Improving Multisensor Precipitation Estimation via Adaptive Conditional Bias–Penalized Merging of Rain Gauge Data and Remotely Sensed Quantitative Precipitation Estimates

ALI JOZAGHI, MOHAMMAD NABATIAN,[a] SEONGJIN NOH,[b] AND DONG-JUN SEO

Department of Civil Engineering, The University of Texas at Arlington, Arlington, Texas

LIN TANG

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

JIAN ZHANG

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 6 June 2019, in final form 27 August 2019)

ABSTRACT

We describe and evaluate adaptive conditional bias–penalized cokriging (CBPCK) for improved multisensor precipitation estimation using rain gauge data and remotely sensed quantitative precipitation estimates (QPE). The remotely sensed QPEs used are radar-only and radar–satellite-fused estimates. For comparative evaluation, true validation is carried out over the continental United States (CONUS) for 13–30 September 2015 and 7–9 October 2016. The hourly gauge data, radar-only QPE, and satellite QPE used are from the Hydrometeorological Automated Data System, Multi-Radar Multi-Sensor System, and Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR), respectively. For radar–satellite fusion, conditional bias–penalized Fisher estimation is used. The reference merging technique compared is ordinary cokriging (OCK) used in the National Weather Service Multisensor Precipitation Estimator. It is shown that, beyond the reduction due to mean field bias (MFB) correction, both OCK and adaptive CBPCK additionally reduce the unconditional root-mean-square error (RMSE) of radar-only QPE by 9%–16% over the CONUS for the two periods, and that adaptive CBPCK is superior to OCK for estimation of hourly amounts exceeding 1 mm. When fused with the MFB-corrected radar QPE, the MFB-corrected SCaMPR QPE for September 2015 reduces the unconditional RMSE of the MFB-corrected radar by 4% and 6% over the entire and western half of the CONUS, respectively, but is inferior to the MFB-corrected radar for estimation of hourly amounts exceeding 7 mm. Adaptive CBPCK should hence be favored over OCK for estimation of significant amounts of precipitation despite larger computational cost, and the SCaMPR QPE should be used selectively in multisensor QPE.

1. Introduction

Accurate real-time quantitative precipitation estimation (QPE) is a prerequisite for accurate water forecasting. With the widespread use of weather radar systems, multisensor QPE using ground-based radar and rain gauge data is now a routine practice in many parts of the world. Given the increasing availability of various real-time satellite QPE products (AghaKouchak et al. 2011; Habib et al. 2009, 2012; Huffman et al. 2007, 2017; Joyce et al. 2004; Okamoto et al. 2005; Sorooshian et al. 2000; Turk and Miller 2005; Vicente et al. 1998), effective utilization of satellite data for multisensor QPE is an increasingly important topic (Ashouri et al. 2015; Gourley et al. 2011; Kalinga and Gan 2010; Prat and Nelson 2013). Numerous efforts have been made to reduce systematic and random errors in multisensor QPE via bias correction and multivariate analysis (Kondragunta et al. 2005; Nelson et al. 2010; Prat et al. 2014, 2015; Seo et al. 2010; Smith and Krajewski 1991;

[a] Current affiliation: ViewTech, Inc., Addison, Texas.

[b] Current affiliation: Department of Civil Engineering, Kumoh National Institute of Technology, Gumi, Gyeongbuk, South Korea.

Corresponding author: Ali Jozaghi, ali.jozaghi@mavs.uta.edu

DOI: 10.1175/JHM-D-19-0129.1

Smith et al. 1996; Vasiloff et al. 2007). Here, by bias correction, we mean multiplicative or additive correction applied to the raw gridded precipitation data at an effective spatiotemporal scale larger than a single grid box or a time step. By multivariate analysis, we mean data assimilation with multivariate observation equations with or without a dynamical model. Many bias correction and multivariate analysis techniques involve minimizing mean square error (MSE) or error variance under unbiasedness (Ciach et al. 2000; Goudenhoofdt and Delobbe 2009; Gandin 1966; Seo and Breidenbach 2002; Seo 1998a,b; Smith et al. 2006).

It is well known in statistics and econometrics that, in the presence of observation error, variance minimization tends to introduce negative and positive biases, or conditional bias (CB), over the upper and lower tails of the predictands, respectively (Fuller 1987; Seber and Wild 1989; Hausman 2001). In regression, the above effect is referred to as regression dilution, which results in attenuation bias in the regression coefficients (Hughes 1993; Frost and Thompson 2000). Such bias may not pose an issue when the prediction is made for the same predictands via the same regression model using new observations from the same observing systems. In QPE, however, the objective is to estimate true precipitation amounts as accurately as possible so that they may be used as observed initial or boundary conditions in a wide range of applications. Because precipitation observations almost always have significant uncertainties, particularly at high spatiotemporal resolutions, presence of CB in QPE is the norm rather than the exception. Therefore, addressing CB is an important topic in multisensor QPE.

One may, in general, differentiate the CB into Types I and II. The Type-I CB, defined as $E[X|\hat{X} = \hat{x}] - \hat{x}$, where X, $\hat{X}$, and $\hat{x}$ denote the unknown truth, the estimate, and the realization of $\hat{X}$, respectively (Joliffe and Stephenson 2003), is associated with false alarm. The Type-II CB, defined as $E[\hat{X}|X = x] - x$, where x denotes the realization of X, is associated with failure to detect an event. Whereas the Type-I CB may be reduced by calibration, the Type-II CB cannot (Wilks 2006; Seo et al. 2018). Ciach et al. (2000) found that minimizing the MSE in radar rainfall estimates increases the CB, and that, when estimating extremes is of interest, there is a trade-off to consider between minimizing the MSE and reducing the CB. To address the detrimental effects of the Type-II CB on estimating extremes, Seo (2013) introduced a new optimal linear estimation method that minimizes the weighted sum of error variance and expectation of the CB squared. When cast in the form of kriging, the method yields CB-penalized kriging (CBPK), which has been shown to broadly outperform conventional kriging in the prediction of high flows and estimation of heavy to extreme rainfall (Brown and Seo 2013; Seo 2013; Seo et al. 2014; Kim et al. 2018). When cast in the form of the Kalman filter (KF), the method yields the CB-penalized KF (Seo et al. 2018; Shen et al. 2019) and, in ensemble form, the CB-penalized ensemble KF (Lee et al. 2019).

Though the CB-penalized estimation techniques significantly improve performance over the tails, they do not minimize the MSE in the unconditional sense. This deterioration in unconditional performance may be reduced by prescribing the weight to the CB penalty adaptively based, for example, on the best available estimate of the unknown true state (Kim et al. 2018; Shen et al. 2019); if the best estimate is near the median or in the tails of the distribution, one may reduce the weight close to zero or to a large value, respectively. When there exist large observational uncertainties or the precipitation field has limited predictability, however, the above approach may not be able to identify the state of the system with sufficient accuracy and consistency to be effective.

In this work, we introduce adaptive CB-penalized optimal estimation for merging rain gauge data and radar-only or radar–satellite-fused QPE, which explicitly optimizes the weight for the CB penalty in real time. The specific optimal linear estimation technique considered is CB-penalized cokriging (CBPCK; Kim et al. 2018). The resulting technique is referred to as adaptive CBPCK. The outcome sought is improved estimation over the tails of the distribution of precipitation while performing comparably to ordinary cokriging (OCK) in the unconditional sense. We then comparatively evaluate adaptive CBPCK with OCK used in the National Weather Service's (NWS) Multisensor Precipitation Estimator (MPE; Habib et al. 2013; Kitzmiller et al. 2013; Nelson et al. 2016; Seo et al. 2010). As part of the evaluation, we also assess the incremental value, in reference to the radar-only QPE, of 1) rain gauge–based mean field bias (MFB) correction of radar-only QPE, 2) merging of rain gauge data and MFB-corrected radar QPE, 3) fusion of MFB-corrected SCaMPR QPE with MFB-corrected radar QPE, and 4) merging of rain gauge data and radar–satellite-fused QPE. The main contributions of this paper are development and comparative evaluation of adaptive CBPCK, assessment of the incremental value of MFB correction, fusion, and merging, and advances in the understanding of the CB in multisensor QPE and its correction. The rest of this paper is organized as follows. Section 2 describes the data used. Section 3

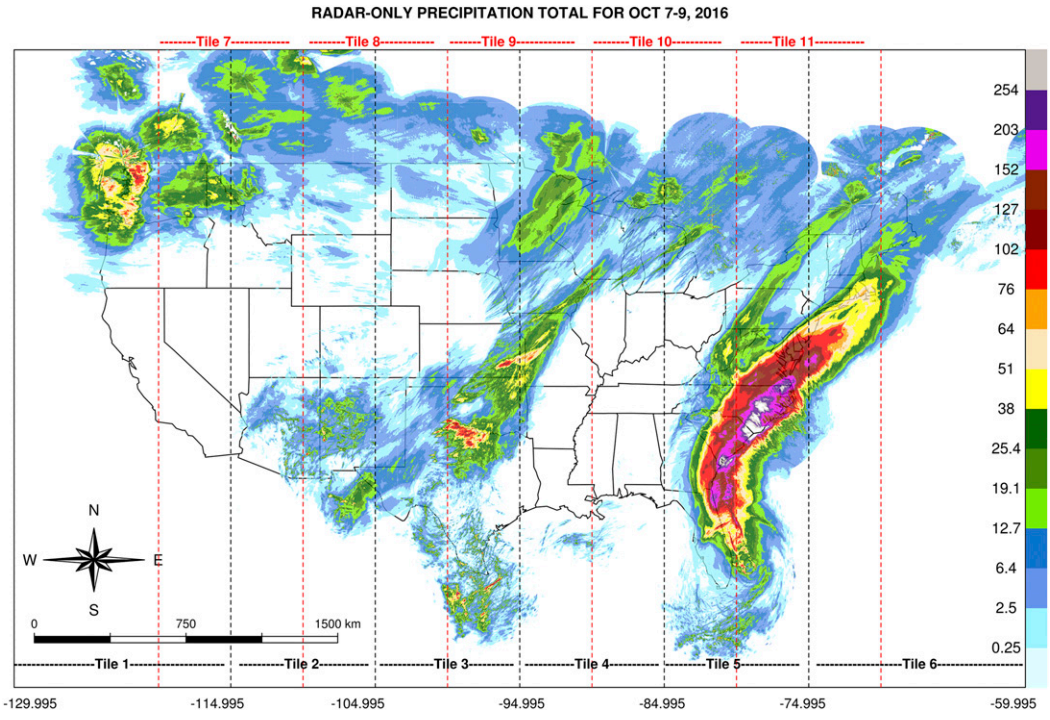

FIG. 1. Radar-only precipitation map for 7–9 Oct 2016. The horizontally overlapping and vertically elongated rectangles, indicated as Tiles 1–11, are the analysis domains used in the MRMS for parallel processing. The first 6 tiles (black dotted lines) cover the full CONUS and the last 5 (red dotted lines) straddle the first 6 to reduce the edge effects. Note that Tiles 1 and 6 are wider than the others. The precipitation map is obtained by averaging all overlapping estimates over the CONUS.

describes the methods used. Section 4 presents the results. Section 5 provides the conclusions and future research recommendations.

