

Investigating the Applicability of Error Correction Ensembles of Satellite Rainfall Products in River Flow Simulations

VIVIANA MAGGIONI,* HUMBERTO J. VERGARA,⁺ EMMANOUIL N. ANAGNOSTOU,*
JONATHAN J. GOURLEY,[#] YANG HONG,[@] AND DIMITRIOS STAMPOULIS*

* *Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut*

⁺ *Hydrometeorology and Remote-Sensing Laboratory, Civil Engineering and Environmental Science, University of Oklahoma, and National Severe Storms Laboratory, NOAA, Norman, Oklahoma*

[#] *National Severe Storms Laboratory, NOAA, Norman, Oklahoma*

[@] *Hydrometeorology and Remote-Sensing Laboratory, Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma*

(Manuscript received 18 May 2012, in final form 5 February 2013)

ABSTRACT

This study uses a stochastic ensemble-based representation of satellite rainfall error to predict the propagation in flood simulation of three quasi-global-scale satellite rainfall products across a range of basin scales. The study is conducted on the Tar-Pamlico River basin in the southeastern United States based on 2 years of data (2004 and 2006). The NWS Multisensor Precipitation Estimator (MPE) dataset is used as the reference for evaluating three satellite rainfall products: the Tropical Rainfall Measuring Mission (TRMM) real-time 3B42 product (3B42RT), the Climate Prediction Center morphing technique (CMORPH), and the Precipitation Estimation from Remotely Sensed Imagery Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS). Both ground-measured runoff and streamflow simulations, derived from the NWS Research Distributed Hydrologic Model forced with the MPE dataset, are used as benchmarks to evaluate ensemble streamflow simulations obtained by forcing the model with satellite rainfall corrected using stochastic error simulations from a two-dimensional satellite rainfall error model (SREM2D). The ability of the SREM2D ensemble error corrections to improve satellite rainfall-driven runoff simulations and to characterize the error variability of those simulations is evaluated. It is shown that by applying the SREM2D error ensemble to satellite rainfall, the simulated runoff ensemble is able to envelope both the reference runoff simulation and observed streamflow. The best (uncorrected) product is 3B42RT, but after applying SREM2D, CMORPH becomes the most accurate of the three products in the prediction of runoff variability. The impact of spatial resolution on the rainfall-to-runoff error propagation is also evaluated for a cascade of basin scales (500–5000 km²). Results show a doubling in the bias from rainfall to runoff at all basin scales. Significant dependency to catchment area is exhibited for the random error propagation component.

1. Introduction

Current runoff prediction systems integrate precipitation measurements into hydrological models that simulate river discharges at the watershed scale either distributed across the basin or as lumped values at the catchment outlet. As observations from rain gauges are nonexistent or sparse over several regions of the globe, remotely sensed rainfall measurements offer a unique and viable alternative source of forcing data for hydrological

models (e.g., Su et al. 2008; Li et al. 2009; Behrangi et al. 2011; Bitew and Gebremichael 2011).

Yilmaz et al. (2010) have shown that a global flood monitoring system that uses satellite-based precipitation was able to simulate the onset of flood events produced by heavy precipitation, but simulation performance was shown to deteriorate in the later stages. Another study by Su et al. (2008) applied precipitation estimates of the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA; 3B42v6) to a semidistributed hydrological model in the La Plata basin, Brazil. They showed that TMPA-driven hydrologic model simulations were able to reproduce the seasonal and interannual streamflow variability, capture daily flooding events, and represent low flows, although peak

Corresponding author address: Emmanouil N. Anagnostou, Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269.

E-mail: manos@enr.uconn.edu

flows tended to be positively biased. Another recent work by Khan et al. (2011) demonstrated that the use of remote sensing data (e.g., precipitation, land cover type, and topography) as inputs to a distributed hydrologic model could potentially improve hydrologic prediction and flood management strategies in ungauged watersheds.

Satellite rainfall retrievals that are typically used for ungauged basin applications are susceptible to errors because of sensor limitations (e.g., sampling and land cover heterogeneity effects), uncertainty in the parameterization of the complex relationship between retrieved variables (e.g., brightness temperature at cloud top) and precipitation, and difficulty in obtaining a global distribution of parameters for the retrieval algorithm (Kummerow et al. 1998; Kidd et al. 2003). Any error in the satellite rainfall retrieval can potentially lead to high uncertainties in the simulation of runoff at the watershed scale. Previous studies have examined the implication of satellite rainfall retrieval and sampling errors on flood prediction uncertainty, showing that relative errors (e.g., biases) in rainfall severely magnify (up to a factor of 3) when transformed to corresponding relative errors in runoff (Guetter et al. 1996; Nijssen and Lettenmaier 2004; Hossain et al. 2004; Nikolopoulos et al. 2010, 2012). These considerations stress the need to better understand the sensitivity of flood prediction uncertainty to the error characteristics of satellite rainfall products. In doing so, progress for a more comprehensive characterization of the error-correction schemes applied to satellite precipitation fields naturally follows.

Several studies have pointed out the necessity of an ensemble prediction system for flood forecasting. Atger (2001) showed that the ensemble prediction system performs better than a single forecast based on the same model. He also demonstrated that the impact of reducing the number of ensemble members is rather small (i.e., differences between 51 members and 21 members are not significant). Moreover, Verbunt et al. (2007) corroborated that probabilistic flood forecasts have advantages compared to the deterministic forecast for a particular flood event. They showed that the deterministic control run underestimated the observed runoff peak, whereas some of the ensemble members correctly predicted the flood peak. Cloke and Pappenberger (2009) pointed out that operational flood forecasting systems are increasingly moving toward the adoption of ensembles of numerical weather simulations to drive their predictions. They also highlighted a few remaining key challenges to successfully use probabilistic flood forecasts. First, they observed that it is rare for any case study to report a false alarm of their particular flood forecasting system if the analysis is based on individual events. Second, many published works include only

qualitative statements on the positive impact of the ensemble prediction systems, lacking the quantitative aspect of the analysis. Third, when a quantitative measure is attempted, model skill is often calculated relative to a reference simulation driven by observed precipitation and not against observed discharge (Pappenberger et al. 2008; Thielen et al. 2009).

An accurate characterization of the satellite rainfall error model in a probabilistic flood prediction system has the potential to enhance flood and flash flood forecasting. In fact, in most of the published works the multidimensional character of the precipitation error at fine space and time scales is ignored (e.g., Pappenberger et al. 2005; Gouweleeuw et al. 2005). Hossain and Anagnostou (2006a) have developed an elaborate two-dimensional satellite rainfall error model (SREM2D) for generating ensembles of satellite rain fields. Unlike traditional error modeling approaches, SREM2D models the joint probability of successful delineation of rainy and nonrainy areas accounting for a spatial structure. Moreover, SREM2D was shown to preserve the satellite retrieval error structure across scales, unlike simpler rainfall error models that revealed significant scale-dependent biases (Hossain and Anagnostou 2006b).

Another fundamental aspect in rainfall-to-runoff error propagation is its dependence on the spatial scale. As shown by Nikolopoulos et al. (2010) for a number of basin scales ranging between 100 and 1200 km², the use of satellite rainfall for flood simulation strongly depends on the catchment area and satellite product resolution.

In this study, we use a stochastic framework for investigating the error propagation from satellite rainfall ensembles streamflow using simulation from the National Weather Service (NWS) Hydrology Laboratory–Research Distributed Hydrological Model (HL-RDHM; Koren et al. 2004). We introduce several novelties that seek to address some of the limitations of earlier works, some of which were pointed out by Cloke and Pappenberger (2009). Specifically, we depict satellite rainfall error in a comprehensive way by applying the complex SREM2D rainfall error model that takes into account probability of detection and false alarms of satellite rainfall retrievals. We compare three different satellite rainfall products of varying resolution and retrieval error characteristics as inputs to the hydrologic model. Moreover, this study provides a quantitative analysis on the rainfall-to-runoff error propagation through performance metrics (such as relative bias, relative RMSE, and exceedance ratios) that are computed against observed discharges and compared to benchmark simulations driven by the most accurate rainfall estimates (i.e., rain gauge–calibrated radar rainfall). We also investigated how the error propagation depends on the catchment

scale by considering a cascade of five basins with spatial resolutions ranging from about 500 to 5000 km².

The manuscript is structured as follows. Section 2 presents a description of the study region and the datasets employed in the study. In section 3 we introduce the methodology by describing the hydrological error model, the satellite rainfall error model, and the experimental setup adopted in this study. In section 4 we present and discuss our results, and in section 5 we summarize the major findings and provide future research directions.

2. Study region and data

The study is conducted on the Tar-Pamlico River basin in North Carolina, which is part of the National Oceanographic and Atmospheric Administration (NOAA)'s Coastal and Inland Flooding Observation and Warning (CI-FLOW) project (Van Cooten et al. 2011). Specifically, this study focuses on the subcatchment of the U.S. Geological Survey (USGS) station located at Tarboro (Streamgauge No. 02083500), which includes the upper Tar River (mean daily flows of about 62 m³ s⁻¹) and Fishing Creek subbasins. The Tarboro basin has a drainage area of 5653 km², and its elevation ranges between 3 and 226 m MSL. The majority of the catchment is classified as forest with pasture/hay, whereas only 7% of the land is urban (Bales et al. 2007). To investigate the basin scale issue on error propagation, we considered a cascade of five subbasins of the Tarboro catchment with different drainage areas, ranging from 500 to 5000 km² (Fig. 1).

The study is based on 2 years of rainfall and streamflow data (2004 and 2006). The precipitation data include three quasi-global satellite products and the NWS's Multisensor Precipitation Estimator (MPE) precipitation dataset that is used as ground reference. Details about each rainfall dataset, including information on algorithms and resolutions, are provided next.

The MPE, developed by the NWS Office of Hydrology, combines rainfall measurements from rain gauges with rainfall estimates from the Next Generation Weather Radar (NEXRAD). There is also an option in MPE to incorporate rainfall estimates from Geostationary Operational Environmental Satellite (GOES) platforms, but this is an atypical practice. In addition to mean-field bias adjustment, the MPE algorithm also includes a local bias correction that accounts for spatially nonuniform biases that may exist in the individual radar products (Seo and Breidenbach 2002). MPE is available at 4 kilometer-hourly spatial-temporal resolution, and it is used in this study as a benchmark to evaluate the three satellite rainfall retrievals. An average over the different basins is considered for both MPE and the satellite products.

