Understanding uncertainty in distributed flash flood forecasting for semiarid regions

Soni Yatheendradas,[1] Thorsten Wagener,[2] Hoshin Gupta,[3] Carl Unkrich,[4] David Goodrich,[4] Mike Schaffner,[5] and Anne Stewart[3]

Received 1 February 2007; revised 25 January 2008; accepted 3 March 2008; published 22 May 2008.

[1] Semiarid flash floods pose a significant danger for life and property in many dry regions around the world. One effective way to mitigate flood risk lies in implementing a real-time forecast and warning system based on a rainfall-runoff model. This study used a semiarid, physics-based, and spatially distributed watershed model driven by high-resolution radar rainfall input to evaluate such a system. The predictive utility of the model and dominant sources of uncertainty were investigated for several runoff events within the U.S. Department of Agriculture Agricultural Research Service Walnut Gulch Experimental Watershed located in the southwestern United States. Sources of uncertainty considered were rainfall estimates, watershed model parameters, and initial soil moisture conditions. Results derived through a variance-based comprehensive global sensitivity analysis indicated that the high predictive uncertainty in the modeled response was heavily dominated by biases in the radar rainfall depth estimates. Key model parameters and initial model states were identified, and we generally found that modeled hillslope characteristics are more influential than channel characteristics in small semiarid basins. We also observed an inconsistency in the parameter sets identified as behavioral for different events, which suggests that model calibration to historical data is unlikely to consistently improve predictive performance for different events and that real-time parameter updating may be preferable.

Citation: Yatheendradas, S., T. Wagener, H. Gupta, C. Unkrich, D. Goodrich, M. Schaffner, and A. Stewart (2008), Understanding uncertainty in distributed flash flood forecasting for semiarid regions, *Water Resour. Res.*, *44*, W05S19, doi:10.1029/2007WR005940.

1. Introduction and Scope

[2] A trend toward a warmer climate has increased global incidences of intense precipitation events [*Trenberth et al.*, 2007]. Consequences of this change in precipitation can be seen in increased chances of flooding [*Milly et al.*, 2002] and in increased flood losses [*Kundzewicz and Kaczmarek*, 2000], despite widespread problems of water scarcity. Arid and semiarid regions, i.e., areas where the annual rain is less than 250 and 250–500 mm/a, respectively, are particularly vulnerable to this change in climate. Additionally, the highly nonlinear nature of the rainfall-runoff relationship in these dry regions makes predictions of the impact of precipitation changes especially difficult.

[3] Arid and semiarid regions span approximately one third of the global land surface, one fourth of the contiguous United States, and more than half of the western United States [*Goodrich et al.*, 2000]. The semiarid southwestern United States experiences extremely localized and intense summertime convective thunderstorms [*Roeske et al.*, 1989], often leading to ''short-fused'' local floods. Within the United States, floods occurring within 6 h of the causative rainfall event are termed flash floods [*National Weather Service*, 2002]. These flash floods kill more Americans annually than any other natural disaster [*American Meteorological Society*, 1985], while accounting for more than 80% of all flood-related deaths, and causing average annual economic losses of one billion U.S. dollars.

[4] Perhaps the most effective way to mitigate the risks due to flash flood occurrence is to implement a real-time flood forecast and warning system [e.g., *Creutin and Borga*, 2003; *Kitanidis and Bras*, 1980a, 1980b]. This nonstructural (or ''soft'') measure is also suitable for climate change adaptation applications [*Kundzewicz*, 2002]. Requirements for such a system include a watershed model of the highly nonlinear hydrological processes occurring at the semiarid land surface [*Pilgrim et al.*, 1988]. Only a high-resolution spatially distributed model structure can appropriately represent the semiarid spatiotemporal variability in the rainfall forcing and the consequent infiltration and runoff processes [*Michaud and Sorooshian*, 1994; *Osborn*, 1964].

[5] In this paper, we investigate the predictive utility of a semiarid flash flood forecasting system based on an event-oriented, physically based, semiarid rainfall-runoff model, KINEROS2 (Kinematic Runoff and Erosion) [e.g., *Smith et*

[1]Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA.
[2]Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, Pennsylvania, USA.
[3]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA.
[4]Southwest Watershed Research Center, Agricultural Research Service, U.S. Department of Agriculture, Tucson, Arizona, USA.
[5]Binghamton Weather Forecast Office, National Weather Service, Johnson City, New York, USA.

0043-1397/08/2007WR005940$09.00

al., 1995; *Semmens et al.*, 2005]. Both gauge and radar estimates of high-resolution rainfall for eight summertime convective thunderstorm events were used over a subbasin of the semiarid Walnut Gulch experimental watershed in southern Arizona [*Renard et al.*, 1993; *Moran et al.*, 2008]. The Generalized Likelihood Uncertainty Estimation framework (GLUE) [*Beven and Binley*, 1992] is used to assimilate incoming event information into KINEROS2 in a forecasting mode. A variance-based and interaction-accounting global sensitivity analysis method is applied to explore the relative influence of uncertainties from sources such as rainfall forcing, initial conditions and model parameters on uncertainty in the estimated streamflow.

2. Sources of Uncertainty in Semiarid Flash Flood Forecasting

[6] Flash flood warnings in the United States are provided by the Weather Forecast Offices of the National Weather Service (NWS). In the southwestern United States, these warnings are typically based directly on estimates of spatially averaged rainfall, or on flow estimates computed using (currently) lumped rainfall-runoff models. The former approach is implemented in the flash flood monitoring and prediction approach that compares temporal accumulations of spatially averaged rainfall for small basins with corresponding flood guidance thresholds. A flood warning is issued if a predefined rainfall threshold is exceeded. Thresholds in arid regions are typically established based on either past experience, or empirical or semiquantitative approaches. While an experienced forecaster might be able to provide good forecasts based on this information and on an in-depth understanding of the forecast region, such thresholds may not properly consider the effects of spatially heterogeneous runoff production or of variability in initial soil moisture conditions.

[7] Models presently used in NWS operational flash flood forecasting are spatially lumped and were originally developed to represent hydrologic processes in humid watersheds, with time steps too coarse to provide adequate simulation of semiarid hydrographs that can peak in 15 min or less. These models include the Sacramento Soil Moisture Accounting model [*Burnash*, 1995] and the Continuous Antecedent Precipitation Index model [*Anderson*, 1993]. Neither of these models incorporates semiarid physics, e.g., channel transmission losses, though the models' flexibility allows for the simulation of a wide range of natural regimes, but not necessarily for the right reasons [e.g., *Van Werkhoven et al.*, 2008]. Additional sources of uncertainty (rainfall forcing, initial soil moisture conditions, model parameters) are discussed in detail in the following sections. The important overall question we attempt to address here is as follows: "How reliable is a semiarid flash flood forecast model under the compounding effects of these uncertainties"? Further research questions are defined on the basis of the reviews provided in the following sections.

2.1. Uncertainty in Rainfall Estimates

[8] Errors in rainfall estimates arising often from biases in observations of the spatial rainfall representation and of the rainfall volumes, can dominate the uncertainty in the modeled semiarid runoff response [*Goodrich et al.*, 1994; *Faurès et al.*, 1995]. The highly nonlinear watershed model can magnify the effects of rainfall estimation errors, particularly as surface runoff rates decrease with increasing basin size due to partial storm coverage and channel infiltration losses [*Goodrich*, 1990; *Goodrich et al.*, 1997].

[9] Flash flood events are typically driven by intense summertime convective thunderstorm cells with limited areal extent, usually less than 10–14 km in diameter. Adequate representation of this extreme rainfall variability even over small watersheds requires high-resolution rainfall estimates obtained from dense rain gauge networks, high-resolution radar, or possibly even high-resolution numerical weather prediction models. Satellite precipitation estimates, while improving continuously, are not yet of sufficiently high spatiotemporal resolution [e.g., *Hong et al.*, 2004], and are not likely to be applicable to flood forecasting at small basin scales [*Yilmaz et al.*, 2005]. Typical rain gauge networks are of insufficient density, hence forecasters in the semiarid United States rely on the comprehensive areal coverage provided by the NWS NEXt generation RADar Weather Surveillance Radar–1988 Doppler (NEXRAD WSR-88D) system [*Maddox et al.*, 2002].

[10] Volume biases in rainfall estimates potentially impact the simulated runoff response more than the biases in spatial variability [*Faurès et al.*, 1995]. Summertime convective rainfall estimates are operationally derived using a standardized NWS Z-R (reflectivity-rain) relationship, which translates the amount of reflected radar waves into rainfall rates [*Fulton et al.*, 1998]. This relationship assumes a power law form: $Z = aR^b$, where the dimensionless empirical parameters are set as NWS standard to $a = 300$ and $b = 1.4$. However, the resulting rainfall estimates are often too high for semiarid convective storms [*Morin et al.*, 2005; *Hardegree et al.*, 2003]. Optimal values for a and b have been shown to vary widely between storms, for example, *Morin et al.* [2005] adjusted parameter a in the NWS relationship to an improved value of 655 for 13 individual storms over the Walnut Gulch, and noted a significant range of 0.58–1.8 in ratios of the resultant storm depth estimates compared to values of a very dense gauge network. Another source of uncertainty is hail contamination, which can severely distort rainfall estimates due to the sixth power dependence of reflectivity on drop diameter, even in semiarid summertime convective storms [*Krajewski and Smith*, 2002]. The discussion above leads us to the second important research question: To what extent are modeled runoff predictions affected by errors in the spatial distribution and the volume of the rainfall input?

2.2. Uncertainty in the Initial Soil Moisture Estimates

[11] Soil moisture conditions control runoff generation through their impact on soil infiltration rates and on the connectivity of surface and subsurface runoff pathways [*Goodrich et al.*, 1994; *Grayson and Bloeschl*, 2000]. The rainfall-runoff response has a strong nonlinear dependence on semiarid antecedent wetness [*Nicolau et al.*, 1996; *Beven*, 2002], influencing both overland and channel flows [*White et al.*, 1997]. While the actual rainfall event response of a semiarid watershed has to be modeled with high space-time resolution, the level of detail in the modeling of soil moisture contents during interstorm periods is likely determined by the sensitivity of the flash flood predictions to the accuracy of the model's initial soil moisture states. The third important research question is therefore as follows: How

sensitive are modeled runoff predictions to the specification of initial soil moisture conditions, and thus, how complex must the between storm soil moisture accounting model be for accurate flash flood forecasting?

2.3. Uncertainty in the Model Parameter Estimates

[12] Semiarid watershed models typically require a higher level of spatial and process complexity compared to humid region models. This spatial complexity arises from the high spatial variability of rainfall and the corresponding highly complex spatiotemporal interaction between watershed and storm geometries, which strongly influences the shape and volume of the flood hydrograph [*Osborn*, 1964; *Michaud and Sorooshian*, 1994]. Regardless of the spatial resolution of the distributed watershed model used, some spatial lumping is unavoidable in representing real-world complexity [e.g., *Wagener and Gupta*, 2005]. This process of lumping requires that the model's parameters are estimated in a way to reflect the integration of the spatially distributed physical watershed characteristics at the model element scale. While a priori values for the spatially distributed model parameters can be estimated from soil and other physical watershed characteristics, it is unlikely that those estimates will accurately represent the watershed characteristics at the model element scale. This is largely due to differences in measurement and model element scale. The a priori estimates can be calibrated if observations of streamflow are available, though the large number of spatially distributed parameters makes it likely that a focus on some key parameters is required. Therefore, the fourth important research question is as follows: ''Which model parameter uncertainties strongly influence the modeled predictions of runoff''?