2. Data used

Two analysis periods are used: 7–9 October 2016 and 13–30 September 2015. Figures 1 and 2 show the radar-only precipitation totals. The first period includes Hurricane Matthew in the East Coast, weakly organized convective storms in the central United States, and a coastal storm in the Pacific Northwest. The second period includes generally convective events in the Pacific Northwest, Midwest, Northeast, and Southeast of the United States. The rain gauge data used are the hourly observations collected through the Hydrometeorological Automated Data System (HADS; Kim et al. 2009) operated by the NWS. Figure 3 shows the gauge locations, which total over 21 000. The HADS is a real-time data acquisition, processing, and distribution system supporting the NWS's Flood and Flash Flood Warning programs. The system acquires raw hydrologic and meteorological observations throughout the United States from the Geostationary Operational Environmental Satellites (GOES) Data Collection Platforms, most of which are owned and/or operated by various federal, state, and local agencies.

The hourly radar QPE used are from the Multi-Radar Multi-Sensor System (MRMS; Zhang et al. 2011, 2016) at 1-km resolution. MRMS is a system of automated algorithms that integrate data from multiple radars, surface and upper air observations, lightning detection systems, and satellite and numerical weather forecast models. The system generates a suite of 2D multisensor products for monitoring and short-term prediction of hail, wind, tornado, QPE, convection, icing, and turbulence. The radar precipitation estimates used in this work are the operational reflectivity-only MRMS QPE referred to as Q3RAD (Cocks et al. 2017). The hourly satellite QPE used in the study was from the Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR; Kuligowski 2002, 2013; Kuligowski et al. 2013) at 4-km resolution. SCaMPR uses GOES infrared data for predictor information, and calibrates them against microwave-based rain rates. The algorithm performs discrimination of rain versus no rain using discriminant analysis, calibrates precipitation rate using regression, optimizes the regression using nonlinear transformation of the predictors, and estimates loss of hydrometeors due to evaporation below the cloud base

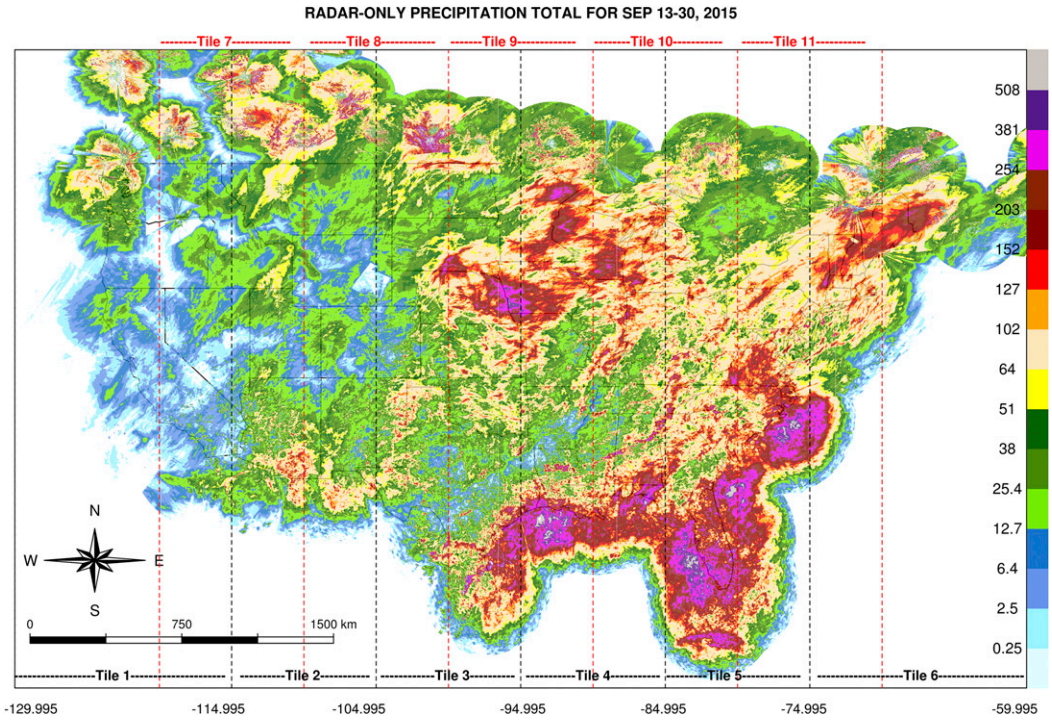

FIG. 2. As in Fig. 1, but for 13–30 Sep 2015.

in dry environments. In this work, all multisensor QPE operations are carried out for each of the 11 horizontally overlapping and vertically elongated rectangular analysis domains, or tiles, used in the MRMS for parallel processing. The first 6 tiles cover the full CONUS and the last 5 straddle the first 6 to reduce the edge effects (see Figs. 1 or 2). The final analysis is obtained by averaging all overlapping estimates over the CONUS.

3. Methods used

This section describes the multisensor QPE methods used in this work. Figure 4 shows the schematic of the estimation process which is described below in detail.

a. Bias correction of gridded QPE

The MRMS radar-only and SCaMPR QPEs (September 2015 only for the latter) are first MFB-corrected for each tile using the respective gridded QPEs and all available gauge data within the tile. The procedure used is the MFB correction algorithm of Seo et al. (1999) which has been in operation in the MPE since the mid-2000s. In the MPE, the algorithm operates for each radar and updates the radar umbrella-wide biases at multiple temporal scales of aggregation ranging typically from hourly to multiannual. The bias estimated is the multiplicative correction factor, $\beta_k = \int_{A_c} G_k(u)\,du / \int_{A_c} R_k(u)\,du$, to be applied to the radar-only QPE spatially uniformly where β_k denotes the MFB in the gridded QPE at the kth hour, A_c denotes the precipitation area, and $G_k(u)$ and $R_k(u)$ denote the gauge and radar precipitation at location u, respectively. To estimate β_k, the procedure uses all available (including posting-delayed) collocated and synchronous pairs of positive gauge and positive radar precipitation observations, and updates β_k in real time at multiple time scales via exponential smoothing (Schweppe 1973). Conceptually, the smoothing operation amounts to recursively estimating age-weighted moving averages of gauge and radar precipitation observations over time windows of different lengths. The updated bias associated with the smallest time scale is then chosen as the final estimate among those with an effective sample size greater than the user-set minimum. In this way, a shorter- and longer-term bias is used in gauge-rich and gauge-poor areas, respectively. For further details, the reader is referred to Seo et al. (1999).

Unlike in the MPE, here we estimate β_k for an entire tile which is much larger than the typical effective coverage of a single S-band radar such as the Weather Surveillance Radar-1988 Doppler version. The above choice is made necessary by the fact that the MRMS radar-only QPE is already a mosaic of data from multiple radars (Zhang et al. 2016). MFB correction for the

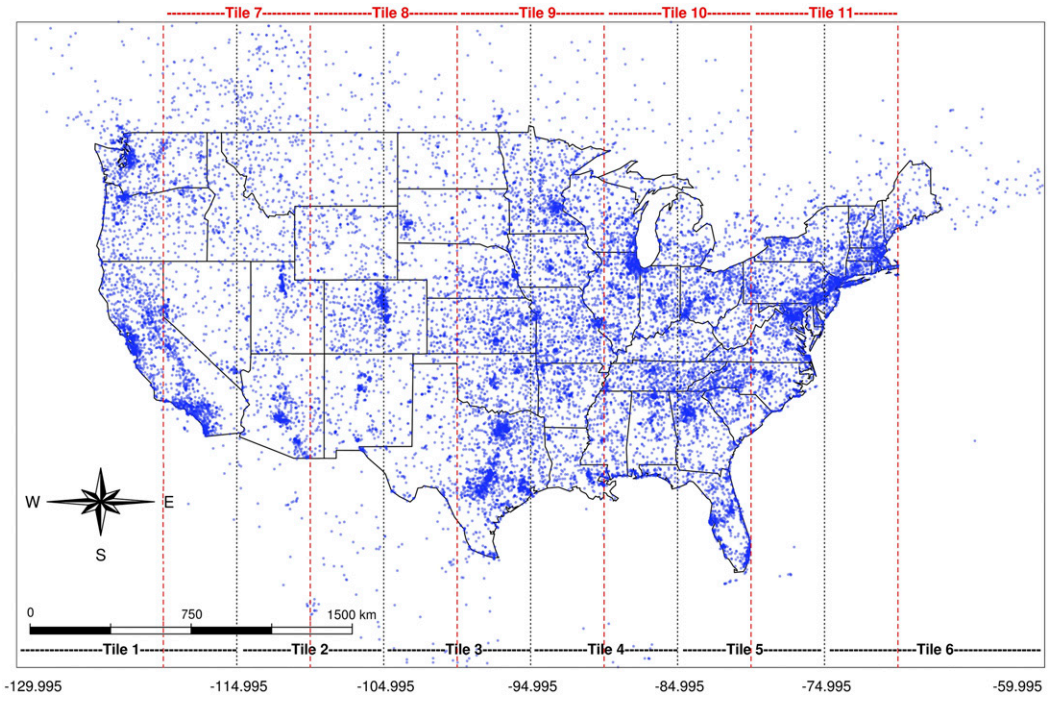

FIG. 3. Rain gauge locations.

SCaMPR QPE is completely analogous. Given the large latitudinal dimension of the tiles, multiple storms with disparate biases may exist in a single tile. In such cases, the effectiveness of MFB correction as implemented in this work is likely to be reduced. Performance assessment under different tile sizes, however, was beyond the scope of this work and is left as a future endeavour.

b. *Merging of rain gauge data and gridded QPE*

We use OCK and CBPCK to estimate the true precipitation amount at an arbitrary location, u_0, using rain gauge observations at u_i, $i = 1, \ldots, n_g$, and remotely sensed QPE at u_i, $i = n_g + 1, \ldots, n_g + n_r$, where n_g and n_r denote the number of rain gauges and remotely sensed QPE, respectively. Though referred to as cokriging, the formulation described here applies to kriging as well except that in the latter all observations are from a single observing system. Dropping the time index for brevity, we write the linear estimator for gauge precipitation at an ungauged location u_0 as

$$Z_0^* = \sum_{i=1}^{n} \lambda_i \frac{m_0}{m_i} Z_i, \tag{1}$$

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{2}$$

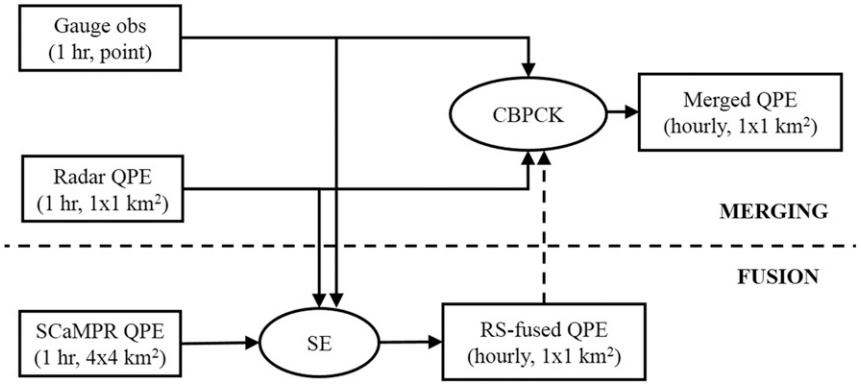

FIG. 4. Schematic of the merging and fusion processes used (see section 3 for details).

