main components: (1) a rainfall ensemble generator that provides maps of rainfall input that mimic the actual radar-rainfall uncertainty and (2) a parsimonious hydrological model whose parameters can be prescribed a priori using physical properties of the watershed, thereby eliminating the need for parameter calibration. The model includes nonlinearities in the system, and its simulated discharge agrees reasonably well with observations. We discuss the main features of these two components below.

[5] Uncertainties in radar rainfall estimates have been studied for more than 30 years [Grayman and Eagleson, 1971; Wilson and Brandes, 1979; Chuckie and Collier, 1991; Ciach and Krajewski, 1999a, 1999b; Seed et al., 1999; Pegram and Clothier, 2001; Jordan et al., 2003; Chumchean et al., 2006; Ciach et al., 2007; Habib et al., 2008; Krajewski et al., 2010; AghaKouchak et al., 2010a, 2010b], and several models have been proposed for the statistical description of radar-rainfall errors (see reviews by Villarini and Krajewski, 2009; Mandapaka and Germann, 2010). Early methods of simulating synthetic radar-rainfall fields, i.e., rainfall fields that are corrupted by radar-like systematic and random errors, were based on a conceptual understanding of the uncertainties involved [Krajewski and Georgakakos, 1985; Krajewski, 1993; Anagnostou and Krajewski, 1997; Carpenter and Georgakakos, 2004]. While these methods attempted to capture the main aspects of the factors that caused uncertainty, they lacked an empirically based quantification of the deviations between the true and radar-estimated rainfall. Therefore, there was no guarantee that actual radar-rainfall uncertainty was realistically represented in studies that used these methods [e.g., Sharif et al., 2002; Sharif et al., 2004]. To overcome this weakness, recent developments focus on the empirically based modeling of radar-rainfall uncertainty

[e.g., Ciach et al., 2007; Germann et al., 2009]. In this study we adopted the radar-rainfall generator proposed by Villarini et al. [2009], which is based on the empirical radar rainfall error model by Ciach et al. [2007]. This model characterizes the statistical structure of radar-rainfall errors with three components: (1) an overall multiplicative bias factor estimated using long-term accumulated radar-rainfall and gauge-rainfall values; (2) a systematic distortion function, conditioned on the radar-estimated rainfall value; and (3) a stochastic factor quantifying residual random errors. The model accounts for range dependency and for spatial and temporal correlation in errors. The generator uses a conditional simulation framework: Given the estimates provided by a specific radar-rainfall estimation algorithm, the method returns rainfall fields that have the same error structure as that observed empirically for that algorithm. In our case, the algorithm is the precipitation processing system [Fulton et al., 1998] that converts the reflectivity data from the WSR-88D weather radars to the hourly accumulation with approximately 4 km by 4 km spatial resolution (see Ciach et al. [2007] and Villarini et al. [2009] for details). The fact that the generator is conditional on actual radar-rainfall fields ensures relevance to practical applications and allows analysis of real events.

[6] Once realistic radar-rainfall ensembles are generated, they are used as input to a hydrological model that simulates streamflow. Each member of the ensemble will result in a somewhat different hydrograph. Does the spread of different characteristics of the resultant hydrographs (such as the peak value or the time-to-peak) depend on the scale of the basin? How does the spread compare to the difference between the model output obtained using the "observed" field (i.e., the one on which the conditional simulation was based) and the streamflow observations? We address these questions herein.

[7] Previous studies have demonstrated, albeit indirectly, that a fair investigation of how rainfall errors affect flood simulation requires a calibration-free hydrological model since calibration camouflages uncertainties related to both the hydrologic model structure and the rainfall input data uncertainty. Carpenter and Georgakakos [2004] investigated the impacts of rainfall input and rainfall-runoff model parametric uncertainty on flow simulation using a calibrated distributed hydrological model. Their results showed that errors due to model parameter estimation are of the same order of magnitude or even larger than the errors due to rainfall uncertainties. Schröter et al. [2011] used a probabilistic model to generate an ensemble of precipitation fields that were then used as input to a hydrological model. The authors calibrated the model based on the different rainfall ensembles and demonstrated that rainfall uncertainties might have a significant impact on the estimated parameter estimates. He et al. [2011] evaluated the impact of radar rainfall uncertainties on water resource modeling using a model that was calibrated based on rain gauge data. According to these authors, if radar precipitation were used to calibrate the model, sensitivity of simulated stream discharge to rainfall input would have changed. Fu et al. [2011] investigated the impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resource model and concluded that, "the effects of precipitation input are so dominant that it could potentially impact the estimates of model parameters when the hydrological model is calibrated." The study further concluded

that model parameters estimated by calibration would be biased in order to compensate for errors introduced by low-resolution precipitation. These studies demonstrate the need for a simulation framework that isolates uncertainties due to parameter estimation, model structure, and various inputs.

[8] For the above data-based simulation framework to provide meaningful results (i.e., with respect to the goal of developing model with predictive skill across scales and at multiple locations), it is necessary to have a hydrologic model with a certain level of skill that results merely from its structure and ability to mimic the dominant processes of hydrologic response to rainfall. The numerical values of the required coefficients (parameters) should be “observable” from the characteristics of the basin, including topography, land cover, land use, soils, etc. The level of skill should be such that given a more accurate input, the model should provide better output. This is not always the case with calibrated models that adjust their parameter values to compensate for uncertainties in the input.

[9] In this study we use a fully distributed, physically based, and calibration-free hydrological model that allows the isolation of errors due to the rainfall input. A large variety of distributed hydrological models exist whose parameters can be derived from physical data. Twelve such models were included in the Distributed Model Intercomparison Project (DMIP) that evaluated the capabilities of existing distributed hydrologic models forced with operational radar-based precipitation [Reed *et al.*, 2004]. In the DMIP project, several research groups compared the results from calibrated and uncalibrated versions of the models and concluded that calibration significantly improves simulation results. In our study, we opted to use the term *calibration-free* since we recognize that calibration is not feasible due to the *dynamic parametric complexity* problem [Gupta, 2004] that arises when highly spatially explicit terrain decomposition is applied. With explicit consideration of spatial variability, the number of parameters increases exponentially. A large number of *error-free* observations that covers all spectra of hydrological functions and spatio-temporal scales would be required to properly estimate these parameters through calibration [Sawicz *et al.*, 2011]. When streamflow at the outlet of the basin is the only information used to calibrate these models, the large number of degrees of freedom results in equifinality, i.e., a limited ability to uniquely identify parameters [e.g., Ebel and Loague, 2006]. In that case, we can never be sure if the model is “right for the right reasons” or if errors in the data and model structure are being compensated for by errors in parameter values [e.g., Refsgaard, 2004; Ajami *et al.*, 2004].

[10] Hossain *et al.* [2004] and Heistermann and Kneis [2011] used Monte Carlo simulation to evaluate the effect of rainfall on hydrological simulations without the need to define optimal parameter sets based on calibration. Both studies demonstrated the viability of the approach for small to medium sized basins; Heistermann and Kneis [2011] identified limitations related to the model’s computation time even for the small catchment adopted in the study (15 km²). Therefore, a similar approach is not computationally feasible for large basins using fine decomposition of the terrain (at the hillslope level). In this study we avoid calibration by adopting a different modeling approach, and instead of calibrating parameters to achieve better results,

we systematically improved the model’s structure. During the iterative process of adding new components and evaluating their impact on the model’s overall performance, we realized the importance of understanding how much of the discrepancy between the model’s simulated and observed responses could be due to uncertainty in the rainfall. This is, in fact, the main motivation behind our study. We have been developing a calibration-free model to enable *prediction everywhere*, in the spirit of the PUB (prediction in ungauged basins) initiative [Sivapalan *et al.*, 2003]. While a full description of our developments is beyond the scope of this paper, we illustrate the expected streamflow prediction uncertainties due to the errors in the radar-provided input across a range of spatial scales.

[11] The model provides an impartial evaluation of how rainfall errors propagate through hydrological models and affect flood prediction across a large range of scales. Calibration is avoided through the use of parameters that are directly or indirectly linked to the physical properties of the watershed (e.g., soil water storage capacity, hillslope shape, and channel flow velocity). Direct methods use spatial information of physical properties (e.g., DEM), while indirect methods adopt a combination of empirical equations (e.g., Manning’s formula) and spatial maps of basin properties. The fine decomposition of the terrain into hillslopes and links results in a realistic representation of the stream and river drainage network, which allows us to apply model equations that represent processes close to the scales as they occur in nature [Mantilla and Gupta, 2005]. Runoff is generated at hillslopes, and water is transported via the drainage network of connected links. Model equations at the hillslope-link scale are based on mass and momentum conservation principles.

[12] Previous studies have also demonstrated that flow simulation uncertainty is strongly dependent on catchment scale, with uncertainty decreasing as basin scale increases [e.g., Carpenter, 2006; Nikolopoulos and Anagnostou, 2010]. This is expected since the river network filters out small-scale variability and uncertainties [e.g., Mandapaka *et al.*, 2009a]. In previous studies [e.g., Carpenter and Georgakakos, 2004; He *et al.*, 2011] this conclusion was reached based on the analysis of midsize basins, including a limited number of points in the watershed for which streamflow observations were available and used to calibrate the model. In order to characterize the scales for which radar data provide good information for flood prediction, a more comprehensive study involving a large number of sites covering a large range of scales (from few to thousands of square kilometers) is required. The basin decomposition method used in this study provides flow simulation for every link in the basin, which allows us to investigate error scale dependency using a large number of locations in the basin (more than 70,000 hillslope links) covering a wide range of spatial scales, from hillslope (~0.1 km²) to large watersheds (~17,000 km²).

[13] The study area and data sources are described in section 2. In section 3 we present methodological aspects of this work, focusing on the rainfall generator and the hydrological model. Section 4 includes the main results of this study, including: (1) the validation of the hydrological model across multiple scales, (2) the rainfall error scenarios and statistics for the generated rainfall ensembles, and (3) flow sensitivity to different rainfall error scenarios. Section 5 presents the main conclusions of this work.

2. Study Area and Data Source

[14] The study is carried out for two adjacent basins, the Iowa and the Cedar River basins, located almost entirely in the state of Iowa. The total drainage area is 7234 km² for the Iowa River (in Marengo) and 16,853 km² for the Cedar River (in Cedar Rapids). The climate in this region is characterized by cold winters, hot summers, and wet springs, with a mean annual precipitation of nearly 900 mm (source: Oregon Climate Service), potential evapotranspiration (PE) of 1060 mm, and actual evaporation of 580 mm (based on MOD16 product [see *Mu et al.*, 2011]). The dominant land use is agriculture, consisting mainly of corn-soybean rotations. The agricultural practice imposes a strong seasonality in land cover dynamics. The growing season usually starts in May and ends in November.