3. Description of the Rainfall-Runoff Modeling Framework and Analytic Techniques

3.1. Modeling Framework

[13] The modeling framework used in this study consists of a semiarid specific watershed model, KINEROS2, and two data preprocessing programs: (1) the GIS-based Automated Geospatial Watershed Assessment tool (AGWA) to estimate the geometric model element attributes (area, slope, mean flow length), model element spatial configuration (e.g., how flow is routed from one model element to the next), and a priori model element parameter values, and (2) the AMBER algorithm that performs the Z-R conversion. Each component is briefly described below.

[14] KINEROS2 [*Woolhiser et al.*, 1990; *Smith et al.*, 1995; *Semmens et al.*, 2005; *Goodrich et al.*, 2006; http://www.tucson.ars.ag.gov/kineros/] is an event-oriented, distributed, physically based model developed to simulate the runoff response in basins that are overland flow-dominated. KINEROS2 simulates interception, dynamic infiltration and infiltration–excess surface runoff, with flow routed downstream using a finite difference solution of the one-dimensional kinematic wave equations over a basin conceptualized as a cascade of planes (hillslopes) and channels. The Parlange three-parameter infiltration equation [*Parlange et al.*, 1982] is implemented as the KINEROS2 dynamic infiltration algorithm interacting with both rainfall and surface runoff in transit, and hence is well suited to model channel transmission losses in semiarid ephemeral streams. Plane microtopographic relief is represented by a plane transverse saw tooth geometry. KINEROS2 was recoded in the context of this study from a procedural version with outer space and inner time looping to an object-oriented one with outer time and inner space looping to facilitate real-time forecasting.

[15] The AGWA tool [*Semmens et al.*, 2005; *Miller et al.*, 2007; http://www.tucson.ars.ag.gov/agwa/] is a GIS framework that creates the parameter file for input to KINEROS2 from spatially distributed physical watershed characteristics (elevation, land cover, soil etc.). The flow duration in Table 2 provides a list of the hillslope and channel hydrologic parameters used in this study and estimated by AGWA. The basin is discretized into plane and channel elements based on a user-selected threshold drainage area, called the Channel (or Contributing) Source Area (CSA) that is expressed as a percent of the watershed area, and applied to a digital elevation model. Geometric shape, hydraulic and infiltration parameters are generated for each KINEROS2 plane using spatially weighted averaging.

[16] The NWS Areal Mean Basin Estimated Rainfall (AMBER) algorithm [*Davis and Jendrowski*, 1996; http://www.erh.noaa.gov/er/rnk/amber/] computes mean accumulated areal rainfall from the digital hybrid reflectivity data of the Weather Surveillance Radar–1988 Doppler (WSR-88D). In this study we customized AMBER to provide radar pixel rainfall information for KINEROS2. AMBER has several customizable parameters, including the reflectivity-rain relationship parameters and the hail threshold.

3.2. GLUE Framework

[17] The Generalized Likelihood Uncertainty Estimation framework (GLUE) [*Beven and Binley*, 1992; *Beven and Freer*, 2001] is a set theoretic approach to uncertainty analysis based on the idea of equifinality [*Beven*, 1993]. For a given model structure, equifinality argues that many parameter sets can typically provide acceptable (or ''behavioral'') simulations of a complex environmental system. In GLUE, each behavioral parameter set is weighted by its relative likelihood (i.e., measure of performance henceforth called likelihood function), to provide model confidence intervals on the resulting forecast ensemble. GLUE implicitly handles model nonlinearity, factor interactions and errors in the model structure and observations. The likelihood values of nonbehavioral parameter sets are set to zero. GLUE was originally inspired by the regional sensitivity analysis method [*Spear and Hornberger*, 1980] which introduced the idea of behavioral/nonbehavioral parameter set classification.

3.3. Variance-Based Global Sensitivity Analysis and Sampling Strategy

[18] Sensitivity analysis (SA) estimates the impact of changes in factor values (e.g., parameters, states or input variables) on the simulated model response of interest [e.g., *Wagener and Kollat*, 2006]. From an uncertainty viewpoint, SA investigates the influence of factor uncertainties on the uncertainty in the model response [e.g., *Saltelli et al.*, 2004], hence, ''sensitivity'' and ''influence'' are used interchangeably. A more influential factor has thus potentially a larger effect on the output uncertainty. An SA approach can be defined as ''global'' if all factors are varied simultaneously, rather than one at a time. The Global SA (GSA) method used here is based on multivariate sampling of the entire

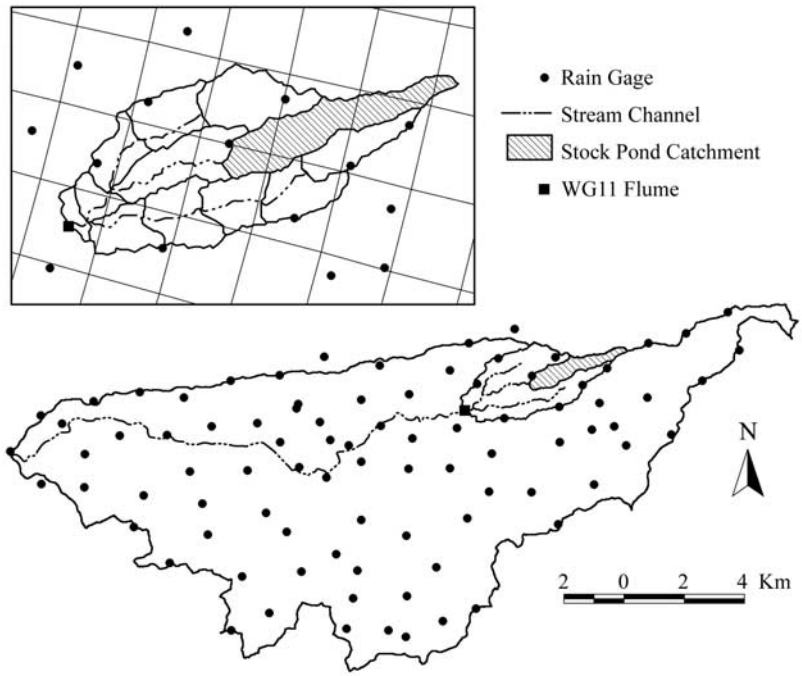

Figure 1. (bottom) The Walnut Gulch Experimental Watershed. (top) WG11 subbasin shown magnified against the radar grid background. Subbasin segmentation has been performed with a discretization level of 7.5% contributing source area (CSA).

feasible factor space to obtain representative sensitivity information. This study used a variance-based GSA method called Sobol' method [*Sobol'*, 1993], which accounts for factor interactions at all orders, and provides resultant quantitative sensitivity information that is both a necessary and sufficient condition to rank the factor influences [*Saltelli et al.*, 2004]. The order of a factor interaction is based on the number of factors involved in that particular interaction. Conceptually, for "n" orthogonal factors, the Sobol method decomposes the total variance (V) of an output variable of interest (y) into components corresponding to those of that variable at different orders:

$$V(y) = \sum_{i=1}^{n} V_i + \sum_{1 \le i \le j \le n} V_{i,j} + V_{1,2,\ldots,n}. \quad (1)$$

Dividing the above equation by $V(y)$ provides the sensitivity terms (S) for all interactions:

$$1 = \sum_{i=1}^{n} S_i + \sum_{1 \le i \le j \le n} S_{i,j} + S_{1,2,\ldots,n} \quad (2)$$

where the subscripts denote the particular interaction, e.g., $S_{1,3,5}$ would denote interaction between factors 1, 3, and 5. The Sobol' method assumes the sufficiency of the second moment (variance) for describing uncertainty. This method was implemented using the SIMLAB software version 2.2 (http://simlab.jrc.cec.eu.int/).

[19] We applied the Sobol' sampling strategy [*Sobol'*, 1993] for both the Sobol' SA and the GLUE analysis [e.g., *Ratto et al.*, 2001]. The enhanced Sobol' method [*Homma and Saltelli*, 1996; *Saltelli*, 2002] computes the first- and total-order (i.e., total influence including all interactions) SA indices directly for any factor without calculating intermediate-order indices. The Sobol' sampling for this method generated 102,400 factor sets in the 24-factor space as a first Monte Carlo run collection, and 98,304 factor sets in the 23-factor space as a second collection. Note that in the presence of interactions, the sum of the total-order indices of the factors can be greater than one. Additional second-order indices also used later in this study were calculated from Sobol' sampling densities of 71,168 and 77,312 for the 23- and 24-factor spaces, respectively.

4. Description of Watershed and Data Sets

4.1. USDA ARS Walnut Gulch Experimental Watershed and the WG11 Study Area

[20] The Walnut Gulch Experimental Watershed (WGEW) is located in southeastern Arizona, primarily in a high foothill alluvial fan portion of the San Pedro River watershed underlain by very deep (>400 m) Cenozoic alluvium regional groundwater aquifer [*Moran et al.*, 2008]. WGEW spans 150 km^2 in area and 1220–1830 m in elevation. Depth to groundwater within the WGEW ranges from ~50 m at the lower end to ~145 m in the central portion of the watershed. The climate over the watershed is semiarid, with substantial interannual and intra-annual variations and an average annual precipitation amount of 324 mm.

[21] The specific area of investigation underlying this study is the WG11 subbasin of the WGEW. WG11 has its outlet at flume 11 and spans an area of 7.9 km^2. From this area, the catchment upstream of pond outlet 16 was excluded as it rarely discharges into the stream, and did not for any of the events used in this study (Figure 1). This reduced the area to 6.4 km^2 with an elevation range of 1430–1525 m. The

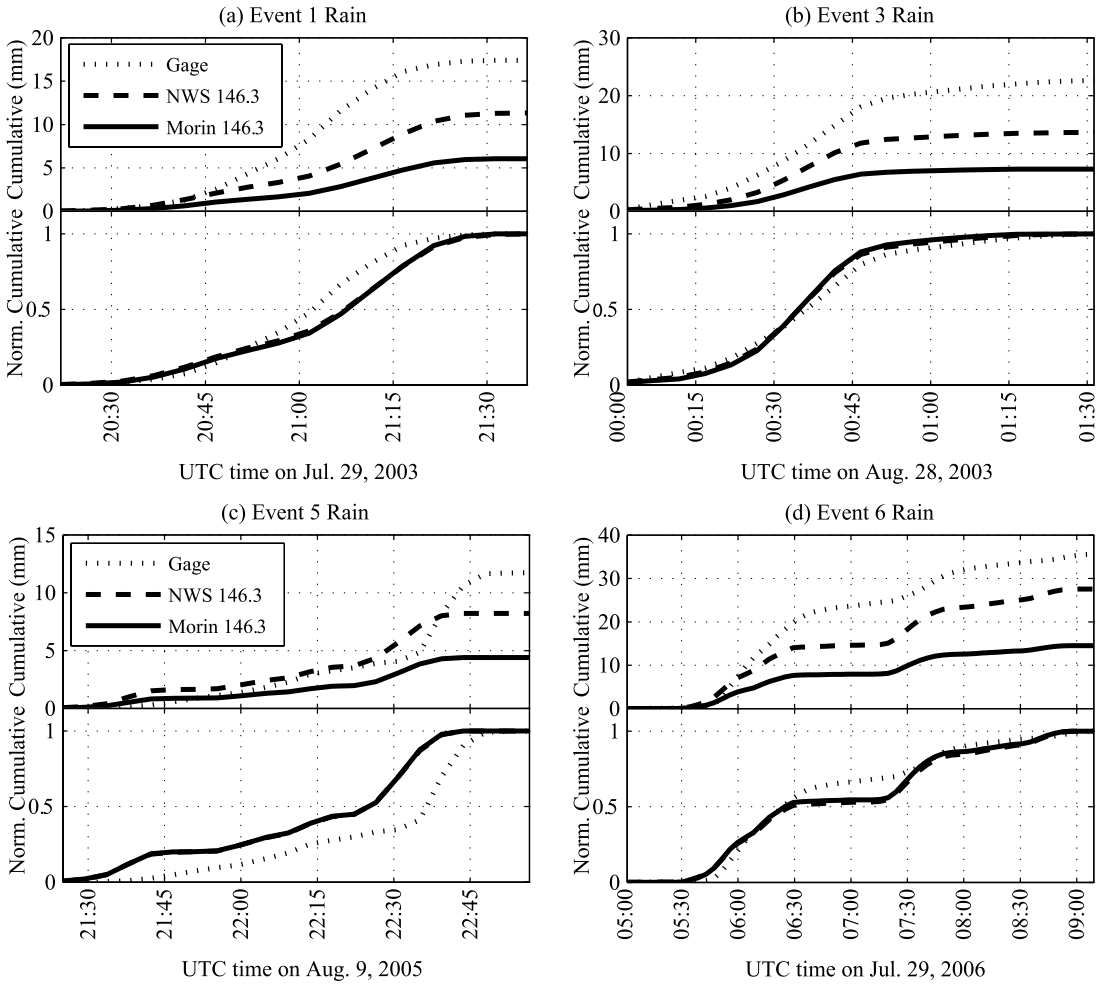

Figure 2. Comparison of rainfall estimates from gauge and from radar (two different Z-R relationships) for four events: (top) Basin mean cumulative rainfall depths. (bottom) Same information after normalizing each series by its total value. The value of 146.3 mm/h indicates the selected hail threshold.

lower end of WG11 is desert brush, while a transition to range grassland occurs approximately midway up the watershed. WG11 has wide sand bed channels which can result in significant channel transmission losses.

