Effects of Retrospective Gauge-Based Readjustment of Multisensor Precipitation Estimates on Hydrologic Simulations

YU ZHANG, SEANN REED, AND DAVID KITZMILLER

National Oceanic and Atmospheric Administration/National Weather Service Office of Hydrologic Development, Silver Spring, Maryland

(Manuscript received 5 August 2009, in final form 23 August 2010)

ABSTRACT

This paper presents methodologies for mitigating temporally inconsistent biases in National Weather Service (NWS) real-time multisensor quantitative precipitation estimates (MQPEs) through rain gauge–based readjustments, and examines their effects on streamflow simulations. In this study, archived MQPEs over 1997–2006 for the Middle Atlantic River Forecast Center (MARFC) area of responsibility were readjusted at monthly and daily scales using two gridded gauge products. The original and readjusted MQPEs were applied as forcing to the NWS Distributed Hydrologic Model for 12 catchments in the domain of MARFC. The resultant hourly streamflow simulations were compared for two subperiods divided along November 2003, when a software error that gave rise to a low bias in MQPEs was fixed. It was found that readjustment at either time scale improved the consistency in the bias in streamflow simulations. For the earlier period, independent monthly and daily readjustments considerably improved the streamflow simulations for most basins as judged by bias and correlation. By contrast, for the later period the effects were mixed across basins. It was also found that 1) readjustments tended to be more effective in the cool rather than warm season, 2) refining the readjustment resolution to daily had mixed effects on streamflow simulations, and 3) at the daily scale, redistributing gauge rainfall is beneficial for periods with substantial missing MQPEs.

1. Introduction

Multisensor quantitative precipitation estimates (MQPEs) are created by the National Weather Service (NWS) River Forecast Centers (RFCs) at near–real time during river and flash flood forecasting operations. The primary basis of MQPEs are precipitation gauge reports and observations of Weather Surveillance Radar-1988 Doppler [WSR-88D; Fulton et al. (1998); see Seo (1998a,b), Seo et al. (1999), and Seo and Breidenbach (2002) for methodology]. At an hourly step and a nominal 4-km mesh length, these products serve as forcing to NWS's hydrologic model and have seen increasing natural resource–related applications (Hardegree et al. 2008; Over et al. 2007).

Despite the promising prospects, these real-time MQPEs remain subject to biases and inaccuracies because of a variety of factors (e.g., Young et al. 1999, 2000; Jayakrishnan et al. 2004; Zhang et al. 2007). The radar precipitation estimates can be compromised by problems ranging from beam blockage to variable drop size distribution (Smith et al. 1996; Krajewski and Smith 2002). The resultant errors can be mitigated by rain gauge–based bias adjustments and multisensor merging (Seo et al. 1999, 2000a,b; Seo and Breidenbach 2002; Seo 1998a,b). Yet, the actual impacts of these measures are limited by the quantity and quality of gauge reports available during operations. Moreover, the software for creating MQPEs contained deficiencies that were corrected over time. Perhaps the most widely known among these is the "truncation error" (TE)—an error in the Next Generation Weather Radar (NEXRAD) precipitation processing system (PPS) that resulted in underestimation of rainfall amounts (Fulton et al. 2003; Zhang et al. 2007). Subsequent software upgrades corrected this error, but these changes, as demonstrated later in this paper, also contributed to temporally inconsistent bias characteristics of the MQPEs.

Hydrologic models are known to be sensitive to the spatiotemporal resolution of the input forcing data (Finnerty et al. 1997; Schaake et al. 1996; Koren et al. 1999) and, therefore, may need to be calibrated specifically against archived MQPEs to take full advantage of their

Corresponding author address: Yu Zhang, National Oceanic and Atmospheric Administration/National Weather Service Office of Hydrologic Development, 1325 East-West Hwy., Silver Spring, MD 20910. E-mail: yu.zhang@noaa.gov

DOI: 10.1175/2010JHM1200.1

higher resolution (see Reed et al. 2004 on benefits of MQPE-based calibration). Yet, the inconsistent bias behaviors of archived MQPEs have since limited their use in model calibration. As such, there has been a great demand for mitigating the inconsistent MQPE bias to facilitate calibration and the application of model-based techniques designed to predict the severity of events relative to modeled historical trends [see, e.g., Reed et al. (2007) for flash flood, Thielen et al. (2008) for river flood, and NCEP-CPC (2005) for soil moisture anomalies].

For removing incoherent temporal bias of the NWS MQPEs arising from software errors and related upgrades, an intuitive approach would be to apply the latest version of the software retrospectively to regenerate the MQPEs. This approach, however, requires human labor and computational power well beyond the current operational capacity of NWS RFCs where the MQPE products are produced. In this paper we propose and evaluate a simpler alternative—that is, retrospectively adjusting the archived real-time MQPEs to match gridded gauge products at and beyond the daily time scale (referred to as ''readjustment approach''). The readjustment approach has the advantage of being able to integrate daily and subdaily records that supplement the operational hourly gauge data involved in creating the MQPEs. It resembles that used in the North American Land Data Assimilation System (NLDAS; Cosgrove et al. 2003), but differs from NLDAS as 1) it uses MQPEs as the primary data source to be bias corrected, rather than as a reference for disaggregating gauge data (as in NLDAS); and 2) it retains the high spatial resolution of NWS MQPEs (NLDAS results are on a $1/8°$ grid).

As the gauge-based readjustment is expected to enhance the quality of archived real-time MQPEs and thereby benefit streamflow simulations, its effectiveness needs to be systematically assessed. Moreover, the appropriate temporal scales for applying the readjustment remain to be determined. This paper addresses these issues through a set of hydrologic simulation experiments in a pilot domain, wherein the comparative impacts of various readjustment schemes on streamflow simulations were evaluated, and these served as the basis for inferring the differential quality of readjusted MQPEs [see similar approaches in Ding et al. (2005a,b)]. These experiments were carried out for the period of 1997–2006 for 12 catchments in central Maryland that lie in the forecast domain of the Middle Atlantic River Forecast Center (MARFC). The readjustment was performed on a monthly and daily basis using 1) 2.5-min Parameter-elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994) outputs and 2) $1/8°$ gridded gauge–based analyses [from the National Centers for Environmental Prediction, Climate Prediction Center (NCEP-CPC)], respectively. The comparative use of these two datasets allows a close examination of the temporal scale dependence in the effects of the readjustments. In addition, by stratifying the streamflow simulations prior to and following the correction, and between the warm and cool seasons, this study examines the seasonal dependence in the efficacy of the methodology and the impacts of the truncation error.

The remainder of the paper is structured as follows. Section 2 offers a brief overview of the NWS quantitative precipitation estimation procedures and associated issues. Section 3 describes study sites, reference datasets, readjustment methodologies, and the hydrologic model. Section 4 presents the simulation results, and section 5 discusses the results and concludes the study.

2. Radar and multisensor precipitation estimation

a. *PPS and MPE*

The NWS precipitation estimation algorithms can be divided into radar-only and multisensor components (the latter include a gauge-only component). The major radar-only component is the PPS (Fulton et al. 1998), which produces real-time estimates of liquid-only precipitation at ground level (Fulton et al. 1998). The PPS outputs are disseminated as digital precipitation array (DPA; also referred to as Stage I products). In the legacy multisensor algorithms, the Stage I products underwent bias correction in the Stage II processing, and the results were quality assured, and mosaicked on a regional basis, to produce Stage III data. The common map grid system for all three stages is the polar stereographic Hydrologic Rainfall Analysis Project (HRAP) mesh (nominal 4-km resolution in the central United States; Greene and Hudlow 1982; Reed and Maidment 1999). The legacy Stage II and III algorithms have since been replaced by the multisensor precipitation estimator (MPE). MPE incorporates new and updated algorithms for mean-field and local bias correction in its products (Seo et al. 1999; Seo and Breidenbach 2002), as well as provisions for creating gauge-only and gauge-radar merged products (Seo 1998a,b). In this paper, Stage III and the later MPE-based products are both referred to as MQPEs.

MQPEs are generated during real-time operations at the RFCs, where forecasters quality control the gauge reports and use the hourly reports to conduct bias correction and multisensor merging (only hourly reports can be directly ingested). To augment the hourly reports, several RFCs now disaggregate daily precipitation reports using collocated radar estimates as reference to create artificial hourly reports, and then lump them with the hourly gauge reports in generating the finalized MQPEs. A large portion of the gauge data are supplied via the Hydrometeorological

Automated Data System (HADS), a real-time platform that acquires and distributes records from a large number of gauges (~6000). Data delivered outside of HADS include those from state and local agencies (i.e., mesonet data), the automated surface observing system (ASOS), the Cooperative (COOP) gauge network, and other networks such as the Community Collaborative Rain, Hail and Snow (CoCoRaHS) Network. It is noted that most automated COOP gauge data are delivered at real time by HADS, whereas the manual ones are not available at real time.

b. TE

Around 2000, analyses of PPS products revealed that the software erroneously truncated scan-to-scan accumulation amounts less than 0.1 mm (Seo et al. 2000a,b, 2002; Fulton et al. 2003). This TE is most pronounced during prolonged light rain events, and could lead to considerable underestimation of the cumulative rainfall in the archived real-time MQPEs (Fulton et al. 2003). The error was corrected in three successive software releases during 2002/03 (Fulton et al. 2003). An independent investigation by MARFC staff, in which MQPEs were compared with reference daily rain gauge reports, found a similar change in bias (Cognitore 2005). In both instances, the bias of MQPEs versus the reference gauge dataset went from low to near neutral.

