Investigating the Optimal Configuration of Conceptual Hydrologic Models for Satellite-Rainfall-Based Flood Prediction

Amanda Harris and Faisal Hossain

***Abstract*—In this letter, we investigated the optimal configuration of conceptual hydrologic models for satellite-rainfall-based flood prediction in the 970-km^2 Upper Cumberland basin of Kentucky. We explored the impact of integrating NASA's real-time global satellite rainfall product (IR-3B41RT), available at 0.25°-hourly resolution, in four conceptual model configurations: three built using the modular Hydrologic Modeling System of the Hydrologic Engineering Center that focused on structural differences in infiltration schemes (i.e., National Resources Conservation Service (NRCS) curve number (CN) method, Green–Ampt infiltration method, and deficit/constant loss method) and the fourth being the topographic-index-based TOPMODEL. For the case presented in this letter, a spatially distributed model application did not appear to yield greater accuracy than lumped approaches when using spatially distributed satellite rainfall data for such a medium-sized basin. In general, the NRCS CN method was found to be most effective in terms of minimizing flood prediction uncertainty, followed by the Green–Ampt infiltration and deficit/constant loss methods.**

***Index Terms*—Floods, model complexity, satellite rainfall.**

I. Introduction

THE GLOBAL importance of satellite-derived rainfall has led to the development of an increasing number of satellite-based rainfall products that are now available to meet the needs of various users (for a summary of currently available products, refer to [3]). However, satellite-estimated rainfall, being only a proxy measurement, has uncertainty that bears significant implications on the hydrologic simulation of land processes. This uncertainty can lead to high uncertainties in runoff simulation [9]. The manifestation of runoff uncertainty due to error in rainfall input is commonly known as "error propagation." If satellite rainfall data, in anticipation of the Global Precipitation Measurement (GPM) mission [11], are to be critically assessed regarding the opportunities for flood monitoring over medium and large basins, it is important that we first understand the error propagation that is associated with satellite-estimated rainfall.

There have been numerous studies on uncertainty of flood prediction, mostly involving radar-estimated rainfall [6], [12], [13]. However, much less is known on the reliability of satellite rainfall for flood monitoring. The structure of satellite precipitation errors can be highly complex at scales relevant for flood modeling. This complexity is now being recognized by hydrologists who are generally not familiar with satellite remote sensing techniques. However, a current knowledge gap that has remained relatively unexplored is how to account for hydrologic model complexity when identifying optimal strategies for integrating satellite rainfall data in operational flood monitoring systems. Recently, calls have been made for hydrologists to become more involved in the satellite rainfall data development process by providing feedback to data producers that can potentially lead to development of better next-generation algorithms for overland applications [5].

In this letter, we investigated the optimal configuration of conceptual hydrologic models for satellite-rainfall-based flood prediction in the Upper Cumberland (UC) basin of Kentucky (KY) that has witnessed periodic catastrophic flooding. We explored the impact of integrating NASA's real-time satellite rainfall data (IR-3B41RT), available at 0.25°-hourly resolution, in four conceptual configurations. Our motivation is driven by the need to understand how model complexity triggers the transformation of the satellite rainfall estimation error to runoff error.

II. Testable Hypothesis

To extract the most hydrologically useful information from satellite rainfall data, we need to understand the streamflow simulation error as a function of complexities in both satellite rainfall estimation error and model configuration.

III. Study Area and Data

Our study area is the 970-km^2 basin of the UC River in southeastern Kentucky bordering with Virginia and Tennessee (Fig. 1). The area is primarily mountainous and forested, and it lies in the Eastern Coal Field physiographic region. The underlying rock formations are primarily sandstone, shale, and siltstone. The percentage distribution of major land-use types is forest land (80.13%), urban (8.20%), cropland and pasture (11.15%), and lakes and reservoirs combined (0.52%). The outlet of the basin is in the town of Loyall, KY.