4.2. Rainfall and Discharge Measurements

[22] The digital hybrid reflectivity (DHR) data provides corresponding rainfall estimates on a polarimetric grid (1° circumferential by 1 km radial resolution). The variable radar rotational scan speed imparts time steps ranging from 4 to 6 min for this study's events. In the DHR product, each individual pixel reflectivity is assigned the respective lowest unobstructed tilt possible (i.e., less than 50% blockage above 150 m above ground level (AGL)). This potentially decreases error in the ground surface rain estimates caused by atmospheric evaporation that occurs when the sampled reflectivity is at a high-elevation AGL. The sampled height had a maximum of around 2 km for WG11 [*Fulton et al.*, 1998]. The Tucson NWS DHR archive started in 2003, but owing to recent dry years over the WGEW, only 8 events were available for the present study (Table 2). Twelve radar pixels completely covered the reduced WG11 area (Figure 1).

[23] The dense rain gauge network of the WGEW [*Goodrich et al.*, 2008] provided rainfall data from 11 gauges in and near WG11, and discharge data for flume 11 (available at http://www.tucson.ars.ag.gov/dap/, see Table 2). The breakpoint instantaneous discharge data (finest resolution being 1 min) was converted to radar time step values by averaging over time windows delineated by the midpoints of the radar time steps; this was judged to represent the original data time series better. Basin-averaged cumulative gauge rainfall calculated for selected events using Thiessen polygons and corresponding observed hydrographs are shown in Figures 2 and 3, respectively.

4.3. A Priori Parameters From AGWA and Field Measurements

[24] Data layers used in AGWA include the digital elevation model from the U.S. Geological Survey (USGS) (at http://data.geocomm.com/), the soils coverage/shapefile from the State Soil Geographic Database (STATSGO), and the land cover grid from the MultiResolution Land Characteristics (MRLC) Consortium–National Land Cover Data (NLCD). All data sets were available at or interpolated to a 10 m resolution. The a priori parameter values obtained for a CSA value of 7.5% (see Table 1) suggest a high spatial uniformity in WG11. These data layers do not (or do not reliably) provide some of the a priori parameters, including

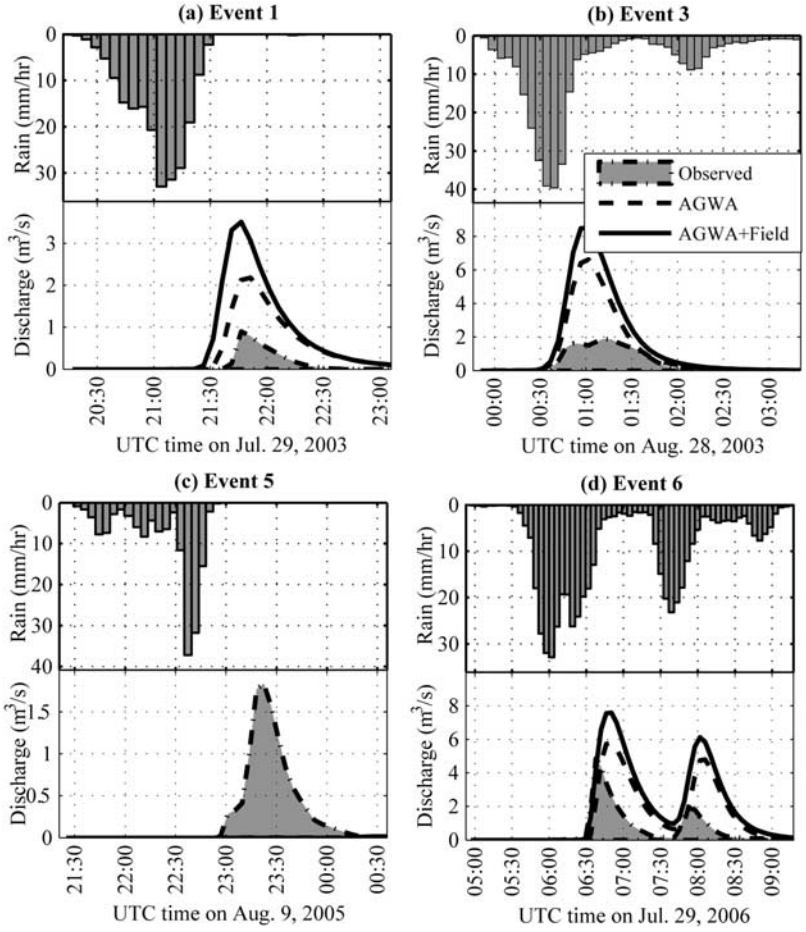

Figure 3. Evaluation of hydrographs obtained from a priori model parameter estimates against observations.

microtopography and channel geometry, which were derived from available field measurements (Table 1). Both the microtopographic roughness and the channel geometry (especially near the outlet) affect the hydrograph, with the microtopography primarily affecting the recession [*Woolhiser et al.*, 1996]. The microtopography parameters rill depth and rill spacing were set from pin meter data [*Bryant et al.*, 2007], with rill depth set as twice the average root mean square height and rill spacing as equal to the correlation length. Channel geometry for selected reaches near the outlet was set to field values from an earlier study [*Goodrich*, 1990].

5. Methodology and Analytic Tools

[25] The four research questions discussed in section 2 were used to develop several model experiments to find answers in the context of the data, model and watershed available for this study. Methods used in these experiments are described below.

5.1. Estimating Rainfall for Model Input

[26] Typically, hail contamination of radar rainfall estimates is simplistically handled by a maximum rain intensity threshold, e.g., 103.8 mm/h in case of the NWS. However, this threshold can vary widely depending on local conditions. For example, *Fulton et al.* [1998] reported values ranging between 75 and 150 mm/h. *Mendez et al.* [2003] analyzed maximum gauge intensities at different durations and frequencies for summer thunderstorms over Walnut Gulch. They found that the maximum 5-min intensities (i.e., at the radar data time step) frequently exceeded 103.8 mm/h, and sometimes were near 250 mm/h. Using radar data, *Morin et al.* [2006] modeled rain cells as circular Gaussian elements with a maximum intensity and a decay factor: these can give theoretical maximum pixel-scale intensities greater than 103.8 mm/h. From the latter two studies, and because flash flood forecasting mainly deals with intense rainfall over short return periods, we raised the hail threshold to a reasonable 146.3 mm/h, which has been estimated as the mean of the maximum 5-min gauge intensities for a 10-year return period [*Mendez et al.*, 2003]. Note that the actual WG11 peak storm intensities are rarely close to, and are mostly less than, this new value (Table 2). However, peak storm intensities much larger than 146.3 mm/h have been observed over the entire WGEW. Raising the hail threshold provided a better match between the cumulative basin-average radar and gauge rainfall series.

[27] The high storm depth biases in the radar estimates (see section 2.1) made the WG11 gauge data with a 0.01″ accuracy [*Keefer et al.*, 2008] more reliable. Following the "bulk adjustment" optimization approach [*Krajewski and Smith*, 2002], the event-specific radar rain estimates were

Table 1. Factors and Their Uncertainty Ranges

Factor	Corresponding uncertainty source type	Factor Notation[a]	Corresponding Uncertainty Source Description	AGWA-Estimated Uncertainty Source Value or Range	User Changes to Uncertainty Source Value/Range	Factor Range
1	Plane element model parameter	PKsM	soil-saturated hydraulic conductivity	8.44 mm/h	-	0.4–2.5
2		PnM	soil surface roughness	0.055–0.057	-	0.4–2.5
3		PCVM	coefficient of variation of PKsM's parameter	0.95	-	0.4–2.0
4		PGM	soil net capillary drive	114.97 mm	regressed against PKsM's parameter	0.67–1.5[c]
5		PRocM	soil volumetric rock fraction	0.43	-	0.75–1.25
6		PIntM	maximum interception depth	2.97–3 mm	-	0.75–1.25
7		PDistM	soil pore size distribution index	0.3	-	0.7–1.3
8		PPorM	soil porosity	0.459	-	0.75–1.25
9		PRillDA	microtopographic rill depth	0	16.8 mm	0–72.5 mm
10		PRillSA	microtopographic rill spacing	0	68.35 mm	0–254 mm
11		PCanM	surface intercepting cover fraction	25%	-	0.75–1.25
12	Channel element model parameter	CKsM	soil-saturated hydraulic conductivity	210 mm/h	-	0.8–1.1
13		CnM	soil surface Roughness	0.035	-	0.4–2.5
14		CCVA	coefficient of variation of CKsM's parameter	0	-	0–2.0
15		CGM	soil net capillary drive	101 mm	regressed against CKsM's parameter	0.4–2.5[c]
16		CRocA	soil volumetric rock fraction	0	-	0–0.1
17		CDistM	soil pore size distribution index	0.545	-	0.7–1.5
18		CPorM	soil porosity	0.44	-	0.85–1.15
19		CWCoM	Woolhiser coefficient	0	0.15	0–3
20		CWidM	bottom width	5.7–14.3 m	selected measurements fixed	0.33–0.7[b]
21		CTortF	channel tortuosity	300–1901 m width, 0.009–0.026 slope	selected measurements fixed	0.95–1.1[b]
22	Plane element initial condition	PSMIA	initial soil moisture	0	event-dependent	0.2–0.6
23	Channel element initial condition	CSMIA	initial soil moisture	0	event-dependent	0.2–0.6
24	Radar rain input	RainM	merged gauge-radar rain	-	-	0.2–1.35

[a]Postfix M, multiplier; postfix A, adder to uncertainty source field values. CTortF (factor 21) affects channel length (multiplied) and slope (divided).
[b]Based on ratio between available information and AGWA values for field-measured sections.
[c]Modifier applied to user-changed, not AGWA-provided values.

improved by scaling to match the basin-average storm totals against those from the gauges, and by adjusting the radar-gauge time lag. These event-specific lags were calculated as the values at which maximum correlation of the WG11 average storm depth series occurred (Table 2). The gauge and improved radar estimates were merged through simple averaging, since there was no reason to prefer one spatial distribution over the other. Figure 2 shows the pattern similarity in some individual and merged cumulative hyetographs.