3. Methodology

a. Reference data and readjustment schemes

The proposed readjustment approach involves modifying the MQPEs, on either a daily or monthly basis, to match corresponding gridded gauge products. These products are assumed to be less biased than the MQPE products as they incorporated gauge reports that were not available in real time. Two such gauge products are employed: 1) the 2.5 arc-minute monthly PRISM and 2) the ⅛ arc-degree gridded daily gauge analysis from NCEP-CPC. Each dataset and the associated readjustment methodology are briefly described below.

1) READJUSTMENT WITH MONTHLY PRISM ANALYSES

PRISM is a knowledge-based approach developed by Daly et al. (1994) that accounts for orographic and other effects in producing gridded products from gauge measurements. In essence, it utilizes a digital elevation model (DEM) to group gauges, and then integrates expert knowledge of other factors (e.g., coastal proximity, boundary layer depth) to determine an elevation–precipitation relationship via regression for each group. The PRISM datasets have been used regularly as a proxy for ground truth for readjusting real-time gauge data at the RFCs. The PRISM monthly dataset on a 2.5 arc-minute grid mesh is employed in this study (available from http://www.prism.oregonstate.edu; see Daly et al. 2004 for methodology). The dataset, which is originally in geographic projection, has been reprojected onto the HRAP grid via a nearest-neighbor method—that is, for each HRAP pixel, the value from the most closely located PRISM pixel (measured under HRAP projection between the HRAP and PRISM pixel centers) is assigned.

The PRISM-based readjustment proceeded as follows. First, hourly MQPE grids were summed within each month to produce a monthly total. To accommodate missing values, the monthly total for each HRAP grid cell was scaled by the ratio of total number of hours to the number of hours without missing values when the latter is no greater than three days (no readjustment is done for cells with more three days of missing values over a month). Then, for each HRAP cell, a multiplicative bias factor was computed as the ratio of PRISM to the MQPE monthly totals. The bias grid thus generated was subsequently applied to scale each hourly MQPE grid for that month to derive the final, readjusted MQPEs, with the assumption that the bias is constant within each month.

2) READJUSTMENT WITH DAILY CPC ANALYSES

The daily CPC gridded gauge-only products were created by NCEP-CPC and are available for the entire continental United States (CONUS). The dataset is produced by interpolating daily and subdaily gauge observations onto ⅛° grids, with each value for a particular day representing areal precipitation accumulated over a 24-hour period ending 1200 UTC of that day. Generating this dataset entails the use of a PRISM scheme in spatial interpolation, and daily reports not ending 1200 UTC were temporally interpolated. On the average, about 11 000 observations per day are involved in creating the archival gridded analysis.

Two schemes were devised for readjusting the hourly MQPE using the daily CPC data. The first one is similar to that for monthly PRISM-based readjustment—that is, gridded daily bias factors were computed from daily totals from MQPE and CPC gauge analyses. Bias is set to unity for any 24-hour period when more than 3 hours of MQPEs were missing (referred to as "gap days"). This scheme performs no readjustment for any of the gap days. A second scheme was developed to address this. This scheme shares the bias computation method with the first scheme, but it distributes the daily gauge rainfall for each grid cell equally among the 24 hours for each of the gap days regardless of the intraday distribution of the gap

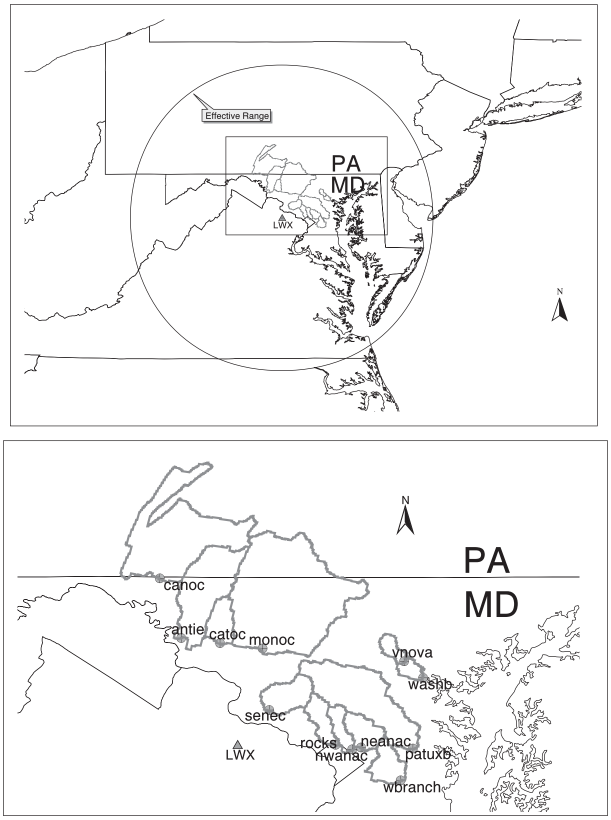

FIG. 1. Study catchments in Maryland and their locations relative to the WSR-88D at Sterling, VA (KLWX). (top) Location of Maryland–Pennsylvania and (bottom) catchments and radar location (LWX).

hours (note the hours with valid values are untouched). These two schemes—referred to as CPC-1 and -2—will be compared to assess the effects of performing the redistribution.

Although the daily CPC data were produced incorporating PRISM-based spatial interpolation, its monthly accumulation often differs from the monthly PRISM (up to 10%), which is likely a result of the difference in gauge reports that were used in creating these datasets. To reduce such discrepancies and facilitate interpretation of the comparisons, the daily CPC data for each HRAP cell were readjusted with a scaling factor constant within a given month, so that its monthly sum matches the corresponding PRISM monthly total for that bin. The corrected, rather than the original, daily CPC data were then used as reference in readjusting the MQPEs. Note that the exact time periods differ slightly for the CPC-based and PRISM monthly totals (i.e., the former starts and ends at 1200 UTC of the last day of a month, whereas the latter starts and ends at 0 local time of the first day of the month, respectively). Despite this difference, the monthly totals from the corrected CPC are very close to the PRISM ones. Also worth noting is that the readjustment of CPC data introduces variations in precipitation values among HRAP cells embedded within a 1/8° box.

b. *Study sites and experimental design*

Twelve catchments in the state of Maryland were selected to evaluate the hydrologic impacts of readjustment. With sizes ranging from 84 to 1200 km, these catchments drain much of Maryland between the Blue Ridge Mountains and the Chesapeake Bay (Fig. 1). Stream gauge locations, drainage area, and related descriptions can be found in Table 1. The study catchments are under the umbrella of the WSR-88D unit located in Sterling, Virginia (KLWX, at 38°58′31″N, 77°208 28′41″W), and each has been gauged continuously at its outlet by the United States Geological Survey (USGS). Geographically, the basins can be divided into three groups. The first group consists of four catchments over or near the Pennsylvania–Maryland border (i.e., canoc, antie, catoc, and monoc; Fig. 1 and Table 1). The second and third groups consist of ones with outlets in south–central Maryland (senec, rocks, nwanac, neanac, wbranch, and patuxb), and in the vicinity of Baltimore (vnova and washb), respectively (Fig. 1).

TABLE 1. Study catchments.