In the above, Z_0^* denotes the precipitation estimate at u_0, n denotes the total number of data used in the estimation, $n = n_g + n_r$; Z_i, $i = 1, \ldots, n_g$, denote the nearest gauge observations; Z_i, $i = n_g + 1, \ldots, n_g + n_r$, denote the nearest remotely sensed precipitation estimates; λ_i, $i = 1, \ldots, n$, denote the weights for Z_i, $i = 1, \ldots, n$; and m_0 and m_i denote the climatological mean precipitation at locations u_0 and u_i, respectively, for which we use the monthly PRISM climatology (Daly et al. 1994). The constraint, Eq. (2), renders the estimate Z_0^* in Eq. (1) unbiased in the mean under the assumption that the gridded QPE is unbiased relative to the gauge observations. Implicit in Eq. (1) is the assumption that the MFB-corrected radar QPE is climatologically unbiased (Seo et al. 2010).

OCK and CBPCK differ mainly in the calculation of the weights, λ_i, $i = 1, \ldots, n$ in Eq. (1). Whereas OCK minimizes the error variance of Z_0^*, $J_{\mathrm{EV}} = E[(Z_0^* - Z_0)^2]$, CBPCK minimizes the weighted sum of J_{EV} and the expectation of the CB squared, $J_{\mathrm{CB}} = E[(E[Z_0^*|Z_0] - Z_0)^2]$, that is, $J = J_{\mathrm{EV}} + \alpha J_{\mathrm{cB}}$ where α denotes the weight given to the CB penalty. If $\alpha = 0$, CBPCK reduces to OCK. One may hence consider CBPCK as generalized OCK. The CBPCK system is given by (Kim et al. 2018)

$$\sum_{j=1}^{n} \lambda_j(\rho_{ij} + \alpha\rho_{i0}\rho_{j0})\sigma_i\sigma_j = (1+\alpha)\rho_{i0}\sigma_i\sigma_0, \quad i = 1, \ldots, n, \tag{3}$$

$$\sum_{i=1}^{n} \lambda_i = 1, \tag{4}$$

where ρ_{ij} denotes the (cross-) correlation between the two variables at u_i and u_j; and σ_i and σ_j denote the standard deviation of the two variables at u_i and u_j, respectively. The objective function J associated with the CBPCK solution from Eqs. (3) and (4) is given by (Kim et al. 2018)

$$J = (1+\alpha)\left(\sigma_0^2 - \sum_{i=1}^{n} \lambda_i\rho_{i0}\sigma_i\sigma_0\right) - \mu, \tag{5}$$

where μ denotes the Lagrange multiplier. Because J reflects not only the error variance but also the CB penalty, it is larger than J_{EV} for $\alpha > 0$. To evaluate J_{EV}, we decompose J into J_{EV} and αJ_{cB} to obtain:

$$J_{\mathrm{EV}} = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j\rho_{ij}\sigma_i\sigma_j - 2\sum_{i=1}^{n} \lambda_i\rho_{i0}\sigma_i\sigma_0 + \sigma_0^2 \tag{6}$$

Because CBPCK does not minimize (unconditional) error variance, Eq. (6) is necessarily larger than the OCK error variance for $\alpha > 0$. The optimal weights, λ_i, $i = 1, \ldots, n$, are functions of the covariance among the gauge observations, covariance among the remotely sensed QPE, and cross covariance between the gauge observations and remotely sensed QPE (see Seo 1998a,b for details). Due to the generally large skewness in hourly precipitation and larger errors near the median than OCK, CBPCK produces negative estimates more frequently than OCK in areas of very light precipitation (see Seo 2013; Seo et al. 2014). To address this, we apply the correction procedure of Kim et al. (2018) in which the negative and positive CBPCK estimates are set to zero and adjusted by a scaling factor, respectively. For details, the reader is referred to Kim et al. (2018) and Seo et al. (2014).

The CBPCK solution depends on α. If there is little or no CB, we have $\alpha \approx 0$, and OCK suffices. If there is a large CB due to small predictability in the precipitation field, sparsity in the rain gauge network, or large uncertainty in the remotely sensed precipitation data (Seo 2013), one may expect CBPCK with a large α to produce more accurate estimates and error variances than OCK. Because there are multiple sources of CB, prescribing α a priori is a difficult proposition. In adaptive CBPCK, the weight α is optimized in real time in a data-driven manner such that the MSE, or any other performance measure or measures of choice, are optimized. To that end, we add the following adaptive steps to CBPCK: 1) discretize the possible range of α, 2) given a value of α, use CBPCK to estimate precipitation at all gauge locations within the tile in a cross-validation mode, 3) repeat step 2 for all values of α, 4) identify the MSE-minimizing α, and 5) perform CBPCK analysis over the entire tile using the ''optimal'' α. Because the sample size from cross validation may vary greatly from hour to hour within the same tile, and from tile to tile for the same hour, the resulting α is subject to potentially large sampling uncertainties. The time series plots of α identified in this way often show unrealistically large fluctuations compared to the characteristic time scales of CB that may be expected from the predictability conditions. For the above reason, we employ exponential smoothing of the error statistics at multiple time scales in a manner analogous to that used in MFB correction.

Because changing α has similar effects to changing the covariance structure [see Eq. (3)], it is possible that optimizing α may not only correct for the CB but also compensate for possibly incorrect covariance structure. Differentiation of the two, however, is not readily possible because the true covariance structure is not known in the real world. Synthetic experiments to

address the above is beyond the scope of this work and is left as a future endeavor. In practice, the lack of differentiation may not pose a significant issue because the uncertainty in the covariance structure is likely to be a contributing factor to the CB as well. In this work, we exponentially smooth the second-order statistics used in CBPCK in a manner analogous to that used in MFB correction. In this way, all statistical parameters used are subject to similar levels of sampling uncertainty. Section 3d describes how the statistical parameters are estimated.

c. *Radar–satellite fusion*

The MRMS radar-only and the SCaMPR QPEs are first MFB-corrected individually as described in section 3a. The resulting gridded QPEs are fused via simple estimation (SE; Rafieeinasab et al. 2015). In this procedure, the 1-km hourly MRMS radar-only estimates are aggregated to the spatial scale of the 4-km hourly SCaMPR estimates. The two estimates are then fused at each SCaMPR grid box according to Eq. (7) where the time index has been dropped for brevity:

$$X_U^* = wR_U + (1-w)S_U. \tag{7}$$

In the above, R_U denotes the upscaled MFB-corrected radar QPE, S_U denotes the MFB-corrected SCaMPR QPE, w denotes the optimal weight to be determined, and X_U^* denotes the hourly fused estimate at 4-km scale. The weight w is obtained via adaptive CB-penalized optimal linear estimation which minimizes the linearly weighted sum of the error variance and the Type-II CB squared:

$$J = E[(X_U^* - X_U)^2] + \alpha_f E[(E[X_U^*|X_U] - X_U)^2], \tag{8}$$

where α_f denotes the weight given to the CB penalty term for fusion. Seo (2013) arrives at the following Fisher-like solution for the CB-penalized optimal linear estimate X_U^* and the error variance Σ:

$$\Sigma = B[\hat{\mathbf{U}}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{U}]^{-1}, \tag{9}$$

$$X_U^* = [\hat{\mathbf{U}}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{U}]^{-1}\hat{\mathbf{U}}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\mathbf{Z} = [w(1-w)]\begin{bmatrix} R_U \\ S_U \end{bmatrix}. \tag{10}$$

In the above, the (2×1) modified unit vector $\hat{\mathbf{U}}$, the modified observation error covariance matrix $\mathbf{\Lambda}$, and the scaling constant B are given by

$$\hat{\mathbf{U}}^{\mathrm{T}} = (1+\alpha_f)\mathbf{U}^{\mathrm{T}}, \tag{11}$$

$$\mathbf{\Lambda} = \mathscr{R} - \alpha_f(\alpha_f+1)\mathbf{U}\sigma_{X_U}^2\mathbf{U}^{\mathrm{T}}, \tag{12}$$

$$B = \alpha_f\sigma_{X_U}^2\hat{\mathbf{U}}^{\mathrm{T}}\mathbf{\Lambda}^{-1}\hat{\mathbf{U}} + (1+\alpha_f). \tag{13}$$

In the above, $\sigma_{X_U}^2$ denotes the variance of X_U, and the observation error covariance matrix $\mathscr{R}$ is given by

$$\mathscr{R} = \begin{bmatrix} \mathrm{Var}[R_U - X_U] & 0 \\ 0 & \mathrm{Var}[S_U - X_U] \end{bmatrix}. \tag{14}$$

If $\alpha_f = 0$, Eqs. (11) and (12) reduce to the classical Fisher (i.e., maximum likelihood) solution (Schweppe 1973). The diagonality of $\mathscr{R}$ reflects the very reasonable assumption that the observation errors in R_U and S_U are independent. To estimate $\mathrm{Var}[R_U - X_U]$ and $\mathrm{Var}[S_U - X_U]$ in Eq. (14), we update the error statistics of the radar and SCaMPR QPEs versus gauge precipitation in real time via exponential smoothing. The weight α_f is optimized in real time analogously to adaptive CBPCK. To obtain the hourly fused estimate at 1-km scale, X_U^* is disaggregated under the assumption that the 1-km hourly radar QPE perfectly captures the spatial variability of precipitation within each grid box of the SCaMPR QPE:

$$X_{i,j}^* = \frac{X_U^*}{R_U}R_{i,j}, \quad R_U > 0; \quad i = 1,\ldots,4; \quad j = 1,\ldots,4. \tag{15}$$

In the above, $R_{i,j}$ and $X_{i,j}^*$ denote the radar and fused estimates at the ijth pixel, respectively. By replacing X_U^* in Eq. (15) with Eq. (7), we may rewrite $X_{i,j}^*$ as

$$X_{i,j}^* = \left[w + (1-w)\frac{S_U}{R_U}\right]R_{ij} = wR_{i,j} + (1-w)S_{i,j}^*, \quad R_U > 0; \quad i = 1,\ldots,4; \quad j = 1,\ldots,4, \tag{16}$$

where

$$S_{i,j}^* = \frac{R_{i,j}}{R_U}S_U, \quad R_U > 0; \quad i = 1,\ldots,4; \quad j = 1,\ldots,4. \tag{17}$$