The first satellite rainfall product is the TMPA, which provides a calibration-based sequential scheme for combining rainfall estimates from various satellites (Huffman et al. 2007). This product uses microwave-calibrated infrared precipitation estimates directly to fill the microwave coverage gaps. It is available at 0.25° × 0.25° and 3-hourly spatiotemporal scales, both as post analysis (3B42v6) and in real time (3B42RT), based on calibration by the TRMM Combined Instrument and TRMM Microwave Imager precipitation products, respectively. In this study, we only focus on the real-time product, henceforth called 3B42RT.

The second satellite rainfall algorithm is the NOAA/Climate Prediction Center morphing technique (CMORPH), which uses motion vectors derived from half-hourly interval geostationary satellite infrared imagery to propagate the less frequent but relatively high-quality rainfall estimates from passive microwave data (Joyce et al. 2004). Specifically, the dynamic morphological characteristics of precipitation features are interpolated between consecutive microwave sensor samples through a time-weighted linear interpolation. This process yields spatially and temporally continuous microwave rainfall fields, guided by infrared imagery but still independent of infrared rain retrievals. The CMORPH product is available half-hourly at 8-km resolution.

The third product is the Precipitation Estimation from Remotely Sensed Imagery Using Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000) Cloud Classification System (CCS; Hong et al. 2004). The algorithm extracts local and regional cloud features from infrared geostationary satellite imagery and used microwave data for training a neural network that assigns precipitation estimates to the infrared temperature. The PERSIANN-CCS dataset, produced by the University of California, Irvine, covers 50°S–50°N and provides estimates of rainfall at spatial and temporal resolutions of 0.04° × 0.04° and 30 min, respectively.

Ground observations of runoff measured at the outlets of each of the five basins are also employed in the study and considered as a "reference" for the model-simulated runoff. Subhourly streamflow observations for Tarboro are in fact available from October 1985 to present at the USGS Instantaneous Data Archive (IDA; USGS 2011). Three-hourly and hourly streamflow observations were derived from this dataset for the purpose of this study.

3. Methodology

a. The hydrologic model

The HL-RDHM was used to simulate streamflow at the five selected locations in Tarboro's catchment. The

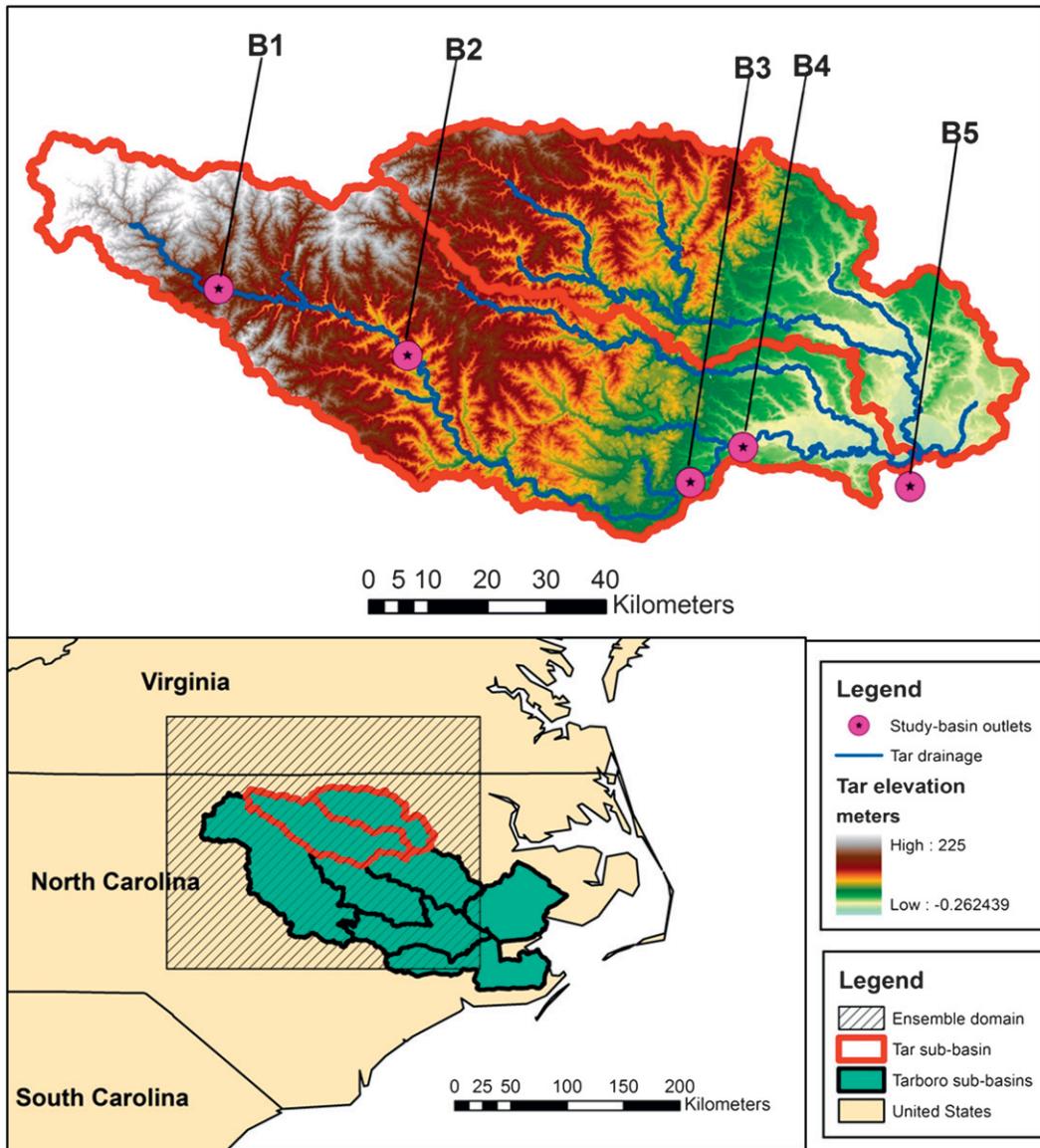


FIG. 1. Study area showing the five basin outlets and the domain that were used to generate the SREM2D ensembles.

HL-RDHM employs the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash et al. 1973) for runoff generation and a kinematic wave model for flow routing. For the latter, a rating curve method is available (Koren et al. 2004), which was selected for this study. SAC-SMA has 17 conceptual parameters while the routing model consists of 2 parameters. Initial spatially distributed estimates for 11 of the SAC-SMA parameters were available from the work of Koren et al. (2000). The remaining 6 parameters were implemented as lumped values established by the NWS from previous experience on different basins (Pokhrel et al. 2008; Yilmaz et al. 2008). As for the routing parameters, initial values were

found from rating curve measurements at eight different gauges within Tarboro’s catchment. These values were then used to generate spatially distributed grids of the two parameters employing HL-RDHM’s built-in tools (NWS et al. 2008). Further details regarding model structure, mechanisms, and other modules available in the HL-RDHM can be found in Koren et al. (2000, 2003, 2004).

The model parameters were adjusted for each of the five selected basins based on an automatic global optimization algorithm [Differential Evolution Adaptive Metropolis (DREAM)] (Vrugt et al. 2009) using MPE rainfall as forcing and streamflow observations. Specifically, DREAM was set to minimize the sum of squared

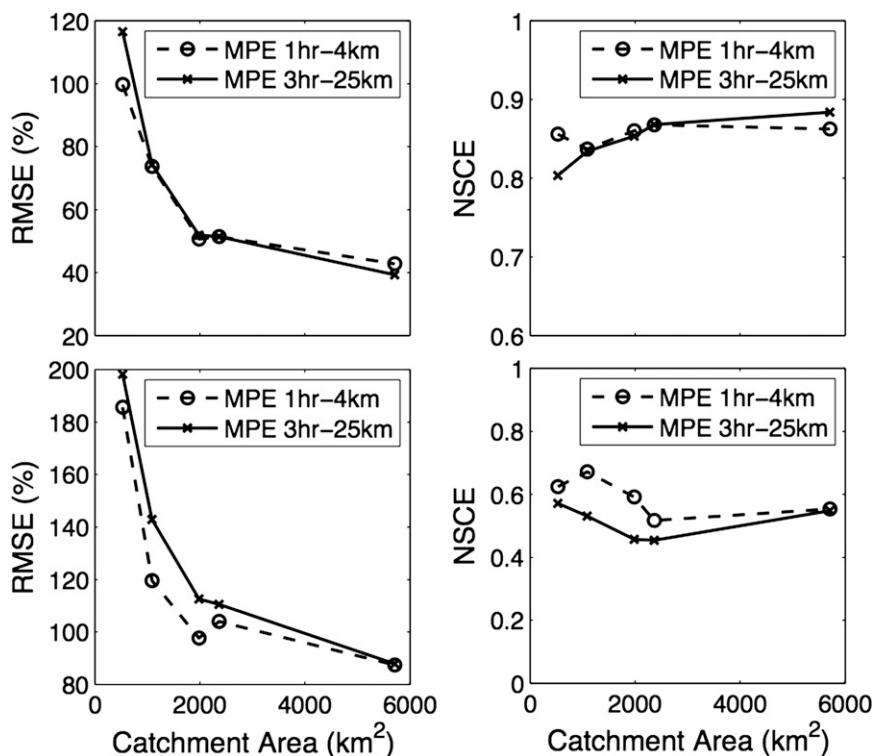


FIG. 2. (top) Calibration and (bottom) validation statistics—(left) RMSE and (right) NSCE—for models forced with MPE at 1 h–4 km and 3 h–25 km temporal–spatial resolution. The statistics are plotted as a function of basin area.

residuals (SSR) with a constraint on the bias to be within $\pm 10\%$ to ensure relatively unbiased simulations. DREAM is based on the well-known Shuffled Complex Evolution algorithm (SCE-UA; Duan et al. 1993). The algorithm estimates the posterior probability density function of the model parameters employing Markov chain Monte Carlo sampling and separates behavioral from nonbehavioral solutions using a cutoff threshold that is based on the sampled probability mass (Vrugt et al. 2008). A period of 3 yr (2004–06) was employed for model calibration and two periods, one from 2002 to 2003 and another from 2007 to 2009, were used for validation of the resulting model. To account for 3B42RT's coarse resolution, the model was also calibrated with MPE resampled to 3-h–25-km resolution. Gourley et al. (2011) have argued about the importance of this latter step when evaluating the hydrologic skill of satellite products.