Station	USGS ID	Latitude (°N)	Longitude (°W)	Area (km^2)	Description
canoc	01614500	39.7164	77.8248	1279	Conococheague Creek at Fairview
antie	01619500	39.4498	77.7302	728	Antietam Creek near Sharpsburg
catoc	01637500	39.4273	77.5562	173	Catoctin Creek near Middletown
monoc	01643000	39.4028	77.3661	2116	Monocacy River at Jug Bridge near Frederick
senec	01645000	39.1281	77.3358	262	Seneca Creek at Dawsonville
rocks	01648000	38.9725	77.0400	161	Rock Creek at Sherrill Drive Washington
nwanac	01651000	38.9523	76.9661	128	NW branch of Anacostia River near Hyattsville
neanac	01649500	38.9603	76.9260	189	NE branch of Anacostia River at Riverdale
wbranch	01594526	38.8142	76.7487	232	Western branch at Upper Marlboro
patuxb	01594440	38.9559	76.6937	901	Patuxent River near Bowie
vnova	01589300	39.3459	76.7332	84	Gwynns Falls at Villa Nova
washb	01589352	39.2715	76.6486	171	Gwynns Falls at Washington Blvd at Baltimore

The research version of the NWS Hydrologic Laboratory Research Distributed Hydrologic Model (HL-RDHM; see descriptions in Koren et al. 2004) was implemented for each of the catchments. The model represents the landscape on the 4 km HRAP grids. The implementation for this study involves three HL-RDHM components. The first one is a snow model (gridded Snow-17) that relies on temperature to distinguish snow versus rain, and computes snow accumulation, ablation, and melt (Anderson 1973, 1976). The second is the Sacramento Soil Moisture Accounting (SAC-SMA) model for computing runoff at each HRAP grid cell for given precipitation, snowmelt, and initial conditions (Burnash et al. 1973; Burnash 1995; Koren et al. 2004). The third is a kinematic wave module for routing overland and channel flows (Koren et al. 2004). The a priori SAC-SMA parameter values from physiographic information and snow model parameters commonly used in RFC operations were adopted without any calibration. These values have been shown to work well, in a relative sense, in other studies (Reed et al. 2004; Anderson et al. 2006). For routing, the cell-to-cell connectivity required for the kinematic routing model was determined from a digital elevation model (Reed 2003). Routing parameters were assigned on the basis of USGS flow measurement data (see Koren et al. 2004 for the method). The reason for using uncalibrated models is that calibration would only be effective when the forcing data were consistently biased from year to year, which is not the case for the RFC MQPEs, given the aforementioned effects of truncation error and its correction. On the other hand, use of uncalibrated models allows us to isolate the impacts of temporal trends in the forcing data, though the associated performance statistics do not necessarily reflect the best results that can be achieved by the model.

Hourly MQPEs for the period of 1997–2006 were retrieved from MARFC archives. The first 5 years of MQPEs were produced using the Stage III algorithm and the rest via the MPE. As indicated earlier, these MQPEs underwent readjustment on the basis of monthly PRISM and daily CPC gauge-only analyses. Both the original and the readjusted MQPEs were then used to drive the hydrologic model. Additional forcing data obtained include 6-hourly 2.5° NCEP gridded temperature reanalysis (needed for the snow model) and monthly climatological potential evapotranspiration (PET) values. In assigning the initial conditions, it was assumed that upper and lower zone free water storage were each at half capacity and that the channels were dry. Under these assumptions, continuous simulations were performed at an hourly time step for each catchment over the 10-yr period. The simulation results were subsequently compared with hourly USGS discharge measurements. To mitigate the uncertainties due to unknown initial conditions, the first year (1997) was treated as the "warm-up" period—following the practice used in phase II of the Distributed Model Intercomparison Project (DMIP-II; Smith et al. 2006)—and the associated results were omitted from subsequent analyses.

According to the Radar Operation Center (ROC), operational MPE with the fixes for the TE went in effect during November 2003 (Daniel Berkowitz 2007, ROC, personal communication). Therefore, the 10-yr period was stratified into two subperiods: 1) precorrection (January 1998–November 2003), and 2) postcorrection (December 2003–December 2006). The analyses of MQPEs entail comparing mean areal precipitation (MAP) and the accuracy of streamflow simulations in the study watersheds. The primary performance metrics for the latter include percentage bias (PB) and linear correlation coefficient (ρ). The definition of PB is provided below:

$$\mathrm{PB} = \frac{\sum_{i=1}^{N} S_i - Q_i}{\sum Q_i} \times 100, \tag{1}$$

where Q_i and S_i are observed and simulated discharge, respectively. Any improvements from the readjustment are judged by the difference in values obtained before and after applying readjustment (the difference in the absolute values, $|\mathrm{PB}_i| - |\mathrm{PB}_j|$, is used for bias).

4. Results

The effectiveness of the three readjustment schemes (PRISM monthly, and CPC-1 and -2) was examined in terms of 1) interannual variations of bias in streamflow simulations, 2) statistics of hourly streamflow simulations before and after the TE correction, and 3) two case studies of major flood events.

a. *Interannual variations of bias*

One of the primary goals of readjustment is to mitigate temporal variations in MQPE bias. The effectiveness of the three schemes in this respect is illustrated by interannual variations in the bias of simulated streamflow over 1998–2006 as characterized by 1) multibasin median of annual MAP and PB of simulated streamflow (Figs. 2a,b), and 2) interannual range of PB for each basin (Fig. 3).

Annual MAPs from the three schemes show appreciable differences from 1998 to 2001, while the values are mostly similar thereafter. For 2000 and 2001, CPC-2 produced considerably higher MAPs than the rest—a consequence of filling the gap days. As for the streamflow bias, one striking feature in Fig. 2b is that, without readjustment, the median annual PB exhibits a distinct

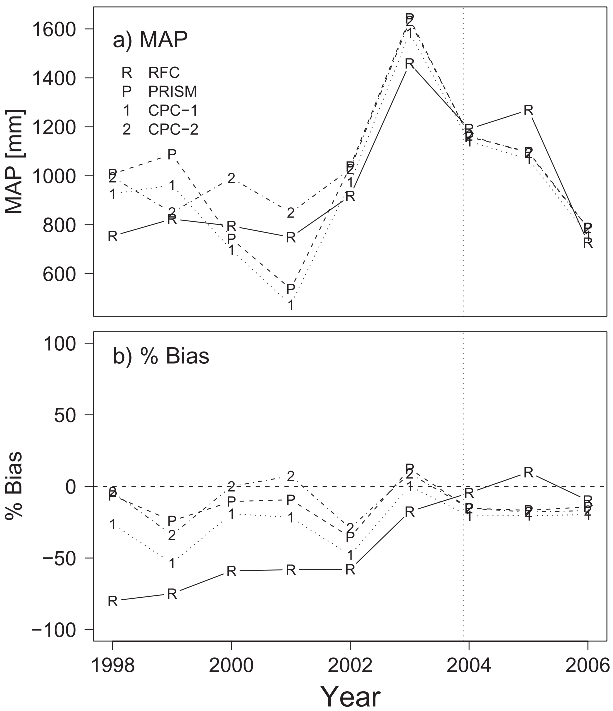

FIG. 2. Interannual variations of annual multibasin median of (a) MAP amounts and (b) the PB in streamflow simulations for MARFC MQPEs (RFC), and readjusted MQPEs using PRISM monthly and CPC daily gauge analysis (PRISM, CPC-1, and CPC-2).

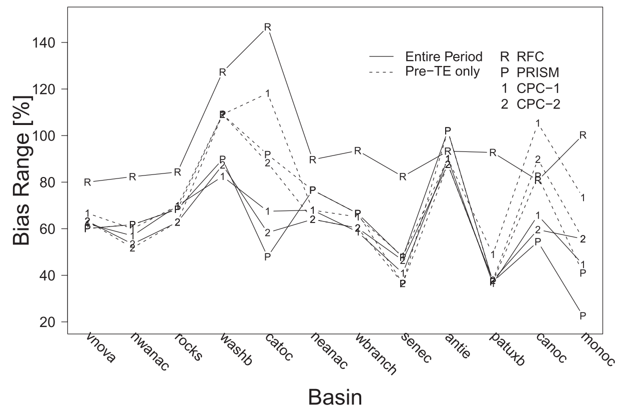

FIG. 3. Interannual range of PB (i.e., $PB_{max} - PB_{min}$) for each basin between 1998 and 2006. The solid lines represent results from applying each readjustment scheme for the entire 9-yr period, and the dotted lines represent those from applying each for the pre-TE period only.

and almost monotonic upward trend, with severe underestimation of runoff in 1998 ($\approx -80\%$) morphing to nearly unbiased values in 2004 and to slight overestimation in 2005. While the presence and the correction of TE was likely behind the underestimation for 1998–2003 and its subsequent improvement, the progressive upward trend in PB values prior to 2003 might reflect enhanced quality of MQPEs due to an increased number of operational gauge observations at MARFC (from 100–200 per hour average over 1998–2000 to around 400–500 over 2001/02, and to 400–600 for 2003–06), an improved data collection and delivery system (the data latency has been much reduced over the period), and a higher level of quality control of gauge and radar data. The three readjustment schemes substantially improved the discharge bias for the earlier period (Fig. 2b). Among the three schemes, PRISM and CPC-2 yielded better (i.e., closer to zero) PB values than CPC-1. By contrast, after the TE correction, each scheme tended to reduce the precipitation (Fig. 2a), and the reduction was severe enough to yield conspicuously negative PB values for 2004 and 2005 (Fig. 2b).