The storm event for this study took place on March 16–20, 2002, with the majority of the rainfall occurring on March 17 and 18. The total storm rainfall volume (i.e., mean areal rainfall), as reported from gages, was about 6.12 in (155.4 mm) over the basin. The observed streamflow was measured at the outlet of the basin in Loyall, KY, at the U.S. Geological Survey streamflow gage #03401000. This gage was operated in cooperation with the U.S. Army Corps of Engineers (USACE).

Manuscript received November 26, 2007; revised March 24, 2008. The work of A. Harris was supported by the Department of Civil and Environmental Engineering, Tennessee Technological University, Cookeville.

The authors are with the Department of Civil and Environmental Engineering, Tennessee Technological University, Cookeville, TN 38505 USA (e-mail: fhossain@tntech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2008.922551

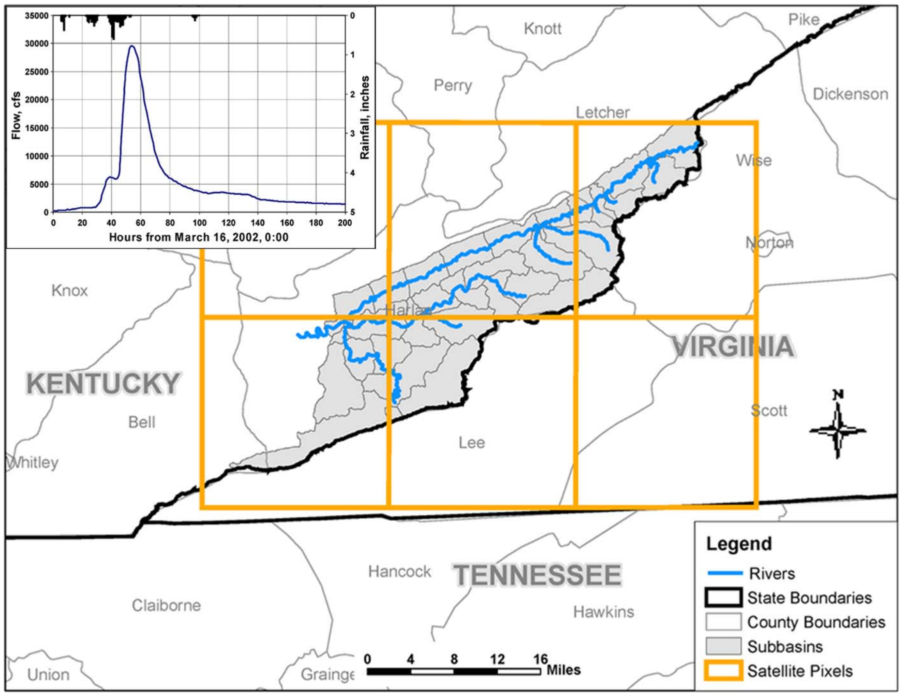

Fig. 1. UC river basin. Yellow lines show locations of 0.25° grids for 3B41RT rainfall data. Rainfall and runoff for March 2002 storm event shown in the inset.

The peak flow for the storm was approximately 29 600 ft^3/s (Fig. 1). The reference rainfall data set for calibrating hydrologic models comprised the National Weather Service Stage III data that have been bias corrected against gage data (hereafter called NEXRAD) [4].

IV. Hydrologic Model Configurations

We selected the Hydrologic Engineering Center's (HEC) Hydrologic Modeling System (HMS) for investigating various hydrologic model configurations. We focused on three particular model configurations related to rainfall–runoff transformation in HEC-HMS. These concerned infiltration scheme conceptualizations to calculate excess rainfall leading to surface runoff. Other components of the model such as base flow, river routing, and evapotranspiration were kept constant. The Muskingum–Cunge routing [1] and ModClark transformation [8] method were used across all model types in HEC-HMS. Required parameters for surface runoff routing were length, energy slope, Manning's n roughness coefficients, and station elevation data. The reader is referred to the *HEC-HMS Technical Reference Manual, March 2000* [14] for more details on these methods. The infiltration schemes considered were the following: 1) deficit/constant loss method; 2) Green–Ampt infiltration; 3) National Resources Conservation Service (NRCS) curve number (CN) method. We used the topographic-index-based model called TOPMODEL, which was first developed by Beven and Kirkby [2], as our fourth model configuration. We provide a very brief description of each rainfall–runoff configuration hereafter.