Figure 2 presents metrics of model performance as functions of basin area for both high-resolution (1 h–4 km) and coarse-resolution (3 h–25 km MPE rainfall forcing). The Nash–Sutcliffe coefficient of efficiency (NSCE) ranges between 0.84 and 0.87 for all of the basins, showing nearly no sensitivity to basin scale. On the

other hand, the relative root-mean-squared error (RMSE) metric manifests high sensitivity to basin scale, particularly for drainage areas under 2000 km². Moreover, it can be seen that the skill of the model run at both resolutions is almost the same, which indicates that rainfall variability lost because of the coarser resolution could be resolved effectively through model calibration. Similar behavior is observed during the validation period, although the skill of the model drops considerably as shown by the increase in RMSE and the reduction in NSCE (Fig. 2). A more detailed examination of the model performance reveals that simulations exhibit a high positive bias for the second interval of the validation period (2007–09; Table 1). We note that the North Carolina Department of Environment and Natural Resources reported unusual dry conditions, classified as a period of extreme drought, particularly from mid-2007 to mid-2008 (NCDWR 2009). The USGS also reported the lowest discharges in more than 110 years for some rivers in the region during this period (USGS 2007). Since the calibration period contains information related to normal and wet conditions in the basins, it is not surprising that the model's performance deteriorates for such extreme drought conditions.

TABLE 1. Relative bias (%) for the simulations forced by MPE at 1 h–4 km resolution.

	2002	2003	2004	2005	2006	2007	2008	2009
B1	−0.6	−15.7	7.5	−1.8	24.5	70.4	125	33.6
B2	−3.9	−15.2	5.7	−8	−2.8	44.7	113.5	28.9
B3	−22	−17.2	7.8	−10	−2.6	33.1	89.9	34.4
B4	−18.2	−16.7	8.2	−5.9	−7.1	29.7	86.4	31.9
B5	−2.8	−17.5	9.2	7.1	−9.2	32.2	84.6	46.1

b. The rainfall error model

SREM2D, developed by Hossain and Anagnostou (2006a), is employed in this study to perturb satellite rainfall fields in order to generate rainfall ensembles that force the hydrologic model. The model was originally developed to use reference rainfall fields as input, and it employs stochastic space–time formulations to characterize the multidimensional error structure of corresponding satellite retrievals. In this study we reverse the original definition of input in SREM2D to derive the multidimensional structure of deviations from the reference (i.e., radar) rainfall with respect to the satellite rainfall estimates, which then become the input to the error model. This process generates ensembles that are closer to the reference (i.e., radar-like) rainfall fields by correcting the different satellite rainfall retrievals. These ensembles are then used to force the hydrologic model, thus generating ensembles of runoff simulations.

SREM2D employs stochastic space–time formulations to characterize the multidimensional error structure of satellite retrievals. It describes the spatial variability of rain rate estimation error and models the joint probability of successful delineation of rainy and nonrainy areas accounting for the spatial structure. Unlike traditional error models, where uncertainty is determined by the multiplicative perturbations, SREM2D may introduce rain in areas where the satellite does not detect rain, and it may assign zero rain where the satellite measures rain. These two cases statistically represent the missed rain detection and false alarms, which are important characteristics of the satellite precipitation retrievals and are often neglected by standard rainfall error models. SREM2D can also account for temporal

correlation. However, the temporal correlation in the passive microwave retrieval error was found to have a negligible effect on runoff prediction (Hossain et al. 2004). For this reason, temporal correlation was set to zero in our experiments.

The input parameters for SREM2D are calibrated for each satellite rainfall product as per previously established protocols and procedures outlined by Hossain and Anagnostou (2006a), and they are presented in Table 2. The error model calibration was performed for the year of 2004 on a gridded domain that includes the study area. Previous studies have shown that, after calibration, SREM2D was able to conserve the satellite retrieval error structure across scales, both in terms of magnitude and variability of the error, unlike simpler error models that revealed significant scale-dependent biases (Hossain and Anagnostou 2006b; Maggioni et al. 2011). Although, SREM2D is computationally more expensive compared to standard multiplicative error models, it provides a holistic representation of satellite retrievals that are susceptible to rain detection and false detection uncertainties (Hossain and Anagnostou 2006b).

c. The experimental setup

The NWS HL-RDHM model is forced with precipitation fields to produce streamflow simulations in two different modes: “deterministic” mode, where input rainfall is not perturbed, and “ensemble” mode, where forced-rainfall fields are perturbed through SREM2D.

We performed four runs of the model in the deterministic mode. First, MPE reference rainfall fields are used to force the hydrologic model. Runoff outputs from this run are compared to ground streamflow observations at the subbasin outlets to evaluate the error introduced by the hydrologic model (i.e., model error), which is conditioned on the MPE input uncertainty assumed to be negligible with respect to the satellite-rainfall algorithms. Next, each of the three unperturbed satellite-rainfall products (i.e., 3B42RT, CMORPH, and PERSIANN-CCS) is used independently to force the hydrologic model. In this study, the runoff output from the model integration forced with the MPE rainfall represents the reference for runoff. Reference error

TABLE 2. SREM2D calibration parameters.

Error parameters	3B42RT	CMORPH	PERSIANN-CCS
Mean (mean Gaussian of log error)	−0.08	0.55	0.20
Sigma (std. dev. Gaussian of log error)	1.00	0.98	1.12
False alarm mean rain rate (mm hr ^{−1})	0.24	0.38	0.48
Probability of detection of no rain	0.83	0.96	0.96
Correlation length _{retrieval error} (km)	100	60	40
Correlation length _{rain det} (km)	280	170	110
Correlation length _{no rain det} (km)	220	60	55

statistics are then computed by contrasting the runoff simulations derived from each satellite rainfall product to the reference runoff.

The ensemble runs are created through Monte Carlo simulations. Each satellite precipitation product is perturbed using the SREM2D rainfall error model to produce ensembles of rainfall fields, which subsequently force the NWS HL-RDHM model. Each ensemble consists of 50 equally likely members and, in turn, generates a runoff ensemble. The three ensembles are then evaluated against the reference runoff simulation, obtained by forcing the model with the unperturbed MPE rainfall, in terms of error statistics. Namely, we analyze the skill of the SREM2D ensembles to represent the uncertainty in runoff with respect to the reference fields and streamflow ground observations.

The temporal and spatial resolution of satellite products and corresponding ensembles used to force the hydrologic model are 3 h–25 km for 3B42RT, 1 h–8 km for CMORPH, and 1 h–4 km for PERSIANN-CCS. The MPE at 3 h–25 km and 1 h–4 km is used as benchmark to the 3B42RT and CMORPH–PERSIANN-CCS satellite products, respectively. Although the CMORPH resolution is 1 h–8 km, we determined that model results using this resolution did not differ from the 1 h–4 km MPE. Then, the rainfall-to-runoff error propagation is studied qualitatively through cumulative rainfall and runoff time series during 2006, as well as a time series of a single storm event that took place in June 2006. A quantitative analysis is also performed by computing error metrics, such as rainfall-to-runoff propagation factors, and exceedance ratio, defined as the frequency of reference runoff exceeding the ensemble runoff prediction envelope. The impact of spatial resolution on the error propagation of the three satellite rainfall products is evaluated based on the five cascade basin scales. A preliminary analysis was also performed to verify the number of ensemble members to be used in this study. Specifically, we performed an analysis of the two statistics employed for the ensemble verification, discussed in the following section, for ensemble sizes ranging from 5 to 50. The metrics converge after about 40 ensemble members for all satellite products and basin scales (not shown here). This gives us confidence that the 50 ensemble members used in this study are sufficient to obtain statistically meaningful results.

4. Results and discussion

a. Qualitative analysis

Figures 3 and 4 illustrate time series of cumulative rainfall and cumulative runoff during 2006 for the smallest

and largest basins, respectively. Rainfall time series from MPE (at the corresponding satellite-product resolution), 3B42RT, CMORPH, and PERSIANN-CCS are shown together with the 50 ensemble members obtained by perturbing each satellite rainfall product with SREM2D. Corresponding runoff simulations are presented in the right panels of Figs. 3 and 4, which are compared with the ground streamflow observations.

By perturbing satellite rainfall fields with SREM2D, we expect to obtain an ensemble of equiprobable radar-like realizations. As a matter of fact, for both small and large basins, SREM2D rainfall ensembles are shifted toward the reference (i.e., MPE) rainfall. In most cases, the SREM2D ensemble is actually able to envelope the MPE rainfall. However, in some other cases (e.g., large basin with perturbed PERSIANN-CCS rainfall), SREM2D cannot entirely correct for the bias between satellite and radar rainfall fields. This can be ascribed to the fact that PERSIANN-CCS is almost unbiased during the calibration period (the entire 2004, not shown here), but during the validation period (2006) the satellite product highly underestimated the reference MPE rainfall and SREM2D is only able to partially correct for this significant bias. Similarly, for the small basin, CMORPH showed a larger bias during 2004 than during 2006, which results in an overestimation of the CMORPH-derived ensemble with respect to MPE during the validation year.

When looking at time series of cumulative runoff, we observe that the ensemble spread relative to the observed runoff simulations is larger than the spread of the SREM2D ensembles relative to observed rainfall. This indicates a magnification of the relative errors from rainfall to runoff. For the small basin case and for all three satellite retrievals the runoff ensemble envelopes the reference runoff simulation, as well as the ground truth, with the only exception of CMORPH between July and November (Fig. 3). This demonstrates that SREM2D is, in fact, able to produce ensemble members that are closer to the reference simulation and to the observed streamflow. On the other hand, for the large basin case, only ensembles obtained by perturbing 3B42RT and CMORPH, and only for part of the time series, produce runoff ensembles that can encapsulate the reference runoff (Fig. 4). By applying SREM2D to PERSIANN-CCS rainfall fields, the corresponding simulated runoff ensemble is placed closer to the reference runoff and to the ground measurements, but it is not able to envelope either. This is related to what is observed in the PERSIANN-CCS rainfall cumulative time series, which shows that SREM2D is only able to partially correct for the strong underestimation of the PERSIANN-CCS rainfall product during 2006.