Despite the differing impacts of the readjustment schemes before and after the TE correction, each scheme—as shown in Fig. 3—was effective in reducing the interannual range of PB (i.e., the difference between the maximum and minimum values). In this respect, readjusting for the entire period appears more effective than for the pre-TE period alone (Fig. 3), and the three schemes appear comparable in effectiveness (Fig. 3).

b. *Statistics before and after TE correction*

1) OVERALL

The basin-specific impacts on the MAP and streamflow simulations are shown in Fig. 4, where the latter are characterized by PB and ρ. The basins are sorted by increasing drainage area and stratified into two periods prior to and following the TE correction. The three schemes are further compared based on the improvement statistics of simulated streamflow in terms of median of difference in PB and ρ before and after readjustment (Table 2). Key observations follow below.

As noted earlier, each scheme elevated MAP values for almost every basin prior to the TE correction but tended to reduce MAP after the TE correction (Figs. 4a,b). For the earlier period, streamflow simulations based on RFC MQPEs show pronounced underestimation (Fig. 4c). Each readjustment scheme increased the MAP and thereby improved the negative bias (Fig. 4c). Among the schemes, PRISM and CPC-2 results are close to bias neutral, whereas CPC-1 ones still show systematic low bias (Fig. 4c). By comparison, after the TE correction, PB values from RFC MQPE-based simulations were more evenly distributed relative to zero (Fig. 4d), reflecting an absence of the systematic underestimation as observed in the earlier period (Fig. 4c). In this context, each readjustment scheme led to a downward shift in precipitation and runoff, and the result is uniformly negative PB and higher |PB| for most of the basins (Fig. 4d; Table 2).

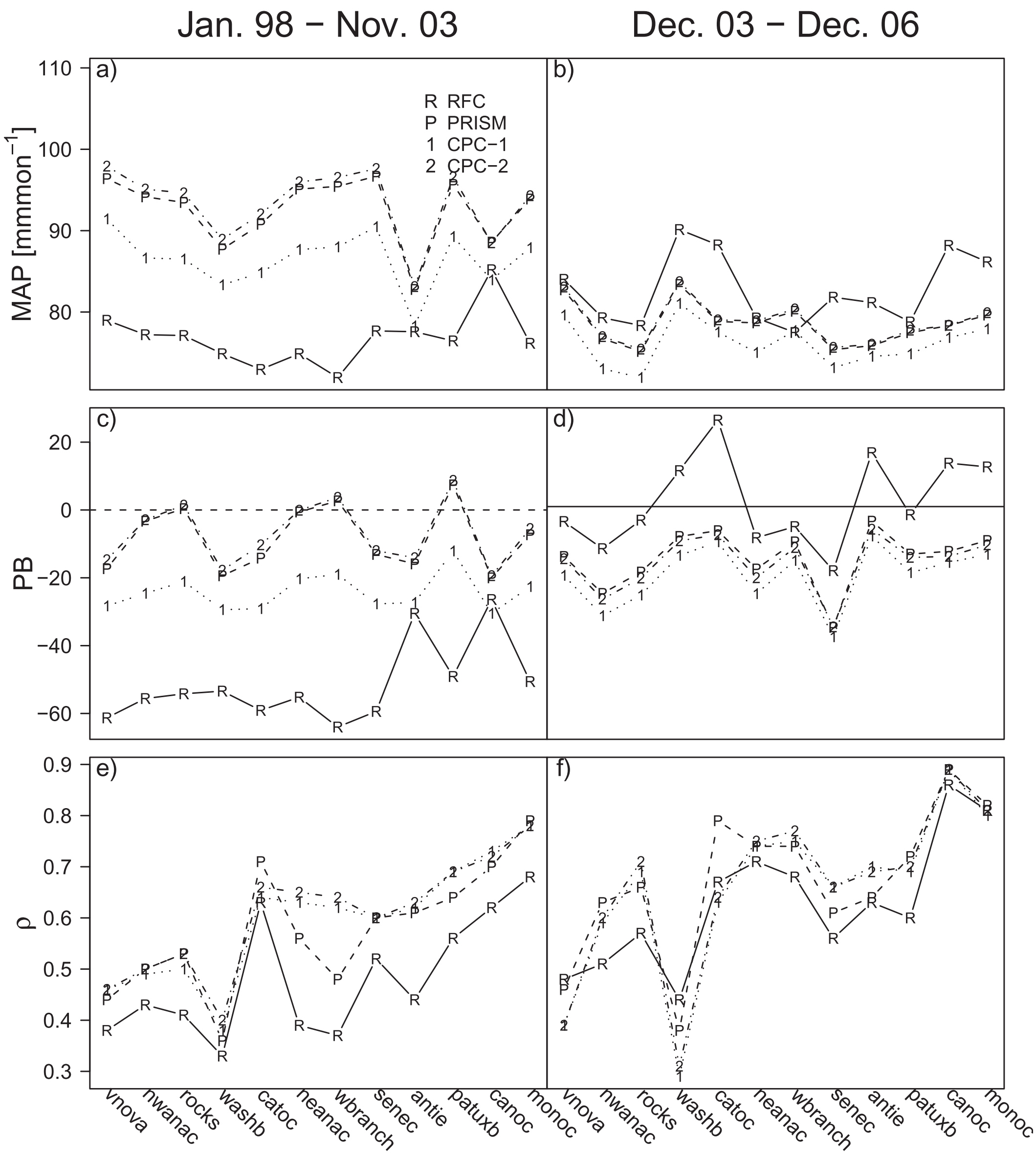

FIG. 4. Statistics of hourly streamflow prior to and after correction: (left) January 1998–November 2003 and (right) December 2003–December 2006 in terms of (a),(b) MAP, (c),(d) PB, and (e),(f) correlation ρ.

The values of correlation, ρ, also contrasted sharply before and after the TE correction. For the earlier period, each scheme greatly improved ρ values for most of the basins (Fig. 4e). In terms of the multibasin median of ρ, CPC-1 and -2 results are comparable, and both are better than PRISM results (Table 2). After the TE correction, each readjustment scheme still rendered considerable, though less pronounced, improvements in ρ values,

TABLE 2. Multibasin median of improvement of degradation in statistics (overall).

	January 1998–November 2003			December 2003–December 2006		
Statistic	PRISM	CPC-1	CPC-2	PRISM	CPC-1	CPC-2
PB	−0.44	−0.31	−0.47	0.07	0.13	0.09
ρ	0.08	0.10	0.10	0.04	0.05	0.05

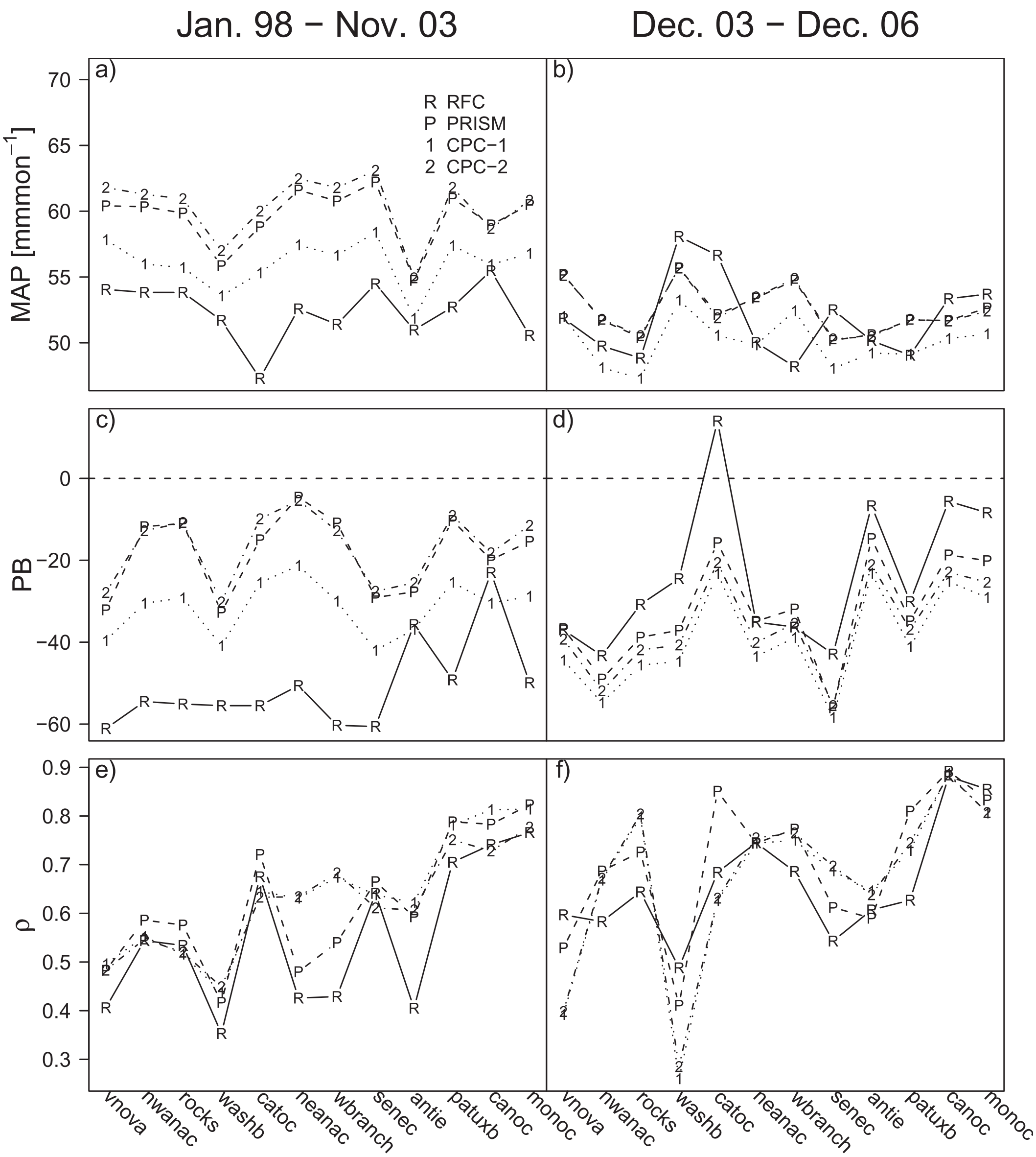

FIG. 5. As in Fig. 4, except for the warm season (April–October).

with CPC daily results again slightly better than PRISM ones (Fig. 4f; Table 2). Among the basins, ρ values in general rises with basin size.