5.2. Determining a Priori Model Parameters and Their Feasible Ranges

[28] Two groups of a priori model runs were conducted. The first group used the AGWA-generated model parameter values, and the second group used additional field measurements of microtopography and channel geometry. In both groups, soil net capillary drive values (G, [inches]) were recomputed from textural soil saturated hydraulic conduc-

Table 2. Observed Discharge Characteristics and Rain Comparison Statistics for the Study Events

Event	Event Date	Flow Start Time, UTC	Flow Duration, min	Number of Peaks	Peak Flows, m^3/s	Flow Volume, m^3	Maximum Rain Intensity, mm/h: Gauge	Maximum Rain Intensity, mm/h: Radar: Morin Z-R	(Radar/Gauge) Rain Depth Ratio	Radar-Gauge Time Step Lag	Derived "a" Value
1	29 Jul 2003	21:38	83	1	0.94	1153	80.8	45.6	0.35	0	149.9
2	25 Aug 2003	19:16	159	2	0.76, 1.14	3006	80.8	74.9	1.33	0	975.3
3	28 Aug 2003	0:00	200	2	1.62, 1.89	5929	76.2	32.6	0.39	0	175.8
4	9 Aug 2005	4:34	44	1	0.10	82	54.8	11.0	0.23	2	82.4
5	9 Aug 2005	22:57	96	1	1.89	2948	106.7	33.3	0.38	2	167.2
6	29 Jul 2006	5:59	197	2	5.08, 1.98	9224	126.5	49.7	0.41	2	186.3
7	30 Jul 2006	14:30	78	2	0.44, 0.48	1054	45.7	23.9	0.34	1	142.0
8	31 Jul 2006	12:05	88	1	1.49	2611	47.3	25.4	0.38	2	168.4

tivity values (K_s^{text}, [inches/h]) via a regression suggested by *Goodrich* [1990]:

$$G = a\left(1/K_s^{text}\right)^b \tag{3}$$

where the empirical parameters are set to a = 4.83 and b = 0.326. The empirical Woolhiser coefficient was set to 0.15 (factor 19 in Table 1): this coefficient reduces the effective wetted infiltration perimeter during low flows when a trapezoidal channel simplification introduces significant error [*Unkrich and Osborn*, 1987]. The initial soil moisture state (Table 1) was set to 0.4, except in events 7 and 8, where it was set to 0.6. These values were chosen on the basis of general antecedent conditions and from previous experience in calibrating KINEROS2 for Walnut Gulch watersheds.

[29] The number of parameter values for which feasible ranges have to be assigned is generally very high in spatially distributed watershed models. This number has to be reduced since such a high dimensional space is infeasible to sampling using a Monte Carlo approach as was utilized here. A common strategy of constraining the dimensionality of the parameter space preserves the relative spatial magnitudes of the a priori field for any parameter by using multipliers, hence giving an overall magnitude adjustment to that parameter's spatial field of a priori values. This aims to preserve the relative spatial magnitudes in the distributed response of the relevant process, under the assumption of linear variation of that process response with the parameter values [e.g., *Canfield and Lopes*, 2004]. In spite of this linearity assumption being potentially questionable [see *Tang et al.*, 2007], particularly when multiple aspects of the system response are considered, we have used this strategy of multipliers, not only on the parameters, but also on the rain estimates and the initial soil moisture. For parameters with a priori values of zero, additive modifiers (adders) are used instead of multipliers, which together with the multipliers (or multiplicative modifiers) constitute the ''factors'' varied in the sensitivity analysis. For the model parameters, the spatial distributions of the a priori fields are preserved by the corresponding factors in the absence of any additional information. Note that the usage of factors assumes that for any parameter, the component of the sensitivity of the modeled response due to uncertainty in the spatial distribution of that parameter (that is ideally close to the spatial distribution of the a priori field), is negligible. In other words, the sensitivity of the modeled response to any parameter is attributed exclusively to the magnitude bias of its spatial field, and not to its spatial distribution.

[30] Table 1 shows the selected factor uncertainty ranges based on information about corresponding parameter ranges from WGEW and from similar Arizona basins, and from previous studies. Factors are generally listed with prefixes P (for plane) and C (for channel), and postfixes M (for multiplier) and A (for adder). The field-measured channel geometry for selected outlet reaches was not varied. The tortuosity factor (CTortF, factor 21 in Table 1) is related to channel lengths and slopes in reciprocally different manners, i.e., channel lengths are magnified commensurately with reductions in channel slopes to conserve the elevation differences between channel end points.

[31] For the Monte Carlo runs, the saturated hydraulic conductivity (K_s) field was further constrained since AGWA 1.5 does not provide textural K_s^{text} values, but only the final KINEROS2-ready K_s^{final} values along with the parameters used in the K_s^{text}-to-K_s^{final} transformation: the volumetric rock content (V_r, [dimensionless]), the porosity (ϕ, [dimensionless]), the vegetation/litter interception (C_I, [%]), and the fraction of surface impervious area (F_I). This transformation initially calculates the gravimetric rock content (G_r, [dimensionless]):

$$G_r = ((1-\phi)*(1-V_r))/(1-\phi*(1-V_r)) \tag{4}$$

This is used to compute the rock-adjusted K_s^{rock}:

$$K_s^{rock} = K_s^{text} * G_r \tag{5}$$

Finally, K_s^{final} is calculated

$$K_s^{final} = K_s^{rock} * \left(e^{0.0105*C_I}\right) * (1-F_I) \tag{6}$$

For WG11, the AGWA F_I values at 0 are not varied. For channel elements, C_I is zero, making K_s^{final} equal to K_s^{rock}. A reverse transformation on the K_s^{final} values provided the a priori K_s^{text} values. The K_s^{final} values were thus implicitly varied through factors to the corresponding ϕ, V_r and F_I values used in the transformation, and were themselves finally subjected to the corresponding factor variation. No factor was applied to K_s^{text}. Note that this differs from similar studies [e.g., *Canfield and Goodrich*, 2006], where the considered spatial K_s field is only subject to its factor uncertainty without any such implicit variations.

5.3. Defining Hydrograph Shape Descriptors, Objective Functions, and Likelihood Functions

[32] From a forecasting perspective, accurate predictions of hydrograph characteristics such as peak magnitude (and volume), peak timing and channel overbank timing on the hydrograph rising limb are critically important. For timely termination of the issued alarm/warning, additional accuracy in the timing of receding flow falling below the overbank level is also desirable. Therefore, five basic hydrograph shape descriptors were considered: peak magnitude, peak volume, start timing, peak timing and end timing. Nonbasic shape descriptors considered included the ''inflection'' point dividing the hydrograph receding limb into nondriven quick and nondriven slow portions [*Boyle et al.*, 2000]. This point was located using an approximate slope-change-based method, as the time-instant point on the receding limb with the maximum vertical depression from a line segment connecting the peak magnitude and peak end point (Figure 4).

[33] With one exception, the measures of simulation error or objective functions (OFs) (later transformed to likelihoods) were formulated in terms of the absolute magnitude of the relative deviation from the observed value, i.e., $F_s = |S_S - S_O|/S_N$, where S_S and S_O denote values for the simulated and observed shape descriptors, respectively. The normalizer (S_N) is either the observed value or a user-selected constant value (e.g., a representative radar time step of 275 s for timing-based OFs like FTP in Table 3). This formulation makes the OFs comparable across different peaks, and for magnitude-based OFs like FQP (Table 3),

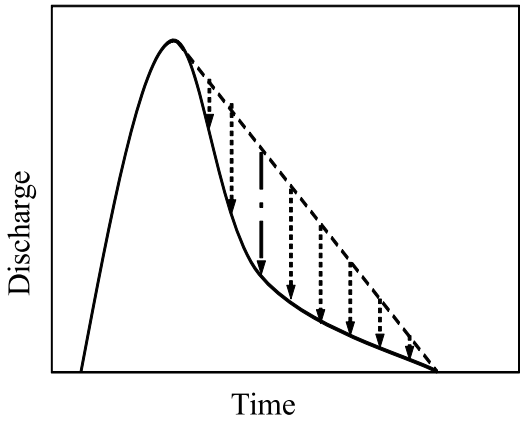

Figure 4. Schematic event hydrograph showing maximum depression (shown as the vertical dash dotted line) method used for selecting the inflection point.

accounts for heteroscedasticity (i.e., nonconstant variance) in the residuals [*Sorooshian and Dracup*, 1980]. Apart from the five basic shape descriptors, other OFs (Table 3) considered flow magnitude, volume, timing, slope, skew and net time series errors. Skew was defined as the hydrograph risetime divided by the total hydrograph time. For the slope-based OFs over different portions of the hydrograph, the magnitude and time terms in the slope equation were normalized using respective observed values to impart a 45° angle to the observed slope. The OFs apply to single peaks, except the last two OFs in Table 3 that apply to the entire event.

[34] The OFs were used to formulate the likelihood functions (LFs) required by both by GLUE and the sensitivity analysis, with the prefix "L" applied in this study when referring to the LF associated with a particular OF (see Table 3). Zero values were assigned to the LFs classified as nonbehavioral, and the remaining LF values were obtained by an additive inversion of the OFs in which they were first rescaled on a 0.1–1 range, and then again rescaled to sum to one. The wide 0.1–1 range in the first rescaling attempted to impart power to the sensitivity analysis of the behavioral region variations, where a narrow range would have essentially devolved to a binary classification. For simulations with insufficient basis to classify the descriptor as nonbehavioral because either the observed or the simulated shape descriptor value did not exist (e.g., connected to the inflection point), the LF values were set to the midpoint of the behavioral LF value range.

[35] A Monte Carlo set membership method [*Van Straten and Keesman*, 1991] was used to establish the overall behavior of a Monte Carlo simulation, with "LFB" denoting the likelihoods based on this binary behavioral/nonbehavioral classification. *Van Straten and Keesman* [1991] defined a behavioral zone with specified upper and lower bounds around the entire observational time series, and simulations were classified as behavioral if they fell fully inside these bounds. Here, we instead defined upper and lower bounds around the five basic shape descriptor values to define a behavioral range. The maximum deviation on peak magnitude (FQP) was set to ±0.15 (i.e., ±15%), while the timing descriptor OFs (FTS, FTP, and FTE, see Table 3) were constrained to within ±2, ±2 and ±3 radar time steps, respectively (i.e., endpoint timing was allowed more flexibility). Because the aforementioned OFs do not explicitly consider hydrograph start and/or end values, the maximum deviation on volume (FVU) was set to ±0.5 to filter out simulations that deviate significantly from observed flows. The behavioral LFB LF only provides information about the binary hydrograph classification, while the "original" shape-descriptor-based LFs represent variations in the respective behavioral portions of the factor space. Hence, we also formulated "enhanced" shape-descriptor-based LFs to represent only their variations in the factor space behavioral to LFB by combining information both from the respective original LFs and from LFB.