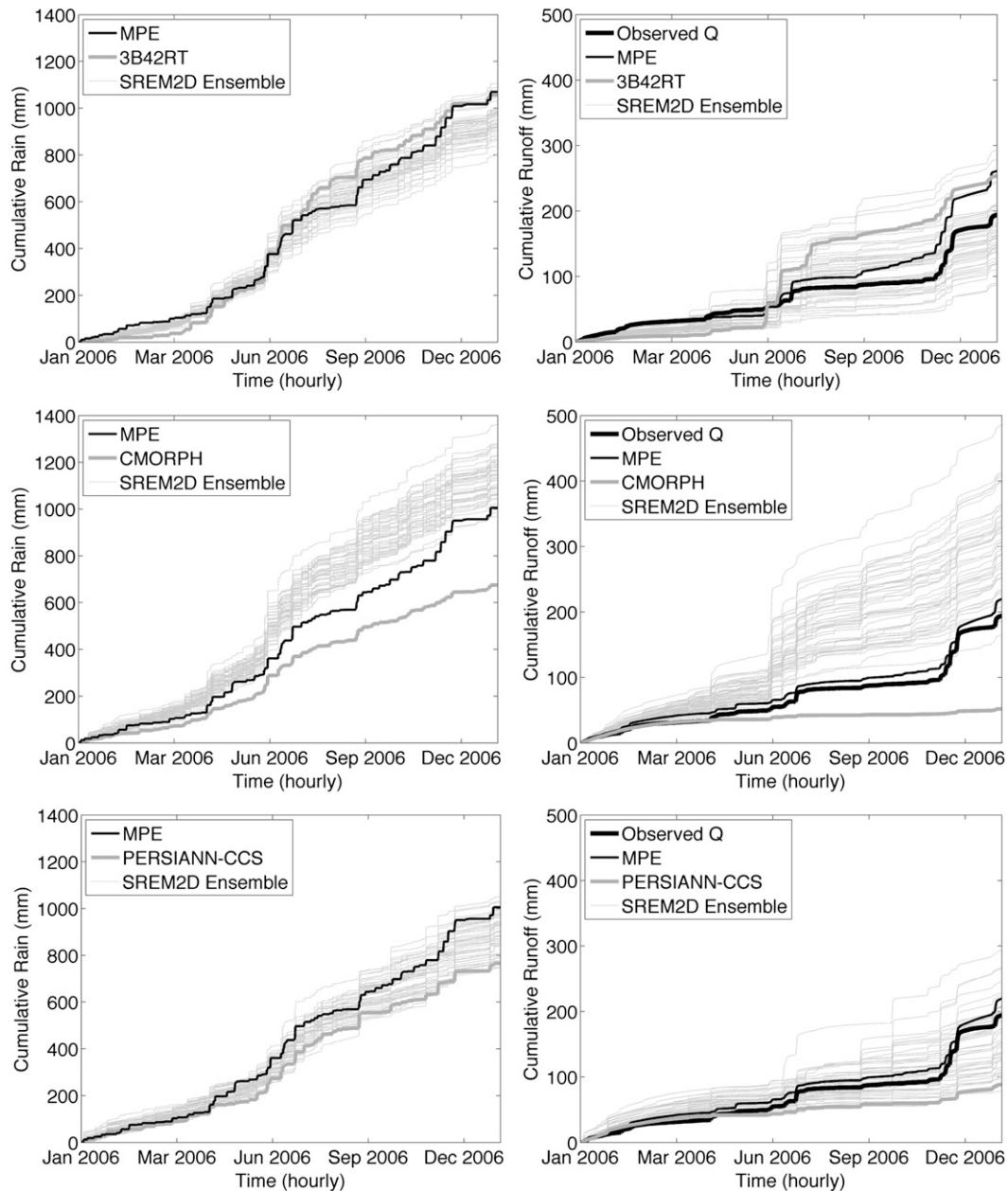


FIG. 3. Time series of (left) cumulative rainfall and (right) runoff for basin B1 during 2006 for (top) 3B42RT, (middle) CMORPH, and (bottom) PERSIANN-CCS rainfall products. MPE rainfall was aggregated at the respective satellite product's temporal-spatial resolution: 3 h–25 km for top panels and 1 h–4 km for middle and bottom panels.

Next, we consider a single event associated with the tropical storm Alberto that hit North Carolina on 14 June 2006. Time series of simulated and observed runoff are shown in Fig. 5 together with the MPE reference rainfall and streamflow simulation. For both unperturbed 3B42RT rainfall fields and SREM2D-perturbed fields the hydrological model simulations overestimated runoff peak in the small basin and slightly underestimated it in the larger basin. On the

other hand, unperturbed CMORPH rainfall forcing did not produce any runoff at both basin scales. However, by perturbing CMORPH with SREM2D, the runoff ensemble is able to predict the peaks and encapsulate the variability. For the PERSIANN-CCS case, neither the unperturbed satellite product nor the corresponding SREM2D rainfall ensemble is able to properly simulate the runoff peak, which is consistent with our previous observation of large bias

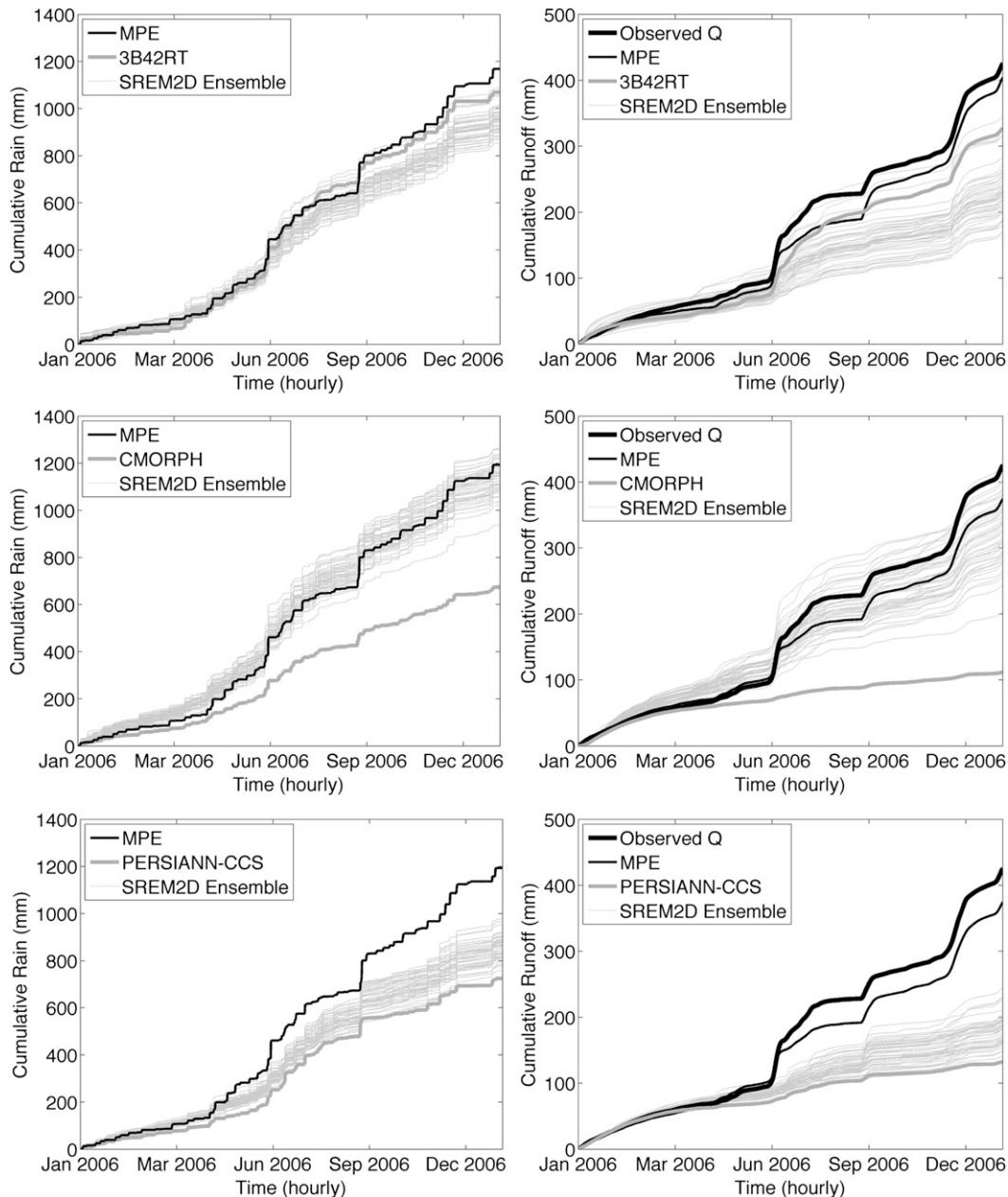


FIG. 4. As in Fig. 3, but for basin B5.

with the PERSIANN-CCS product. These effects are quantitatively evaluated in the next section.

b. Quantitative analysis

To further investigate the propagation of error statistics from precipitation to runoff, we present four error metrics for each basin scale: the relative bias (rBIAS), the bias propagation factor, the relative RMSE (rRMSE), and the RMSE propagation factor. We first focus on the runoff error by examining rBIAS and rRMSE, reported in Figs. 6 and 7, respectively. Specifically, we start by

computing these two statistics between runoff simulated by forcing the model with MPE rainfall and observed runoff to study the model error. These statistics, shown with triangular symbols in Figs. 6 and 7, are defined as follows:

$$\text{rBIAS} = \frac{\sum_{k=1}^N (\hat{R}_k - R_k)}{\sum_{k=1}^N (R_k)}, \quad (1)$$