2) Seasonal variations

The hourly statistics were computed for the warm and cool seasons separately, with the former extending from April to October and the latter from November to March. The results are summarized in Figs. 5 and 6 and Tables 3 and 4.

The warm season results are similar to the overall ones. Yet, several distinct features are also evident. For the period after the TE correction, RFC MQPE-based simulations were biased almost uniformly low in the warm season (Fig. 5d), whereas in the overall results positive bias was seen in half of the basins (Fig. 4d). For both pre- and post-TE correction periods, the negative PB shown in the simulated streamflow was corrected to a lesser extent than in the overall results (Fig. 5d; see Fig. 4d for the latter). Lastly, in terms of median ρ, PRISM performed comparably with the CPC schemes for the earlier period (only slightly worse than CPC-1) and slightly better for the later period (Table 3).

The cool-season season results contrast sharply with the warm season ones (Fig. 6; Table 4). First, for the earlier period, PRISM and CPC-2 schemes rendered the simulated streamflow almost bias neutral (Fig. 6c), whereas for the warm season, the bias was negative despite the

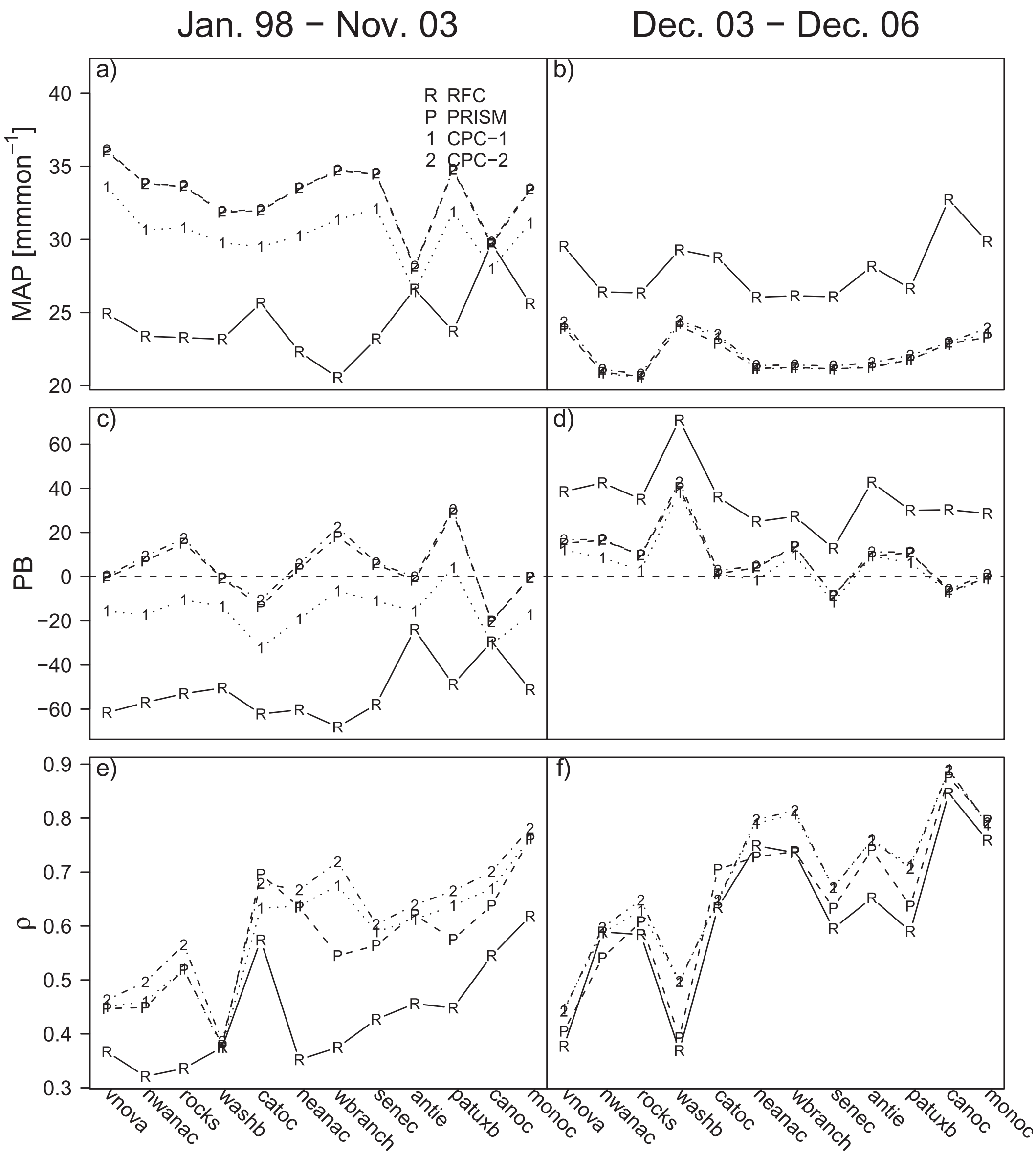

FIG. 6. As in Fig. 4, except for the cool season (November–March).

readjustment (Fig. 6c). The former is evidently a result of greater increases in the cool season MAP values with readjustment (Fig. 6a). For the period following the TE correction, it is evident that all three schemes still led to considerable, albeit less pronounced, improvements relative to the earlier period (Fig. 6). Among the three schemes, CPC-1 yielded the least bias reduction for the earlier period and most bias increase for the later period (Table 4). In terms of median ρ, CPC-2 fared better than the other two schemes (Table 4).

c. *Case studies*

The effectiveness of each readjustment scheme was more closely examined through studies of two major flood events for Monocacy River at Jug Bridge (monoc). The first one occurred prior to the TE corrections—over late May and early June of 2003. The second one took place in late September of 2004 after the TE correction. The MAP and runoff time series for the May–June 2003 event are shown in Fig. 7. This event contained multiple flood episodes, among which the ones on 17 May, 4 June, and 8 June were associated with peaks near or exceeding 400 $m^3\ s^{-1}$. For these episodes, using RFC MQPEs as model forcing resulted in peaks of less than half the observed magnitude. Applying each of the three readjustment schemes considerably improved the simulation results. Yet, even with the improvements, simulated peaks via all three schemes were still much below observed

TABLE 3. Multibasin median of improvement or degradation in statistics (warm season).

	January 1998–November 2003			December 2003–December 2006		
Statistic	PRISM	CPC-1	CPC-2	PRISM	CPC-1	CPC-2
PB	−0.38	−0.24	−0.41	0.04	0.12	0.09
ρ	0.06	0.07	0.02	0.06	0.03	0.03

Values. Another notable feature for this event is that the simulation based on the RFC MQPEs missed the smaller runoff episode between 25 and 30 May. For this episode, PRISM-based readjustment rendered only a slight increase in the precipitation and runoff values. By contrast, CPC-1 and -2 yielded much greater precipitation amounts on 25 May, and therefore were able to better reproduce this runoff episode. The lower MAP values based on RFC MQPEs can be partly explained by the presence of two missing hours at 0300 and 0400 UTC on 25 May, when precipitation was assumed to be zero by the model. In this situation, PRISM-based monthly readjustment yielded only a slight increase in precipitation at adjacent hours, as it spreads missing precipitation amounts throughout month. CPC daily schemes, by contrast, elevated the precipitation over only the remaining 22 hours, and therefore were able to better resolve this episode.