5.4. Combining Likelihood Functions

[36] Likelihood functions (LFs) were combined to evaluate the overall performance of individual factor sets using two different methods. For individual events, the multiobjective LF was obtained by multiplication of the individual shape-based LFs. However, for updating across events, the LFs were combined by addition thereby allowing for an inconsistency in the behavioral factor sets across events, i.e., a factor set was able to simulate one event *or* the other,

Table 3. Objective Functions Considered in This Study[a]

OF	OF Notation	OF Type	OF Description
1	FQF	magnitude-based	relative deviation from observed ratio of inflection to peak magnitude
2	**FQP**	magnitude-based	**relative deviation from observed peak magnitude**
3	FQI	magnitude-based	relative deviation from observed inflection magnitude
4	FSDQ	slope-based	relative deviation from slope angle of observed driven part
5	FSNQ	slope-based	relative deviation from slope angle of observed nondriven quick part
6	FSNS	slope-based	relative deviation from slope angle of observed nondriven slow part
7	**FTS**	based on time relative to rain event	**relative deviation from observed time to start**
8	**FTP**	based on time relative to rain event	**relative deviation from observed time to peak**
9	**FTE**	based on time relative to rain event	**relative deviation from observed time to end**
10	FTR	hydrograph time duration-based	relative deviation from observed time of rise
11	FTF	hydrograph time duration-based	relative deviation from observed time of fall
12	FS	hydrograph time duration-based	relative deviation from observed hydrograph skew
13	**FVU**	volume-based	**relative deviation from observed peak volume if multipeak**
14	FV	volume-based	relative deviation from observed event volume
15	NSC	observation series error-based	1-(Nash-Sutcliffe efficiency, i.e., NSE)

[a]Objective functions (OFs), basic OFs as mentioned in section 5.3 are shown in bold. Corresponding likelihood functions have an L prefix.

Table 4. Evaluation Results of a Priori Model Runs[a]

	AGWA-Generated Parameters			AGWA Plus Field Estimates			
OFs	Mean	Min	Max	Mean	Min	Max	Change[b]
FQP	1.10	0.20	2.52	1.42	0.07	3.53	−0.32
FTS	2.79	0.92	5.50	1.46	0.00	3.67	1.33
FTP	3.21	0.00	10.35	1.46	0.00	2.75	1.75
FTE	4.44	2.16	7.59	3.31	1.08	5.41	1.13
FVU	1.42	0.66	3.03	1.97	0.57	5.08	−0.55
FV	1.23	0.80	3.01	1.72	0.05	5.11	−0.49
NSE	−2.00	−5.43	−0.45	−6.33	−21.15	0.15	−4.33

[a]Mean, Min, and Max refer to the mean, minimum, and maximum value across events.
[b]Positive indicates improvement.

but not both [see also *Zak et al.*, 1997]. For the latter, each hydrograph-likelihood was weighted by the number of behavioral simulations obtained for that hydrograph before a combined LF was calculated.

6. Results and Discussion

6.1. A Priori Parameter Model Runs

[37] Both groups of a priori runs (as defined in section 5.2) did not match the observations well (see Figure 3 for selected events). Table 4 lists the ranges and means of the basic objective functions (OFs), the volume-based FV OF, and the Nash-Sutcliffe efficiencies (NSEs) for all events, along with the changes observed in moving from the first parameter set (AGWA generated parameters only) to the second parameter set (AGWA generated with additional field observations for some parameters). Since the OFs are normalized, a positive difference in mean values over all runoff events represents improvement. This improvement is seen to occur mostly in the OFs related to hydrograph timing (FTS, FTP, and FTE, see Table 3), while peak (FQP) and volume (FVU and FV) deteriorate. The overall NSE also gets worse (note that for NSE, higher values are better). Further, most of the NSE values are less than the behavioral threshold selected to be zero. This result clearly highlighted the need for adjustment of the a priori parameter values based on available observations of the watershed streamflow response.

6.2. Intercomparison of Rainfall Estimates and Their Streamflow Simulations

[38] Gauge, radar, and merged rainfall estimates show complete WG11 storm coverage for almost all events. We intercompared all three estimates by first examining the corresponding areal average cumulative rain depth series. Figures 2a (top)–2d (top) show the data for four events. As expected, the NWS and the *Morin et al.* [2005] radar estimates match in shape, since the revised hail threshold did not cut off the peak intensities here. However, discrepancies in magnitude and in shape are seen between gauge and radar estimates, likely due to evaporation impacts and the radar-gauge time lag, or both. These discrepancies may also be due to radar operational problems, such as the clutter suppression effect wherein the radar algorithm may mistake a highly stationary rain field as background clutter and filter it out. The reflectivity (Z) estimates for event 1 are believed to be contaminated with this type of error (E. Pytlak, NOAA NWS, personal communication, 2006). Table 2 also shows a significant range in the ratio of the *Morin et al.* [2005] radar reflectivity-rain (Z-R) estimates to gauge values.

[39] The effect of errors in the spatial rainfall representation on the streamflow predictions was examined in a limited way by comparing the behavioral simulations from

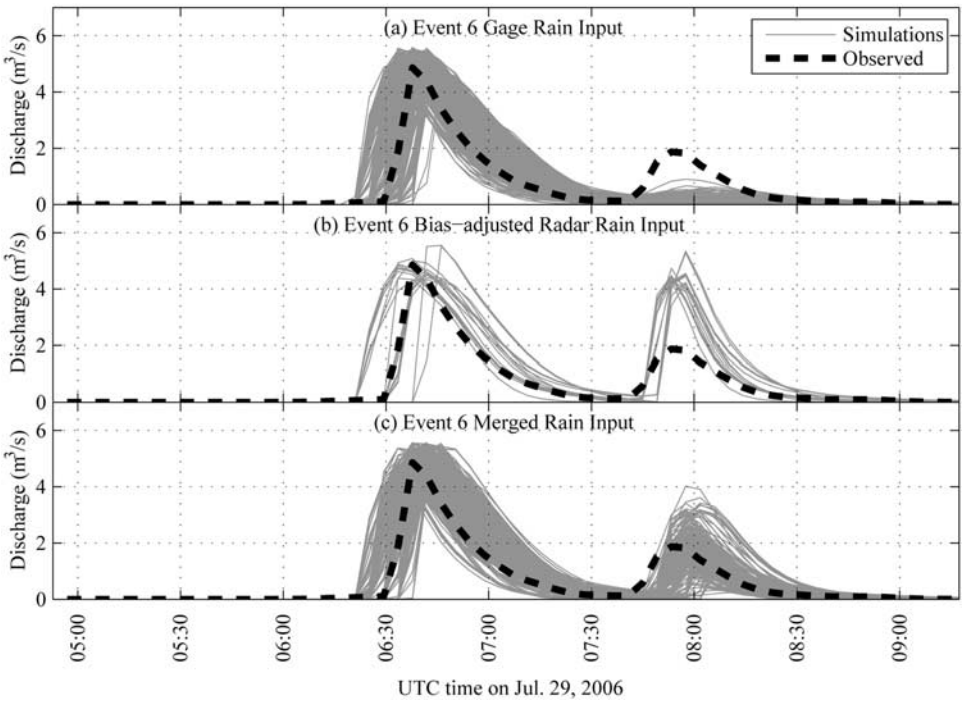

Figure 5. Simulations selected as behavioral to the first (left) peak of this two peak event hydrograph (event 6). Performance on second peak differs with rainfall source used to drive model.

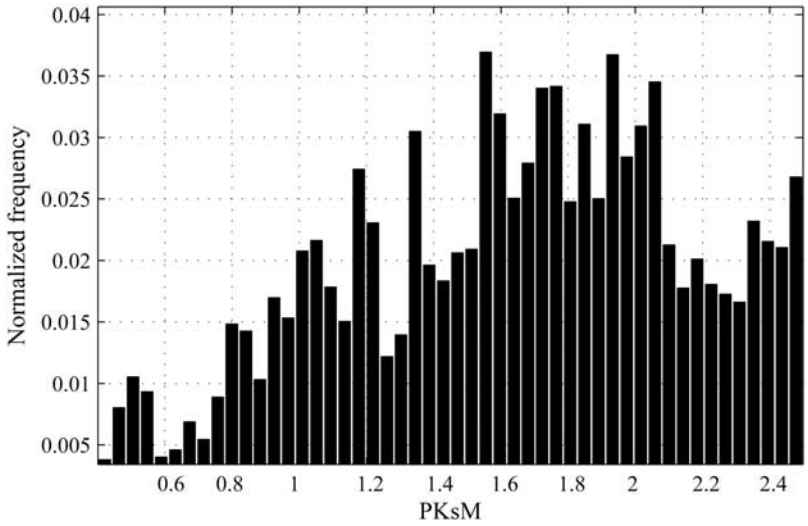

Figure 6. Example posterior distribution of the all event behavioral region over the range of the PKsM factor (corresponding to the plane soil saturated hydraulic conductivity).

three different 23-factor Monte Carlo sample populations driven by the merged rain input, the gauge input and the *Morin et al.* [2005] radar estimates, respectively. This comparison indicated the general superiority of the merged estimates in two ways: they provided better consistency in the behavioral factor sets across hydrograph peaks (even in multipeak events), and they replicated the observations better by providing behavioral simulations for a larger number of the observed peaks. To demonstrate the first improvement, Figure 5 shows, for each rainfall input, the simulations behavioral with respect to the first peak of event 6. For the second peak, these simulations are seen to be nonbehavioral for both gauge (consistent underestimation) and radar (consistent overestimation), but are mostly behavioral for the merged case. To exemplify the second improvement, the number of simulations behavioral to the second peak of event 6 was examined: these were 262, 69, and 0 for the merged, gauge, and improved radar estimates, respectively, thus again suggesting that the merged product is preferable.

[40] However, even considering only merged rainfall estimates, a high degree of variability in the behavioral factor sets was still observed across events. This suggests that updating-based methods for parameter estimation might be better suited than optimization-based methods for modeling ephemeral streams in dry regions. Increasing the acceptable likelihood ranges to increase the population of behavioral models did not reduce the problem of variability of behavioral parameter sets between events.

6.3. Behavioral Ranges and Correlation Structures in Factors and Likelihood Functions

[41] Behavioral solutions could be found throughout the a priori factor ranges for either of the Monte Carlo (MC) collections (refer to section 3.3). However, some restructuring tendency of the behavioral distributions away from the initial uniformly sampled pattern was seen for factors PKsM, PnM, CnM, CWCoM, and RainM (refer to Table 1). PKsM showed higher frequencies in the 1.5–2 range for the posterior distribution plots (Figure 6), suggesting a preference for larger hillslope-scale soil saturated hydraulic conductivities, and correspondingly higher infiltration. Similarly, factors PnM, CnM and CWCoM showed higher frequencies for values less than 1.5. RainM showed a clear tendency toward values greater than 0.5 with a maximum near 1, due to the low radar-gauge rainfall depth ratios (Table 2). Apart from these five factors, no tendency toward a deviation of factors from uniform distributions was found.

[42] Also, no significant correlation was found using a two-factors-at-a-time analysis. However, moderate correlations were observed for some individual peaks: PKsM versus RainM in the first MC run collection (0.6–0.7 range), and PKsM versus PRocM in the second MC run collection (0.45–0.62 range; refer to Table 1 for PRocM). These positive correlations can be explained by the co-occurrence of these factor pairs either as quotients or differences in the model equations. RainM affects the rain rates at the land surface that occur only in the model infiltrability equations after being scaled using K_s^{final} (refer to section 5.2), i.e.,

$$r_* = r/K_s^{final} \tag{7}$$

where r is the surface rain rate. This division by K_s^{final} results in a positive correlation. PRocM affects the vadose zone behavior only through model equations that calculate K_s^{final} from the G_r manifestation of V_r (see equations (4) to (6)), where the presence of K_s^{final} and G_r in the numerators on both sides of equation (6) (i.e., after substituting equation (5)) results in a positive correlation. Aside, a principal components analysis also showed no strong directional components in the factor space.