[15] We chose this study area for two main reasons. First, the region is rich in hydrologic information. Four NEXRAD weather radars (in Des Moines and Davenport in Iowa, La Cross, Wisconsin, and Minneapolis, Minnesota) cover the two basins; also 24 USGS streamflow gauges collect data at the outlet of drainage areas ranging from approximately 22 to 16,853 km². Figure 1 presents a map of the study area with the location of the weather radars and the USGS streamflow sites. The second reason for selecting the area is its frequent flooding. The region was the center of an extreme flood event in 2008 that affected approximately 4800 km² of Iowa's agricultural land and

many communities throughout the state. About 1300 blocks in the city of Cedar Rapids were flooded, which affected more than 5000 homes and 900 businesses. In Iowa City, 16 buildings on The University of Iowa campus were reported flooded and the total cost of recovery from the flood reached \$750M. We will use the 2008 flood event to demonstrate the hydrological model's ability to simulate floods across scales and to investigate the effect of rainfall uncertainty on the prediction of floods.

[16] We used numerous sources of data in this study. Table 1 summarizes the data sets used as model input, for parameter estimation, and for model validation. Model input includes rainfall and potential evapotranspiration (PET). For rainfall we used the stage IV rainfall product, which represents hourly accumulation on approximately 4 km by 4 km grids. Stage IV products are provided by the National Weather Service. *Fulton et al.* [1998] describe the methods and data used to produce the different multisensor precipitation products provided by the NWS. Stage IV is a postprocessed product based on the merging of radar and rain gauge data, in particular to remove the mean-field bias in the radar-only estimates. Hereafter, we refer to these overall bias corrected fields as reference rainfall. We used NEXRAD's stage IV products as a reference in this study to compare how different rainfall error sources and scenarios affect hydrological prediction across scales.

[17] For PET we use the product produced by the North American Regional Reanalysis NLDAS-2 with 1 h resolution in time and 0.125° (~13 km) resolution in space. PET is used to estimate the actual evaporation from the surface and from the unsaturated and saturated layers of the soil. *Dingman* [2002] defines PET as "the rate at which evapotranspiration would occur from a large area completely and uniformly covered with growing vegetation which has access to an unlimited supply of soil water and without advection or heat-storage effect." Using PET to calculate actual evaporation presumes that the only factor limiting evapotranspiration is water availability. No limitations are imposed on the evaporation of water from the surface storage.

[18] We used the National Elevation Data set (NED) with 90 m resolution to extract the stream and river network. Our procedure [*Mantilla and Gupta*, 2005] resulted in 78,503 hillslopes for the Cedar River basin with an average hillslope area of 0.20 km² and 44,527 hillslopes for the Iowa River basin with an average hillslope area of 0.16 km². We used the USGS hydraulic measurements to estimate the parameters of hydraulic geometry equations based on the formulation proposed by *Dodov* [2004], *Paik and Kumar* [2004], and *Mantilla* [2007]. To estimate soil hydraulic properties and water storage availability, we used the Soil Survey Geographic (SSURGO) Database. To estimate the roughness coefficient (Manning's) for the overland water transport equations, we used the high-resolution national land cover data set [*Homer et al.*, 2007].

[19] Finally, we used two data sets to evaluate the model's performance. First, we used the USGS streamflow data for a total of 24 sites to compare observed and simulated discharge. It is important to point out that we did not use this information to calibrate the model's parameters; therefore, the comparison constitutes an independent evaluation of the model's performance. We also compared the streamflow simulations from

Figure 1. Study area with radar coverage and streamflow sites.

Table 1. Data Set Used for the Simulations (Model Input, Estimation of Model Parameters, Setting Soil Initial Condition, and Model Validation)

Physical Processes and Properties	Data Sets ^a
Rainfall	Radar Stage IV, NWS ($\Delta T = 1$ h and $\Delta S = 0.05^\circ$, almost real time)
Potential Evapotranspiration	PE from North American Regional Reanalysis, NLDAS-2 Model Input ($\Delta T = 1$ h; $\Delta S = 0.125^\circ$, ~ 13 km)
DEM	National Elevation Data Set, $\Delta S = 90$ m
Hydraulic Measurements	USGS Hydraulic Measurements
Soil Properties	SSURGO, $\Delta S =$ polygons with 1 to 10 acres (~ 0.05 km ²) size map delineation
Land Cover	National Land Cover Database (NLCD), $\Delta S = 30$ m, 2001
Soil Moisture (Initial Condition)	Total soil column wetness (0–200 cm), NLDAS-2 Model Output ($\Delta T = 1$ h; $\Delta S = 0.125^\circ$, ~ 13 km)
Streamflow	USGS streamflow measurements for a total of 24 sites in the study area ($\Delta T = 1/4$ or 1 h)
	Simulated Streamflow, SAC-SMA Model

^a ΔS = spatial resolution in minutes (′), arc-seconds (″), degrees (°), or km. ΔT = temporal resolution in minutes, hours, days, or years.

CUENCAS with those from the Sacramento Soil Moisture Accounting (SAC-SMA) model [Burnash, 1995], which is used by the National Weather Service as the main component of their flood forecasting system [Welles *et al.*, 2007].

3. Methodology

[20] To demonstrate our model, we selected a month-long period with rainfall and streamflow data immediately preceding the 2008 flood in Iowa. Our flood prediction model is a physically based rainfall-runoff model that simulates response to rainfall forcing at a wide range of scales, with the smallest being at hillslope scale. The geometry of the hillslope elements of the model is irregular, but each hillslope is connected to a channel link whose location is determined by analysis of the high-resolution topography data, as described above. The model uses the information on land cover and land use, soil types, and topography based on readily available data that are mapped to the scale of the hillslopes. This mapping is one of the possible sources of uncertainty in the model, and investigating these uncertainties will be the focus of future studies. The function of the hillslopes is to partition the rainfall input into surface runoff, infiltration, and evapotranspiration. This is accomplished by using empirically based parameterizations of the relevant processes documented in the literature. We do not perform local adjustments (calibration) of the coefficients to force better agreement with the observed streamflow. Surface runoff and subsurface quick flow feed the channels, and the water is transported down the drainage network. The discharge is aggregated as the water flows to higher order streams. The aggregated discharge is attenuated by the selected water velocity model. The model has a power law functional form with the velocity dependent on the magnitude of the discharge and the upstream drainage area. We estimated the coefficients of the velocity model by multivariate regression using the USGS collected and published discharge and water velocity data [Hirsch and Costa, 2004]. We did this independently of our rainfall-runoff model.

[21] The fact that the model's parameters are not calibrated to fit the discharge data implies that we have avoided favoring the scales at which we have available discharge observations. Our model is, in principle, suitable for

prediction at any scale provided that the velocity model is able to correctly simulate transport through the river network. For more information on the validity of the velocity model applied in this study, please refer to Paik and Kumar [2004], Mantilla [2007], and Mandapaka *et al.* [2009a].

[22] We evaluated the model performance prior to performing our simulation experiment. Clearly if the model does not provide accurate results, it should not be used to study the effects of uncertainty in the input data. On the other hand, if the model's parameters are fitted to a particular input (rainfall) product, it would, by construction, bias the performance of a competing input product. For a calibrated model it would be possible to perform better when forced by an inferior input product. Moreno *et al.* [2012] recently presented a study in which calibration, performed based on a specific data set, biases model performance when forced by a different input. The author calibrated the model based on the NWS stage II radar rainfall data set and concluded that the model forced by rain gauge data tends to overestimate simulated streamflow. This is expected since these authors also pointed out that stage II products underestimated rainfall compared to the rain gauge (bias equal to 0.9).

[23] Following our methodology, we forced our model with the radar-rainfall input generated by a recently developed error model of the NEXRAD-based hourly rainfall maps [Ciach *et al.*, 2007; Villarini and Krajewski, 2009]. The model is based on a large sample of empirical data and has a flexible structure, where the total error in radar-rainfall estimate is separated into systematic and random errors. Therefore, the model is suitable for studying the impact of systematic and random errors separately. Using the error model, we generated multiple realizations (equally probable) of radar-rainfall fields that were conditional on the observed (reference) field. Each of the generated realizations has the same statistical error structure as the reference field.

[24] We applied the generated ensemble of the radar-rainfall input to the hydrologic rainfall-runoff model and thus simulated multiple realizations of the discharge. We then calculated several performance measures to evaluate the spread of the obtained discharge values and compared it with the discrepancy (errors) between the discharge obtained using the reference input and the observed discharge. We calculated these measures at a range of spatial

scales from very small to fairly large ($\sim 20,000 \text{ km}^2$). We then repeated the above ensemble simulation for a different scenario of the radar-rainfall model.

[25] In the following sections we will focus on the description of the two main components of this work: (1) the distributed physically based hydrological model and (2) the radar-rainfall error model and the ensemble generator used to produce equally probable rainfall fields based on the reference rainfall (stage IV) and the radar-rainfall error structure scenarios (SCs).

3.1. Hydrological Model

[26] We adopted a hillslope-linked based model that compartmentalizes the landscape into small areas where runoff generation occurs (hillslopes) and that are connected by the river network (links). Hillslope-link models provide a more accurate representation of the river basins and have lower computational requirements than grid-based models [e.g., Yang *et al.*, 2002]. The river network plays an important role in flood generation by delineating how peak flow changes across scales [e.g., Gupta *et al.*, 1996; Gupta, 2004; Menabde *et al.*, 2001]. Therefore, a correct representation of the river network is essential for simulating floods across multiple scales.

[27] We use a set of four coupled nonlinear ordinary differential equations that describe changes in mass balance in each control volume to represent hydrological and hydraulic processes at hillslopes and links. The four equations account for water balance in the (1) link; (2) surface hillslope storage; (3) saturated soil zone, and (4) unsaturated soil layers. All state variables are solved simultaneously using a time-adaptive numerical method to avoid numerical inconsistencies [Kavetski, 2011]. The solutions to these equations provide hydrographs at each channel junction in the network and continuously account for moisture states in the surface, saturated and unsaturated soil layers, and channels over the entire basin domain.