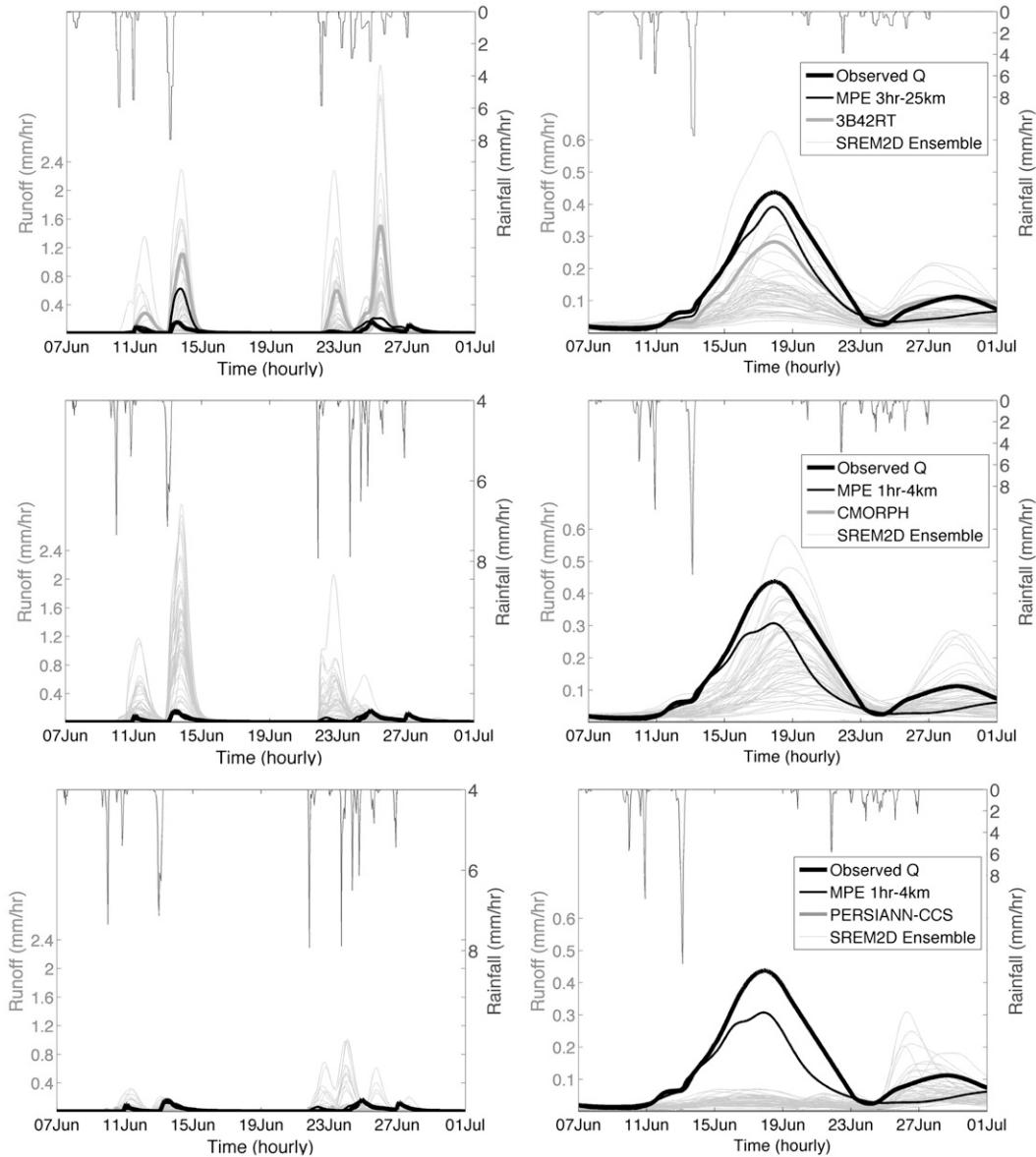


FIG. 5. Time series of MPE satellite product resolution rainfall and simulated runoff for (left) B1 and (right) B5 basins during Tropical Storm Alberto (June 2006). Runoff simulations are based on (top) 3B42-RT, (middle) CMORPH, and (bottom) PERSIANN-CCS retrievals, and the respective SREM2D ensembles and MPE rainfall are aggregated at the respective satellite product’s temporal–spatial resolution: 3 h–25 km for top panels and 1 h–4 km for middle and bottom panels.

$$rRMSE = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{R}_k - R_k)^2}}{\frac{1}{N} \sum_{k=1}^N (R_k)}, \quad (2)$$

where N is the total number of time steps, \hat{R} represents MPE-forced model runoff, and R is the observed runoff. The reference model simulation (i.e., runoff obtained by forcing the hydrologic model with MPE rainfall) is

almost unbiased with respect to ground observations (relative bias is within ± 0.1), as the triangular symbols lie on the zero-line bias (Fig. 6). The rRMSE between the reference simulation and observations is also low, even when the 3 h–25 km model calibration is performed (i.e., representing the 3B42RT resolution case) (Fig. 7).

Next, we compute reference values for these two statistics that measure the errors between runoff simulated by forcing the model with each unperturbed satellite precipitation and the reference MPE-forced model

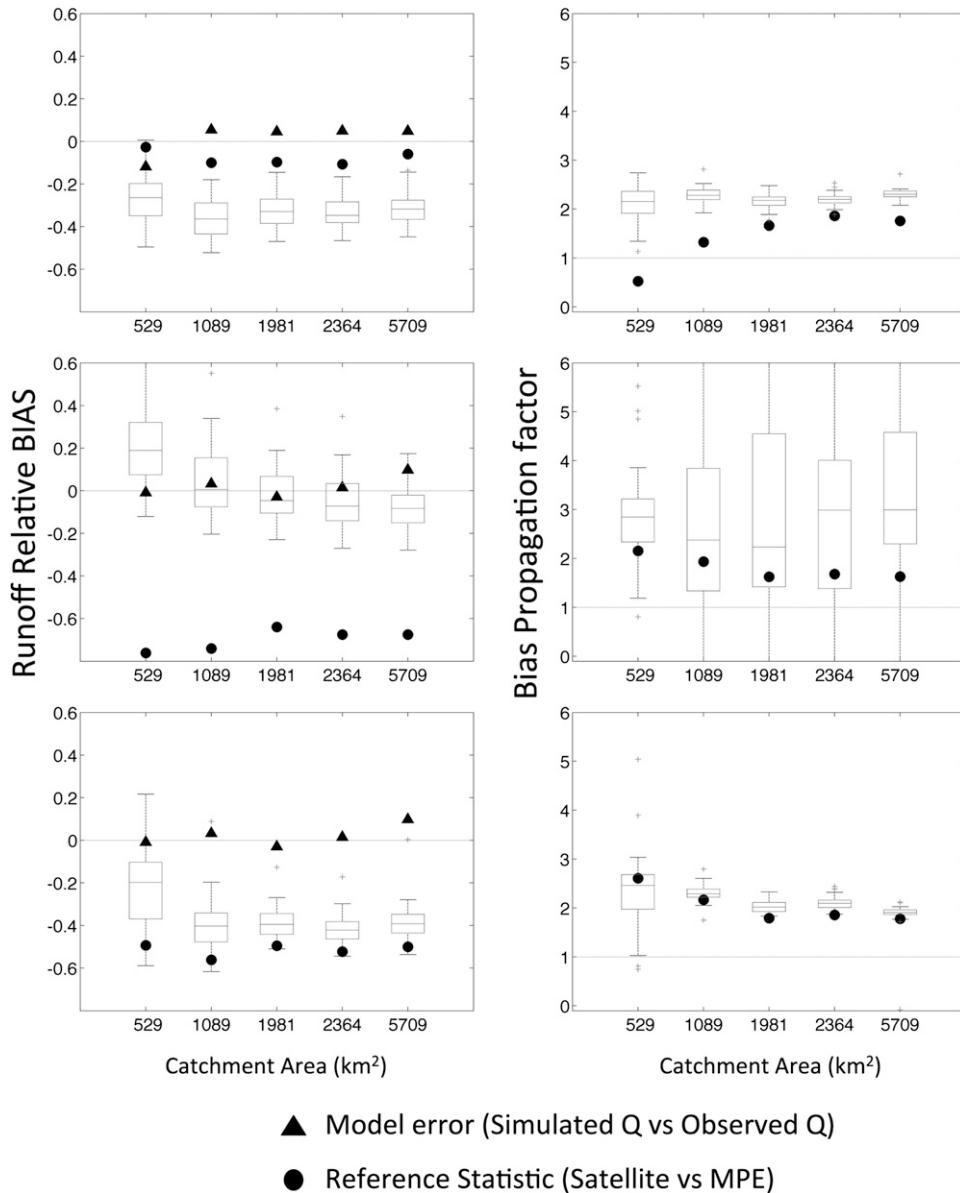


FIG. 6. Runoff rBIAS and bias propagation factor for (top) 3B42RT, (middle) CMORPH, and (bottom) PERSIANN-CCS as a function of catchment area. Model errors (triangle symbols) are based on model simulations with MPE rainfall forcing at the respective satellite product's temporal-spatial resolution: 3 h–25 km for top panels and 1 h–4 km for middle and bottom panels.

runoff. In this case, \hat{R} in Eqs. (1) and (2) represents satellite-forced model runoff and R is the MPE-forced model runoff. The reference statistics are shown in Figs. 6 and 7 as dots. We compute the same statistics for each individual member of the SREM2D ensemble for simulated runoff (again versus the reference MPE-forced runoff simulation). In Eqs. (1) and (2) \hat{R} becomes runoff from a single member of the SREM2D ensemble and R is the MPE-forced model runoff. The rBIAS and rRMSE metrics for the individual ensemble members are shown

in Figs. 6 and 7 as boxplots and can be compared to the reference statistics. For each box in the figures, the central mark is the median, the edges of the box are the 25th and the 75th quantiles, the whiskers are the most extreme data points that the algorithm considers not to be outliers, and the outliers are plotted individually with cross symbols.