The time series for the September 2004 event are shown in Fig. 8. In this event, remnants of Hurricane Jeanne brought heavy precipitation and extensive flooding to Maryland. At monoc, observed peak hourly discharge exceeded 540 $m^3 s^{-1}$ at 1900–2000 UTC on 30 September—the highest during the period after the TE correction. The model simulation using RFC MQPEs produced a comparable peak (601 $m^3 s^{-1}$) at 0800 UTC on 30 September. Applying the readjustments improved the timing of the peak, possibly because lowered peak magnitude tends to decrease routing velocity in kinematic wave routing (401 $m^3 s^{-1}$ for PRISM and 337 $m^3 s^{-1}$ for CPC-1 and -2). While daily CPC schemes led to severe underestimation of peaks, the associated peak timing estimates were markedly improved (at 1700–1800 UTC). Further examination of the 24-hour RFC and CPC precipitation accumulations ending on 29 September revealed considerable underrepresentation of the magnitude and spatial variations of precipitation by the latter (Fig. 9). For this period, 24-hour accumulation from RFC MQPEs features a sharp southeastward gradient within monoc, with values ranging from 46 to 157 mm and a standard deviation of 23 mm. By comparison, the corresponding CPC gauge analysis shows a much lower and narrower distribution of precipitation values within the basin—that is, between 55 and 82 mm with a standard deviation of 6 mm. The inability of the gauge-only analysis in resolving the spatial gradient likely contributed to a much lower daily MAP (69 mm versus 88 mm by the RFC MQPEs), which in turn translated to the low bias in streamflow simulations.

5. Discussions and conclusions

a. *Effectiveness of readjustments and the determining factors*

This study highlights some of the potential benefits, as well as the pitfalls, of gauge-based readjustment of operational MQPEs to hydrologic simulations. As our analyses indicated, precipitation readjustments, conducted on either a monthly or daily scale, may considerably improve the quality of real-time MQPEs and, in turn, benefit streamflow simulations. The amounts of improvements, however, can vary depending on the intrinsic bias characteristics of the MQPEs, season, magnitude of rainfall events, and readjustment methods. It appears that the truncation error in the NEXRAD PPS, until its correction in November 2003, is a major factor contributing to the low bias in earlier MQPEs, which forms the most compelling reason for performing the readjustments. At least partly because of the TE, the model simulated streamflow with MARFC MQPEs as forcing systematically underestimated runoff for the 12 Maryland catchments between 1998 and November 2003. By contrast, for the post-TE correction period, streamflow simulations based on RFC MQPEs were close to bias neutral—arguably a reflection of a fundamental

TABLE 4. Multibasin median of improvement or degradation in statistics (cool season).

	January 1998–November 2003			December 2003–December 2006		
Statistic	PRISM	CPC-1	CPC-2	PRISM	CPC-1	CPC-2
PB	−0.48	−0.38	−0.47	−0.24	−0.22	−0.23
ρ	0.13	0.15	0.17	0.01	0.03	0.03

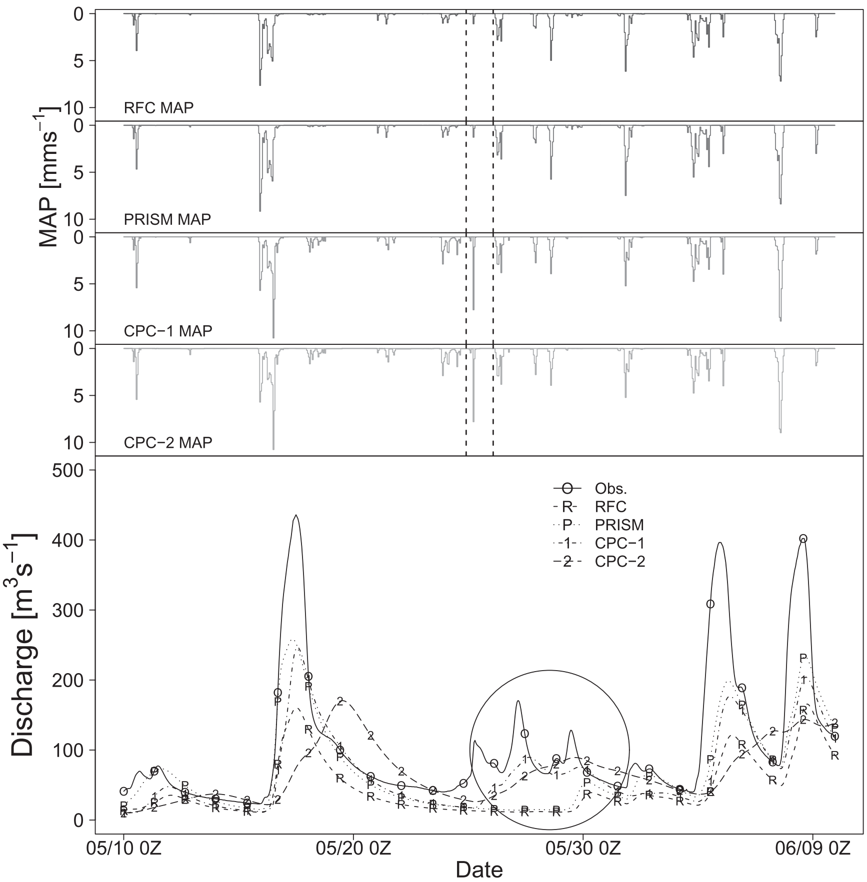

FIG. 7. (top) Time series of MAP from original RFC MQPEs and from readjusted MQPEs via PRISM, CPC-1 and -2 schemes from 20 May to 9 Jun of 2003. (bottom) Observed and simulated discharge using each of the aforementioned MQPEs. The dotted lines delineate the time segment where CPC-1 and -2 schemes yielded appreciably higher MAP values, which led to subsequent higher simulated discharges that are closer to observed values (highlighted by the circle).

improvement in the intrinsic bias of MQPEs. While the low bias in the earlier period can also be explained by model structural and parametric errors along with the lack of model calibration, such explanations are implausible given a lack of such low bias after the TE correction.

As evidenced by the contraction in the interannual range of streamflow PB over 1998–2006 (Fig. 3), each readjustment scheme helps reduce the temporal inconsistency in MQPE. The actual effects on the MQPE quality, again, contrast between the two periods. For the earlier period, all three readjustment schemes substantially and consistently improved the accuracy in streamflow simulation. For the later period, however, the results were mixed with increased |PB| present for some basins.

The performance of the readjustment schemes also exhibits strong seasonal dependence. Both before and after the TE correction, each scheme was much more effective in improving the accuracy of streamflow simulations for the cool rather than the warm season, as judged by the two indicators (PB and ρ). For the pre-TE correction period, the improvement led by the readjustments was much greater for the cool season, where postadjustment bias values approached zero, and much less so for the warm season, where bias remained systematically negative despite the readjustments. After the TE correction, each scheme tended to reduce precipitation for both seasons. This reduction mitigated the existing positive bias for the cool season, but it proved excessive for the warm season as it turned some positive biases to negative for some basins and exacerbated the existing underestimation for others. One possible explanation of the seasonal contrasts in the effects of readjustments is the frequent presence of bright band effects due to low freezing level that artificially elevates the radar rain rate estimates. For the warm season, gauge-only analysis may underestimate areal rainfall as the sparsity of the gauge network makes it difficult to resolve convective precipitation with limited spatial

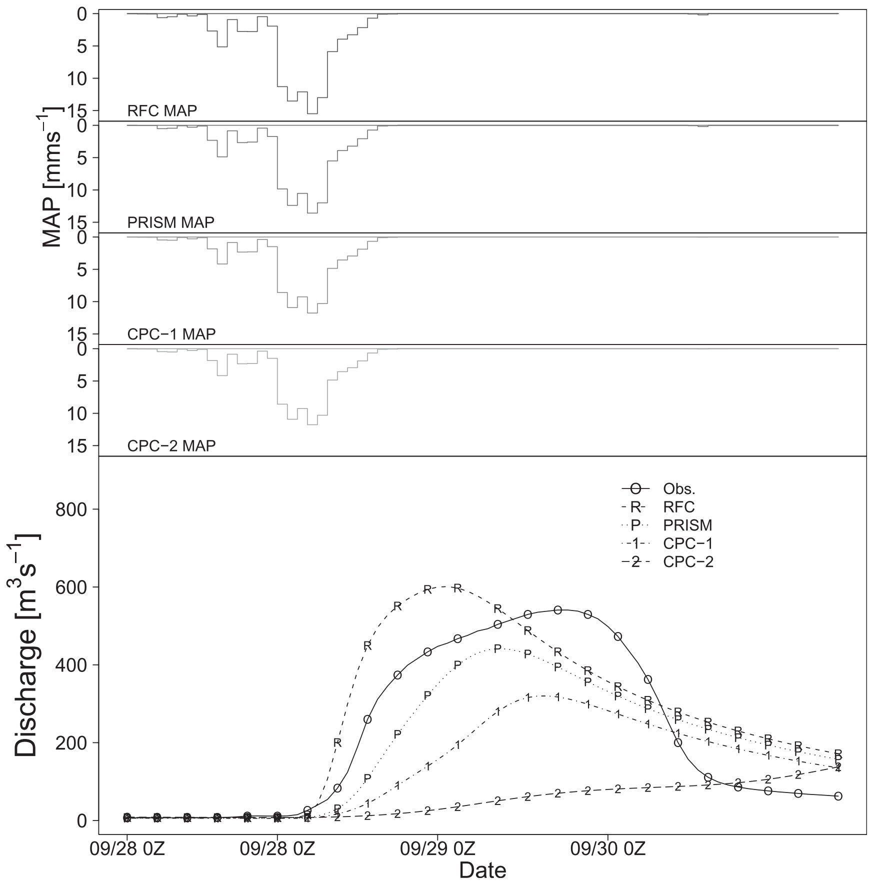

FIG. 8. As in Fig. 7, but from 28 Sep to 1 Oct of 2004.