[28] We reduce complexities on the hillslope dynamics by parameterizing known macroscale hydrological behavior [McGlynn *et al.*, 2002; Graham and McDonnell, 2010]. Since we estimate parameters based on observations of relevant processes, complexities in the model's formulation are limited by data availability. We acknowledge that some important processes are not represented in the model (e.g., soil macropores). For the reasons previously mentioned, we opted for a simplified model formulation that does not require calibration instead of using a complex conceptualization that suffers from equifinality. In this section we provide a general description of the model. A more detailed description including all the equations is provided by Cunha [2012].

[29] Figure 2a presents a schematic representation of all the fluxes in the model. Weather radar rainfall is the main input to the model. We remap the radar rainfall grid to the hillslope-link structure adopted in the model. Once rainfall reaches the surface, it is transformed into surface ponding. Ponding water infiltrates, evaporates, or runs to the channel. The model accounts for infiltration excess (Hortonian), overland flow (HOF), and saturation excess overland flow (SEOF). HOF occurs when ponding water exceeds the infiltration capacity at the areas of the basin where soil is not saturated (the permeable area, represented by green in Figure 2). In this region, the percentage of surface water that

infiltrates depends on the deficit of water in the soil (soil volumetric moisture) and on the soil infiltration capacity. Soil properties are provided by SSURGO. SEOF occurs in the areas of the basin where soil is saturated (impermeable areas, represented by gray in Figure 2 [Freeze, 1980]). As in nature, HOF and SEOF might occur simultaneously in the same hillslope. The transport of the ponding water is simulated using Manning's equations with roughness parameters estimated based on land cover data.

[30] We adopt a simplified conceptualization of the hillslope subsurface geometry to estimate impermeable area as a function of hillslope shape and water volume in the saturated soil layer. In Figure 2 we present a schematic representation of this geometry for convex (b-1) and concave (b-2) hillslopes. Hillslope shape is an important factor affecting rainfall-runoff partitioning. Convex hillslopes exhibit a higher transport power (due to higher slope) and a lower saturation capacity than concave hillslopes [Sabzevari *et al.*, 2010] (see Figures 2(b-1) and 2(b-2)). We assume that the bedrock surface is parallel to the hillslope surface, so the water table is calculated as a function of the volume of water in the saturated soil layer. We use the topography data (DEM) to estimate a relationship between the water table and the percentage of impermeable area (for an example, see plots in Figures 2(b-1) and 2(b-2)).

[31] Once the water enters into the soil, it percolates down to recharge the groundwater reservoir. We use the extension of Darcy's law, proposed by Buckingham [1907], to simulate flow in the partially saturated layer of the soil. The author postulated that Darcy's law is also valid for soils that are partially saturated and, in this case (unsaturated), hydraulic conductivity is a function of water content. The rate of change of unsaturated hydraulic conductivity with soil water content is a function of soil properties [Davidson *et al.*, 1969]. Water in the saturated layer of the soil flows to the channel as base flow. We estimated unsaturated hydraulic conductivity based on soil water content and soil properties [Davidson *et al.*, 1969]. The flux from the saturated layer to the channel is also simulated by Darcy's equation, using saturated hydraulic conductivity and terrain slope.

[32] Part of the water is removed from the surface, and from the unsaturated and saturated soil layers as evapotranspiration. Evapotranspiration strongly affects soil moisture content, especially between rainfall events in the summer and spring seasons, and consequently plays an important role in runoff generation. We use potential evapotranspiration (PE) to estimate the actual evaporation from the surface and soil. Dingman [2002] defines PE as "the rate at which evapotranspiration would occur from a large area completely and uniformly covered with growing vegetation which has access to an unlimited supply of soil water and without advection or heat-storage effect." Using PE to calculate actual evaporation corresponds to assuming that the only limitation for evaporation is water availability. We impose no limitations on the evaporation of water from the surface storage. Evaporation from the soil's unsaturated layers is limited by the soil's volumetric moisture and relative depth of water compared to the total depth of soil. When a large amount of water is available, PE limits maximum evaporation. As input to the model, we use potential evapotranspiration provided by NLDAS-2 [Xia *et al.*, 2012], with hourly resolution in time and $1/8^\circ$ resolution in space.

Figure 2. Schematic representation of the fluxes among (a) control volume and the simplified conceptualization hillslope geometric properties for (b-1) convex and (b-2) concave hillslope. In the figure, the following are properties of the hillslope: H_T is the total hillslope relief, H_{wmax} is the maximum water table, H_b is the depth to the bedrock, S_h is the hillslope average slope, and A_H is the hillslope area. The following are state variables: h_w is the water table level and a_i is the hillslope impermeable area.

[33] We implemented the model's equations as part of the code built into CUENCAS, initially introduced by Mantilla and Gupta [2005] as a research tool to investigate the morphological characteristics of the river networks and their roles in peak flow scaling. The model has been used in previous studies. Mandapaka *et al.* [2009a] applied a simplified version of the model to investigate the effect of rainfall variability on the statistical structure of peak flow. To investigate land cover changes' effects on flood risk, Cunha *et al.* [2011] applied a more complete version of the model that includes surface and subsurface processes. Gupta *et al.* [2010] refer to CUENCAS as an important tool to perform theoretical investigations that would explain the link between peak flow versus area scaling parameters and physical properties of the inputs and the watershed.

[34] As we mentioned earlier, the model is not calibrated to fit the observed discharge. The fact that, in the process of solving the governing equations, we have to track the water transport in all links and hillslopes implies that the model's structure is amenable to *prediction everywhere*. Similarly, there is no imposed temporal scale. In principle, the rainfall rate input could be provided and integrated with any temporal

resolution. At this point, the temporal scale is being defined by the sampling frequency of the observations.

3.2. Radar Rainfall Error (RRE) Model and Ensemble Generator

[35] We employed the RRE generator proposed by Villarini *et al.* [2009] to produce ensembles of equally probable rainfall fields. The method uses an empirically derived error model to generate synthetic probable rainfall fields conditioned on a given rainfall map. The uncertainty model adopted in the generator was proposed by Ciach *et al.* [2007]. In this model, the true pixel-scale average rainfall $R_A(x,y)$ at a location (x,y) within the basin is expressed as the product of a deterministic component and a random component, both conditioned on reference rainfall, $R_R(x,y)$,

$$R_A = h(R_R)\varepsilon(R_R). \quad (1)$$

[36] The deterministic distortion function $h(R_R)$ accounts for the conditional (i.e., on rainfall magnitude) biases, whereas the stochastic factor $\varepsilon(R_R)$ describes the random

deviations from the true but unknown rainfall. Note that both components are a function of the reference rainfall rate (R_R). This feature has important implications that we reveal later. The deterministic distortion function is approximated by the two-parameter power law

$$h(R_R) = a[B_0 \cdot R_R]^b \quad (2)$$

and the random component is approximated by a Gaussian distribution with the mean of 1 and its standard deviation modeled as a rapidly decreasing hyperbolic function of the reference rainfall:

$$\sigma[R_R(x, y)] = c + d[B_0 R_R(x, y)]^e. \quad (3)$$

[37] In equations (2) and (3) B_0 is the overall (time integrated) bias, defined as the ratio of time-integrated gauge estimates to time-integrated radar estimates. The parameters a and b describe the deterministic distortion function, and c , d , and e are the parameters of the random component. The parameters were estimated by *Ciach et al.* [2007] using six-years of level-II hourly rainfall data from the Oklahoma City WSR-88D radar site (KTLX) and rain gauge records from the Oklahoma Mesonet [*Brock et al.*, 1995] and the Agricultural Research Service (ARS) Micro-net. The parameters vary with season and distance from the radar site. In addition, the authors used the radar- and gauge-rainfall data to estimate the spatial and temporal correlation of random errors. *Villarini and Krajewski* [2009] extended this work by proposing three-parameter exponential functions to describe the correlations in the random component as

$$\rho(\Delta s) = c_1 \exp \left[-\frac{\Delta s}{c_2} \right]^{c_3}, \quad (4)$$

where c_1 is the nugget effect that characterizes the small-scale variability of the process and/or measurement errors, c_2 represents the correlation length, c_3 is the shape parameter that controls the shape of the fitted correlation function at the origin, and x is the separation distance. *Villarini et al.* [2009] noted difficulties in generating random error fields that simultaneously preserve spatiotemporal correlations and the dependence of error standard deviation on the rain rates. Consequently, they simplified the generator by neglecting temporal correlations. In this study we also neglected temporal correlations in the random field generator. This is acceptable if you note that the temporal errors decorrelate much faster than the response scale of the basins we study. One could also argue that the spatial correlation has a more direct effect on the scale-dependent model performance. More details about the generator are provided by *Villarini et al.* [2009].

[38] The generation of ensemble radar-rainfall fields can be summarized as follows: (1) the reference rainfall field was corrected for conditional bias using equation (2); (2) an ensemble of spatially correlated Gaussian fields with a unit mean and standard deviation conditional on the reference rainfall field was then generated using the Cholesky decomposition method; and (3) the correlated Gaussian fields were then multiplied with the bias corrected reference field

from step 1 to obtain an ensemble of equally probable rainfall fields. Note that for very weak rain rates, because of the hyperbolic nature of the random error standard deviation (equation (3)), there might be negative values in the random error fields (step 2). Any such pixels with negative values are set to zero to avoid unrealistic rainfall values in step 3. Setting negative values to zero introduces slight bias into the generated error fields. However, this bias is negligible and therefore would not significantly affect the general conclusions of this study.

[39] In this work, we use as “reference” the stage IV rainfall map products provided by the National Weather Service [*Fulton et al.*, 1998]. Since, by construction, the radar-rainfall error model is specific to a radar-rainfall product, this raises the issue of whether the parameter values of the ensemble generator are appropriate for this product. Since the stage IV data set is corrected to match rain gauge accumulations, we considered it to be free of overall bias; therefore, we set B_0 to 1.0. We can also argue that the overall character of the microphysical rainfall processes in Iowa, especially in the summer convective storms, is similar to that in Oklahoma. In support of this argument, we cite a study by *Seo and Krajewski* [2011], which shows a similar behavior of the random error dependence on rainfall magnitude. Another potential issue with the transferability of the *Ciach et al.*, [2007] results is the fact that their model is valid for a single radar, whereas stage IV is a multiradar product. While, in principle, combining data from multiple radars reduces the random radar error, in our case, all radars that cover the Cedar and Iowa River basins do this at a far range, where errors are significant. To mitigate, at least partially, the problem of strict applicability of the radar-rainfall model to our study area, we define and consider 12 different scenarios of RRE and generate 50 ensembles for each one.