[43] The histograms of the original and the enhanced Likelihood Functions (LFs; refer to Table 3) showed no peaked distribution, hence showing no loss in the discriminatory power of the sensitivity analysis due to a good variation in the LF values. In the 2-LF correlation plots, the five basic original LFs showed strong correlations (i.e., >0.75) for LFTS versus LFTP (i.e., start timing versus peak timing), and for LFQP versus LFVU (i.e., peak magnitude versus volume). Hence, for a multiobjective analysis, the three LFs selected were LFQP, LFTP and LFTE (i.e., end timing). We henceforth

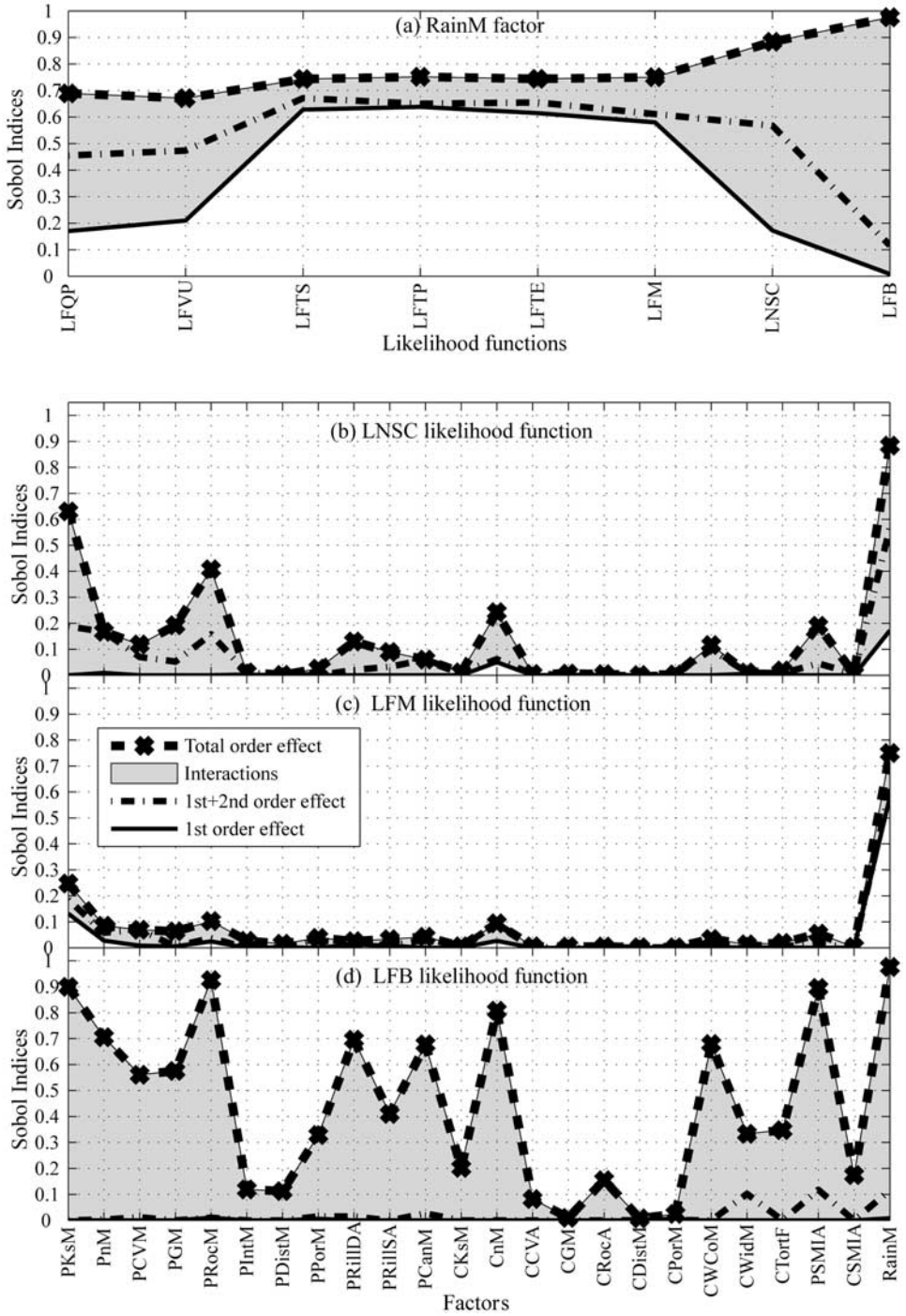

Figure 7. Sobol' indices showing sensitivity of model output to variations (a) in rainfall multiplier, RainM, through selected likelihood functions (LFs) and showing sensitivity of three different LFs to all factors: (b) Nash Sutcliffe efficiency based LNSC, (c) Multiobjective LFM, and (d) Behavioral LFB. RainM is rightmost factor in Figures 7b–7d.

denote the combined multiobjective LF of these three LFs as "LFM." The five basic enhanced LFs showed no strong mutual correlation, suggesting that all of them should be used for formulating an enhanced multiobjective LF.

6.4. Sensitivity Analysis

[44] Using the enhanced likelihood functions (LFs) resulted in almost the same sensitivity indices as the behavioral classification-based LFB. This showed the loss of power to extract sensitivity information using the enhanced LFs which were supposedly representative of the variations in the behavioral regions. This problem was due to the dominance of the binary behavioral classification on the sensitivity indices over these variations in the calculation of these sensitivity indices when the percentage of behavioral simulations is very low (<1% here). Hence, the enhanced LFs were not considered in further analyses.

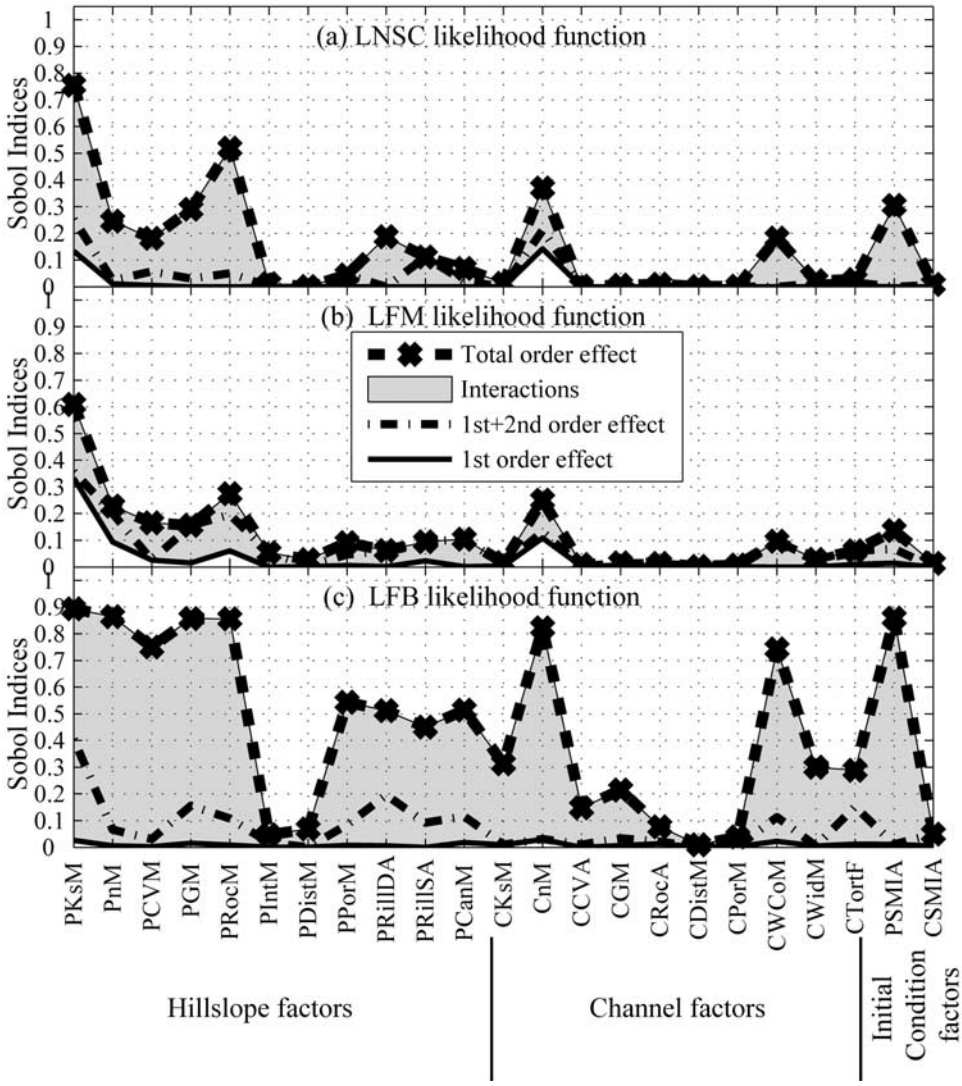

Figure 8. Sobol' indices showing relative influence of model parameters and initial conditions on the model response. Individual plots show fractional influence on: (a) Nash Sutcliffe efficiency based LNSC, (b) multiobjective LFM, and (b) behavioral LFB.

[45] In the first Monte Carlo (MC) run collection, the merged rain estimates had a factor range of 0.2 to 1.35 (see factor 24 in Table 1), reflecting the radar depth bias across the study events (Table 2). Figure 7a shows the all-event influence of the rain depth bias factor (RainM; refer to Table 1). The total-order indices show the dominant influence of RainM, highlighting the importance of depth bias-free rain inputs for accurate forecasts. LFB has a second-order effect, though most of the individual peaks have negligible second- and first-order effects (not shown). This would have made it difficult to detect the rain influence from the behavioral classification using sensitivity analysis (SA) methods which do not explicitly consider interaction at all orders (e.g., RSA or regional SA with a second-order factor correlation analysis). The rain influence using RSA is probably better detected from LFs like the multiobjective LFM on which RainM is seen to have a significant first-order effect. Also, use of identifiability plots may help to constrain the RainM uncertainty in regions without rain gauge measurements, potentially allowing the universal application of radar rain estimates under the current large depth biases.

[46] Figures 7b and 7c show the high relative rainfall estimate influence as compared to other factors for the Nash-Sutcliffe efficiency-based LNSC and the multiobjective LFM LFs, respectively. LNSC seemed slightly more sensitive to the model factors compared to LFM (LNSC similarly compared against LFM's single-objective components LFQP, LFTP, and LFTE; refer to Table 3 for these LFs). The influence of the PKsM factor (refer to Table 1) on the LNSC LF was seen to approach that of RainM: Sobol' SA results showed a significant 37% RainM-PKsM second-order effect out of the total PKsM index (see section 6.3 regarding this interaction). Figure 7d for the behavioral LFB LF (refer to section 5.3) shows the almost complete influence of RainM, with generally very small first-order effects. This happens because LFB is an implicit combination of the basic hydrograph shape descriptors, and hence of factors which influence each of the descriptors, with consequent high-order interactions in the LFB combination. Hence, we

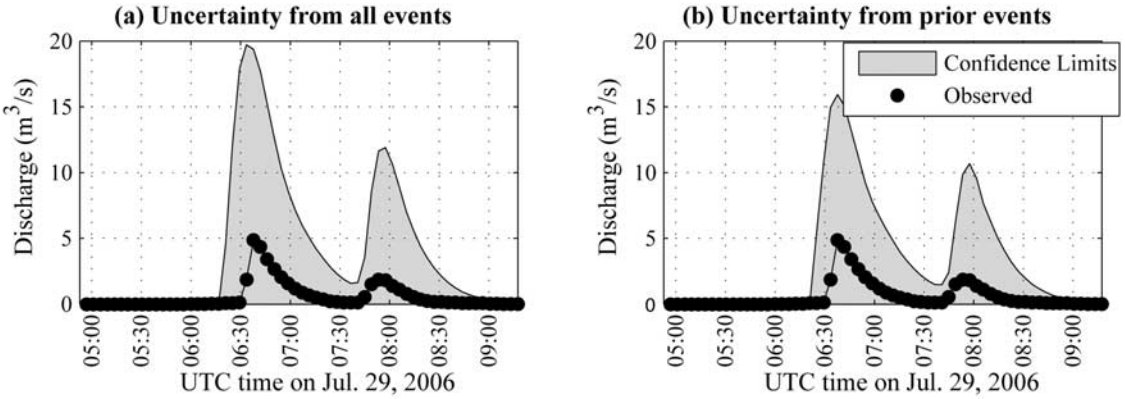

Figure 9. The 90% confidence intervals incorporating radar rainfall uncertainty for prediction of the 29 July 2006 event using behavioral factor sets from (a) all events and (b) prior events only.

answer the second research question about rain influence as follows: While merging of high spatial resolution gauge and radar estimates might potentially provide better simulations than either one by itself, the influence of rain depth/volume bias uncertainty almost completely dominates the model response.