Unlike in merging, the weight w does not vary in space owing to the gridded nature of the input QPEs, which renders the fusion algorithm extremely simple. It is important to note that, whereas rain gauge data are used to estimate the observation error statistics, they are not used in the fusion itself unlike in OCK or CBPCK. As such, the timeliness of

the rain gauge data is not as critical to fusion as it is to merging.

d. Estimation of statistical parameters

As implemented in this work, both OCK and CBPCK require modeling spatial covariance structures of intermittency and inner variability of precipitation (Seo 1998a,b). Under the assumption of local homogeneity, we equate the mean fractional coverage of precipitation with the probability of precipitation (Seo and Smith 1996). As in Seo (1998b), we assume $\sigma_g = \sigma_r$; $m_{Ig} = m_{Ir}$; $m_g = m_r$; $\rho_g(|h|) = \rho_c(|h|) = \rho_r(|h|)$; and $\rho_{Ig}(|h|) = \rho_{Ic}(|h|) = \rho_{Ir}(|h|)$, where m_{Ig}, m_g, σ_g, $\rho_{Ig}(|h|)$, and $\rho_g(|h|)$ denote the mean fractional coverage, unconditional mean, standard deviation, indicator correlation, and conditional correlation of gauge precipitation, respectively, with h being the separation distance, the same five statistics for MFB-corrected remotely sensed precipitation are subscripted by r, and $\rho_c(|h|)$ and $\rho_{Ic}(|h|)$ denote the conditional and indicator cross correlation, respectively. The indicator and conditional correlation model the inner variability (i.e., variability of positive precipitation) and intermittency of precipitation (i.e., variability of precipitation versus no precipitation), respectively.

The correlograms are estimated using the hourly radar-only QPE (Seo 1998b). It is impractical to model fully spatiotemporally varying covariance structures in real time due to insufficient data, modeling complexity, and large computing requirements. In the current implementation of OCK in the MPE, the correlation structures are not estimated in real time, but are based on climatological estimates (Seo 1998b; Seo and Breidenbach 2002). In this work, we assume the exponential model (Journel and Huijbregts 1978) for both intermittency and inner variability, and estimate the parameters for the correlation model for each hour for each tile. To reduce computing time, we estimate the correlation coefficients only for the first few lags, from which the nugget effect and the range are estimated (Journel and Huijbregts 1978). Though the exponential model very often provides the best fit for hourly radar and rain gauge precipitation among widely used correlogram models (Seo and Breidenbach 2002; Seo et al. 2014), it may not be reasonable for stochastically highly-regular (i.e., mean-square differentiable) precipitation fields (Vanmarcke 1983) or orographic precipitation (Chua and Bras 1982). In addition, the correlation structure may not be homogeneous within a tile given the large latitudinal dimension (see Fig. 1). As such, the covariance models are subject to potentially significant uncertainties, which may also contribute to the CB. Additional research is needed to improve real-time modeling of the spatiotemporally varying covariance structure of precipitation over large areas.

Because we are using ordinary, rather than simple, cokriging (Journel and Huijbregts 1978), specifying the radius of influence requires additional care. In this work, we set the default radius of influence for locating the neighboring gauge observations to be 2.5 times larger than the larger of the indicator and conditional correlation scales. In this way, weakly correlated rain gauge observations may also be included in the estimation process. It is well known that the predictability of precipitation depends strongly on the magnitude of precipitation. In general, the predictability peaks around the median and decreases toward the tail ends of the distribution of precipitation amount (Seo 1996). Because high-resolution positive precipitation is highly skewed, the predictability tends to decrease rather quickly as the precipitation amount increases. The above dependence may be modeled explicitly using nonlinear estimation techniques such as indicator cokriging (Brown and Seo 2013). They are, however, computationally very expensive, and require large amounts of data for parameter estimation. In this work, we model the above dependence by parameterizing the radius of influence with the radar precipitation amount at u_0 as follows:

$$\mathrm{ROI} = \mathrm{ROI}_{\mathrm{def}} \exp\left[-\frac{Z_R(u_0)}{L}\right], \tag{18}$$

where ROI denotes the radius of influence, $\mathrm{ROI}_{\mathrm{def}}$ denotes the default radius of influence, $Z_R(u_0)$ denotes the radar precipitation at u_0, and L denotes the characteristic precipitation amount (mm). If radar precipitation does not exist at u_0, the default radius of influence is used. The maximum number of neighboring gauge observations used in the estimation process is 30 to limit the amount of computation. The actual number of rain gauge observations used (i.e., those within ROI), however, is generally much smaller, particularly in gauge-sparse areas. The fractional coverage of precipitation is estimated by dividing the number of positive observations by the total number of observations within $\mathrm{ROI}_{\mathrm{def}}$.

e. Evaluation

For comparative evaluation of the different QPEs considered, we carried out true validation, which means that, for 13–30 September 2015, we randomly selected 4%–5% of all available gauge observations within each tile for each hour, withheld them for validation, and used the rest for parameter estimation. The gauge network density varies greatly from

one tile to another, and within each tile (see Fig. 3). In gauge-poor areas, setting aside more gauges would keep the neighboring gauges too distant to provide a predictive skill that may actually exist. The choice of 4%–5% represents a compromise between retaining as much of the actual gauge network as possible and increasing the sample size for validation. For large-sample validation, the above experiment was repeated 5 times, each time withholding previously unwithheld gauges. The total number of data points obtained in this way for validation is over 20 000 for the 11 tiles. For 7–9 October 2016, 3%–4% of the gauges were withheld. The experiment was then repeated 10 times using previously unwithheld gauges. The total number of data points obtained for validation is over 19 000 for this analysis period. Comparative evaluation for the October 2016 case is focused on addressing the following questions: 1) How much does the MFB-corrected radar QPE improve over the radar-only? 2) How much do the OCK estimates from merging gauge data and MFB-corrected radar QPE improve over the latter alone? and 3) How much does adaptive CBPCK improve over OCK? Comparative evaluation for the September 2015 case is focused on addressing: 1) How much do the SE estimates from fusing the MFB-corrected SCaMPR and radar QPEs improve over the latter alone? 2) How much do the OCK estimates from merging gauge data and MFB-corrected gridded QPE improve over the latter alone? and 3) How much does adaptive CBPCK improve over OCK? For both cases, we also address how the above comparative performance may vary according to the magnitude of precipitation being estimated.

For evaluation metrics, we use the root-mean-square error (RMSE), its decomposition, and percent reduction in RMSE relative to reference QPE. The RMSE collectively measures biases in the mean and in standard deviation, and strength of correlation. Both OCK and CBPCK are unbiased estimators, and their performance with respect to the three attributes above has also been reported in Seo et al. (2010), Seo et al. (2014), and Kim et al. (2018). The percent reduction in RMSE, or PRiRMSE($\mathrm{QPE}_{\mathrm{eval}}$), is defined as

$$\mathrm{PRiRMSE}(\mathrm{QPE}_{\mathrm{eval}}) = \frac{\mathrm{RMSE}(\mathrm{QPE}_{\mathrm{ref}}) - \mathrm{RMSE}(\mathrm{QPE}_{\mathrm{eval}})}{\mathrm{RMSE}(\mathrm{QPE}_{\mathrm{ref}})} \times 100, \tag{19}$$

where $\mathrm{RMSE}(\mathrm{QPE}_{\mathrm{eval}})$ and $\mathrm{RMSE}(\mathrm{QPE}_{\mathrm{ref}})$ denote the RMSEs of the QPE under evaluation and the reference QPE, respectively. MSE decomposition (Murphy and Winkler 1987; Nelson et al. 2010) is given by

$$\mathrm{MSE} = (m_e - m_o)^2 + (\sigma_e - \sigma_o)^2 + 2\sigma_e\sigma_o(1 - \rho_{e,o}), \tag{20}$$

where m_e and m_o denote the mean of the estimate and verifying observation, respectively, σ_e and σ_o denote the standard deviation of the estimate and verifying observation, respectively, and $\rho_{e,o}$ denotes the correlation between the estimate and verifying observation.

MFB correction and merging address the first- (i.e., systematic) and second-order (i.e., random) errors, respectively. One may consider the CB a 1.5th-order error in that it is systematic but exists only over the tails of the distribution. First-order errors impact the accuracy across the board. As such, one may expect MFB correction to have the largest impact if significant first-order errors exist. The focus of this paper is on addressing the CB. We are hence interested in assessing the accuracy of the merged QPE relative to that of the MFB-corrected radar, in addition to that of the radar-only QPE. For this reason, we present the results of both MFB correction and merging relative to radar-only QPE so that the improvement due solely to each may be ascertained.

4. Results

This section presents the merging results for 7–9 October 2016, and 13–30 September 2015, and the fusion and merging results for 13–30 September 2015.

a. Radar–gauge merging

The performance of the procedures described above depends on the predictability of the precipitation field, the skill in the remotely sensed QPE, and the gauge network density. To assess the first two factors between the two periods, we first examine the indicator and conditional spatial correlation scales of radar QPE, and the (spatial) lag-0 indicator and conditional cross correlation between gauge observation and radar QPE as estimated in the simulated real-time mode. Figure 5 shows the histograms of the above four parameters for the entire analysis periods over the CONUS. Because these estimates reflect different types of precipitation events in different phases of