scale and pronounced spatial variation in rain rates, as shown in the September 2006 event. Another plausible explanation of the seasonal contrasts lies in the seasonal change in hydrologic model physics—that is, the snow model functioned only during the cool season and its function likely induced a change in the bias characteristics of streamflow simulations.

b. *Daily versus monthly readjustments*

Ideally, daily scale readjustments—by accounting for the interday variation of precipitation and suppressing magnitude- and time-dependent bias—would yield more accurate precipitation information than monthly scale readjustments. This was expected to be reflected in the resultant streamflow simulations. Our analyses, however, yielded mixed results. For the cool season, daily CPC schemes indeed led to tangibly higher correlation. Yet, during the warm season, daily statistics were comparable and sometimes slightly worse relative to the PRISM monthly results, and the bias was evidently worse in the daily readjustment results unless redistribution was allowed (as in CPC-2).

These mixed results underscore complex trade-offs between a refined depiction of time–magnitude variations of bias when the readjustment window moves from monthly to daily and a potential loss of accuracy in the readjusted precipitation values due to elevated uncertainties in the reference gauge data at the finer scale. Naturally, gauge measurements are subject to greater random errors with refining time resolution. In addition, it must be noted that the CPC daily data were constructed by interpolating reports not ending 1200 UTC (not an issue for the PRISM data)—such a practice inevitably introduced additional uncertainties. Finally, as our analysis indicated, the ability of gauge reports to represent daily spatial rainfall fields remains severely constrained by the density of the gauges and the grid resolution of the gauge products. Given the limited data source and coarser grid resolution of CPC daily data, daily readjustments are susceptible to errors arising from small-scale variations in precipitation, and especially so during warm season convective storms.

Despite the aforementioned deficiencies, daily readjustments did show substantial value under certain

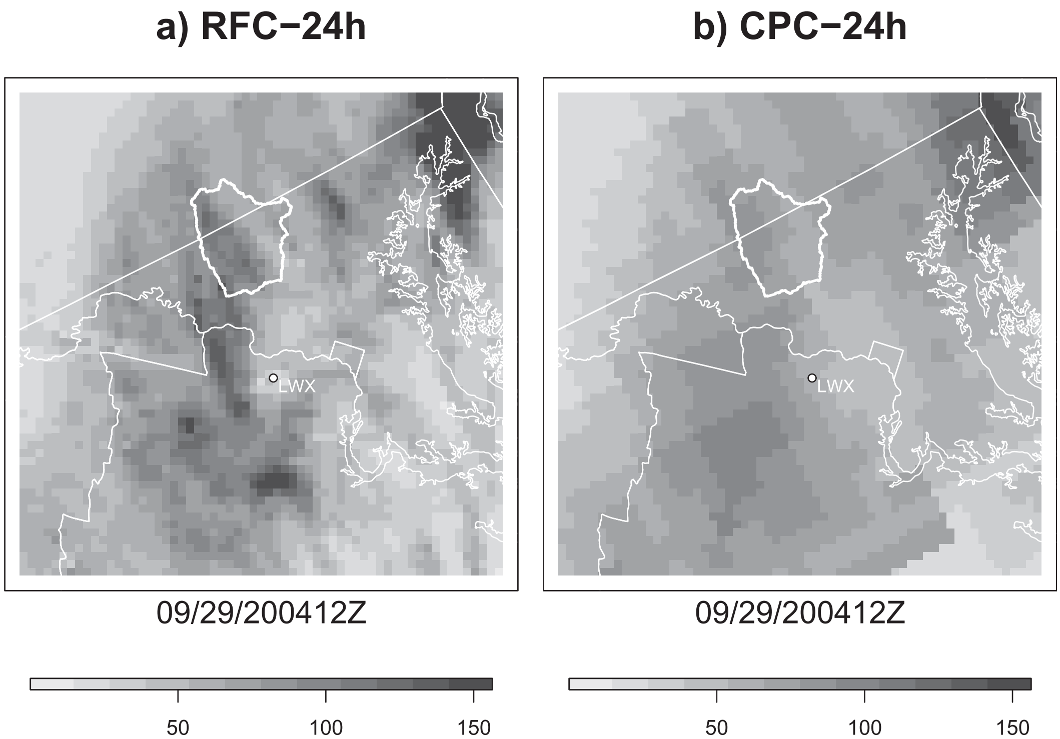

FIG. 9. Cumulative precipitation (mm) for the 24 h ending at 1200 UTC on 29 Sep 2004 based on (a) RFC MQPEs and (b) CPC gauge-only analysis.

circumstances. In particular, it could help resolve flood timing and peak magnitude when corresponding MQPEs are missing, as it did in the May–June 2003 flood event. Even after the correction of TE, daily readjustments proved valuable in eliminating spurious precipitation in MQPEs. Though excessive missing data (i.e., exceeding 72 hours for a HRAP bin) are relatively rare in MARFC MQPEs (less than 1% out of the 9-yr period), daily readjustment with redistribution can be particularly helpful when such a situation does arise (in the Mid-Atlantic and other RFCs).

c. Limitations, implications, and future directions

This study, while yielding valuable insights on issues related to MQPE readjustments and their hydrologic effects, was constrained by its premises and methodologies. First, this study employed hydrologic model simulation results as a proxy to infer the quality of MQPEs, as ''true'' precipitation cannot be determined regardless of the quality of sensors. Yet, hydrologic models—being inexact analogs of natural processes—are subject to structural and parametric errors; simplifications, such as the use of climatic PET that are interannually invariant, introduced additional uncertainties. Second, the study was performed for only 12 basins within a limited geographic area. It is yet unknown how MQPEs are biased in other areas outside of the study domain and whether the readjustments can yield similar improvements elsewhere. Third, the analyses involved a relatively short post-TE correction period (3 years). Longer records for this period will help illuminate the effects of readjustments over a wider range of rainfall-runoff conditions. Fourth, the study focuses on an area with a relatively dense gauge network, whereas in less-populated areas of the United States, the effects of readjustments using PRISM and CPC data may be further limited by the underlying gauge density of these products. Finally, the readjustment methods, especially the daily CPC-based ones, need to be further analyzed and refined to overcome some of the shortcomings such as underrepresenting spatial rainfall variability during heavy rainfall events.

These limitations notwithstanding, the results of the model simulations offer strong evidence on the TE-related variations in temporal bias characteristics of MQPEs and help define the anticipated benefits of retrospective readjustments. It was shown that the model, though imperfect and uncalibrated, was able to generate streamflow simulations of reasonable accuracy both before and after the TE correction with readjusted MQPEs (median ρ around 0.6 for streamflow). In addition, as demonstrated in this study, hydrologic model simulations can help detect patterns otherwise difficult to capture in MAP analysis alone, such as the progressive upward trend in the PB values that was likely linked to a corresponding upward shift in the MQPE bias (Fig. 2).

To summarize, our results indicate that readjusting archived real-time MQPEs can substantially improve the overall temporal consistency in the bias of the dataset, despite the presence of mixed results for the period after the TE correction. Therefore, it is expected that the readjustments can provide immediate benefits by expanding the usable MQPE archive for calibrating hydrologic models. Among the readjustment schemes, the monthly scheme appears to be adequate in this respect. Yet, daily readjustments are helpful and sometimes necessary for accommodating excessive missing values in MQPEs and clearing spurious ones. As gauge-based analyses are likely to underrepresent precipitation totals for both periods, readjusting for the entire period tends to be more effective in suppressing the interannual bias range than for the pre-TE period alone. For model calibration, readjusting for the entire period might be preferable, as temporally consistent bias can be relatively easily removed by adjusting model parameters. A caution here is that any calibration data should be as similar to the operational data as possible. There is a trade-off in using a longer, consistently biased MQPE dataset versus using a shorter one with bias characteristics closer to the real-time operational data. The readjustment will definitely be helpful for calibration for the pre-TE correction period. However, for the post-TE correction period, modelers need to examine simulations using both original and readjusted datasets to determine which one to use, particularly when refining the parameters affecting quick response runoff.