4. Results and Discussion

4.1. Radar Rainfall Error Scenarios

[40] Table 2 presents the 12 rainfall error scenarios considered in this study. We conceptualized these scenarios to allow the independent evaluation of the impacts that different aspects of the rainfall error structure, including systematic errors, standard deviation of random errors, and spatial correlation of random errors, have on flood simulation. Since we do not know much about the error properties for the radar-rainfall for eastern Iowa for June 2008, we can get a better idea about the range of sensitivities involved by exploring multiple scenarios. As we mentioned in section 3.1, our rainfall ensemble is obtained by imposing Gaussian random error fields on the bias corrected reference rainfall fields. For each scenario, we used the generator to produce an ensemble of 50 equally probable rainfall fields.

[41] *Ciach et al.* [2007] and *Villarini et al.* [2009] provided values of the RRE parameters for Oklahoma that were empirically estimated based on a dense rain gauge network over a six year period. Since rainfall in Oklahoma during summer presents patterns similar to rainfall in Iowa, we use those parameters to represent a realistic scenario of radar rainfall error. The estimation of the RRE model parameters for Iowa would require data from a dense network of rain gauges and is outside of the scope of this study. The

Table 2. Rainfall Error Scenarios

Par.	Deterministic Distortion and Bias			Bias								
	SC 1	SC 2	SC 3	SC 4	SC 5	SC 6	SC 7	SC 8	SC 9	SC 10	SC 11	SC 12
<i>Deterministic Bias</i>												
<i>a</i>	1.34	1	1	1	1	1	1	1	1	1	1	1
<i>b</i>	0.84	0.85	1.15	1	1	1	1	1	1	1	1	1
<i>Random Error, SD</i>												
<i>c</i>	0.45	0.45	0.45	0.45	0.45	0.2	0.45	0.45	0.45	0.45	0.45	0.45
<i>d</i>	0.59	0.59	0.59	0.59	0	0	0.59	0.59	0.59	0.59	0.59	0.59
<i>e</i>	−0.62	−0.62	−0.62	−0.62	0	0	−0.62	−0.62	−0.62	−0.62	−0.62	−0.62
<i>Random Error, Spatial Correlation</i>												
<i>C</i> ₁	1	1	1	1	1	1	0	1	1	1	1	1
<i>C</i> ₂	37	37	37	37	37	37	0	15	37	60	100	100
<i>C</i> ₃	0.39	0.39	0.39	0.39	0.39	0.39	0	1	1	1	1	0.39

Oklahoma parameters provide a realistic description of the expected error in radar rainfall data obtained with the NEXRAD system and can be applied for the scope of this study. The radar rainfall error parameters for Oklahoma are presented in SC 1, Table 2. To obtain the remaining scenarios, we systematically altered the Oklahoma parameter values to obtain RRE scenarios that isolated the effect of different components of the RRE model. Note that when we altered a particular component in the model, the parameter values for other components remain fixed to the Oklahoma values.

[42] In section 3.1 we have seen that the systematic errors are characterized by the deterministic distortion function, which is a power law with a coefficient *a* and an exponent *b* (equation (3)). The effect of coefficient *a* on simulated hydrographs is trivial because it is just a multiplicative factor. Therefore, we focused on the exponent *b* and altered its values such that we have one scenario with systematic overestimation (SC 2; *b* = 0.85) and another scenario with underestimation by radars (SC 3, *b* = 1.15). We emphasize that in all three scenarios (SC 1–3), the systematic errors are nonlinearly related to the radar-rainfall estimate, as demonstrated in Figure 3a.

[43] For the remaining scenarios we isolate the effects of random errors by setting parameters *a* and *b* to 1 and varying the parameters that control the standard deviation and spatial correlation structure of the random errors. For SC 4 we adopted the empirically estimated parameters for random error, and for SC 5 and 6 we removed the dependence of standard deviation on radar-rainfall estimates by setting parameters *d* and *e* to zero. While for SC 5 the value of parameter *c* was fixed to the Oklahoma value of 0.45, for SC 6 it was changed to 0.2. Therefore, SC 4, 5, and 6 together quantify the effect of standard deviation of random errors both when it is conditional and when it is not conditional on the radar-rainfall estimate. We then gradually increased the spatial correlation distance (*C*₂) of random errors from 0 to 100 km to obtain SC 7–11. Scenarios 7–11 therefore quantify the effect of spatial coherence of random errors on the simulated peak hydrographs. For SC 12 we changed the shape parameter (*C*₃) that controls the rate of decrease of the spatial correlation with distance. Smaller values of *C*₃ result in a faster decrease in the correlation close to the origin. Therefore, the effect of small-scale spatial structure

of random errors can be quantified by comparing simulations from SC 11 and 12. Figure 3b presents the standard deviation of the random error, and Figure 3c presents the spatial correlation of the random error adopted for the different scenarios.

[44] Figure 4 shows the time series of the accumulated mean areal precipitation (AMAP) for all ensemble members and scenarios for the 2008 event, along with the reference AMAP (black line). Different uncertainty scenarios translate into a different spread of the ensemble AMAP and different biases between the median of the ensemble and the reference value. Note that in Figure 4 we are looking at the spread in the time series of (space-time) accumulated precipitation fields (or AMAPs). Therefore, the deterministic distortion function also affects the spread of AMAPs. This is why SC 3 appears to have larger spread than SC 2 even though both scenarios have the same random error structure. To reduce the effect of the deterministic distortion function on the spread of the ensemble, we calculated statistics for the time series of spatial averages (without temporal accumulation). In the bottom right of each plot, in blue font, we present a measure of the dispersion (ΔP) and a measure of the bias (P^*) calculated based on the time series of rainfall:

$$\Delta P = \frac{1}{nt} \sum_{i=1}^{nt} \frac{P_{95}(t) - P_5(t)}{P_{50}(t)} \quad \text{for all } t \text{ for which } P_{50}(t) > 0, \quad (5)$$

$$P^* = \frac{\sum_{i=1}^t P_{50}(t) - P^{\text{ref}}(t)}{\sum_{i=1}^t P^{\text{ref}}(t)}, \quad (6)$$

where $P_5(t)$, $P_{50}(t)$, and $P_{95}(t)$ are the 5th, 50th, and 95th percentiles of the ensemble of AMAP over Cedar Rapids, and $P^{\text{ref}}(t)$ is the accumulated mean area precipitation for the reference data set. Note that from here on we use the term *dispersion* as a synonym for the spread of the ensemble. In equation (5) we first normalize the time series of dispersion by the median value to avoid the effect of bias and then we calculate the average value for the dispersion.

Figure 3. For all the considered scenarios of RRE: (a) Deterministic distortion as a function of rain rate, (b) standard deviation of the random error as a function of rain rate, and (c) spatial correlation of the random error as a function of the distance in km.

In this case, we just include values for which $P_{50}(t)$ are larger than zero. In equation (6) we focus on quantifying bias so we estimated the accumulated bias normalized by the reference accumulated rainfall values.

[45] Scenarios 1–4 have the same error structure (Table 2), but the dispersion statistic differs slightly (from 1.43 to 1.48). This slight difference is mainly due to the stochastic nature of the error fields. In general, equation (5) effectively captures the variation in spread across different scenarios. For instance, recall that the standard deviation of errors in SC 5 is independent of the radar-rainfall estimate. This results in a decrease in the spread of the ensemble, which is clearly captured by the dispersion statistic (it dropped from 1.44 for SC 4 to 1.21 for SC 5). When the standard deviation was further decreased to 0.2 in SC 5, ΔP decreased from 1.21 to 0.42. When the random errors were uncorrelated (SC 7), the dispersion dropped from 1.44 to 0.32. Scenarios 7 to 12 evaluate the effect of the spatial correlation of the random error. From Scenarios 7 to 11 we increase the correlation length from 0 to 100 km and observe a systematic increase in the dispersion of the ensemble (increased from 0.32 to 1.80). From scenario 11 to 12, we modified the rate of decrease of the correlation with distance at the origin. Since correlation decreases faster for scenario 12, the observed dispersion of the ensemble is lower for this scenario.

[46] The effect of the deterministic distortion in SC 1–3 can be clearly seen in the bias statistic P^* shown in Figure 4. For the scenarios with no deterministic distortion (SC 4–12), we would expect the median of the AMAP of the ensemble to be equal to that of reference rainfall ($P^* = 0$). However, there is slight overestimation (P^* varies from 0.02 to 0.11), which is due to: (1) the effect of the non-negative rainfall condition imposed on the random error fields; (2) the interplay between the spatial correlation scale of the random error and the basin scale and shape; and (3) the skewness of the rainfall distribution combined with the nonlinear character of the error model since, for some of the scenarios, random error is a nonlinear function of the rain rate.

[47] In Figure 4 we also included dispersion (ΔP_p) and bias (P_p^*) statistics for the ensemble maximum areal average rainfall intensity (AARI) for Cedar Rapids, calculated based on equations (7) and (8), respectively:

$$\Delta P_p = \frac{P_{p95} - P_{p5}}{P_{p50}}, \quad (7)$$

$$P_p^* = \frac{P_{p50} - P_p^{\text{ref}}}{P_p^{\text{ref}}}. \quad (8)$$

[48] For all the scenarios, the spread in the AARI is larger than the one observed for AMAP. The same statistics

Figure 4. Accumulated mean areal precipitation for the 12 rainfall error scenarios for the Cedar River basin. The values indicated in brackets in the bottom right of each plot are the ΔP_A and P_A^* (see equations (5) and (6)).

will be used to quantify dispersion and bias for simulated peak flow in section 4.3.2, where we evaluate how uncertainty in rainfall propagates through the hydrological model and affects simulated peak flow.

[49] While it is probable that none of the scenarios we have considered (except SC 1, whose parameters were empirically estimated) represents real radar-rainfall error structure, collectively they allow us to develop a better understanding of the radar-rainfall uncertainty on basin-scale rainfall. The results demonstrate that the rainfall generator employed in this work is a useful tool to develop rainfall uncertainty scenarios for hydrologic error propagation studies. The generated rainfall fields will allow us to map rainfall uncertainty to the response of the hydrological system.