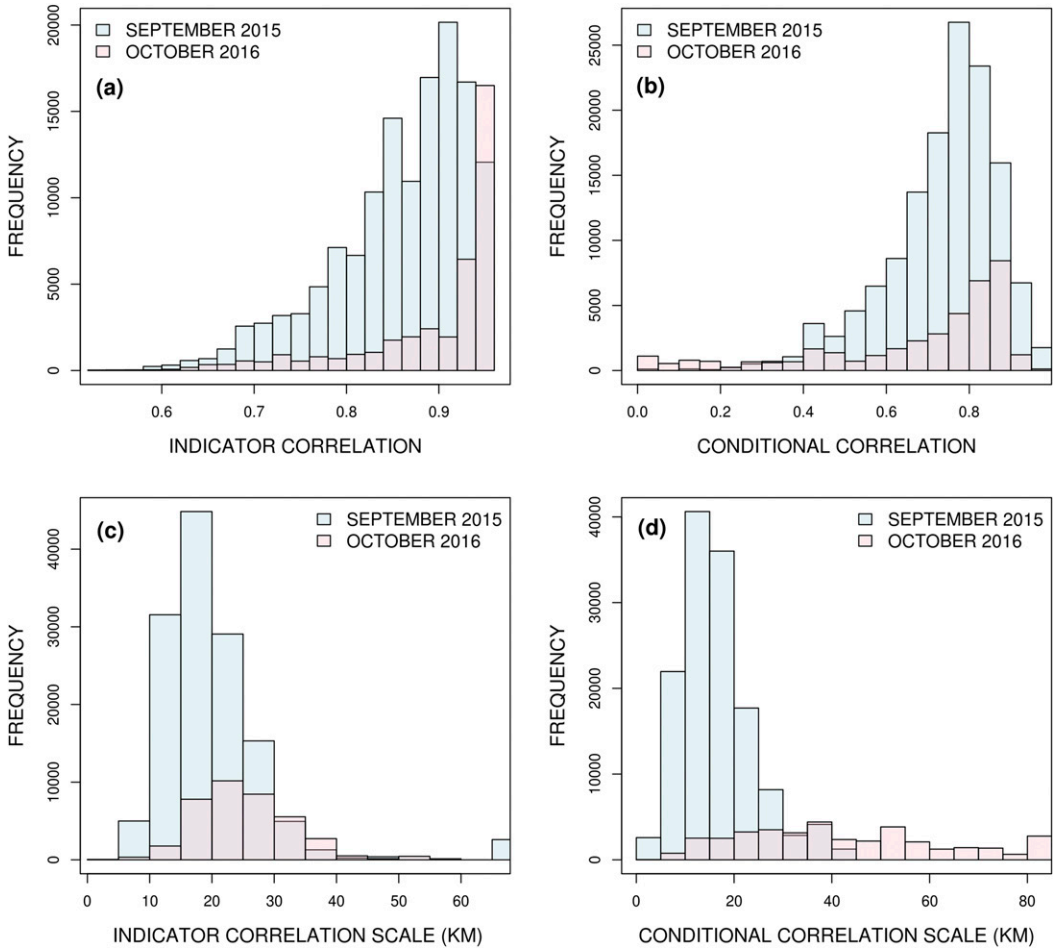

FIG. 5. Histograms of (a) lag-0 indicator correlation, (b) lag-0 conditional correlation, (c) indicator correlation scale, and (d) conditional correlation scale for 13–30 Sep 2015, and 7–9 Oct 2016.

development and dissipation under a wide range of conditions for radar observation of precipitation, the sample statistics necessarily vary greatly. Nonetheless, it is readily seen that all four parameters for September 2015, particularly the conditional correlation and conditional correlation scale, are significantly smaller than those for October 2016, an indication that the precipitation fields in the first period are characterized by significantly smaller predictability. One may hence expect merging to be more potent for October 2016, and the CB to be larger for September 2015.

1) 7–9 OCTOBER 2016

This period includes precipitation from storms in three largely different regions: extreme amounts from Hurricane Matthew along the Atlantic Coast, significant amounts from a relatively weakly organized convective storm in the central United States, and significant amounts from a coastal storm in the Pacific Northwest (see Fig. 1). Figure 6a shows the RMSE of the radar-only, MFB-corrected radar, OCK, and adaptive CBPCK estimates over the CONUS conditional on the verifying observed hourly precipitation exceeding the amount shown on the x axis. The values on the y axis at $x = 0$ represent the RMSE conditional on the verifying observation being nonzero, which is very close to the unconditional RMSE. As such, we refer to the results for $x = 0$ and $x > 0$ as unconditional and conditional performances, respectively. Also shown in the figure is the sample size whose axis is shown at the right end of the plot, also on a logarithmic scale. Figure 6b shows the percent reduction in RMSE of the MFB-corrected radar, OCK, and adaptive CBPCK estimates relative to the radar-only QPE. In both figures, the conditioning threshold is cut off at about 34 mm, above which the results are very noisy due to the small sample size. Figure 6b may be

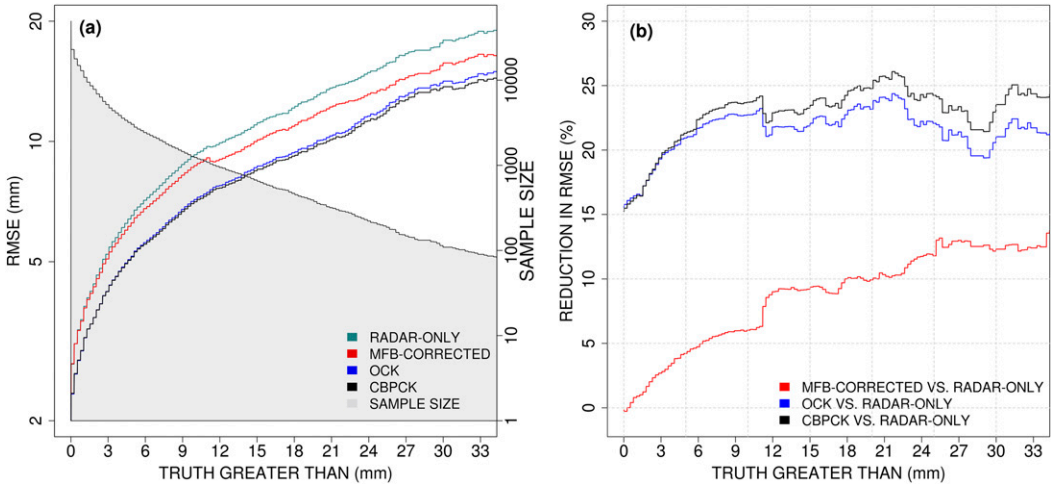

FIG. 6. (a) RMSE of radar-only, MFB-corrected radar, OCK, and adaptive CBPCK estimates for 7–9 Oct 2016, conditional on truth exceeding the amount on the *x*-axis; (b) as in (a), but for percent reduction in RMSE of radar-only QPE by MFB-corrected radar, OCK, and adaptive CBPCK estimates.

summarized as follows. Both OCK and adaptive CBPCK reduce the unconditional RMSE of radar-only QPE by about 16% whereas MFB correction provides no reduction. Adaptive CBPCK performs comparably to OCK unconditionally, an indication that adaptively optimizing α does produce the desired effect of improving the unconditional performance of CBPCK. For precipitation amounts exceeding 1.5 mm, adaptive CBPCK improves over OCK. At the conditioning amount of 24.5 mm, MFB correction, OCK, and adaptive CBPCK reduce the conditional RMSE of radar-only QPE by about 13%, 22%, and 25%, respectively. The relatively modest improvement by adaptive CBPCK over OCK is a reflection that the CB is generally not very large in this period owing to the large spatial predictability in the well-developed precipitation systems in the east.

To show how the different estimates may compare, we show in Fig. 7 the scatterplots of radar-only (upper left), MFB-corrected radar (upper right), OCK (lower left), and adaptive CBPCK (lower right) estimates versus the verifying gauge precipitation. The data points are color-coded by region so that the performance for different storms within the CONUS may be examined. The generally positive impact of MFB correction of radar QPE is readily seen in the figure as reflected by a nearly diagonal quantile–quantile (Q–Q) plot, but at the expense of inflating large, overestimated radar-only precipitation. Such miscorrections occur because biases in radar-only QPE may be spatially nonuniform (Seo and Breidenbach 2000) or nonlinear, that is, precipitation magnitude dependent, which cannot be addressed by MFB correction alone. The lower panel shows that both OCK and adaptive CBPCK greatly reduce the scatter around the diagonal. Overall, the OCK and adaptive CBPCK estimates are very similar as Fig. 6 would suggest. For the eastern region, which encompasses Hurricane Matthew, the two estimates show little difference, an indication that there is little CB present owing to the large predictability, dense gauge networks, and generally favorable conditions for radar observation of precipitation. Though small, noticeable differences are seen in the plotting areas of $40 < \text{truth} < 60$ (mm) and $20 < \text{estimate} < 40$ (mm) for the central region, and $0 < \text{truth} < 20$ (mm) and $0 < \text{estimate} < 20$ (mm) for the western region, where a number of adaptive CBPCK estimates are closer to the diagonal than the OCK estimates.

2) 13–30 SEPTEMBER 2015

This analysis period includes multiple mostly convective events of relatively low predictability in the Pacific Northwest, Midwest, Northeast, and Southeast of the United States (see Fig. 2). Figure 8 is the same as Fig. 6 but for September 2015. Figures 8a and 8b may be summarized as follows. MFB correction reduces the unconditional RMSE of radar-only QPE by about 9%. Both OCK and adaptive CBPCK are able to increase the margin of reduction to about 17%. For these lower predictability events, however, the effectiveness of merging is reduced as the conditioning amount increases. For truth exceeding about 18 mm, the OCK estimates are no longer more accurate than the MFB-corrected radar in the mean squared error sense. The adaptive CBPCK estimates, on the other hand, perform better than the better of the OCK and the MFB-corrected radar estimates for all conditioning amounts.

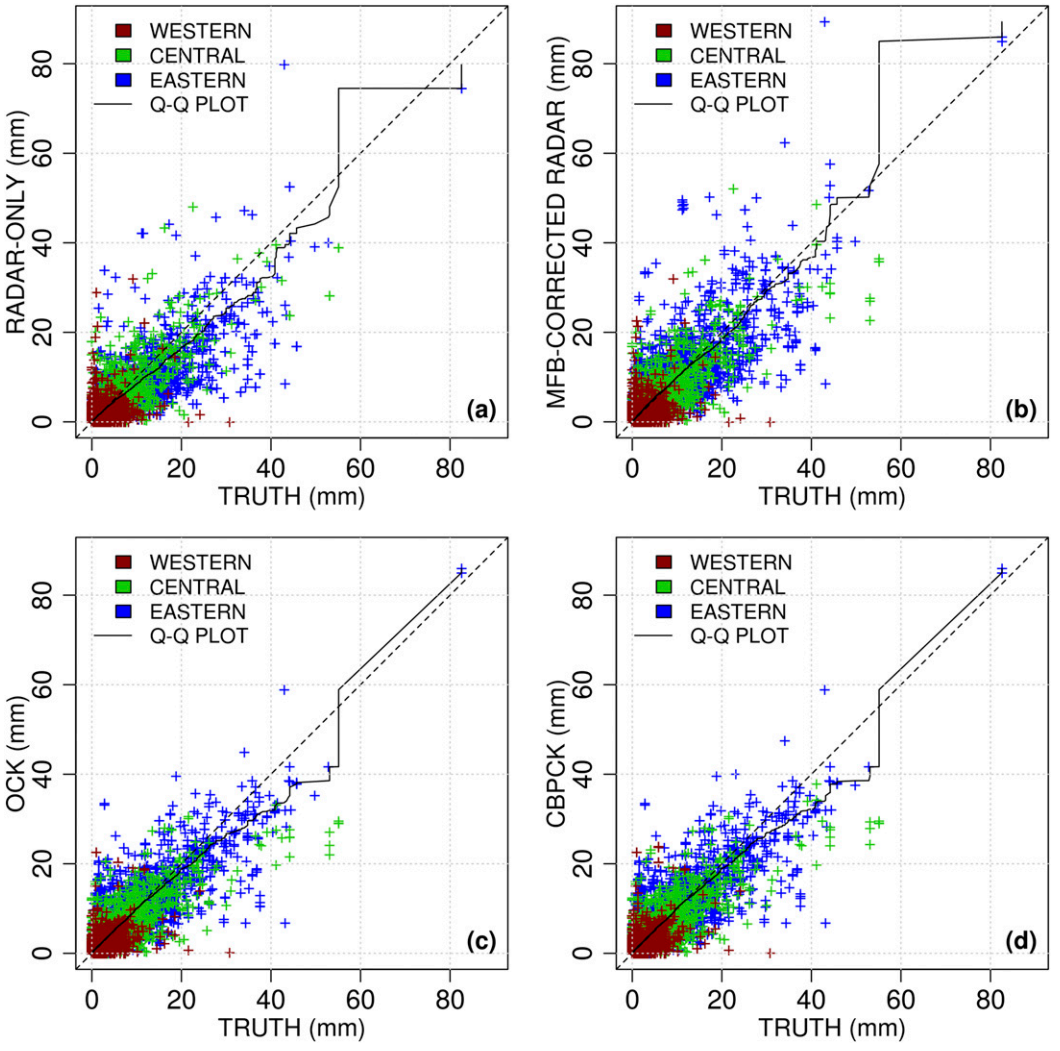

FIG. 7. Scatterplots of (a) radar-only, (b) MFB-corrected radar, (c) OCK, and (d) adaptive CBPCK estimates for 7–9 Oct 2016 vs truth.