To address aforementioned limitations, particularly those related to sample size and duration, similar studies with data from other areas are under way, wherein more recent MQPE data will be incorporated and examined for a larger number of catchments. Other focal points of future studies include possible enhancement of the daily readjustment scheme and assessing the effects of readjustments on threshold frequency of simulated discharge. The former will be investigated by a combination of analysis of the quality of gauge reports, performing daily readjustments at coarser spatial scales where areal precipitation might be more accurate, and possibly incorporating a Kalman filter or a simplified version being used in the MPE (Seo et al. 1999).

Acknowledgments. We would like to express our gratitude to Peter Anhert and Joseph Ostrowski at MARFC for providing the MQPE data and offering insightful comments on the manuscript. We would also like to thank Yun Fan at NCEP-CPC and Christopher Daly of the PRISM group for data and support, and Brian Cosgrove at NWS for reviewing the manuscript.

REFERENCES

Anderson, E. A., 1973: National Weather Service River Forecast System—Snow Accumulation and Ablation Model. NOAA Tech. Memo. NWS HYDRO-17, 271 pp.

——, 1976: A point energy and mass balance model of a snow cover. NOAA Tech. Rep. 19, 150 pp.

Anderson, R. M., V. I. Koren, and S. M. Reed, 2006: Using SSURGO data to improve Sacramento Model a priori parameter estimates. *J. Hydrol.,* **320,** 103–116.

Burnash, R. J. C., 1995: The NWS river forecast system—Catchment modeling. *Computer Models of Watershed Hydrology,* V. P. Singh, Ed., Water Resources Publications, 311–366.

——, R. L. Ferral, and R. A. McGuire, 1973: A generalized streamflow simulation system—Conceptual modeling for digital computers. Joint Federal and State River Forecast Center, U.S. National Weather Service and California Department of Water Resources Tech. Rep., 204 pp.

Cognitore, P., 2005: An investigation of multisensor precipitation estimates (MPE) and operational use of MPE at the Middle Atlantic River Forecast Center (MARFC). NOAA/National Weather Service, Middle Atlantic River Forecast Center Tech. Rep. 2005-03, 16 pp.

Cosgrove, B. A., and Coauthors, 2003: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.,* **108,** 8842, doi:10.1029/2002JD003118.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.

——, W. P. Gibson, M. Doggett, J. Smith, and G. Taylor, 2004: Up-to-date monthly climate maps for the conterminous United States. *Proc. 14th Conf. on Applied Climatology,* Seattle, WA, Amer. Meteor. Soc., P5.1. [Available online at http://ams.confex.com/ams/84Annual/techprogram/paper_71444.htm.]

Ding, F., D. Kitzmiller, D. Riley, K. Shrestha, F. Moreda, and D.-J. Seo, 2005a: Evaluation of the range correction algorithm and convective stratiform separation algorithm for improving hydrological modeling. Preprints, *32nd Conf. on Radar Meteorology,* Albuquerque, NM, Amer. Meteor. Soc., 13R.2. [Available online at http://ams.confex.com/ams/32Rad11Meso/techprogram/paper_96543.htm.]

——, ——, D.-J. Seo, D. Riley, C. Dietz, C. Pham, and D. Miller, 2005b: A multi-site evaluation of the range correction and convective-stratiform separation algorithms for improving WSR-88D rainfall estimates. Preprints, *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., P2.11. [Available online at http://ams.confex.com/ams/Annual2005/techprogram/paper_86418.htm.]

Finnerty, B. D., M. B. Smith, D.-J. Seo, V. Koren, and G. E. Moglen, 1997: Space–time sensitivity of the Sacramento model to radar-gage precipitation inputs. *J. Hydrol.,* **203,** 21–38.

Fulton, R. A., J. P. Breidenbach, D. J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

——, F. Ding, and D. Miller, 2003: Truncation errors in historical WSR-88D rainfall products. Preprints, *31st Int. Conf. on Radar Meteorology,* Seattle, WA, Amer. Meteor. Soc., P2B.7. [Available online at http://ams.confex.com/ams/32BC31R5C/techprogram/paper_63261.htm.]

Greene, D., and M. Hudlow, 1982: Hydrometeorologic grid mapping procedures. *Int. Symp. on Hydrometeorology,* Denver, CO, AWRA, 20 pp.

Hardegree, S. P., S. S. Van Vactor, D. H. Levinson, and A. Winstra, 2008: Evaluation of NEXRAD radar precipitation products for natural resource applications. *Rangeland Ecol. Manage.,* **61,** 346–353.

Jayakrishnan, R., R. Srinivasan, and J. G. Arnold, 2004: Comparison of raingage and WSR-88D Stage III precipitation data over the Texas–Gulf basin. *J. Hydrol.,* **292,** 135–152.

Koren, V., B. D. Finnerty, J. C. Schaake, M. B. Smith, D.-J. Seo, and Q.-Y. Duan, 1999: Scale dependencies of hydrologic models to spatial variability of precipitation. *J. Hydrol.,* **217,** 285–302.

——, S. M. Reed, M. Smith, Z. Zhang, and D.-J. Seo, 2004: Hydrology Laboratory Research Modeling System (HL-RMS) of the U.S. National Weather Service. *J. Hydrol.,* **291,** 297–318.

Krajewski, W., and J. Smith, 2002: Radar hydrology: Rainfall estimation. *Adv. Water Resour.,* **25,** 1387–1394.

NCEP-CPC, cited 2005: Soil moisture outlooks. [Available online at http://www.cpc.noaa.gov/soilmst/forecasts.shtml.]

Over, T. M., E. A. Murphy, T. W. Ortel, and A. L. Ishii, 2007: Comparisons between NEXRAD radar and tipping-bucket gage rainfall data: A case study for DuPage County, Illinois. *Proc. World Environmental and Water Resources Congress 2007,* Tampa, FL, 1–14.

Reed, S. M., 2003: Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modeling. *Water Resour. Res.,* **39,** 1238, doi:10.1029/2003WR001989.

——, and D. R. Maidment, 1999: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. *J. Hydrol. Eng.,* **4,** 174–182.

——, and Coauthors, 2004: Overall distributed model intercomparison project results. *J. Hydrol.,* **298,** 27–60.

——, J. Schaake, and Z. Zhang, 2007: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *J. Hydrol.,* **337,** 402–420.

Schaake, J. C., V. Koren, and Q. Duan, 1996: Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.,* **101,** 7461–7475.

Seo, D.-J., 1998a: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol.,* **208,** 37–52.

——, 1998b: Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. *J. Hydrol.,* **208,** 25–36.

——, 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.

——, ——, and E. R. Johnson, 1999: Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol.,* **223,** 131–147.

——, ——, R. Fulton, D. Miller, and T. O'Bannon, 2000a: Real-time adjustment of range-dependent biases in WSR-88D rainfall estimates due to nonuniform vertical profile of reflectivity. *J. Hydrometeor.,* **1,** 222–240.

——, and Coauthors, 2000b: Final report, interagency memorandum of understanding among the NEXRAD Program, the WSR-88D Operational Support Facility, and the National Weather Service Office of Hydrology. Hydrologic Research Laboratory, Office of Hydrology, NOAA/NWS Tech. Rep.

——, and Coauthors, 2002: Final report, interagency memorandum of understanding among the NEXRAD Program, the WSR-88D Operational Support Facility, and the National Weather Service Office of Hydrology. Hydrologic Research Laboratory, Office of Hydrology, NOAA/NWS Tech. Rep.

Smith, J. A., D.-J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.,* **32,** 2035–2045.

Smith, M., V. Koren, S. Reed, Z. Zhang, D.-J. Seo, F. Moreda, and Z. Cui, 2006: The distributed model intercomparison project: Phase 2 science plan. Hydrology Laboratory, Office of Hydrologic Development, NOAA/NWS, 46 pp. [Available online at http://www.weather.gov/oh/hrl/dmip/2/docs/dmip_2_plan_march10_06_update.pdf.]

Thielen, J., J. Bartholmes, M.-H. Ramos, and A. de Roo, 2008: The European flood alert system—Part 1: Concept and development. *Hydrol. Earth Syst. Sci. Discuss.,* **5,** 257–287.

Young, C. B., B. Nelson, A. Bradley, J. Smith, C. Peters-Lidard, A. Kruger, and M. Baeck, 1999: An evaluation of NEXRAD precipitation estimates in complex terrain. *J. Geophys. Res.,* **104,** 19 691–19 703.

——, A. A. Bradley, W. F. Krajewski, A. Kruger, and M. L. Morrissey, 2000: Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. *J. Hydrometeor.,* **1,** 241–254.

Zhang, Y., T. Adams, and J. V. Bonta, 2007: Subpixel-scale rainfall variability and the effects on separation of radar and gauge rainfall errors. *J. Hydrometeor.,* **8,** 1348–1363.