[47] Figure 8 shows similar SA results from the second collection of MC runs. Figures 8a, 8b, and 8c show factor influences using the LNSC, LFM and LFB LFs, and are similar to the respective Figures 7b, 7c, and 7d. Small differences occur due to the removal of the rain depth bias uncertainty. Among themselves, all the original LFs uniformly show sensitivity to almost the same few factors in the same order, which hinders a detailed analysis about which factors affect which specific hydrograph shape descriptors. The parameters corresponding to the three most influential hillslope factors are mostly the hillslope soil saturated hydraulic conductivity (i.e., PKsM factor), the hillslope soil volumetric rock fraction (i.e., PRocM), and the hillslope soil surface roughness (i.e., PnM). Similarly, the parameters corresponding to the most influential channel factors are the channel soil surface roughness (i.e., CnM) and the channel Woolhiser coefficient (i.e., CWCoM). The influence of CWCoM, undetected in earlier studies, shows the importance of estimating the effective channel cross-sectional wetted perimeter through the Woolhiser coefficient for accurate infiltration calculations. Note that the five factors observed to be identifiable in section 6.3 are also the most sensitive.

[48] Overall, hillslope factors dominate the model response more than the channel factors, showing the importance of detailed field measurement and scaling studies of hillslope parameters for reducing corresponding uncertainties in semiarid flash flood forecasting in mechanistic models of small basins. Over larger basins, channel factors will likely become more influential due to the greater impact of channel infiltration losses [*Goodrich et al.*, 1997]. Hence, we answer the fourth research question about model parameter influence as follows: In general, uncertainties in the hillslope model parameters impact predictions more than those in the channel parameters for small basins. The most influential hillslope parameter uncertainties stem from the soil saturated hydraulic conductivity, the soil volumetric rock fraction, and the soil surface roughness, while those for the channel are from the soil surface roughness and the Woolhiser coefficient.

[49] Figures 7 and 8 consistently show that initial soil moisture specification in channels (CSMIA factor) is redundant for the model response. However, the hillslope initial soil moisture (PSMIA factor) can be important, and for the behavioral LFB LF which is particularly relevant for flash flood forecasting, it is almost as influential as PKsM. Hence, the third research question about the initial soil moisture influence can be answered as follows: Initial hillslope soil moisture can have a dominant effect on the predicted response in flash flood forecasting. Hence, more sophisticated interstorm model components are required to track hillslope soil moisture with a high degree of accuracy.

6.5. Model Predictive Uncertainty

[50] Considering the high behavioral response inconsistency across events, and the associated differences in the model responses, a high predictive uncertainty is expected. Equispaced values over the feasible RainM factor range (refer to Table 1) were combined with the behavioral factor sets from the second Monte Carlo run collection to obtain Figures 9a and 9b, which show the predicted 90% confidence intervals for event 6 using behavioral factor sets from all and from prior events only, respectively. These uncertainties shown are higher than desired for flash flood forecasting (e.g., more than 10 times in the peak magnitude here). Hence, the answer to the first research question posed earlier is as follows: Given current uncertainties in the radar rain estimates, model parameters and initial conditions, the predictive uncertainties in flash flood forecasting can be much higher than desired.

7. Conclusions and Future Work

[51] There is a clear and pressing need for improvements in our ability to generate operational flash flood forecasts in the semiarid southwestern United States and other dry regions around the world. This study investigated the predictive uncertainty of the physically based, distributed semiarid rainfall-runoff model KINEROS2 driven by high-resolution radar rainfall input for this purpose. Uncertainty sources considered were rainfall estimates, model parameters, and initial moisture conditions. The variance-based Sobol' global sensitivity analysis method was used to

investigate dominant sources of uncertainty. The flash flood forecasting system was implemented using the GLUE methodology to facilitate operational assimilation of incoming event information. The approach is applicable to any model amenable to a Monte Carlo framework, and can be implemented operationally for a fixed number of factor sets without requiring resampling or optimization. The Monte Carlo framework utilized factors applied to the rainfall, initial soil moisture and model parameters.

[52] The methodology was applied to simulate the runoff response for eight summertime convective thunderstorm events occurring over the WG11 subbasin of the semiarid Walnut Gulch Experimental Watershed. The Walnut Gulch runoff regime is characterized by ephemeral flow with long periods of no-flow between storm events. Hydrograph fits were characterized by shape descriptors that represent magnitude, volume, shape, and timing. In decreasing order of importance, the dominant uncertainties were found to originate from the bias in the radar rainfall depth estimates, the model parameters, and the initial soil moisture conditions.

[53] In particular, we found the following:

[54] 1. The uncertainty due to depth/volume bias in the radar rainfall estimates almost completely dominated the uncertainty in the modeled response.

[55] 2. Merging of high spatiotemporal resolution rainfall estimates from gauge and radar has the potential to improve the model forecasting ability.

[56] 3. Uncertainties in the hillslope model parameter values had greater impact on the predictions than uncertainties in the channel parameters. For hillslopes, the three most influential parameters were soil saturated hydraulic conductivity, the soil volumetric rock fraction and the soil surface roughness. For channels, most influential were soil surface roughness and the Woolhiser coefficient. The latter parameter reduces the effective wetted perimeter for infiltration during low flows when a trapezoidal channel simplification introduces significant error.

[57] 4. The initial channel soil moisture does not significantly influence the modeled response, whereas the initial hillslope soil moisture can have a strong effect at least in small basins as analyzed here.

[58] 5. Given the typical level of uncertainty in currently available radar rainfall estimates, model parameters and initial conditions, the predictive uncertainty in the modeled flash flood response is often likely to be much higher than what would be considered acceptable for accurate flash flood forecasting. This result is likely to be similar for virtually any flash flood forecasting model due to the dominance of rainfall uncertainty, and improved rainfall estimates are the most needed step to improve forecasting skills.

[59] 6. Behavioral parameter sets of one event do often not remain behavioral for another event, resulting in a corresponding inconsistency across events. The inconsistency of behavioral parameter sets between events suggests that a real-time parameter updating procedure is likely to be more useful than a calibration approach which optimizes the model with respect to past events.

[60] This study illustrates the considerable difficulty involved in the identification of models for flood forecasting in semiarid regions. The most pressing concern is that improved real-time bias-free rainfall estimates are required for achieving reduced uncertainty in flood forecasts. Given the capabilities of the current generation of weather radars, such real-time bias removal may only be possible when used in conjunction with sophisticated numerical weather prediction models. The influential uncertainties found for poorly defined parameters indicate the need for better parameter estimates in the mechanistic model, either through detailed scaling-related field studies predominantly of the hillslope processes, and/or through improved strategies in investigating high-dimensional parameter spaces that consider potential errors in the spatial distribution. Our findings regarding the dominant effect of initial hillslope soil moisture states indicates a need for improved continuous soil moisture accounting. In general, the complexity of semiarid hydrometeorological processes that is represented in correspondingly complex semiarid watershed models, and the high degree of observational error (introduced through the actual observation and through the data processing), still make studies of the kind presented here challenging. Improvements in all aspects are needed to advance operational prediction accuracy and to enable hypotheses testing in scientific studies.

[61] **Acknowledgments.** Partial financial support for this research was provided by the UCAR/NOAA/COMET Program (S0344674), the NOAA/NWS-HL grant (NA04NWS462001), the SAHRA NSF Science and Technology Center (EAR-9876800), and the Salt River Research Fellowship program. Soren Scott (University of Nebraska) and Shea Burns (USDA-ARS) are gratefully acknowledged for the AGWA 1.5 customization. Stefano Tarantola (EC/JRC Italy) is gratefully acknowledged for providing the SIMLAB software.

References

American Meteorological Society (1985), Flash floods—A statement of concern by the American Meteorological Society, *Bull. Am. Meteorol. Soc.*, *66*(7), 858–859.

Anderson, E. A. (1993), A continuous, incremental antecedent precipitation index (API) model for use with the National Weather Service River Forecast System, in *Engineering Hydrology, Proceedings of the Symposium Sponsored by the Hydraulics Division of ASCE*, edited by C. Y. Kuo, pp. 934–939, Am. Soc. of Civ. Eng., New York.

Beven, K. J. (1993), Prophecy, reality and uncertainty in distributed hydrological modeling, *Adv. Water Resour.*, *16*, 41–51, doi:10.1016/0309-1708(93)90028-E.

Beven, K. (2002), A perceptual model for runoff generation in semi-arid areas, in *Dryland Rivers: Hydrology and Geomophology of Semi-arid Channels*, edited by L. J. Bull and M. J. Kirkby, pp. 57–105, John Wiley, New York.

Beven, K. J., and A. M. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298, doi:10.1002/hyp.3360060305.

Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, *J. Hydrol.*, *249*, 11–29, doi:10.1016/S0022-1694(01)00421-8.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrological models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*(12), 3663–3674, doi:10.1029/2000WR900207.

Bryant, R., M. S. Moran, D. P. Thoma, C. D. Holifield-Collins, S. Skirvin, M. M. Rahman, K. Slocum, P. Starks, D. Bosch, and M. P. Gonzalez-Dugo (2007), Measuring surface roughness to parameterize radar backscatter models for retrieval of surface soil moisture, *IEEE Geosci. Remote Sens. Lett.*, *4*(1), 137–141, doi:10.1109/LGRS.2006.887146.

Burnash, R. J. C. (1995), The NWS River Forecast System—Watershed modeling, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 311–366, Water Resour. Publ., Highlands Ranch, Colo.

Canfield, H. E., and D. C. Goodrich (2006), The impact of parameter lumping and geometric simplification in modelling runoff and erosion in the shrublands of southeast Arizona, *Hydrol. Processes*, *20*(1), 17–35, doi:10.1002/hyp.5896.

Canfield, H. E., and V. L. Lopes (2004), Parameter indentification in a two-multiplier sediment yield model, *J. Am. Water Resour. Assoc.*, *40*(2), 321–332.

Creutin, J.-D., and M. Borga (2003), Radar hydrology modifies the monitoring of flash-flood hazard, *Hydrol. Processes*, *17*, 1453–1456, doi:10.1002/hyp.5122.

Davis, R. S., and P. Jendrowski (1996), The Operational Areal Mean Basin Estimated Rainfall (AMBER) module, paper presented at 15th Conference on Weather Analysis and Forecasting, Am. Meteorol. Soc., Norfolk, Va.

Faurès, J.-M., D. C. Goodrich, D. A. Woolhiser, and S. Sorooshian (1995), Impact of small-scale rainfall variability on runoff simulation, *J. Hydrol.*, *173*, 309–326, doi:10.1016/0022-1694(95)02704-S.

Fulton, R. A., J. P. Breidenbach, D. J. Seo, and D. A. Miller (1998), The WSR-88D rainfall algorithm, *Weather Forecast.*, *13*, 377–395, doi:10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2.

Goodrich, D. C. (1990), Geometric simplification of a distributed rainfall-runoff model over a range of basin scales, Ph.D. dissertation, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson.

Goodrich, D. C., T. J. Schmugge, T. J. Jackson, C. L. Unkrich, T. O. Keefer, R. Parry, L. B. Bach, and S. A. Am (1994), Runoff simulation sensitivity to remotely sensed initial soil water content, *Water Resour. Res.*, *30*(5), 1393–1405, doi:10.1029/93WR03083.