4.2. Validation of Hydrological Simulations

[50] Before we used our hydrologic rainfall-runoff model to study the effects of uncertainty propagation of the rainfall input, we needed to establish that the model accurately represents the hydrologic response to heavy rainfall. We did this by evaluating the model's performance using observed streamflow data for 24 sites in Iowa and by comparing the simulated streamflow with that generated by a semidistributed version of the SAC-SMA. Since 1960, the Sacramento soil moisture accounting model [Burnash, 1995] has been used by most of the NWS River Forecast Centers as the main flood prediction model [Welles et al., 2007]. The model is

classified as semidistributed, deterministic, continuous, and nonlinear. SAC-SMA contains parameters that describe the rainfall-runoff and evaporation dynamics as well as parameters that describe channel flow transport between two subbasins [Ajami et al., 2004]. In the semidistributed version, the watershed is decomposed in multiple subbasins, and one set of parameters is calibrated for each subbasin using 6 h mean areal precipitation (MAP) products produced by the National Weather Service [Johnson et al., 1999]. The model uses 14 parameters to describe the rainfall-runoff processes and two parameters to describe channel routing, if the kinematic wave method is used. The results presented in this study were generated based on the parameters and SAC-SMA model version currently in use by the North Central River Forecast Center of the NWS for flood prediction in the study area.

[51] In Table 3 we present goodness-of-fit statistics for both models. We begin with the mean streamflow values based on observations, CUENCAS simulations, and SAC-SMA simulations. We then present the correlation coefficient and the Nash-Sutcliffe coefficient for both models [Nash and Sutcliffe, 1970]. We present these statistics since they are commonly used to evaluate the results of hydrological models [Krause et al., 2005]. However, they do not provide a comprehensive view of how well model simulations represent reality. Based on these evaluation statistics, we conclude that both models represent a very similar level of performance.

Table 3. Statistics for CUENCAS and the Sacramento Model

Sites Information					Model Evaluation						
					Mean			Correlation		Nash-Sutcliffe	
Basin	ID Map	USGS ID	Site Name	Area	Obs.	CUENCAS	SAC	CUENCAS	SAC	CUENCAS	SAC
CR	1	5457000	Cedar River near Austin, MN	1033	45.78	46.08	41.02	0.92	0.93	0.85	0.83
CR	2	5457700	Cedar River at Charles City, IA	2,728.7	169.31	196.74	147.63	0.80	0.98	0.56	0.95
CR	3	5458000	Little Cedar River near Ionia, IA	792.2	65.44	68.09	54.84	0.90	0.85	0.79	0.72
CR	4	5458300	Cedar River at Waverly, IA	4005	283.95	298.34		0.92		0.85	
CR	5	5458500	Cedar River at Janesville, IA	4,300.1	319.31	319.38	245.64	0.94	0.91	0.89	0.78
CR	6	5459500	Winnebago River at Mason City, IA	1,361.8	79.60	91.50	56.18	0.82	0.76	0.0	0.47
CR	7	5462000	Shell Rock River at Shell Rock, IA	4,520.2	274.08	319.98	224.46	0.76	0.96	0.4	0.9
CR	8	5458900	West Fork Cedar River at Finchford, IA	2,190.2	164.74	153.40	151.75	0.62	0.69	−0.10	−0.33
CR	9	5463000	Beaver Creek at New Hartford, IA	898.3	74.98	69.31	53.06	0.77	0.4	0.56	−0.74
CR	10	5464220	Wolf Creek near Dysart, IA	774.1	67.20	52.51		0.80		0.58	
CR	11	5463500	Black Hawk Creek at Hudson, IA	784.4	64.01	47.86	45.01	0.80	0.93	0.58	0.76
CR	12	5464000	Cedar River at Waterloo, IA	13,322.4	883.17	933.55	754.15	0.94	0.98	0.85	0.91
CR	13	5464500	Cedar River at Cedar Rapids, IA	16,853.6	1331.7	1178.3	980.4	0.93	0.98	0.83	0.81
IC	1	5451210	S. Fork Iowa River NE of N. Prov, IA	579.9	51.57	40.82		0.82		0.24	
IC	2	5451500	Iowa River at Marshalltown, IA	3,966.2	286.33	268.34	275.70	0.86	0.95	0.24	0.76
IC	3	5451700	Timber Creek near Marshalltown, IA	305.5	41.01	23.46	23.68	0.91	0.86	0.68	0.55
IC	4	5451900	Richland Creek near Haven, IA	145.2	8.09	7.46	11.05	0.59	0.65	0.24	0.18
IC	5	5452200	Walnut Creek near Hartwick, IA	183.6	9.33	9.23	12.88	0.57	0.73	0.27	0.45
IC	6	5453000	Big Bear Creek at Ladora, IA	489.3	31.64	22.54	30.29	0.68	0.72	0.41	0.45
IC	7	5453100	Iowa River at Marengo, IA	7,233.3	612.91	470.11	513.42	0.77	0.73	0.20	0.43
IC	8	5454090	Muddy Creek at Coralville, IA	22.5	1.61	1.50		0.63		0.30	
IC	9	5454000	Rapid Creek near Iowa City, IA	65.5	3.41	3.91		0.40		−0.20	
IC	10	5454220	Clear Creek near Oxford, IA	151.2	10.45	10.55		0.88		0.53	
IC	11	5454300	Clear Creek near Coralville, IA	254	21.84	14.69	6.74	0.70	0.67	0.13	−0.37

[52] Figures 5 and 6 present the observed and simulated hydrographs produced by CUENCAS and SAC-SMA for all streamflow sites in the Cedar River and Iowa River basins, respectively. We normalized the hydrographs by the mean annual flood that is used as an approximation for the bankfull discharge (red line) [Leopold *et al.* 1964]. We calculated the mean annual flood as a function of drainage area using USGS historical data for the Cedar and Iowa Rivers. Values above this line approximate flow levels above the riverbank. The area of each site is shown in the left corner, while the statistics and the site number (refer to Figure 1) are shown in the right corner. In the plots, dark blue lines represent the observed hydrographs; black lines show discharge simulated by CUENCAS using the reference rainfall; light blue lines show the hydrographs simulated by SAC-SMA; and red lines indicate mean annual flood. Differences between observed and predicted streamflow are due to uncertainties in the input, output, and model formulation and parameterization. In this study we focus on evaluating uncertainties that arise particularly from rainfall input errors.

[53] If rainfall input and observed streamflow are used to calibrate model parameters, uncertainties due to the different factors cannot be isolated. Since we did not calibrate our model's parameters, we will demonstrate in this study that, depending on the basin area and on the error structure of the radar rainfall data, uncertainties due to input uncertainty can be as large as the discrepancies between simulated and observed streamflow.

[54] It is important to point out that CUENCAS was not calibrated to fit the hydrographs at specific locations. In this sense, sometimes the model is able to reproduce the peak of the event rather well, but it is not able to correctly reproduce the time to peak (e.g., CR 7). In such cases, the

performance statistics are reduced dramatically. When the model is calibrated to fit a specific hydrograph, the timing of the response is probably the most important criterion since it has a significant impact on the commonly used mean square error objective functions. Another important aspect is the presence of uncertainties in the streamflow observations. Many of the streamflow time series presented long periods of missing data (e.g., CR 8, CR 2, and CR 5). To estimate the statistics using observed time series containing missing values, we adopted a simple linear interpolation procedure to fill the gaps in the time series.

[55] Values of the performance statistics are higher for the Cedar Rapids basin for both models. The lower values for the Iowa River basin are probably due to a large number of sites for small drainage areas and a large number of sites that experienced the backwater effect, which is not accounted for in the formulation of channel dynamics of both hydrological models. Another reason could be due to the spatially correlated errors in the input data that have a large effect over small basins. Better performance is expected for large basins since small-scale error and variability are averaged out by the effect of the river network. In general, we deem the results of the noncalibrated model acceptable in the context of this work.

[56] Throughout this work we use simulated peak flow values to evaluate the effects of rainfall uncertainty on flood prediction across scales. The simulated peak flow values obtained with stage IV data will be referred to as reference value (Q_p^{ref} , gray dots). In Figure 7 we demonstrate the model's ability to simulate peak flow across scales. The figure presents the observed and simulated peak flows versus drainage area for the Cedar (Figure 7a) and the Iowa River (Figure 7b) basins. The orange line represents the mean annual flood. Note that the observed values are above

Figure 5. Observed (dark blue line) and simulated hydrographs produced by the CUENCAS (gray line) and SAC-SMA (light blue line) models for the Iowa River basin. The basin drainage area is presented on the left side, and the site number (refer to Figure 1) and Nash coefficients are presented on the right side [CUENCAS, SAC-SMA].

this line for all sites with a drainage area larger than approximately 100 km², which indicates that water levels were above river banks throughout the watershed. The line in red corresponds to a nonparametric regression between peak flow and the drainage area. This figure demonstrates that our model is able to capture the scaling of peak flows for the 2008 event for the Cedar and Iowa River basins since the simulated and observed peak flow values follow the same scaling patterns. Streamflow observations present many missing data, especially during the time of highest flow (e.g., Figure 6, sites CR 3 and CR 8). Therefore, the direct comparison of the observed and simulated results based solely on

the peak flow values can be misleading. Unfortunately, there are no observed data for small scale sites to confirm the validity of the scaling relationships for small areas. The site with the smallest area is located in the Iowa River (inferior rectus (IR) 1 with approximately 22 km² of drainage area).

4.3. Analysis of Error Propagation

[57] The hydrological model is executed using each rainfall error scenario listed in Table 2 as input to obtain an ensemble of streamflow hydrographs. We evaluate the rainfall error effects on hydrological prediction by comparing the ensemble of simulated hydrographs and the corresponding

Figure 6. Same as Figure 3 but for the Cedar River basin.

peak flows, with the streamflows obtained using the reference rainfall. We also evaluate peak flow variability for each rainfall scenario ensemble and how it changes with basin drainage area.

4.3.1. Comparisons With the USGS Streamflow Observations

[58] We propagate the resulting ensemble rainfall maps through the hydrological model to produce ensembles of flood hydrographs. Results obtained for the Iowa River and Cedar River basins are presented in Figures 8 and 9, respectively. Both figures present results for error rainfall scenarios 2 (Figures 8a and 9a) and 7 (Figures 8b and 9b). SC 2 presents random errors correlated in space and deterministic bias, and SC 2 consists of random errors not correlated in space. For the sake of simplicity and brevity, we opted to include only four hydrographs for the Iowa River and six hydrographs for the Cedar River out of a total of 24 simulated hydrographs. The selected hydrographs cover a wide spectrum of basin areas and are spread around the study area. The location of the sites is represented in Figure 1 by the blue dots.