A. Deficit/Constant Loss Method

The deficit/constant loss method assumes that, for a storm event, the precipitation loss potential is constant. Until the precipitation accumulation becomes greater than the initial loss value, runoff does not occur. The method requires two input parameters to describe the physical characteristics of the watershed: the initial loss and the constant rate of loss.

B. Green–Ampt Infiltration Method

The Green and Ampt soil loss method is physically conceptualized based on the law of conservation of mass and Darcy's equation for flow in porous media. The advantage of this configuration is that its parameters can be estimated by knowledge of the soil type and are physically meaningful. When the initial loss is filled, the Green and Ampt equations are used to determine the excess precipitation. The required parameters for the Green and Ampt method include initial loss, volumetric moisture deficit, wetting front suction, conductivity, and percent imperviousness.

C. NRCS CN Method

In the NRCS CN method, the antecedent moisture (AM), soil type, land use, and cumulative precipitation are used to determine the excess precipitation. Like the deficit/constant loss method, no runoff occurs until the initial abstraction has been filled. The CNs for the basins in this letter were determined by delineating the land use from aerial photography taken by USACE. Assuming AM group 2, good hydrologic conditions, and soil group B, the CN for each land-use type were obtained from the *HEC-HMS Technical Reference Manual, Appendix A*. When a basin contained more than one type of land use, the CN was determined using the area-weighted method. The input parameters required for the NRCS method were initial abstraction and the CN.

D. TOPMODEL Method

This model may be considered the most complex among the four configurations assessed in this letter. A topographic index

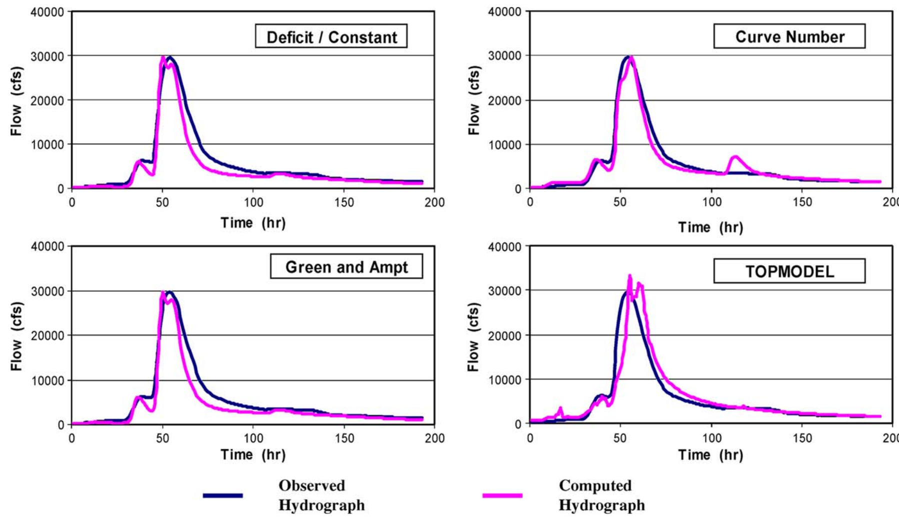

Fig. 2. Observed versus NEXRAD simulated streamflow for the four hydrologic model configurations (time starts from March 16, 2002, 0000 UTC).

$\ln(a/\tan\beta)$ is used as an index of hydrologic similarity, where a is the area draining through a point and $\tan\beta$ is the local surface slope. In this letter, the topographic index was derived from a 10-m-resolution digital elevation model for the UC watershed by using a multiple-flow-direction algorithm [10]. Further details of TOPMODEL can be found in [2].

Using NEXRAD radar rainfall data as the baseline, it was possible to derive very accurate streamflow simulations for the various model configurations (see Fig. 2). Due to the very negligible difference observed between baseline simulation and observed streamflow across all four model types (Fig. 2), all subsequent assessment of satellite-derived streamflow simulation was therefore performed with respect to the observed streamflow hydrograph.