Goodrich, D. C., L. J. Lane, R. A. Shillito, S. N. Miller, K. H. Syed, and D. A. Woolhiser (1997), Linearity of basin response as a function of scale in a semiarid watershed, *Water Resour. Res.*, *33*(12), 2951–2965, doi:10.1029/97WR01422.

Goodrich, D. C., et al. (2000), Preface paper to the Semi-Arid Land-Surface-Atmosphere (SALSA) program special issue, *Agric. For. Meteorol.*, *105*(1–3), 3–20, doi:10.1016/S0168-1923(00)00178-7.

Goodrich, D. C., C. L. Unkrich, and R. E. Smith (2006), KINEROS2—New features and capabilities, paper presented at the Joint 8th Federal Interagency Sedimentation Conference and 3rd Federal Interagency Hydrologic Modeling Conference, Adv. Comm. on Water Inf., Reno, Nevada, 2–6 April.

Goodrich, D. C., T. O. Keefer, C. L. Unkrich, M. H. Nichols, H. B. Osborn, J. J. Stone, and J. R. Smith (2008), Long-term precipitation database, Walnut Gulch Experimental Watershed, Arizona, United States, *Water Resour. Res.*, *44*, W05S04, doi:10.1029/2006WR005782.

Grayson, R., and G. Bloeschl (2000), Spatial Processes, Organizations and patterns, in *Spatial Patterns in Catchment Hydrology*, edited by R. Grayson and G. Bloeschl, pp. 3–16, Cambridge Univ. Press, Cambridge, U. K.

Hardegree, S., et al. (2003), Multi-watershed evaluation of WSR-88D (NEXRAD) radar-precipitation products, in *Proceedings of the 1st Interagency Conference on Research in the Watersheds*, edited by K. G. Renard et al., pp. 486–491, Benson, Ariz., 27–30 Oct.

Homma, T., and A. Saltelli (1996), Importance measures in global sensitivity analysis of model output, *Reliab. Eng. Syst. Safety*, *52*(1), 1–17, doi:10.1016/0951-8320(96)00002-6.

Hong, Y., K.-L. Hsu, X. Gao, and S. Sorooshian (2004), Precipitation estimation from remotely sensed imagery using artificial neural network—Cloud classification system, *J. Appl. Meteorol.*, *43*(12), 1834–1853, doi:10.1175/JAM2173.1.

Keefer, T. O., C. L. Unkrich, J. R. Smith, D. C. Goodrich, and J. R. Simanton (2008), An event-based comparison of two types of automated-recording weighting bucket rain gauges, *Water Resour. Res.*, doi:10.1029/2006WR005841, in press.

Kitanidis, P. K., and R. L. Bras (1980a), Real-time forecasting with a conceptual hydrologic model: 1. Analysis of uncertainty, *Water Resour. Res.*, *16*(6), 1025–1033, doi:10.1029/WR016i006p01025.

Kitanidis, P. K., and R. L. Bras (1980b), Real-time forecasting with a conceptual hydrologic model: 2. Applications and results, *Water Resour. Res.*, *16*(6), 1034–1044, doi:10.1029/WR016i006p01034.

Krajewski, W. F., and R. A. Smith (2002), Radar hydrology: Rainfall estimation, *Adv. Water Resour.*, *25*, 1387–1394, doi:10.1016/S0309-1708(02)00062-3.

Kundzewicz, Z. W. (2002), Non-structural flood protection and sustainability, *Water Int.*, *27*(1), 3–13.

Kundzewicz, Z. W., and Z. Kaczmarek (2000), Coping with hydrological extremes, *Water Int.*, *25*(1), 66–75.

Maddox, R. A., J. Zhang, J. J. Gourley, and K. W. Howard (2002), Weather radar coverage over the contiguous United States, *Weather Forecast.*, *17*, 927–934, doi:10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2.

Mendez, A., D. C. Goodrich, and H. B. Osborn (2003), Rainfall point intensities in an air mass thunderstorm environment: Walnut Gulch, Arizona, *J. Am. Water Resour. Assoc.*, *39*(3), 611–621, doi:10.1111/j.1752-1688.2003.tb03679.x.

Michaud, J., and S. Sorooshian (1994), Effects of rainfall sampling errors on simulation of desert flash floods, *Water Resour. Res.*, *30*, 2765–2775, doi:10.1029/94WR01273.

Miller, S. N., D. J. Semmens, D. C. Goodrich, M. Hernandez, R. C. Miller, W. G. Kepner, and D. P. Guertin (2007), The automated geospatial watershed assessment tool, *Environ. Model. Software J.*, *22*(3), 365–377, doi:10.1016/j.envsoft.2005.12.004.

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

Moran, M. S., J. J. Stone, K. G. Renard, P. Heilman, D. C. Goodrich, and T. O. Keefer (2008), Preface to special section on Fifty Years of Research and Data Collection: U.S. Department of Agriculture Walnut Gulch Experimental Watershed, *Water Resour. Res.*, doi:10.1029/2007WR006083, in press.

Morin, E., R. A. Maddox, D. C. Goodrich, and S. Sorooshian (2005), Radar Z-R relationship for summer monsoon storms in Arizona, *Weather Forecast.*, *20*(4), 672, doi:10.1175/WAF878.1.

Morin, E., D. C. Goodrich, R. A. Maddox, X. Gao, H. V. Gupta, and S. Sorooshian (2006), Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response, *Adv. Water Resour.*, *29*(6), 843–860, doi:10.1016/j.advwatres.2005.07.014.

National Weather Service (2002), Advanced hydrologic prediction services—Concept of services and operations, report, U.S. Dep. of Commer., NOAA, Silver Spring, Md.

Nicolau, J. M., A. Solé-Benet, J. Puigdefàbregas, and L. Gutiérrez (1996), Effects of soil and vegetation on runoff along a catena in semi-arid Spain, *Geomorphology*, *14*, 297–309, doi:10.1016/0169-555X(95)00043-5.

Osborn, H. B. (1964), Effect of storm duration on runoff from rangeland watersheds in the semiarid southwestern United States, *Bull. Int. Assoc. Sci. Hydrol.*, *9*(4), 40–47.

Parlange, J.-Y., I. Lisle, R. D. Braddock, and R. E. Smith (1982), The three-parameter infiltration equation, *Soil Sci.*, *133*(6), 337–341.

Pilgrim, D. H., T. G. Chapman, and D. G. Doran (1988), Problems of rainfall–runoff modelling in arid and semiarid regions, *Hydrol. Sci. J.*, *33*(4), 379–400.

Ratto, M., S. Tarantola, and A. Saltelli (2001), Sensitivity analysis in model calibration: GSA-GLUE approach, *Comput. Phys. Commun.*, *136*, 212–224, doi:10.1016/S0010-4655(01)00159-X.

Renard, K. G., L. J. Lane, J. R. Simanton, W. E. Emmerich, J. J. Stone, M. A. Weltz, D. C. Goodrich, and D. S. Yakowitz (1993), Agricultural impacts in an arid environment: Walnut Gulch case study, *Hydrol. Sci. Technol.*, *9*, 145–190.

Roeske, R. H., J. M. Garrett, and J. H. Eychaner (1989), Floods of October 1983 in southeastern Arizona, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *98-4225-c*.

Saltelli, A. (2002), Making best use of model evaluations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*, 280–297, doi:10.1016/S0010-4655(02)00280-1.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto (2004), *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, 219 pp., John Wiley, New York.

Semmens, D., D. C. Goodrich, C. L. Unkrich, R. E. Smith, D. A. Woolhiser, and S. N. Miller (2005), KINEROS2 and the AGWA modeling framework, paper presented at the GWADI International Modeling Workshop, UNESCO, Roorkee, India.

Smith, R. E., D. C. Goodrich, D. A. Woolhiser, and C. L. Unkrich (1995), KINEROS2—A KINematic Runoff and EROSion Model, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 697–732, Water Resour. Publ., Highlands Ranch, Colo.

Sobol', I. M. (1993), Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput.*, *1*(4), 407–414.

Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, *16*, 430–442, doi:10.1029/WR016i002p00430.

Spear, R. C., and G. M. Hornberger (1980), Eutrophication in Peel Inlet–II. Identification of critical uncertainties via generalized sensitivity analysis, *Water Res.*, *14*, 43–49, doi:10.1016/0043-1354(80)90040-8.

Tang, Y., P. Reed, K. van Werkhoven, and T. Wagener (2007), Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis, *Water Resour. Res.*, *43*, W06415, doi:10.1029/2006WR005813.

Trenberth, K. E., et al. (2007), Observations: Surface and atmospheric climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the*

Intergovernmental Panel on Climate Change, edited by S. Solomon et al., pp. 235–336, Cambridge Univ. Press, New York.

Unkrich, C. L., and H. B. Osborn (1987), Apparent abstraction rates in ephemeral stream channels, *Hydrol. Water Resour. Ariz. Southwest*, *17*, 35–42.

Van Straten, G., and K. J. Keesman (1991), Uncertainty propagation and speculation in projective forecasts of environmental change, *J. Forecast.*, *10*, 163–190, doi:10.1002/for.3980100110.

van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*, W01429, doi:10.1029/2007WR006271.

Wagener, T., and H. V. Gupta (2005), Model identification for hydrologic forecasting under uncertainty, *Stochast. Environ. Res. Risk Assess.*, *19*, 378–387, doi:10.1007/s00477-005-0006-5.

Wagener, T., and J. B. Kollat (2006), Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox, *Environ. Model. Software*, *22*(7), 1021–1033, doi:10.1016/j.envsoft.2006.06.017.

White, S., J. M. Garcia-Ruiz, and A. Gomez-Villar (1997), The 1996 Biescas campsite disaster in the central Spanish Pyrenees, and its temporal and spatial context, *Hydrol. Process.*, *11*(14), 1797–1812, doi:10.1002/(SICI)1099-1085(199711)11:14<1797::AID-HYP605>3.0.CO;2-7.

Woolhiser, D. A., R. E. Smith, and D. C. Goodrich (1990), *A Kinematic Runoff and Erosion Model: Documentation and User Manual*, ARS 77, U.S. Dep. of Agric., Tucson, Ariz.

Woolhiser, D. A., R. E. Smith, and J.-V. Giraldez (1996), Effects of spatial variability of saturated hydraulic conductivity on Hortonian overland flow, *Water Resour. Res.*, *32*(3), 671–678, doi:10.1029/95WR03108.

Yilmaz, K., T. S. Hogue, K.-L. Hsu, S. Sorooshian, H. V. Gupta, and T. Wagener (2005), Evaluation of rain gauge, radar and satellite-based precipitation estimates with emphasis on hydrologic forecasting, *J. Hydrometeorol.*, *6*(4), 497–517, doi:10.1175/JHM431.1.

Zak, S. K., K. J. Beven, and B. Reynolds (1997), Uncertainty in the estimation of critical loads: A practical methodology, *Water Air Soil Pollut.*, *98*, 297–316.

D. Goodrich and C. Unkrich, USDA-ARS-SWRC, 2000 East Allen Road, Tucson, AZ 85719, USA.

H. Gupta and A. Stewart, Department of Hydrology and Water Resources, University of Arizona, Harshbarger Building Room 314, 1133 East North Campus Drive, P.O. Box 210011, Tucson, AZ 85721, USA.

M. Schaffner, Binghamton Weather Forecast Office, National Weather Service, 32 Dawes Drive, Johnson City, NY 13790, USA.

T. Wagener, Department of Civil and Environmental Engineering, Pennsylvania State University, 226B Sackett Building, University Park, PA 16802-1408, USA.

S. Yatheendradas, Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology, 801 Leroy Place, MSEC, Socorro, NM 87801, USA. (soni@nmt.edu)