[59] Figures 8 and 9 use the same symbol convention as Figures 4 and 5, with the addition of gray shadows that represent the range of values simulated by CUENCAS using the 50 ensembles of rainfall. The range of values represented by the gray shadowed region represents uncertainty in stream-

flow simulation as a result of uncertainty in rainfall. Recall that the streamflow ensemble presented in these plots were generated using hypothetical scenarios of rainfall error constructed to isolate known aspects of the error (deterministic or random) and to help us understand how each component contributes to error in streamflow across scales. Therefore, one should expect that the ensemble does not always capture the observed flow. Moreover, differences between observed and simulated streamflow might be due to errors in the parameterization of the model, model structure, or uncertainties in streamflow time series, the last one being especially large during extreme events due to extrapolation of rating curves and hysteresis. Errors can also arise from the mapping of land cover and land use, soil types, and topography features to the scale of the hillslopes. Investigating these uncertainties will be the focus of future studies.

[60] For the Cedar River basin in Cedar Rapids, the hydrograph ensemble reflects the findings of the evaluation in terms of AMAP for SC 7: small spread and a bias between the median of the ensemble and the reference data set. For scenario 2, which considers the radar overestimate rainfall, peaks were reduced across scales, and some spread is still observed at the scale of the Cedar River basin. Comparing both scenarios, we can see that the spread for the Cedar River is much smaller for scenario 7 than for scenario 2.

Figure 7. Peak flow scaling for the Cedar and Iowa Rivers. Gray dots are simulated by CUENCAS, dark blue represent observed, and light blue represent those simulated by SAC-SMA. The red line was obtained by nonparametric regression between the drainage area and peak flow simulated by CUENCAS. The orange line represents mean annual flood.

However, when we look at the small-scale sites, the spread for scenario 7 is larger. This demonstrates that scale dependence of streamflow uncertainty changes for different characteristics of rainfall error structure.

[61] In Figure 10 we present the relative difference between the ensemble and the reference AMAP (plots a and b) and maximum rainfall intensity (plots c and d), normalized by the reference values, for scenarios 2 and 7, for all the USGS sites in the Iowa and Cedar River basins. The spread for AMAP is small since, in this case, the data were aggregated in space and time, and random errors were canceled out. On the other hand, we observe significant variability for the maximum rainfall intensity for each watershed. Due to the purely random character of scenario 7, variability in AMAP and maximum rainfall intensity decreases with drainage area. Rainfall variability also decreases with drainage area for scenario 2, but some spread is still observed for the largest basin due to the spatial correlation of the random error. These errors are propagated through the hydrological model. In Figures 10e and 10f, gray dots represent the relative difference between the ensemble rainfall (gray dots) and observed peak flow for the same scenarios (SC 2 and 7), normalized by the observed peak flow. The variability observed in simulated peak flow is closer to the variability observed for maximum rainfall intensity than it is to the one observed for AMAP.

[62] The black dots in Figures 10e and 10f represent the relative difference between the simulated peak flow using the reference rainfall and observed peak flow, normalized by the observed peak flow. For some of the basins, the difference is around 50%. It is important to point out that a considerable part of these differences might result from inconsistencies in the series of observed streamflow. Due to the extreme magnitude of the event, many of these stream gauges malfunctioned or even stopped working during the periods of large flow. Another source of error in streamflow is due to the extrapolation of the rating curves. Some of the sites experienced water levels much higher than the maximum water levels used to define the rating curve. Model structural and parametric uncertainties also contribute to the differences between observed and simulated values.

[63] The results presented in this section demonstrate the dependence of the error on basin area. However, due to the limited number of points, it is difficult to understand how errors change with scale for different scenarios. In the next section we present a more comprehensive analysis that involves results for all links in the river network and a relative measure of the dispersion of the flow ensemble.

4.3.2. Peak Flow Structure in the River Network

[64] In this section we evaluate the range of simulated peak flow for all the links in the river network using a normalized measure of peak flow spread. For each link in the watershed (1) we calculate the difference between the 95th $[Q_{P95}(i)]$ and 5th $[Q_{P5}(i)]$ percentiles of the ensemble flow values normalized by the median $[Q_{P50}(i)]$ ensemble flow value:

$$\Delta Q_P(i) = \frac{Q_{P95}(i) - Q_{P5}(i)}{Q_{P50}(i)}. \quad (9)$$

[65] We also calculated the normalized difference between the median ensemble value and the reference peak discharge $Q_P^{\text{ref}}(i)$:

$$Q_P^*(i) = \frac{Q_{P50}(i) - Q_P^{\text{ref}}(i)}{Q_P^{\text{ref}}(i)}. \quad (10)$$

[66] Results for equations (9) and (10) are presented in Figures 11 and 12, respectively, for all of the rainfall error scenarios and basin scales ranging from hillslopes ($<0.1 \text{ km}^2$) to the total basin area ($\sim 16,000 \text{ km}^2$). $\Delta Q_P(i)$ (Figure 11) represents the extension of the streamflow uncertainty band at the time of peak for each site. Bias does not affect this variable since it is normalized by the median value. Values equal to one mean that the error band extension is as large as the median of the peak flow simulation. Values equal to zero mean that rainfall data is “error-free” or that errors do not affect streamflow prediction for that specific site.

[67] We also added a nonparametric regression line between $\Delta Q_P(i)$ and the basin area to the plots in order to demonstrate the tendency of $\Delta Q_P(i)$ values to decrease with basin area (blue for SC 1 and gray for the remaining SCs). This means that as the basin area increases, small-scale uncertainties are filtered out by the aggregation effect of the river network. To help the visual comparison among SCs, the nonparametric regression for SC 1 (blue line) is included in all the plots. The rate at which $\Delta Q_P(i)$ decreases as the basin area increases is not the same for all the scenarios. We included in the top right of each plot the

Figure 8. Hydrographs for four sites on the Iowa River showing results of the simulation for rainfall error scenarios (a) 2 and (b) 7. The discharge is normalized by mean annual flood. The dark blue line represents observed hydrographs; the black line is simulated by CUENCAS using the reference rainfall; the gray shadow represents the lowest and highest values for the ensemble hydrographs simulated by CUENCAS using the 50 ensembles of rainfall; the dash gray line is the median of the ensemble; the light blue line represents the hydrographs simulated by SAC; and the red line reflects mean annual flood.

values of $\Delta Q_P(i)$ for the outlet of Cedar River basin to facilitate a direct comparison with the values included in Figure 4 for AMAP and AARI.

[68] Scenarios 1–4 have the same random error structure, and therefore the spreads in the values of normalized dispersion are almost identical (Figure 11). There is a slight variation in the dispersion statistic among these scenarios, which could be due to different deterministic distortion functions (although the dispersion statistic is normalized with the median of the ensemble's values). Scenario 5, whose random errors are not conditional on the radar-rainfall estimate, displayed similar spread in normalized dispersion as scenarios 1–4. This is because of the hyperbolic dependence of the error standard deviation on the radar-rainfall estimate (see equation (3)) and demonstrates that the conditional dependence of the error standard deviation does not have a major impact on the hydrologic response of the basin. When the standard deviation is decreased to a constant value of 0.2 (SC 6), the values of normalized dispersion and their spread decreased considerably. Overall,

from scenarios 4 to 6, it can be inferred that the absolute value of standard deviation (either conditional or unconditional) is a key factor in controlling the behavior of peak flows across scales. For scenarios 7 to 12 we systematically increase the correlation length from 0 to 100 km. The values of normalized differences (gray dots in Figure 11) and their spread are larger for the uncorrelated error scenario (SC 7, $c_2 = 0$ km) at small scales. The spread decreased slowly between 0.1 and 10 km² and rapidly thereafter to reach a value close to zero at the largest scales. The mean value of dispersion (i.e., the non-parametric regression line in Figure 11) drops from ~ 1.4 at the smallest scale to ~ 0.1 at the largest scale. This indicates that the impact of uncorrelated radar-rainfall errors is mainly for the watersheds with a size < 10 km², and the river basin filters out the errors for larger scales. The behavior changed significantly with the introduction of correlations in the radar errors. Even for a smaller correlation length of 15 km (SC 8), the normalized differences failed to reach zero at the larger scales. However, the spread in errors is lower compared to the uncorrelated scenario. In other words, it is beyond the

Figure 9. Same as Figure 6 but for six sites in the Cedar River basin.

“capacity” of the Cedar River basin to filter out radar-rainfall errors if the errors are correlated. The effect of correlations on the simulated peak flows is more evident for c_2 values between 0 and 37 km (scenarios 7–9) and less clear for c_2 values between 37 and 100 km (scenarios 9–11). These results show that the spatial correlation of precipitation errors reduced the rate at which errors decreased with drainage area, as has been demonstrated by *Nijssen* [2004].

[69] For small scales, $\Delta Q_P(t)$ presents a high range of values for all scenarios. For basin areas close to the hillslope area, it can go from very low values (approximately 0.4) to

values of more than 3. This means that small-scale basins can go from receiving almost no water to receiving a large amount of water. The uncertainty for small scale basins will depend on the rainfall characteristics for the specific area, the rainfall error structure, and the properties of the basin.

[70] The same convention of Figure 11 is adopted for Figure 12, which presents the measure of bias $Q_P * (t)$. The bias in peak flow caused by the deterministic error is presented for SC 1 to 3. For SC 1, the deterministic error affects only large values of rainfall, and the effect for the Cedar River is negligible. For SC 2 and 3, the deterministic

Figure 10. (a and b) Differences between the ensemble and the reference areal mean accumulated precipitation. (c and d) Maximum rainfall intensity normalized by the reference values. (e and f) Difference between simulated peak flow using the reference rainfall (black dots) as input and the ensemble rainfall (gray dots) and observed peak flow, normalized by the observed peak flow values. Figures 1a, 1c, and 1e are for scenario 2, while Figures 1b, 1d, and 1f are for scenario 5, Cedar River basin. In the bottom left of each line, we added the equations used to generate each plot.

error shifts the median of the sample, depending on whether the radar is expected to overestimate or underestimate rainfall. For SC 4 to 12, we would expect no bias. However, some bias is introduced due to the reasons we mentioned in section 4.1. For all the scenarios, the effects of the bias decreases with drainage area and is negligible at the outlet for SC 4 to 6 and SC 8 to 12. The bias is still significant ($\sim 11\%$) at the outlet for SC 7, which presents random errors not correlated in space.

[71] It is important to include here a discussion about the possible effect of ignoring the temporal correlation of the radar estimation error on peak flow across scales. Villarini *et al.* [2009] estimated the temporal correlation of the radar error for Oklahoma for the hot season. The authors obtained correlation equal to 0.5 for 1 h and 0.35 for 2 h. Based on this information, the temporal correlation is not significant for the generation of ensemble radar rainfall maps with a temporal resolution of 1 h. Moreover, the temporal scale of hydrological response is often larger than 1 h. The inclusion of temporal correlations in radar-rainfall errors would solely affect time to peak of very small scales that present a concentration time on the order of hours. The effect of temporal correlation

would probably be negligible for larger basins since the magnitude of this effect depends on the relation between the decorrelation time of temporal errors (lower than 2 h) and the time to peak of the hydrograph (that is in the order of days for medium to large basins).