The larger margin of improvement by adaptive CBPCK over OCK reflects the presence of a larger CB due to smaller spatial predictability in this period. About 65% of the true-validated estimates in Fig. 8 are associated with $\alpha = 0$, that is, the adaptive CBPCK estimates are the same as the OCK estimates.

To assess the quality of the adaptive CBPCK estimates exclusively, it is necessary to consider only those estimates associated with nonzero α. To that end, we plot in Fig. 9 the RMSEs of the radar-only, MFB-corrected radar, OCK, and adaptive CBPCK estimates (upper panels), and the percent reduction in RMSE of the radar-only QPE by the MFB-corrected radar, OCK and adaptive CBPCK estimates (lower panels) for different ranges of positive α and conditioning amounts of truth. Also shown in the upper panels are the 90% confidence intervals for the OCK and adaptive CBPCK estimates obtained via bootstrapping. Figure 9 indicates that, when the CB is present (i.e., $\alpha > 0$), the adaptive CBPCK estimates are superior to the OCK estimates for truth exceeding 25.4 mm at a significance level of 0.10, but that, when all amounts of truth are considered, the improvement is not statistically significant. The latter is not at all surprising given the fact that smaller amounts of precipitation, for which little CB exists, far outnumber large amounts. Figure 10 shows the scatterplots of the estimates versus the observed. Reduction of bias and scatter due to bias correction and merging, respectively, is readily seen. For this period, larger differences are seen between the OCK and adaptive CBPCK estimates than in October 2016, particularly in the eastern region. The general direction of change is that the CBPCK estimates tend to decrease and increase very small and larger OCK estimates, respectively.

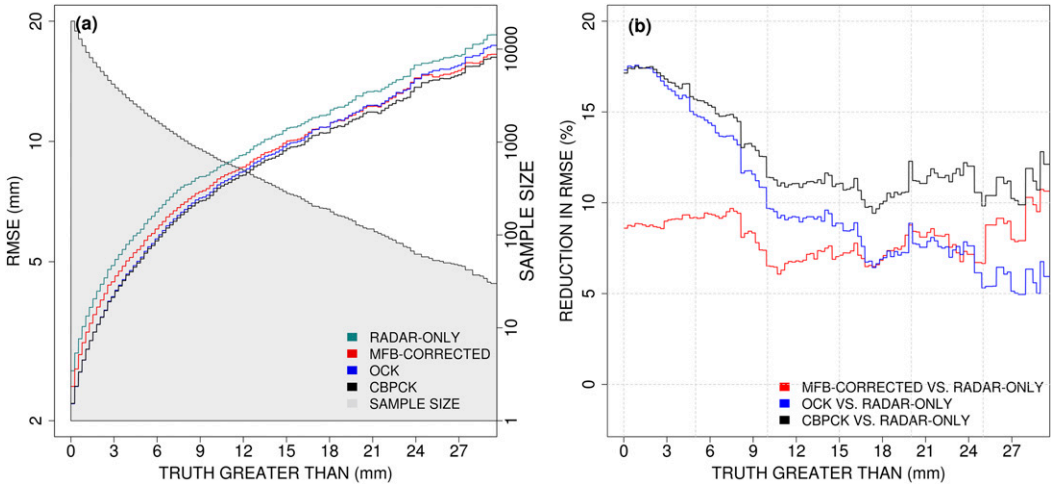

FIG. 8. As in Fig. 6, but for 13–30 Sep 2015.

Figures 11a, 11b, and 11c show the $(m_e - m_o)^2$, $(\sigma_e - \sigma_o)^2$, and $\rho_{e,o}$ terms in the MSE decomposition of Eq. (20), respectively. They are for the radar-only, MFB-corrected, OCK, and adaptive CBPCK estimates conditioned on the truth exceeding 0, 6.4, 12.7, and 25.4 mm. In each figure, the sample statistics for the September 2015 and October 2016 periods are plotted on the x and y axes, respectively. For each conditioning threshold, the sample statistics for the 4 different estimates are connected with a dashed line. The most desirable estimates would place the sample statistics of $(m_e - m_o)^2$ and $(\sigma_e - \sigma_o)^2$ closest to the lower-left corner in Figs. 11a and 11b, respectively, and that of $\rho_{e,o}$ closest to the upper-right corner in Fig. 11c. Figure 11 indicates that the MFB-corrected estimates are least biased in the mean and standard deviation, followed by the CBPCK estimates, and that the CBPCK estimates are most strongly correlated with the truth, followed by the OCK estimates. The first result is not very surprising in that the sole purpose of MFB correction is to reduce bias in the mean. Also, being conditional expectation operators, OCK and adaptive CBPCK necessarily reduce variability due to averaging. One may avoid such smoothing by performing conditional simulation (Deutsch and Journel 1992; Seo et al. 2000) using OCK or adaptive CBPCK in an ensemble framework. Such an approach, however, is computationally too expensive to be practical for real time QPE. Figure 11 indicates that OCK and adaptive CBPCK reduce MSE over MFB-corrected radar by a combination of smoothing and improved correlation with the truth [i.e., reduced σ_e and increased $\rho_{e,o}$ in Eq. (20), respectively], and that the adaptive CBPCK estimates are superior to the OCK estimates in all categories except in unconditional bias in the mean and conditional bias in the standard deviation for the 25.4-mm threshold.

b. *Fusion of SCaMPR QPE*

Figure 12 is the same as Fig. 8 but for the radar-only, MFB-corrected radar, SE-fused, OCK, and adaptive CBPCK estimates for September 2015 over the CONUS. The SE results are based on fusing the MFB-corrected radar and MFB-corrected SCaMPR estimates as described in section 3c (see also Fig. 4). The OCK and adaptive CBPCK results are based on merging the gauge data with the fused QPE. Figure 12 indicates that the MFB-corrected radar and adaptive CBPCK provide significantly better conditional performance than the others. Figure 12b shows that the percent reduction in unconditional RMSE by the MFB-corrected radar, SE-fused, OCK, and adaptive CBPCK over the radar-only QPE is about 8%, 12%, 18%, and 18%, respectively. As the conditioning amount increases, however, the accuracy of the OCK estimates deteriorates and, at the conditioning amount of about 15 mm, it falls below that of the fused estimates. The adaptive CBPCK estimates, on the other hand, remain better than the MFB-corrected radar for conditioning amounts of up to 25 mm. The above results demonstrate the adaptive CBPCK's ability to improve conditional performance over OCK while performing comparably to OCK in the unconditional sense.

We now assess the impact of the MFB-corrected SCaMPR QPE in radar–satellite fusion by comparatively evaluating the SE-fused versus the MFB-corrected radar estimates. Note that, in this comparison, we are not assessing the relative value of the radar- and SCaMPR-only estimates, but that of the MFB-corrected radar and MFB-corrected SCaMPR estimates. Accordingly, additional factors such as the rain gauge network

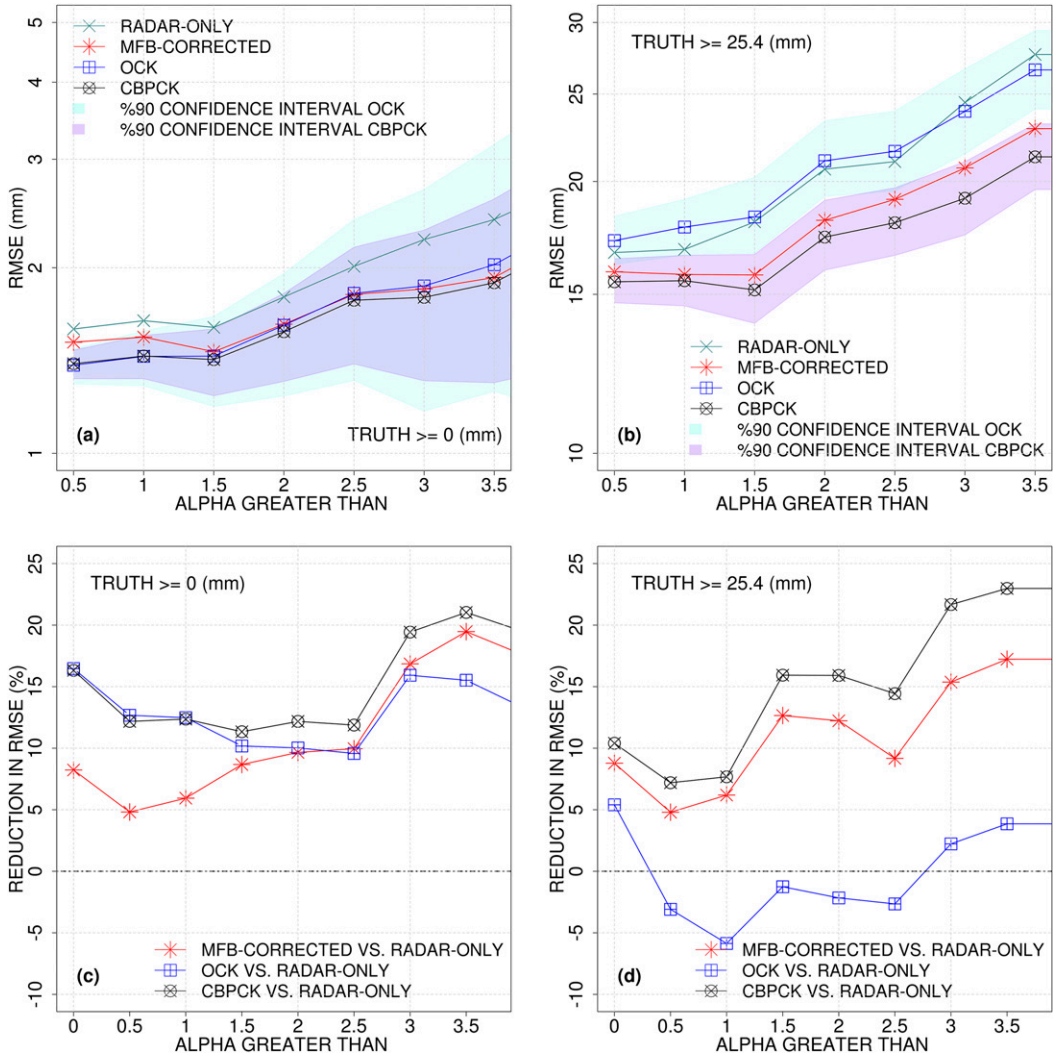

FIG. 9. (a) RMSE of radar-only, MFB-corrected radar, OCK, and adaptive CBPCK estimates and 90% confidence interval of the OCK and adaptive CBPCK estimates, for 13–30 Sep 2016, conditional on α exceeding the value on the x-axis for truth exceeding zero; (b) as in (a), but for truth exceeding 25.4 mm; (c) as in (a), but for percent reduction in RMSE of radar-only QPE by MFB-corrected radar, OCK, and adaptive CBPCK estimates; and (d) as in (c), but for truth exceeding 25.4 mm.