V. Satellite Data and Satellite Rainfall Error Model

We employed a 2-D (i.e., spatial) satellite rainfall error model (SREM2D) [6] for ensemble generation of synthetic satellite rainfall data for error propagation experiments. The purpose of SREM2D was to mimic satellite-like representation of the rainfall fields using, as input, the reference rainfall of higher accuracy (i.e., NEXRAD Stage III data in this letter). This can be achieved through error corruption of the input rain fields in a space-and-time stochastic framework. The major dimensions of error structure in satellite estimation, which are modeled by SREM2D, are as follows: 1) the joint probability of successful delineation of rainy and nonrainy areas accounting for a spatial structure; 2) the temporal dynamics of the conditional rainfall estimation bias (rain > 0 unit); and 3) the spatial structure of the conditional (rain > 0 unit) random deviation. In this letter, error parameters were derived for a real-time satellite rainfall data product produced by NASA–3B41RT [7], using the NEXRAD rainfall as our reference for a nearby region over Oklahoma [6]. This satellite rainfall product is produced at 0.25°-hourly resolution and is globally available on a near-real-time basis from the World Wide Web. Five pixels was directly over the basin area (Fig. 1).

VI. Simulation Methodology for Error Propagation

Because the most current version of HEC-HMS lacks an automatic scheme for Monte Carlo (MC) simulation, the error propagation using SREM2D was carried out manually. Hence, only a 50-member MC ensemble of rain time series was produced. These simulated ensembles were propagated one by one through the three HEC-HMS-model and TOPMODEL configurations.

In this letter, the following two modes of SREM2D were used for the analysis: 1) spatially lumped mode over the basin and 2) spatially distributed mode using each satellite grid box (Fig. 1). The SREM2D error modeling approach was applied to the actual satellite rainfall data. This is a more pragmatic approach for ungauged basins where ground validation rainfall data (such as NEXRAD) is not available. The roles of reference and estimated rain fields were therefore reversed in SREM2D. This means that ensembles of time-varying rainfall fields were generated from satellite rainfall data [7]. In essence, this approach tried to "correct" available satellite rainfall data according to the error structure known *a priori*. For the spatially lumped case, a 1-D (spatially lumped) version of the error model was used.

For this letter, we limited our uncertainty estimation in runoff (i.e., delineation of error bounds) to calculation of the standard deviation of simulated streamflow at every time step. Two contrasting issues were considered in the error propagation. If either the uncertainty limits of simulation runoff were too narrow or the whole ensemble envelope was biased (i.e., the observed streamflow is consistently outside the prediction error bounds), then comparison with *in situ*/reference measurements would suggest that the model complexity structure was invalid for the satellite rainfall data. If, on the other hand, the uncertainty limits were too wide, then it could be concluded that the hydrologic modeling structure had little predictive capability. The dichotomous nature of "structural validity" and "predictive ability" was quantified by the exceedance probability (EP) and uncertainty ratio (UR) in (1) and (2), respectively, found at the bottom of the next page.

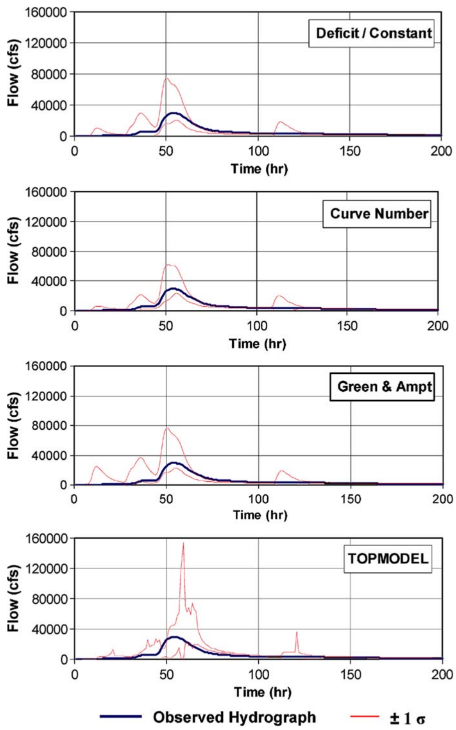

Fig. 3. Streamflow error bars using satellite rainfall ensembles in MC simulations in the lumped mode (i.e., basin is considered lumped).