5. Conclusions

[72] This study was motivated by the need to better understand the impacts of radar rainfall uncertainties on flood forecasting that emerged during the process of developing a data intensive hydrological model that does not rely on calibration. In the processes of building, modifying, evaluating, and improving a hydrological model with these features and that focuses on flood prediction across scales, logical questions arise: how much of the discrepancies between simulated and observed discharge are due exclusively to rainfall errors? And how well can we simulate observed values, at different basin scales, based on the best currently available rainfall products that cover large areas? This study is an initial attempt to answer these questions.

Figure 11. Gray dots $[\Delta Q_P(i)]$, see equation (9): the difference between the 95th (Q_{P95}) and 5th (Q_{P5}) percentile ensemble flow values normalized by the median (Q_{P50}) ensemble flow value. The values indicated in brackets in the top right of each plot are the $\Delta Q_P(\text{outlet})$ for the outlet of the Cedar River basin (see equation (7)).

[73] Our methodology consisted of propagating radar derived rainfall uncertainties through a fully distributed hydrological model. We used a calibration-free hydrological model and evaluated different sources of error that are likely to occur when precipitation is estimated using radar observations of rainfall. We found that flood peak uncertainty is sensitive to both the deterministic and random components of rainfall error.

[74] The effect of systematic biases on radar-rainfall estimates is to shift the peak flows up or down depending on the type of bias (overestimation versus underestimation). The effect of the random component on simulated peak flows is more complex. In general, the random errors are filtered out by the aggregation effect of the river network, and uncertainty decreases as catchment area increases. When the errors are not correlated in space, uncertainty decreases considerably with basin area, and the peak flow error at the outlet of the Cedar River basin is practically negligible (this result was consistent for all 50 ensemble members). The dispersion in the simulated peak flow in this case was $\sim 140\%$ (of reference value) at the smallest scale and $\sim 8\%$ at the largest scale ($\sim 17,000 \text{ km}^2$) for the Cedar River in Cedar Rapids.

[75] However, we found the efficiency of the river basin in filtering out random errors to be highly sensitive to the presence of spatial correlation in errors. When rainfall errors are correlated in space, the process of aggregation

and attenuation by the river network is not as effective in filtering out uncertainties. We found that the Cedar River basin could not filter out errors even with a correlation distance as short as 15 km. Compared to uncorrelated errors, the bias and dispersion decreased at smaller scales and increased at larger scales. As we further increased the correlation length to 100 km, the dispersion was found to be $\sim 40\%$ for the largest scale. The shape parameter also has a strong effect since it controls the rate of decrease of the spatial correlation with distance.

[76] The deterministic component of the error tends to produce a shift in the value of simulated peak flow up or down, but the width of the error band is still determined by the magnitude of the random error component and its spatial correlation. It is important to point out that these results improve our understanding of how rainfall errors propagate through a hydrological model and affect flood prediction across a large range of scales. Some conclusions are general and can be applied qualitatively for any radar with an error structure similar to the one applied here. However, quantitative error assessment depends on the specific radar-rainfall estimation algorithm and, thus, error model parameters for the radar and area of interest.

[77] Our results are limited to the effect of hourly rainfall accumulation given on a 4 km by 4 km grid. Higher resolution of radar rainfall is unlikely to change our conclusions

Figure 12. Gray dots [Q_p^* (1), see equation (10)]: the difference between the median ensemble value and the reference peak discharge (Q_p^{ref}) normalized by the reference peak discharge. The values indicated in brackets in the top right corner of each plot are the Q_p^* (outlet) for the outlet of the Cedar River basin.

for the larger scales, but it may influence the scale at which the errors begin to rapidly decrease. Much work remains to be done before we can comprehensively understand the propagation of uncertainty in the hydrologic prediction of floods.

[78] Based on these findings, we conclude that rainfall uncertainties strongly affect hydrological predictions across scales. Errors are scale dependent and decrease as basin scales increase. While modern radar networks provide rainfall maps over vast regions and with high resolution is space and time, rainfall uncertainties remain high, and for the area and flood event investigated in this study, they affect streamflow simulations for scales up to thousands of square kilometers. Our results demonstrate the need to continuously improve our capabilities to quantify rainfall and our knowledge about the errors in the currently available rainfall datasets. This essentially depends on the availability of suitable ground-reference data with the appropriate space-time coverage, configuration, and quality.

[79] This work accomplishes the primary goal of demonstrating the importance of using a calibration-free model to investigate the effect of input uncertainty on hydrological prediction by showing that a consistent modeling implementation provides an unbiased interpretation of the effect of errors in input data. The study also demonstrates that

radar-rainfall uncertainty should be estimated independently of hydrologic models. Specialized networks of rain gauges and other in situ instruments are needed to provide an independent reference against which radar-rainfall estimates can be evaluated [e.g., Mandapaka et al. 2009b; Seo and Krajewski, 2011]. In developing new and improved rainfall products, the first priority should be on minimizing spatial bias. When forced by the “correct” volume of rainfall, the river network takes care of filtering out radar rainfall random errors. We have shown, by performing an evaluation of the results using a large number of nested sites within the main basin, that streamflow sensitivity to input errors varies as a function of basin scale. The magnitude and spatial correlation of the radar rainfall random error define the rate for which errors in simulated streamflow decreases with scale.

[80] In this study we investigated the marginal distribution of errors in streamflow prediction caused exclusively by rainfall errors. However, uncertainty in streamflow prediction has many other sources that propagate through the hydrological model, including parameter uncertainty, model structural uncertainty, and observation uncertainty of input and output variables other than rainfall (e.g., landscape characterization, river discharge, evapotranspiration). Investigating how these uncertainties affect flood simulation will be the subject of our future studies.

[81] **Acknowledgments.** The authors acknowledge support from the IA Flood Center and the NSF Award DMS-1025483 to conduct this study. Luciana Cunha was supported by NASA Headquarters under the Earth Science Fellowship grant NNX10AI46G. The third author acknowledges the support of the Rose & Joseph Summers Chair endowment. The authors also acknowledge many useful comments and suggestions of two anonymous reviewers and Alan Seed from the Australian Bureau of Meteorology on the original version of this paper. This revised version hopefully communicates more effectively the findings of our study.