density, spatiotemporal variability of precipitation, the skill in radar-only QPE, and the efficacy of MFB correction come into play. Figure 12b shows that MFB correction reduces the RMSE of the radar-only QPE by about 8% over the CONUS, and the fused QPE additionally reduces the RMSE by about 4%. The tile-specific results indicate that, for the western half of the CONUS (i.e., Tiles 1–3), the additional reduction by the fused QPE is over 6% whereas for the eastern half it is only about 3%. Compared to the western half of the CONUS, the eastern half has significantly denser rain gauge networks (see Fig. 4), and is generally more favorable for radar QPE. The latter point may be seen in the tile-specific correlation coefficient of radar-only QPE with rain gauge observations; for September 2015, the correlation is 0.64, 0.44, 0.66, 0.77, 0.77, and 0.77 for Tiles 1, 2, 3, 4, 5, and 6, respectively. It is surmised that, for the eastern half, the combination of the skillful radar QPE and the dense rain gauge networks is able to produce significantly more accurate MFB-corrected radar QPE, and that the comparative skill of the MFB-corrected SCaMPR estimates is too small to improve on the former significantly. The western region, on the other hand, has a substantially lower density of rain gauges and are not very favorable for radar QPE. As such, the MFB-corrected SCaMPR is able to provide larger improvement. Co-examination of Figs. 8 and 12 indicates that, whereas

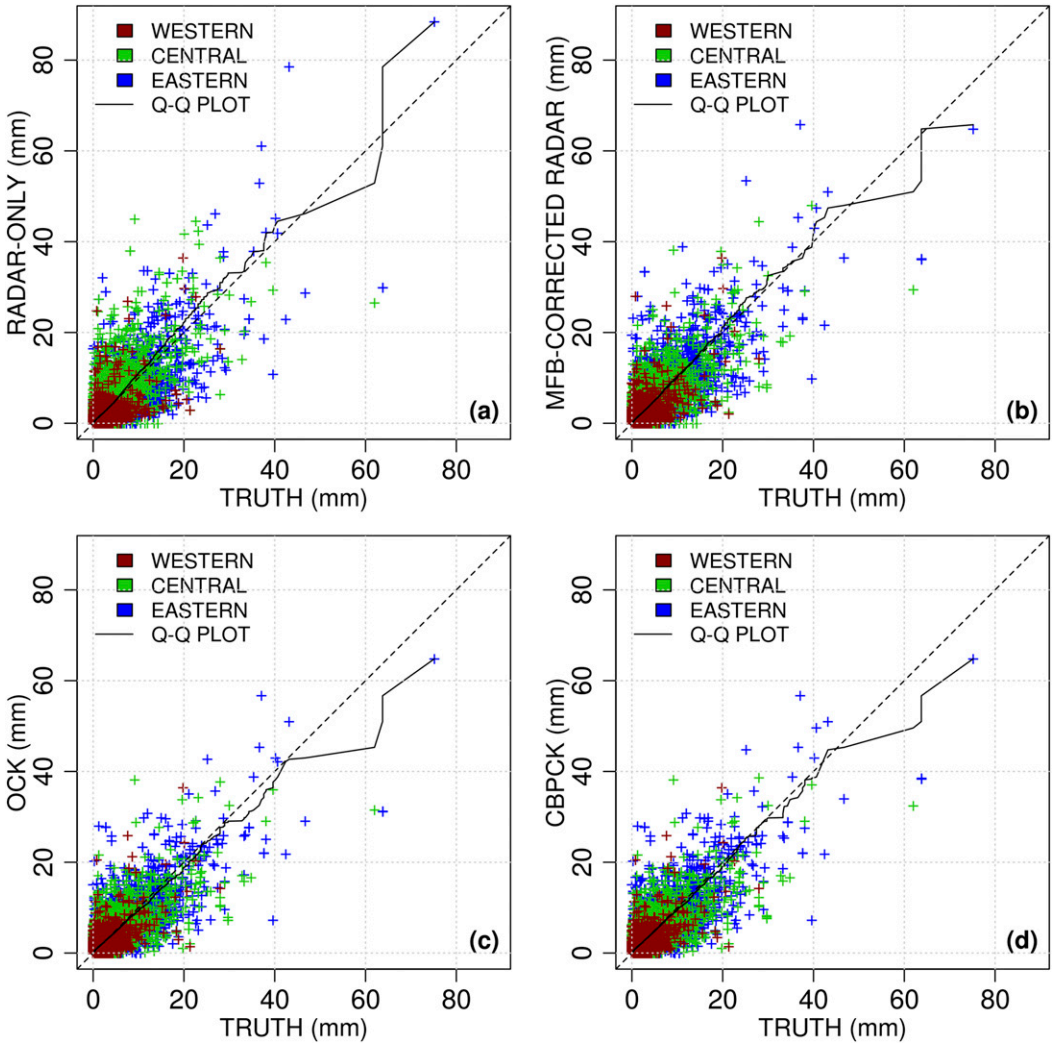

FIG. 10. As in Fig. 7, but for 13–30 Sep 2015.

the addition of the SCaMPR QPE does improve the accuracy of the final merged QPE in the unconditional sense, in particular in the western half of the CONUS, the conditional performance deteriorates. Figures 8b and 12b show that the added value of the SCaMPR estimates is lost when the precipitation amount exceeds about 1.5 mm, above which merging rain gauge data and MFB-corrected radar QPE via adaptive CBPCK is superior. The above findings suggest that the SCaMPR product should be used selectively in multisensor QPE.

Computationally, adaptive CBPCK requires solving an $(n_g + n_r)$-dimensional linear system multiple times whereas OCK requires solving a comparable system only once. In the above, n_g and n_r are usually on the orders of 10 and 1, respectively, and α may range from 0 to 4 incremented by 1. With the naïve optimization of α used in this work, adaptive CBPCK is hence several times more expensive than OCK. To improve understanding of the dependence of the CBPCK solution on α, and to develop a more effective and computationally efficient approach for its optimization, additional research is needed.

5. Conclusions

The principal conclusion of this work is that, to produce multisensor estimates that are more accurate than the ingredient quantitative precipitation estimates (QPE) for all precipitation amounts, it is necessary to address the conditional bias (CB), and that adaptive conditional bias–penalized cokriging (CBPCK) described in this paper improves estimation of significant amounts of precipitation by explicitly considering the CB. It is shown that, beyond the reduction in root-mean-square error (RMSE)

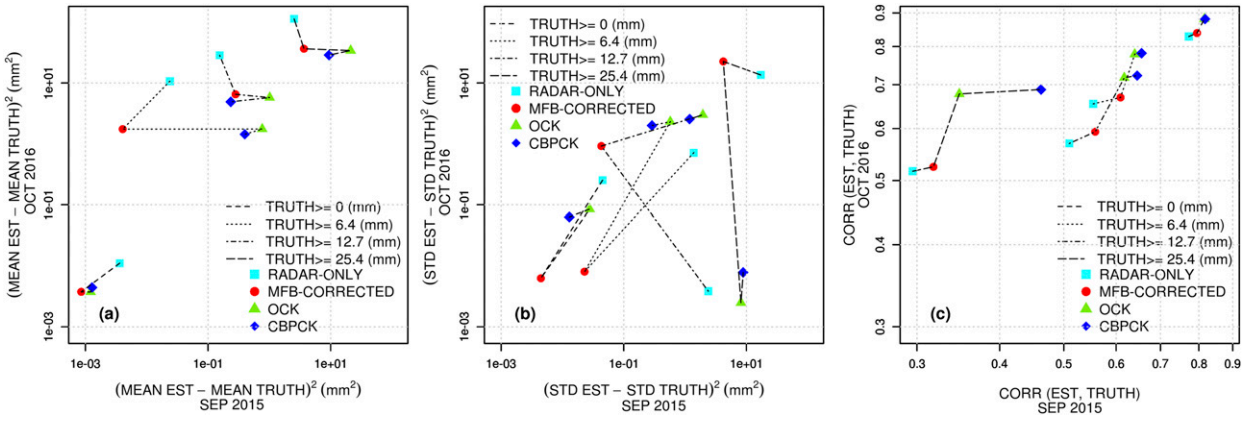

FIG. 11. Decomposition of MSE of radar-only, MFB-corrected, OCK, and adaptive CBPCK estimates for 13–30 Sep 2015 (x axis) and 7–9 Oct 2016 (y axis) into (a) $(m_e - m_o)^2$, (b) $(\sigma_e - \sigma_o)^2$, and (c) $\rho_{e,o}$ conditioned on truth exceeding 0, 6.4, 12.7, and 25.4 mm.

due to mean field bias (MFB) correction, both ordinary cokriging (OCK) and adaptive CBPCK additionally reduce the unconditional RMSE of radar-only QPE by 16% and 9% over the continental United States (CONUS) for the more and less predictable 7–9 October 2016, and 13–30 September 2015, events, respectively, and that adaptive CBPCK improves over OCK for estimation of hourly precipitation exceeding about 1 mm. Jointly, MFB correction and adaptive CBPCK reduce the RMSE of the radar-only QPE by about 16%–26% for the more predictable 7–9 October 2016, events and by about 10%–17% for the less predictable 13–30 September 2015, events for all ranges of precipitation amounts.