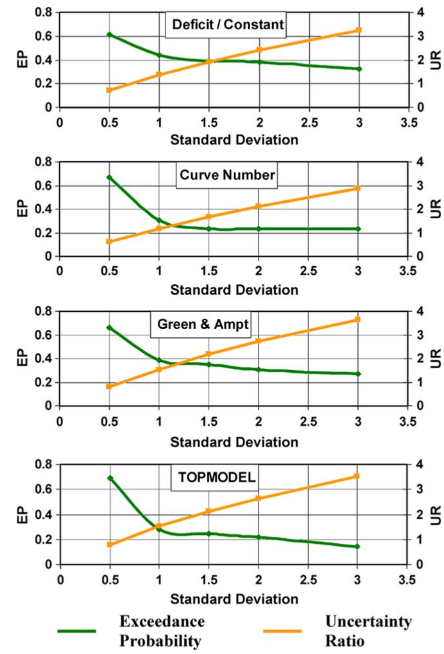

Fig. 4. EP and UR versus standard deviation of streamflow simulation uncertainty for lumped MC simulations. The x-axis represents the width of uncertainty limit in terms of standard deviation of distribution.

VII. Results and Discussion

A. Streamflow Error Propagation in Lumped Mode

In Fig. 3, the streamflow error propagation in the MC framework using the SREM2D error model in the lumped mode [6] is shown. The error bars of the four models shown represent the uncertainty width in streamflow simulation for two standard deviations ($\pm 1\sigma$ from the mean) of the derived distribution. For a normal distribution, this would represent about 67% of "total" uncertainty. It is interesting to note that while TOPMODEL yields the narrowest uncertainty bars during the rising and receding limbs of the hydrograph, its simulation uncertainty is the highest during peak flows.

A more critical understanding of the dichotomous nature of simulation uncertainty (in terms of precision UR and accuracy EP) is obtained from Fig. 4. The NRCS CN method stands out as the most optimal one because it generally has both the narrowest uncertainty limits and the greatest proportion of observed streamflow values within those limits. The EP drops steeply while maintaining a moderate level in UR [note: the UR is the lowest among all four models (Fig. 4, second panel from top)].

B. Streamflow Error Propagation in Spatially Distributed Mode

When the error propagation experiments are repeated for the spatially distributed configuration using SREM2D, the simulation uncertainty is found to be considerably lower than that from the lumped configuration of application (compare Fig. 5 with Fig. 3). However, the nature of error propagation is much more complex. First, a false peak flow manifests more prominently during the rising limb of the hydrograph. Second, the true peak runoff is almost invariably not captured within the simulation error bars of any of the four models. While the magnitude of the peak is within two standard deviations of uncertainty, the peak is simulated late, and hence, the observation falls outside the envelope. Possible reasons for this may be attributed to the spatial mismatch between satellite and true rainfall data in terms of successful detection of rain and no rain. It is well known that satellite rainfall at these hydrologic scales can often suffer from imperfect detection capabilities for delineation of nonstationary rainy and nonrainy areas. Consequently, this may result in some subbasins overestimating the surface runoff and vice versa.