References

- AghaKouchak, A., E. Habib, and A. Bardossy (2010a), Modeling radar rainfall estimation uncertainties: Random error model, *J. Hydrolog. Eng.*, **15**(4), 265–274.
- AghaKouchak, A., E. Habib, and A. Bardossy (2010b), Copula-based uncertainty modeling: Application to multi-sensor precipitation estimates, *Hydrol. Processes*, **24**(15), 2111–2124.
- Ajami, N. H., H. Gupta, T. Wagener, and S. Sorooshian (2004), Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *J. Hydrol.*, **298**, 112–135, doi:10.1016/j.jhydrol.2004.03.033.
- Anagnostou, E. N., and W. F. Krajewski (1997), Simulation of radar reflectivity fields: Algorithm formulation and evaluation, *Water Resour. Res.*, **33**(6), 1419–1429.
- Arnaud, P., J. Lavabre, C. Fouchier, S. Diss, and P. Javelle (2011), Sensitivity of hydrological models to uncertainty in rainfall input, *Hydrolog. Sci. J.*, **56**(3), 397–410, doi:10.1080/02626667.2011.563742.
- Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts (1995), The Oklahoma Mesonet—A Technical Overview, *J. Atm. Ocean. Technol.*, **12**(1), 5–19.
- Buckingham, E. (1907), *Studies on the movement of soil moisture*, Bull. 38, USDA, Bureau of Soils, Washington, DC.
- Burnash, R. J. C. (1995), The NWS river forecast system—Catchment modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resources Publications, Littleton, CO.
- Carpenter, T. M. (2006), Discretization scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modeling, *J. Hydrol.*, **328**(1–2), 242–257, doi:10.1016/j.jhydrol.2005.12.008.
- Carpenter, T. M., and K. Georgakakos (2004), Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model, *J. Hydrol.*, **298**(1–4), 202–221, doi:10.1016/j.jhydrol.2004.03.036.
- Chumchuan, S., A. Sharma, and A. Seed (2006), An integrated approach to error correction for real-time radar-rainfall estimation, *J. Atm. Ocean. Technol.*, **23**(1), 67–79.
- Ciach, G. J., and W. F. Krajewski (1999a), Radar-rain gauge comparisons under observational uncertainties, *J. Appl. Meteorol.*, **38**(10), 1519–1525.
- Ciach, G. J., and W. F. Krajewski (1999b), On the estimation of radar rainfall error variance, *Adv. in Water Resour.*, **22**(6), 585–595.
- Ciach, G. J., W. F. Krajewski, and G. Villarini (2007), Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data, *J. Hydrometeorol.*, **8**(6), 1325–1347, doi:10.1175/2007JHM814.1.
- Cluckie, I. D., and C. G. Collier (1991), *Hydrological Applications of Weather Radar*, World Meteorological Organization Bulletin, Chichester.
- Cunha, L. K. (2012), Exploring the benefits of Satellite Remote Sensing for flood prediction across scales, Ph.D. thesis, The University of Iowa, Iowa City, IA.
- Cunha, L. K., W. F. Krajewski, and R. Mantilla (2011), A framework for flood risk assessment under nonstationary conditions or in the absence of historical data, *J. Flood Risk Manage.*, **4**(1), 3–22.
- Davidson, J. M., L. R. Stone, D. R. Nielsen, and M. E. Larue (1969), Field measurement and use of soil-water properties, *Water Resour. Res.*, **5**(6), 1312–1321, doi:10.1029/WR005i006p01312.
- Dingman, S. L. (2002), *Physical Hydrology*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ.
- Dodov, B. (2004), Generalized hydraulic geometry: Derivation based on a multiscale formalism, *Water Resour. Res.*, **40**, W06302, doi:10.1029/2003WR002082.
- Ebel, B. A., and K. Loague (2006), Physics-based hydrologic-response simulation: Seeing through the fog of equifinality, *Hydrolog. Process.*, **20**, 2887–2900.
- Freeze, R. (1980), A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope, *Water Resour. Res.*, **16**(2), 391–408.
- Fu, S., T. O. Sonnenborg, K. H. Jensen, and X. He (2011), Impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resources model, *Vadose Zone J.*, **10**(1), 25–36, doi:10.2136/vzj2009.0186.
- Fulton, R., J. Breidenbach, D. Seo, and D. Miller (1998), The WSR-88D rainfall algorithm, *Weather Forecasting*, **13**, 377–395.
- Germann, U., M. Berenguer, D. Sempere-Torres, and M. Zappa (2009), REAL-ensemble radar precipitation estimation for hydrology in a mountainous region, *Q. J. R. Meteorol. Soc.*, **135**, 445–456.
- Graham, C. B., and J. J. McDonnell (2010), Hillslope threshold response to rainfall: (2) Development and use of a macroscale model, *J. Hydrol.*, **393**, 77–93, doi:10.1016/j.jhydrol.2010.03.008.
- Grayman, W. M., and P. S. Eagleson (1971), Evaluation of radar and rain-gage systems for flood forecasting, *R. M. Parsons Lab. Rep. 138*, 324 pp., Mass. Inst. of Technol., Cambridge, MA.
- Gupta, V. K. (2004), Emergence of statistical scaling in floods on channel networks from complex runoff dynamics, *Chaos, Soliton Fract.*, **19**(2), 357–365, doi:10.1016/S0960-0779(03)00048-1.
- Gupta, V. K., S. Castro, and T. Over (1996), On scaling exponents of spatial peak flows from rainfall and river network geometry, *J. Hydrol.*, **187**, 81–104.
- Gupta, V. K., R. Mantilla, B. M. Troutman, D. Dawdy, and W. F. Krajewski (2010), Generalizing a nonlinear geophysical flood theory to median-sized river networks, *Geophys. Res. Lett.*, **37**(11), L11402, doi:10.1029/2009GL041540.
- Habib, E., A. V. Aduvala, and E. A. Meselhe (2008), Analysis of radar-rainfall error characteristics and implications for stream flow simulation uncertainty, *Hydrol. Sci. J.*, **53**, 568–587.
- He, X., J. C. Refsgaard, T. O. Sonnenborg, F. Vejen, and K. H. Jensen (2011), Statistical analysis of the impact of radar rainfall uncertainties on water resources modeling, *Water Resour. Res.*, **47**(9), W09526, doi:10.1029/2011WR010670.
- Heistermann, M., and D. Kneis (2011), Benchmarking quantitative precipitation estimation by conceptual rainfall-runoff modeling, *Water Resour. Res.*, **47**(6), W06514, doi:10.1029/2010WR009153.
- Hirsch, R. M., and J. E. Costa (2004), U.S. stream flow measurement and data dissemination improve, *Eos Trans. AGU*, **85**(20), 197–203.
- Homer, C., J. Dewitz, J. Fry, and M. Coan (2007), Completion of the 2001 National Land Cover Database for the Conterminous United States, *Photogram. Eng. Remote Sensing*, **73**, 337–341.
- Hossain, F., E. Anagnostou, and T. Dinku (2004), Sensitivity analyses of satellite rainfall retrieval and sampling error on flood prediction uncertainty, *IEEE Trans. Geosci. Remote Sensing*, **42**, 130–139.
- Johnson, D., M. Smith, V. Koren, and B. Finnerty (1999), Comparing mean areal precipitation estimates from NEXRAD and rain gauge networks, *J. Hydrol. Eng.*, **4**(2), 117–124.
- Jordan, P. W., A. W. Seed, and P. E. Weinmann (2003), A stochastic model of radar measurement errors in rainfall accumulations at catchment scale, *J. Hydrometeorol.*, **4**, 841–855.
- Kavetski, D. (2011), Numerical troubles in conceptual hydrology: Approximations, absurdities and impact on hypothesis testing, *Hydrol. Processes*, **25**, 661–670, doi:10.1002/hyp.7899.
- Krajewski, W. F. (1993), Hydrology, in *Encyclopedia of Applied Physics*, Vol. 7, pp. 563–579, VCH, Berlin.
- Krajewski, W. F. and K. Georgakakos (1985), Synthesis of radar rainfall data, *Water Resour. Res.*, **21**(5), 764–768.
- Krajewski, W. F., G. Villarini, and J. A. Smith (2010), Radar-rainfall uncertainties: Where are we after thirty years of efforts?, *Bull. Am. Meteorol. Soc.*, **91**(1), 87–94.
- Krause, P., D. P. Boyle, and F. Bäse (2005), Comparison of different efficiency criteria for hydrological model assessment, *Adv. Geosci.*, **5**, 89–97.
- Leopold, L. B., M. G. Wolman, and J. P. Miller (1964), *Fluvial Processes in Geomorphology*, W.H. Freeman, New York.
- Mandapaka, P. V., and U. Germann (2010), Radar-rainfall error models and ensemble generators, in *Rainfall: State of Science*, edited by M. Gebremichael and F. Testik, Geophysical Monograph Series, 191, pp. 247–264, American Geophysical Union, Washington, DC.
- Mandapaka, P., W. F. Krajewski, and R. Mantilla (2009a), Dissecting the effect of rainfall variability on the statistical structure of peak flows, *Adv. Water Resour.*, **32**, 1508–1525.
- Mandapaka, P. V., W. F. Krajewski, G. J. Ciach, G. Villarini, and J. A. Smith (2009b), Estimation of radar-rainfall error spatial covariance, *Adv. Water Resour.*, **32**(7), 1020–1030.
- Mantilla, R. (2007), Physical basis of statistical scaling in peak flows and stream flow hydrographs for topologic and spatially embedded random

- self-similar channel networks, Ph.D. thesis, University of Colorado, Boulder, CO.
- Mantilla, R., and V. K. Gupta (2005), A GIS numerical framework to study the process basis of scaling statistics in river networks, *IEEE Geosci. Remote Sensing Lett.*, 2(4), 404–408, doi:10.1109/LGRS.2005.853571.
- McGlynn, B., J. McDonnell, and D. Brammer (2002), A review of the evolving perceptual model of hillslope flowpaths at the Maimai catchments, New Zealand, *J. Hydrol.*, 257, 1–26.
- Menabde, M., S. Veitzer, and V. Gupta (2001), Tests of peak flow scaling in simulated self-similar river networks, *Adv. Water Resour.*, 24, 991–999.
- Moreno, H. A., E. R. Vivoni, and D. J. Gochis (2012), Utility of quantitative precipitation estimates for high resolution hydrologic forecasts in mountain watersheds of the Colorado Front Range, *J. Hydrol.*, 438–439, 66–83, doi:10.1016/j.jhydrol.2012.03.019.
- Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing Environ.*, 115, 1781–1800.
- Nash, J. E. and J. Sutcliffe (1970), Flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, 10, 282–290.
- Nijssen, B. (2004), Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites, *J. Geophys. Res.*, 109, D02103, doi:10.1029/2003JD003497.
- Nikolopoulos, E., and E. Anagnostou (2010), Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model, *J. Hydrometeorol.*, 11, 520–532.
- Paik, K., and P. Kumar (2004), Hydraulic geometry and the nonlinearity of the network instantaneous response, *Water Resour. Res.*, 40, W03602, doi:10.1029/2003WR002821.
- Pegram, G. and A. Clothier (2001), Downscaling rainfields in space and time using the string of beads model in time series mode, *Hydrol. Earth Syst. Sci.*, 5, 175–186.
- Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D. Seo, and DMIP Participants (2004), Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60.
- Refsgaard, J. (2004), Modelling guidelines—Terminology and guiding principles, *Adv. Water Resour.*, 27, 71–82.
- Sabzevari, T., A. Talebi, R. Ardakanian, and A. Shamsai (2010), A steady-state saturation model to determine the subsurface travel time (STT) in complex hillslopes, *Hydrol. Earth Syst. Sci.*, 14(6), 891–900, doi:10.5194/hess-14-891-2010.
- Sawicz, K., T. Wagener, M. Sivapalan, and P. Troch (2011), Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 4495–4534.
- Schröter, K., X. Llort, C. Velasco-Forero, M. Ostrowski, and D. Sempere-Torres (2011), Implications of radar rainfall estimates uncertainty on distributed hydrological model predictions, *Atm. Res.*, 100(2–3), 237–245, doi:10.1016/j.atmosres.2010.08.014.
- Seed, A. W., R. Srikanthan, and M. Menabde (1999), A space and time model for design storm rainfall, *J. Geophys. Res.*, 24, 31623–31630.
- Seo, B.-C., and W. F. Krajewski (2011), Investigation of the scale-dependent variability of radar-rainfall and rain gauge error correlation, *Adv. Water Resour.*, 34(2), 152–163.
- Sharif, H. O., F. L. Ogden, W. F. Krajewski, and M. Xue (2002), Numerical simulations of radar-rainfall error propagation, *Water Resour. Res.*, 38(8), 1140, doi:10.1029/2001WR000525.
- Sharif, H. O., F. L. Ogden, W. F. Krajewski, and M. Xue (2004), Statistical analysis of radar-rainfall error propagation, *J. Hydrometeorol.*, 5(1), 199–212.
- Sivapalan, M. (2003), Prediction in ungauged basins: A grand challenge for theoretical hydrology, *Hydrol. Processes*, 17, 3163–3170.
- Villarini, G., and W. F. Krajewski (2009), Empirically based modelling of radar-rainfall uncertainties for a C-band radar at different time-scales, *Q. J. R. Meteorol. Soc.*, 135(643), 1424–1438, doi:10.1002/qj.454.
- Villarini, G., and W. F. Krajewski (2010), Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall, *Surveys Geophys.*, 31(1), 107–129.
- Villarini, G., W. F. Krajewski, G. J. Ciach, and D. L. Zimmerman (2009), Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates, *Water Resour. Res.*, 45(1), W01404, doi:10.1029/2008WR006946.
- Welles, E., S. Sorooshian, G. Carter, and B. Olsen (2007), Hydrologic verification: A call for action and collaboration, *Bull. Am. Meteorol. Soc.*, 88, 503–511.
- Wilson, J. W., and E. A. Brandes (1979), Radar measurement of rainfall—A Summary, *Bull. Am. Meteorol. Soc.*, 60, 1–11.
- Xia, Y., et al., (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, *J. Geophys. Res.*, 117, D03109, doi:10.1029/2011JD016048.
- Yang, D., S. Herath, and K. Musiake (2002), A hillslope-based hydrological model using catchment area and width functions, *Hydrolog. Sci. J.*, 47(1), 49–65, doi:10.1080/02626660209492907.