Runoff ensemble simulations reduce the bias in both CMORPH- and PERSIANN-CCS-forced simulations when compared to the corresponding reference statistics,

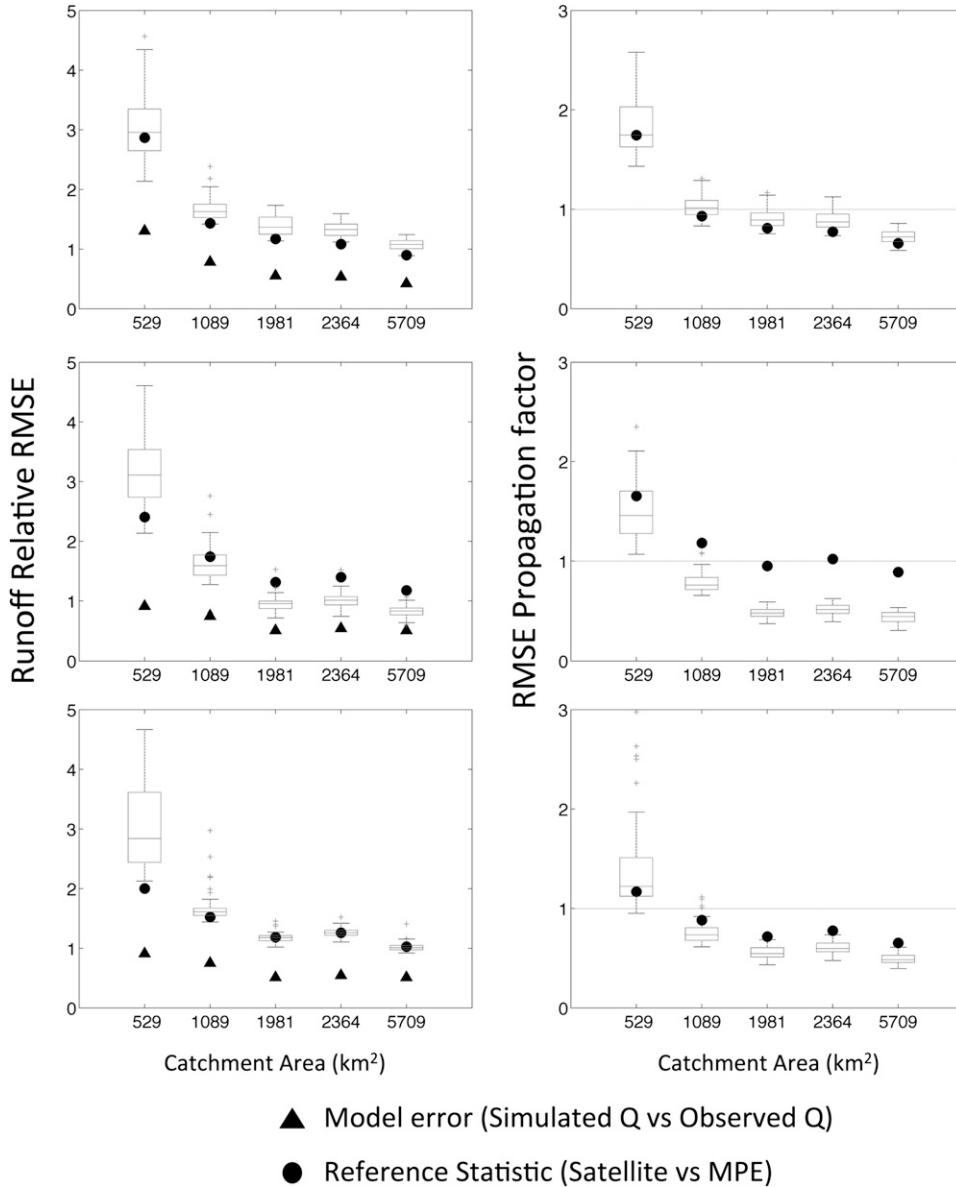


FIG. 7. As in Fig. 6, but for rRMSE and RMSE propagation factor.

showing that SREM2D is able to correct for the bias in the forcing rainfall (Fig. 6). The 3B42RT product yields nearly unbiased streamflow estimates. However, SREM2D introduces a slight bias in the predicted runoff ensemble. No catchment area dependency is observed for the relative bias for any for the satellite rainfall product (Fig. 6).

When looking at the random error analysis in Fig. 7, in most cases reference statistics are within the corresponding ensemble metrics (i.e., dots are within the boxplot bounds). This indicates that SREM2D ensembles are able to replicate the reference random error. Moreover, the relative RMSE decreases as a function of catchment area consistently for all ensembles as well as

for the reference statistics. As expected, the spread of the rRMSE ensembles (shown by the boxplot width) decreases with increasing basin area, indicating that the high variability of the random error observed at small basin scales is smoothed out for larger catchments.

Figures 6 and 7 also present the bias and the RMSE propagation factors, respectively. In this study, we define the bias (RMSE) propagation factor as the ratio between the runoff relative bias (or relative RMSE) to the corresponding rainfall relative bias (or relative RMSE). Analogously to what we presented for the runoff metrics, we compute the propagation factors as both reference statistics and ensemble statistics.

The first point to note is that the relative bias doubles from rainfall to runoff consistently across all basin scales and for the three satellite retrievals (Fig. 6), which confirms what was observed in previous studies (e.g., Nikolopoulos et al. 2010). The second point to note is that CMORPH exhibits more ensemble member variability in the rainfall-to-runoff bias propagation compared to the other two satellite products (3B42RT and PERSIANN-CCS), as shown by the wider boxplots and already highlighted by the time series (Figs. 3, 4).

Similar to what is shown for the runoff rRMSE, the RMSE propagation factor decreases as a function of catchment area for all ensembles to about 0.5 for PERSIANN-CCS and CMORPH and to 0.8 for 3B42RT products (Fig. 7). The CMORPH and PERSIANN-CCS RMSE propagation factors cross the ratio 1 line at $\sim 1000\text{-km}^2$ basin scale. In other words, for basin scales less than 1000 km^2 (or 2000 km^2 for 3B42RT) the rainfall-forcing random error is enhanced by the hydrological model in the streamflow prediction, whereas by increasing the basin area the error in rainfall is dampened by the model integration.

The exceedance ratio (ER) metric is introduced to evaluate the ability of the ensemble simulations to represent the runoff prediction uncertainty for the different satellite rainfall products. This metric is computed for each basin scale and satellite product by determining the fraction of reference runoff values that lie outside the quantile range of the ensemble runoff simulations:

$$\text{ER} = \frac{N_{\text{exceedance}}}{N_t}, \quad (3)$$

where $N_{\text{exceedance}}$ is the number of instances the reference streamflow falls outside a specified quantile range of the ensemble envelope and N_t is the total number of time steps. Reference runoff simulations are obtained by forcing the hydrological model with the corresponding satellite-product resolution MPE rainfall data (3 h–25 km for 3B42RT and 1 h–4 km for CMORPH and PERSIANN-CCS). We considered different quantile ranges [(0.05–0.95) to (0.45–0.55)] that represent different exceedance probabilities (EPs) (0.1–0.9). Comparing the quantile EP values to the corresponding ER values calculated from the simulated ensemble provides an evaluation of the error variability predicted by SREM2D. An ER value higher (lower) than EP would indicate that the ensemble runoff simulations underestimate (overestimate) the actual variability of the error propagation of the satellite products. Close agreement between ER and EP values would indicate an accurate representation of satellite error propagation. Three runoff thresholds (representing the 20th, 50th, and 95th

quantiles of the reference simulation) are used to evaluate the ensemble simulations.

Results are presented in Fig. 8 that show the ER/EP ratio versus EP for the three different satellite products, basin scales, and runoff thresholds [20th (left), 50th (middle), 95th (right)]. A point to note is that in all products SREM2D-based ensembles give ER values consistently higher than the EP values (ratios > 1), which indicates underestimation of the error variability. CMORPH ensembles exhibit the lowest underestimation compared to the other two retrievals for all catchment areas examined in this study. This is particularly significant at low EP values (< 0.2), which translates into better accuracy for encapsulating the reference simulations. The ratios do not show strong dependency on catchment area for CMORPH and PERSIANN-CCS ensembles, particularly for the 20th and 50th threshold values. However, when the 3B42RT ensemble is used to force the model, stronger ER/EP ratios are observed for the larger basins, which indicates basin-scale dependence.

To further investigate the runoff ensembles obtained by forcing the hydrological model with the three perturbed satellite rainfall products, we introduce the critical success index (CSI), or threat score (Gourley et al. 2012), defined as follows:

$$\text{CSI} = \frac{c}{a + b + c}, \quad (4)$$

where c is the number of times the reference-simulated runoff and the ensemble member exceed a certain runoff threshold, a is the number of times that the reference runoff exceeds the threshold but the ensemble member does not, and b is the number of times that the ensemble member exceeds the threshold but the reference does not. CSI ranges between 0 and 1, and a perfect performance corresponds to an index of 1. For the three satellite rainfall products, the index is computed for each of the 50 ensemble members. Figure 9 shows the mean with error bars, expressed in terms of standard deviations, across all the ensemble members for each rainfall product as a function of the catchment area. Three runoff thresholds corresponding to the 20th, 50th, and 95th quantiles of the reference simulation are computed and shown for the large basin. Figure 9 also shows CSI of MPE-forced runoff simulations compared against USGS streamflow observations as a function of basin area. Considering that MPE is an accurate rainfall estimate, CSI values of MPE-driven runoff simulations represent the modeling uncertainty, which is contrasted in this study to the different CSI values representing the satellite rainfall uncertainty alone.