Although the width of the error bars may be lower than that from lumped application, the uncertainty in terms of accuracy, i.e., EP, is found to be higher. Fig. 6 shows that the sensitivity of EP to width of simulation uncertainty (or spread of derived distribution) is lower and continues to remain at higher levels than the lumped approach. Thus, if flood warnings were to be issued with wider levels of error bars (from $\pm 1\sigma$ or $\pm 2\sigma$ error bars) using spatially distributed rainfall, the benefit in terms of observations more frequently falling within the stated error bounds is likely to remain unimproved compared to the lumped application. Overall, it is observed that the NRCS CN method using lumped rainfall is the configuration that performs at the

$$\mathrm{EP} = \frac{\text{Number of times that observed streamflow exceeds the uncertainty limits}}{\text{Total number of time steps}} \tag{1}$$

$$\mathrm{UR} = \frac{\text{Uncertainty in runoff volume simulation (between uncertainty limits)}}{\text{Observed runoff volume}} \tag{2}$$

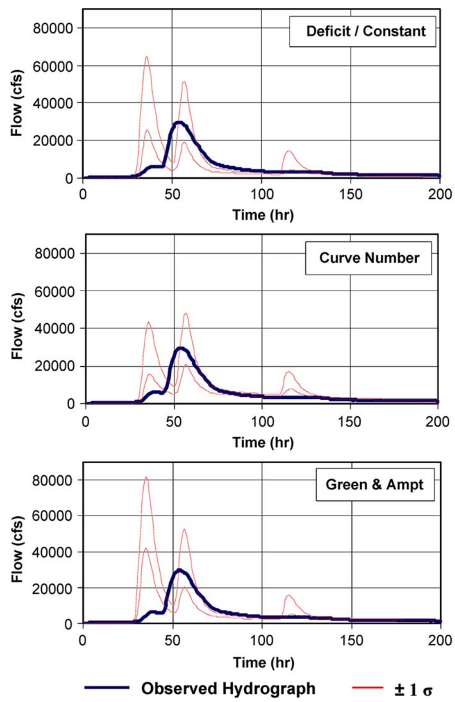

Fig. 5. Streamflow error bars using spatially distributed satellite rainfall ensembles in MC simulations (time starts from March 16, 2002, 0000 UTC).

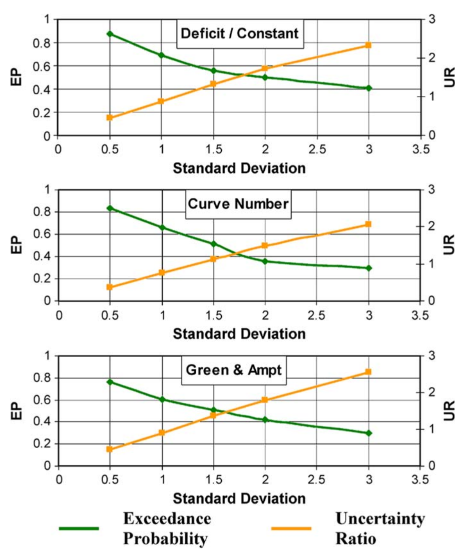

Fig. 6. EP and UR versus standard deviation of runoff simulation uncertainty for distributed MC simulations (i.e., model was run using distributed rainfall data).

optimal level in terms of EP and UR (note: both EP and UR are the lowest; TOPMODEL is not shown, as it was run only in the lumped mode).

VIII. Conclusion

We explored NASA's real-time global satellite rainfall product (IR-3B41RT), available at 0.25°-hourly resolution, for four conceptual model configurations: three built using the modular HMS of HEC that focused on structural differences in infiltration schemes and the fourth being the topographic-index-based TOPMODEL. The spatially distributed model application did not appear to yield greater accuracy than lumped approaches when using spatially distributed satellite rainfall data for such a medium-sized basin. In general, the NRCS CN method was found to be most effective in terms of minimizing flood prediction uncertainty, followed by the Green–Ampt infiltration and deficit/constant loss methods.

The complex-natured uncertainty of satellite rainfall at hydrologically relevant scales may warrant simpler and lumped rainfall–runoff modeling schemes for modeling dynamic flood events when the fundamental scale of satellite rainfall data is relatively large compared to the overall size of the basin. This, however, should not be construed as our universal endorsement of lumped conceptual models over fully distributed physically based hydrologic models. We believe that the spatial averaging of satellite rainfall over a medium-sized basin tends to minimize the magnitude of the various dimensions of error (such as error variance, falsely detected rain and no rain, etc.) during a dynamic flood event. Consequently, this results in less error magnification during the rainfall–runoff transformation process for the lumped application of satellite rainfall data.