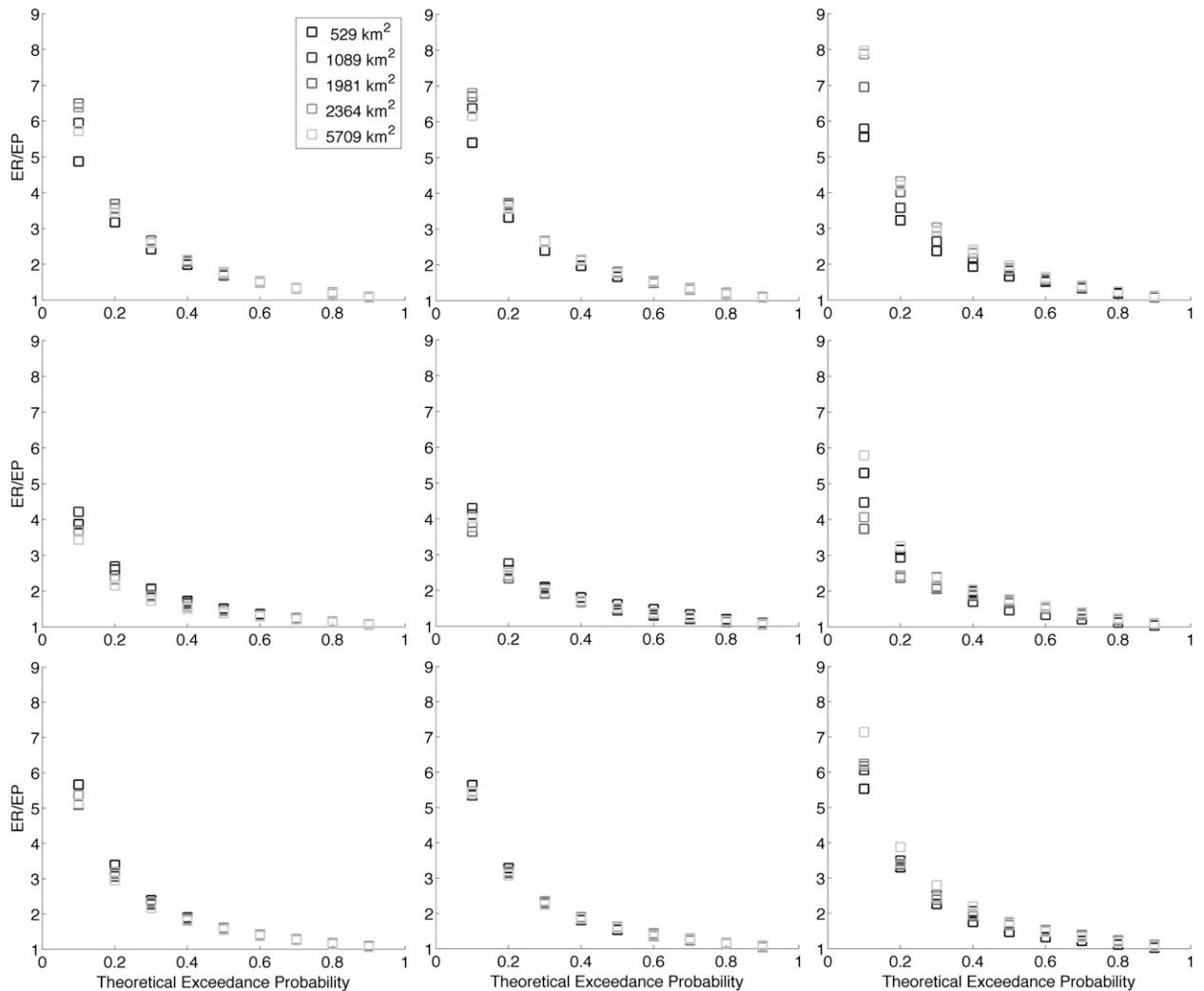


FIG. 8. Ratio of the ensemble-derived ERs to the quantile-range EPs vs EPs for the different catchment areas, satellite products [(top) 3B42RT, (middle) CMORPH, and (bottom) PERSIANN-CCS], and runoff thresholds associated with different quantiles: (left) 20th, (middle) 50th, and (right) 95th.

Analogously to what was found for the exceedance ratio, the highest CSIs are observed for the CMORPH rainfall product ensemble-derived runoff, whereas 3B42RT and PERSIANN-CCS perform similarly to each other. As expected, by increasing the threshold (i.e., by accounting for more extreme events) CSI decreases. The best performances are obtained when we consider all excess runoff values (i.e., using the 20th quantile as a threshold to remove baseflow values). Moreover, a dependence on the basin area is noted for the three satellite-derived runoff ensembles, particularly when the threshold is set at the 20th and 50th quantiles. Comparing the satellite rainfall to modeling uncertainty in streamflow simulations, we note similar CSI values in the case of CMORPH rainfall product ensemble-derived runoff for the 20th and 50th quantile

thresholds, whereas 3B42RT and PERSIANN-CCS exhibit poorer performances. However, when only high runoff values are considered (i.e., 95th quantile threshold), the runoff ensemble performances degrade for all satellite rainfall datasets (~ 0.3) and do not exhibit any dependence on the catchment area, which indicates a modest capability of these products in predicting flood events.

5. Conclusions

In this study, we investigated uncertainties associated with satellite rainfall products and the subsequent impact on streamflow model predictions. We used an ensemble-based satellite rainfall error model (SREM2D) to examine the error propagation of three satellite rainfall

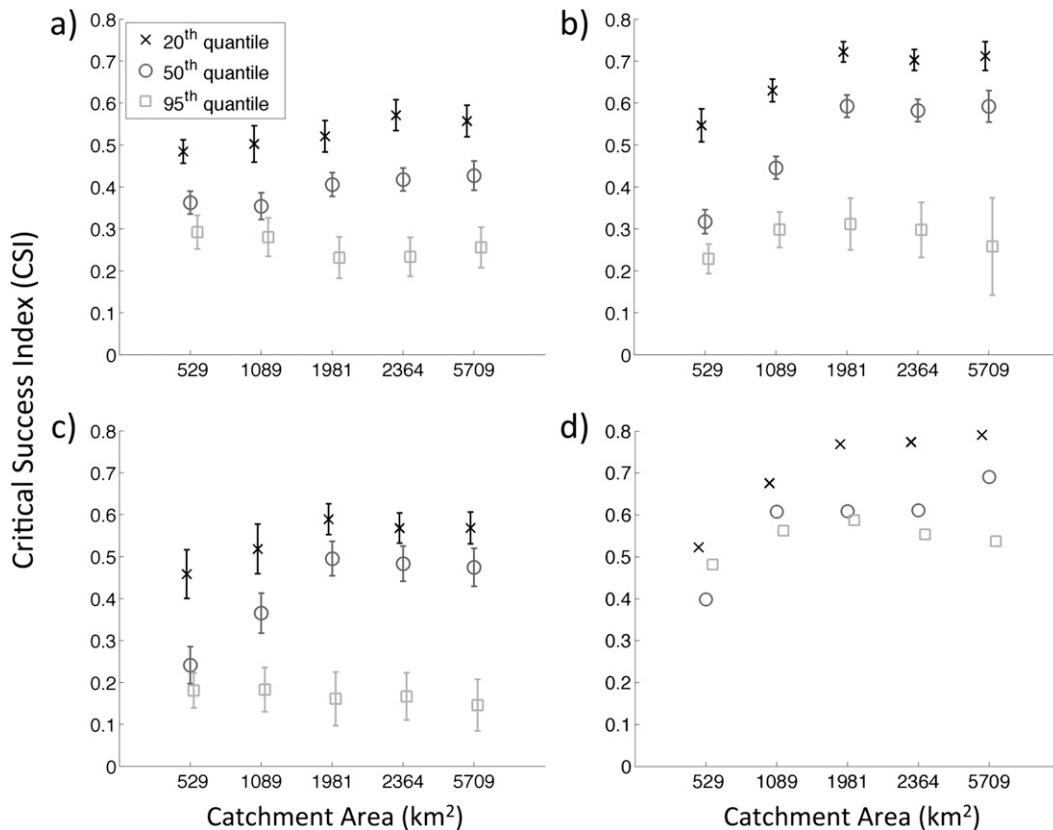


FIG. 9. CSI (dimensionless) for (a) 3B42RT, (b) CMORPH, and (c) PERSIANN-CCS ensemble-derived runoff as a function of catchment area. (d) CSI for the MPE rainfall-forced simulation computed against streamflow observations.

products in flood simulation across a range of basin scales in the Tar-Pamlico River basin located in the southeastern United States. The three satellite rainfall products are TRMM 3B42RT, CMORPH, and PERSIANN-CCS, and they were evaluated against the reference MPE precipitation dataset. We forced the NWS HL-RDHM hydrologic model with precipitation fields to produce streamflow in a deterministic mode, where input satellite and radar rainfall was not corrected by the error model, and then in an ensemble mode, where the three satellite rainfall products were independently corrected using SREM2D. Ground-measured runoff and the MPE-driven runoff simulations (i.e., the reference simulation) were used as benchmarks to evaluate the ensemble streamflow simulations obtained by forcing the model with SREM2D-perturbed satellite rainfall.

First, we conducted a qualitative analysis of the rainfall-to-runoff error propagation. Time series of cumulative rainfall and cumulative runoff during 2006 for two basins (the smallest and largest) were examined. We observed that for both small and large basins, SREM2D rainfall ensembles were closer to the reference MPE

rainfall than the uncorrected deterministic satellite products, and that in most cases, SREM2D ensembles enveloped the MPE rainfall. Regarding runoff simulations, at the small basin scale the runoff ensembles were able to envelope both the reference runoff simulation as well as observed streamflow for all three satellite rainfall products. However, at the largest basin scale the PERSIANN-CCS-derived runoff ensemble did not envelope either of the benchmark streamflows. Furthermore, an enhancement of the rainfall-to-runoff relative errors is exhibited by the ensemble spread of the runoff simulations compared to the spread of the corresponding rainfall ensembles. From tropical storm Alberto, both unperturbed and perturbed 3B42RT-derived runoff simulations overestimated peak runoff on the small basin and slightly underestimated in the larger basin. CMORPH yielded no considerable runoff at any basin scale, but the use of SREM2D perturbations provided runoff simulations that encapsulated the observed peaks at both basin scales. However, SREM2D perturbations were not able to resolve the biases in PERSIANN-CCS rainfall and associated runoff simulations.

The quantitative error analysis using rBIAS and rRMSE metrics exhibited good agreement between reference model simulations and ground observations. Then, reference statistics, defined as errors between runoff simulated by forcing the model with each satellite precipitation product (unperturbed and SREM2D perturbed) and the reference MPE-forced runoff simulations, were evaluated together with ensemble statistics. SREM2D was shown to correct for the bias in the rainfall forcing, as the rBIAS in the runoff ensembles was shown to be lower than the corresponding reference statistics for both CMORPH and PERSIANN-CCS-forced simulations. However, for the 3B42RT product, which produced almost unbiased streamflow in the validation period, SREM2D introduced a slight bias in the runoff ensemble. No dependency on catchment area was observed for rBIAS, while the relative RMSE decreased as a function of basin area for all ensembles and for the reference statistics. The spread of the ensemble random error also decreased with increasing catchment area.

We defined and then evaluated bias and RMSE rainfall-to-runoff propagation factors to investigate how satellite rainfall error propagates to runoff simulations through the hydrological model. We observed that the relative bias doubles from rainfall to runoff across all basin scales and for all three satellite retrievals, with CMORPH exhibiting more ensemble variability in the rainfall-to-runoff bias propagation. The RMSE propagation factor decreased as a function of catchment area for all ensembles. Specifically, for small watersheds ($\sim 500 \text{ km}^2$) the random error increases in the propagation from rainfall to runoff, whereas by increasing the basin area the rainfall random error is dampened by the model. Error metrics, evaluated for each satellite products, were usually within the corresponding SREM2D-ensemble error metrics boxplot, which demonstrates the SREM2D's capability to reproduce the satellite rainfall-to-runoff error propagation.

The ability of the SREM2D-generated runoff ensemble to characterize the error variability was evaluated through the comparison of ensemble-generated ERs to the EPs associated with specified quantile ranges. We showed that simulated runoff ensembles from SREM2D gave ER/EP ratios systematically above 1, which indicates underestimation of the actual error variability. The CMORPH-derived runoff ensemble exhibited ER/EP ratios closer to 1, which indicates the better capability to encapsulate the reference simulations, compared to 3B42RT and PERSIANN-CCS. Moreover, low ER/EP ratio dependency on the catchment area is noted for CMORPH and PERSIANN-CCS ensembles, while the 3B42RT ensemble yielded notably higher ER/EP ratios for the larger basins, particularly at low exceedance probability values.

Finally, we presented the critical success index (CSI) to evaluate the performance of the runoff ensembles obtained by forcing the model with the three SREM2D-perturbed satellite products. The highest CSI values were observed for the CMORPH ensemble-derived runoff, whose performance was found comparable to the one of the reference MPE-forced simulation. However, when only extreme events were taken into account, the ensemble performances degraded to values lower than 0.3 for all the satellite rainfall products, which suggests that the ability of these satellite products in the prediction of extreme events (above the 95th quantile) is modest.

A caveat to our results is that we tested the rainfall-to-runoff error propagation using a single hydrologic model in the Tarboro basin based on a short period (2 yr) of satellite rainfall data. Furthermore, we acknowledge the fact that most of the satellite rainfall products are currently undergoing revisions and improvements that would not fully represent the performance statistics presented herein. To generalize the findings of this study, more systematic investigations are needed to involve longer time series with the latest product versions, as well as additional flood cases associated with different hydroclimatic regimes. However, we are confident that our general findings about the error propagation and the SREM2D ensemble-based error variability remain valid if other hydrological models were employed and in regions of the world with climatology and terrain that are not too different from that of the study area, even if some details of the error statistics might change. Overall, we believe the results obtained from this study provide useful information about the application of satellite rainfall products and SREM2D ensemble error model to simulate hydrologic processes. This provides valuable feedback not only to future hydrologic missions, for example, NASA Global Precipitation Measurement (<http://gpm.gsfc.nasa.gov>), but also to the development or implementation of ensemble satellite rainfall for flood prediction.

Acknowledgments. The research for this paper was supported by NASA Precipitation Science Team Grant NNX07AE31G (E. Anagnostou and Y. Hong). V. Maggioni was supported by a NASA Earth System Science Graduate Fellowship.

REFERENCES

- Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlinear Processes Geophys.*, **8**, 401–417.
- Bales, J., C. R. Wagner, K. C. Tighe, and S. Terziotti, 2007: LiDAR-derived flood-inundation maps for real-time flood-mapping

- applications, Tar River Basin, North Carolina. North Carolina Floodplain Mapping Program, USGS, Reston, VA, 52 pp. [Available online at <http://pubs.usgs.gov/sir/2007/5032/pdf/SIR2007-5032.pdf>.]
- Behrangi, A., B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian, 2011: Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.*, **397**, 225–237.
- Bitew, M. M., and M. Gebremichael, 2011: Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands. *Hydrol. Earth Syst. Sci.*, **15**, 1147–1155.
- Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, 1973: A generalized streamflow simulation system—Conceptual modeling for digital computers. Tech. Rep. to the Joint Federal and State River Forecast Center, U.S. National Weather Service and California Department of Water Resources, Sacramento, 204 pp.
- Cloke, H. L., and F. Pappenberger, 2009: Ensemble flood forecasting: A review. *J. Hydrol.*, **375**, 613–626.
- Duan, Q., V. K. Gupta, and S. Sorooshian, 1993: Shuffled complex evolution approach for effective and efficient global minimization. *J. Optim. Theory Appl.*, **76**, 501–521.
- Gourley, J. J., Y. Hong, Z. L. Flamig, J. Wang, H. Vergara, and E. N. Anagnostou, 2011: Hydrologic evaluation of rainfall estimates from radar, satellite, gauge, and combinations on Ft. Cobb basin, Oklahoma. *J. Hydrometeorol.*, **12**, 973–988.
- , J. M. Erlingis, Y. Hong, and E. B. Wells, 2012: Evaluation of tools used for monitoring and forecasting flash floods in the United States. *Wea. Forecasting*, **27**, 158–173.
- Gouweleeuw, B. T., J. Thielen, G. Franchello, A. P. J. de Roo, and R. Buizza, 2005: Flood forecasting using medium-range probabilistic weather prediction. *Hydrol. Earth Syst. Sci.*, **9**, 365–380, doi:10.5194/hess-9-365-2005.
- Guetter, A. K., K. P. Georgakakos, and A. A. Tsonis, 1996: Hydrologic applications of satellite data: 2. Flow simulation and soil water estimates. *J. Geophys. Res.*, **101** (D21), 26 527–26 538.
- Hong, Y., K. L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteorol.*, **43**, 1834–1852.
- Hossain, F., and E. N. Anagnostou, 2006a: A two-dimensional satellite rainfall error model. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1511–1522.
- , and —, 2006b: Using a multi-dimensional satellite rainfall error model to characterize uncertainty in soil moisture fields simulated by an offline land surface model. *Geophys. Res. Lett.*, **32**, L15402, doi:10.1029/2005GL023122.
- , —, and T. Dinku, 2004: Sensitivity analyses of satellite rainfall retrieval and sampling error on flood prediction uncertainty. *IEEE Trans. Geosci. Remote Sens.*, **42**, 130–139, doi:10.1109/TGRS.2003.818341.
- Huffman, G. J., and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55.
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503.
- Khan, S. I., and Coauthors, 2011: Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria basin: Implications for hydrologic prediction in ungauged basins. *IEEE Trans. Geosci. Remote Sens.*, **49**, 85–95, doi:10.1109/TGRS.2010.2057513.
- Kidd, C. K., D. R. Kniveton, M. C. Todd, and T. J. Bellerby, 2003: Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeorol.*, **4**, 1088–1104.
- Koren, V., M. Smith, D. Wang, and Z. Zhang, 2000: Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. Preprints, *15th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., 2.16. [Available online at <https://ams.confex.com/ams/annual2000/webprogram/Paper6074.html>.]
- , —, Q. Duan, Q. Duan, H. Gupta, S. Sorooshian, A. Rousseau, and R. Turcotte, 2003: Use of *a priori* parameter estimates in the derivation of 9 spatially consistent parameter sets of rainfall-runoff models. *Calibration of Watershed Models*, Q. Duan et al., Eds., Water Science and Application, Vol. 6, Amer. Geophys. Union, 239–254.
- , S. Reed, M. Smith, Z. Zhang, and D. J. Seo, 2004: Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. *J. Hydrol.*, **291**, 297–318.
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15**, 809–817.
- Li, L., and Coauthors, 2009: Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Nat. Hazards*, **50**, 109–123, doi:10.1007/s11069-008-9324-5.
- Maggioli, V., R. H. Reichle, and E. N. Anagnostou, 2011: The effect of satellite-rainfall error modeling on soil moisture prediction uncertainty. *J. Hydrometeorol.*, **12**, 413–428.
- NCDWR, cited 2009: Drought monitor history. [Available online at http://www.ncwater.org/Drought_Monitoring/dmhistory/?startdate=2007-01-02&enddate=2009-12-29&label=false.]
- Nijssen, B., and D. P. Lettenmaier, 2004: Effect of precipitation sampling error on simulated hydrological Fluxes and states: Anticipating the Global Precipitation Measurement satellites. *J. Geophys. Res.*, **109**, D02103, doi:10.1029/2003JD003497.
- Nikolopoulos, E. I., E. N. Anagnostou, F. Hossain, M. Gebremichael, and M. Borga, 2010: Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model. *J. Hydrometeorol.*, **11**, 520–532.
- , —, and M. Borga, 2012: Using high-resolution satellite rainfall products to simulate a major flash flood event in northern Italy. *J. Hydrometeorol.*, **14**, 171–185.
- NWS, OHD, HL, and HSMB, 2009: Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) User's Manual V. 2.4.3. CBRFC Internal Doc., 108 pp. [Available online at http://amazon.nws.noaa.gov/articles/HRL_Pubs_PDF_May12_2009/zc/RDHM_User_Manual.pdf.]
- Pappenberger, F., K. J. Beven, N. M. Hunter, P. D. Bates, B. T. Gouweleeuw, J. Thielen, and A. P. J. de Roo, 2005: Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrol. Earth Syst. Sci.*, **9**, 381–393.
- , J. Bartholmes, J. Thielen, H. L. Cloke, R. Buizza, and A. de Roo, 2008: New dimensions in early flood warning across the globe using grand-ensemble weather predictions. *Geophys. Res. Lett.*, **35**, L10404, doi:10.1029/2008GL033837.
- Pokhrel, P., H. V. Gupta, and T. Wagener, 2008: A spatial regularization approach to parameter estimation for a distributed

- watershed model. *Water Resour. Res.*, **44**, W12419, doi:10.1029/2007WR006615.
- Seo, D. J., and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor.*, **3**, 93–111.
- Sorooshian, S., K. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046.
- Su, F., Y. Hong, and D. P. Lettenmaier, 2008: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *J. Hydrometeor.*, **9**, 622–640.
- Thielen, J., J. Bartholmes, M.-H. Ramos, and A. de Roo, 2009: The European flood alert system—Part 1: Concept and development. *Hydrol. Earth Syst. Sci.*, **13**, 125–140.
- USGS, 2007: Lowest streamflows in more than 110 years for some North Carolina rivers as drought worsens. USGS News Release, 31 August. [Available online at <http://www.usgs.gov/newsroom/article.asp?ID=1767>.]
- , cited 2011: Instantaneous Data Archive. [Available online at <http://nwis.waterdata.usgs.gov/nwis/>]
- Van Cooten, S., and Coauthors, 2011: The CI-FLOW Project: A system for total water level prediction from the summit to the sea. *Bull. Amer. Meteor. Soc.*, **92**, 1427–1442.
- Verbunt, M., A. Walser, J. Gurtz, A. Montani, and C. Schar, 2007: Probabilistic flood forecasting with a limited-area ensemble prediction system: Selected case studies. *J. Hydrometeor.*, **8**, 897–909.
- Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson, 2008: Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environ. Res. Risk Assess.*, **23**, 1011–1026, doi:10.1007/s00477-008-0274-y.
- , —, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon, 2009: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlinear Sci. Numer. Simul.*, **10**, 273–290.
- Yilmaz, K. K., H. V. Gupta, and T. Wagener, 2008: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, **44**, W09417, doi:10.1029/2007WR006716.
- , R. F. Adler, Y. Tian, Y. Hong, and H. F. Pierce, 2010: Evaluation of a satellite-based global flood monitoring system. *Int. J. Remote Sens.*, **31**, 3763–3782.