We therefore recommend that considerable caution be exercised in satellite rainfall data application when the scale of available satellite rainfall data is comparable to the overall size of the watershed.

References

[1] D. A. Barry and K. Bajracharya, "On the Muskingum–Cunge flood routing method," *Environ. Int.*, vol. 21, no. 5, pp. 485–490, 1995.

[2] K. J. Beven, R. Lamb, P. Quinn, R. Romanowicz, and J. Freer, "TOPMODEL," in *Computer Models of Watershed Hydrology*, V. P. Singh, Ed. Fort Collins, CO: Water Resource Publication, 1995, pp. 627–668.

[3] E. E. Ebert, J. E. Janowiak, and C. Kidd, "Comparison of near-real-time precipitation estimates from satellite observations and numerical models," *Bull. Amer. Meteorol. Soc.*, vol. 88, no. 1, pp. 47–64, Jan. 2007.

[4] R. A. Fulton, J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, "The WSR-88D rainfall algorithm," *Weather Forecast.*, vol. 13, no. 2, pp. 377–395, Jun. 1998.

[5] F. Hossain and D. P. Lettenmaier, "Flood prediction in the future: Recognizing hydrologic issues in anticipation of the global precipitation measurement mission," *Water Resour. Res.*, vol. 44, p. W11 301, Nov. 2006.

[6] F. Hossain, E. N. Anagnostou, T. Dinku, and M. Borga, "Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty," *Hydrol. Process.*, vol. 18, no. 17, pp. 3277–3291, Dec. 2004.

[7] G. J. Huffman, R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff, "The TRMM multi-satellite rainfall analysis: Quasi-global, multi-year, combined-sensor rainfall estimates at fine scale," *J. Hydrometeorol.*, vol. 8, pp. 38–55, 2007.

[8] D. W. Kull and A. D. Feldman, "Evolution of Clark's unit graph method to spatially distributed runoff," *J. Hydrol. Eng.*, vol. 3, no. 1, pp. 9–19, 1998.

[9] B. Nijssen and D. P. Lettenmaier, "Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the global precipitation measurement satellites," *J. Geophys. Res.*, vol. 109, p. D02 103, Jan. 2004.

[10] P. R. Quinn, K. J. Beven, and R. Lamb, "The $\ln(\alpha/\tan\beta)$ index: How to calculate it and how to use it within the TOPMODEL framework," *Hydrol. Process.*, vol. 9, no. 2, pp. 161–182, 1995.

[11] E. Smith, G. Asrar, Y. Furuhama, A. Ginati, C. Kummerow, V. Levizzani, A. Mugnai, K. Nakamura, R. Adler, V. Casse, M. Cleave, M. Debois, J. Durning, J. Entin, P. Houser, T. Iguchi, R. Kakar, J. Kaye, M. Kojima, D. P. Lettenmaier, M. Luther, A. Mehta, P. Morel, T. Nakazawa, S. Neeck, K. Okamoto, R. Oki, G. Raju, M. Shepherd, E. Stocker, J. Testud, and E. F. Wood, "The international Global Precipitation Measurement (GPM) program and mission: An overview," in *Measuring Precipitation From Space: EURAINSAT and the Future*, V. Levizzani and F. J. Turk, Eds. Norwell, MA: Kluwer, 2007, pp. 611–654.

[12] H. O. Sharif, F. L. Ogden, W. F. Krajewski, and M. Xue, "Numerical simulations of radar rainfall error propagation," *Water Resour. Res.*, vol. 38, no. 8, p. 1140, Aug. 2002.

[13] H. O. Sharif, F. L. Ogden, W. F. Krajewski, and M. Xue, "Statistical analysis of radar rainfall error propagation," *J. Hydrometeorol.*, vol. 5, no. 1, pp. 199–212, Feb. 2004.

[14] *Hydrologic Modeling System HEC-HMS: Technical Reference Manual*, U.S. Army Corps Eng., Hydrologic Eng. Center, Davis, CA, 2000.