It is shown that, for the September 2015 events, fusing the MFB-corrected radar QPE with the MFB-corrected SCaMPR QPE reduces the unconditional RMSE of radar-only QPE by about 12% over the CONUS whereas the reduction by MFB-corrected radar QPE alone over radar-only QPE is about 8%. For the western half of the CONUS, where the rain gauge network is sparser and the radar QPE is less skillful, the margin of reduction increases to 6% from the above 4%. The conditional performance of the fused QPE, however, falls below that of the MFB-corrected radar QPE as the conditioning amount exceeds about 7 mm of hourly precipitation. The above suggests that the SCaMPR product should be used selectively in multisensor QPE.

The error variance estimates from adaptive CBPCK were not used in this work. Coutilizing both the estimate and the error variance is likely to improve the optimization of α. Additional evaluation is needed to assess the performance of adaptive CBPCK further, and to

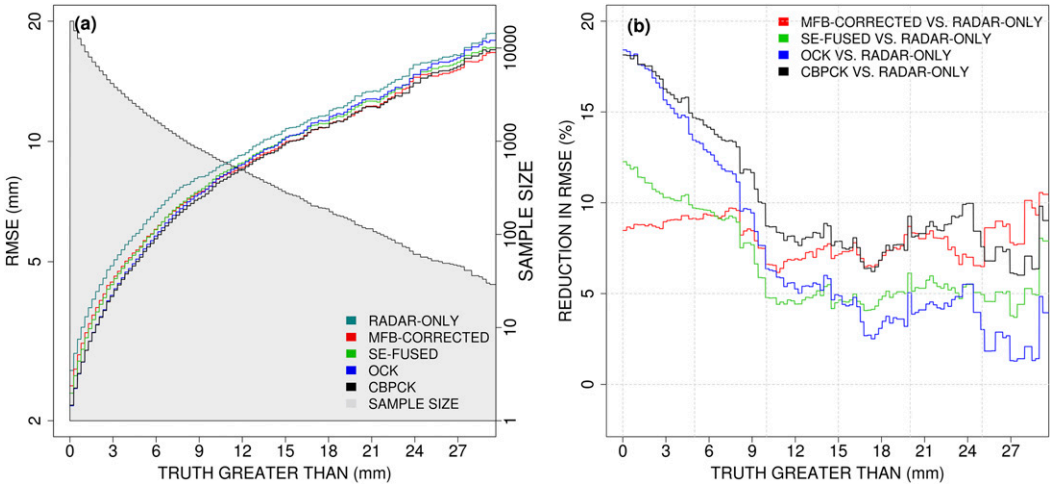

FIG. 12. As in Fig. 8, but for radar-only, MFB-corrected radar, SE-fused, OCK, and adaptive CBPCK estimates. SE fuses the MFB-corrected radar and MFB-corrected SCaMPR. OCK and adaptive CBPCK merge the SE-fused estimates and rain gauge data.

optimize its parameters. We note here that the algorithm described in this work is being comparatively evaluated against Stage IV (Nelson et al. 2010) in real time for possible operational implementation in the Multi-Radar Multi-Sensor (MRMS) system (Tang et al. 2019), and the results are to be reported in the near future.

Acknowledgments. This material is based upon work supported in part by the NOAA's Joint Technology Transfer Initiative Program (Grant NA16OAR4590232) and by the NSF (Grant CyberSEES-1442735). These supports are gratefully acknowledged. We thank Mr. David Kitzmiller of NOAA/NWS/Office of Water Prediction (OWP) for valuable input and feedback throughout the course of this work, and careful review of the manuscript. We are also grateful to Mr. Greg Fall of NOAA/NWS/OWP for helpful input at various stages of this work, and Dr. Robert Kuligowski of NOAA/NESDIS/Center for Satellite Applications and Research for providing the SCaMPR QPE products used in this work.

APPENDIX

List of Acronyms and Abbreviations

CB	Conditional bias
CBPCK	Conditional bias–penalized cokriging
CBPK	Conditional bias–penalized kriging
CONUS	Continental United States
GOES	Geostationary Operational Environmental Satellites
HADS	Hydrometeorological Automated Data System
KF	Kalman filter
MFB	Mean field bias
MSE	Mean square error
MPE	Multisensor Precipitation Estimator
MRMS	Multi-Radar Multi-Sensor System
NWS	National Weather Service
OCK	Ordinary cokriging
QPE	Quantitative precipitation estimate
Q–Q	Quantile–quantile
RMSE	Root-mean-square error
SCaMPR	Self-Calibrating Multivariate Precipitation Retrieval
SE	Simple estimation

REFERENCES

AghaKouchak, A., A. Behrangi, S. Sorooshian, K. Hsu, and E. Amitai, 2011: Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *J. Geophys. Res.*, **116**, D02115, https://doi.org/10.1029/2010JD014741.

Ashouri, H., K. L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, and O. P. Prat, 2015: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.*, **96**, 69–83, https://doi.org/10.1175/BAMS-D-13-00068.1.

Brown, J. D., and D.-J. Seo, 2013: Evaluation of a nonparametric post-processor for bias correction and uncertainty estimation of hydrologic predictions. *Hydrol. Processes*, **27**, 83–105, https://doi.org/10.1002/hyp.9263.

Chua, S. H., and R. L. Bras, 1982: Optimal estimators of mean areal precipitation in regions of orographic influence. *J. Hydrol.*, **57**, 23–48, https://doi.org/10.1016/0022-1694(82)90101-9.

Ciach, G. J., M. L. Morrissey, and W. F. Krajewski, 2000: Conditional bias in radar rainfall estimation. *J. Appl. Meteor.*, **39**, 1941–1946, https://doi.org/10.1175/1520-0450(2000)039<1941:CBIRRE>2.0.CO;2.

Cocks, S. B., J. Zhang, S. M. Martinaitis, Y. Qi, B. Kaney, and K. Howard, 2017: MRMS QPE performance east of the Rockies during the 2014 warm season. *J. Hydrometeor.*, **18**, 761–775, https://doi.org/10.1175/JHM-D-16-0179.1.

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, https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

Deutsch, C. V., and A. G. Journel, 1992: *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, 340 pp.

Frost, C., and S. G. Thompson, 2000: Correcting for regression dilution bias: Comparison of methods for a single predictor variable. *J. Roy. Stat. Soc.*, **163A**, 173–190, https://doi.org/10.1111/1467-985X.00164.

Fuller, W. A., 1987: *Measurement Error Models*. Wiley, 440 pp.

Gandin, L. S., 1966: *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translations, 242 pp.

Goudenhoofdt, E., and L. Delobbe, 2009: Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol. Earth Syst. Sci.*, **13**, 195–203, https://doi.org/10.5194/hess-13-195-2009.

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. Hydrometeor.*, **12**, 973–988, https://doi.org/10.1175/2011JHM1287.1.

Habib, E., A. Henschke, and R. F. Adler, 2009: Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmos. Res.*, **94**, 373–388, https://doi.org/10.1016/j.atmosres.2009.06.015.

——, A. T. Haile, Y. Tian, and R. J. Joyce, 2012: Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates. *J. Hydrometeor.*, **13**, 1784–1798, https://doi.org/10.1175/JHM-D-12-017.1.

——, L. Qin, D.-J. Seo, G. J. Ciach, and B. R. Nelson, 2013: Independent assessment of incremental complexity in NWS multisensor precipitation estimator algorithms. *J. Hydrol. Eng.*, **18**, 143–155, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000638.

Hausman, J., 2001: Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *J. Econ. Perspect.*, **15**, 57–67, https://doi.org/10.1257/jep.15.4.57.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear,

combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

——, D. T. Bolvin, and E. J. Nelkin, 2017: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. NASA Tech. Doc., 54 pp., https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf].

Hughes, M. D., 1993: Regression dilution in the proportional hazards model. *Biometrics*, **49**, 1056–1066, https://doi.org/10.2307/2532247.

Joliffe, I. T., and D. B. Stephenson, Eds., 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 254 pp.

Journel, A. G., and Ch. J. Huijbregts, 1978: *Mining Geostatistics*. Academic Press, 600 pp.

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. Hydrometeor.*, **5**, 487–503, https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Kalinga, O. A., and T. Y. Gan, 2010: Estimation of rainfall from infrared-microwave satellite data for basin-scale hydrologic modelling. *Hydrol. Processes*, **24**, 2068–2086, https://doi.org/10.1002/HYP.7626.

Kim, B., D.-J. Seo, S. Noh, O. P. Prat, and B. R. Nelson, 2018: Improving multisensor estimation of heavy-to-extreme precipitation via conditional bias-penalized optimal estimation. *J. Hydrol.*, **556**, 1096–1109, https://doi.org/10.1016/j.jhydrol.2016.10.052.

Kim, D., B. Nelson, and D.-J. Seo, 2009: Characteristics of reprocessed Hydrometeorological Automated Data System (HADS) hourly precipitation data. *Wea. Forecasting*, **24**, 1287–1296, https://doi.org/10.1175/2009WAF2222227.1.

Kitzmiller, D., D. Miller, R. Fulton, and F. Ding, 2013: Radar and multisensor precipitation estimation techniques in National Weather Service hydrologic operations. *J. Hydrol. Eng.*, **18**, 133–142, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000523.

Kondragunta, C., D. Kitzmiller, D.-J. Seo, and K. Shrestha, 2005: Objective Integration of satellite, rain gauge, and radar precipitation estimates in the multisensor precipitation estimator algorithm. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 2.8, https://ams.confex.com/ams/Annual2005/techprogram/paper_86219.htm.

Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeor.*, **3**, 112–130, https://doi.org/10.1175/1525-7541(2002)003<0112:ASCRTG>2.0.CO;2.

——, 2013: GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for rainfall rate (QPE). NOAA/NESDIS/STAR Algorithm Theoretical Basis Doc., version 2.6, 46 pp., https://www.star.nesdis.noaa.gov/goesr/documents/ATBDs/Baseline/ATBD_GOES-R_Rainrate_v2.6_Oct2013.pdf.

——, Y. Li, and Y. Zhang, 2013: Impact of TRMM data on a low-latency, high-resolution precipitation algorithm for flash-flood forecasting. *J. Appl. Meteor. Climatol.*, **52**, 1379–1393, https://doi.org/10.1175/JAMC-D-12-0107.1.

Lee, H., S. Noh, S. Kim, H. Shen, D.-J. Seo, and Y. Zhang, 2019: Improving flood forecasting using conditional bias-penalized ensemble Kalman filter. *J. Hydrol.*, **575**, 596–611, https://doi.org/10.1016/j.jhydrol.2019.05.072.

Murphy, A. H., and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.*, **115**, 1330–1338, https://doi.org/10.1175/1520-0493(1987)115<1330:AGFFFV>2.0.CO;2.

Nelson, B. R., D.-J. Seo, and D. Kim, 2010: Multisensor precipitation reanalysis. *J. Hydrometeor.*, **11**, 666–682, https://doi.org/10.1175/2010JHM1210.1.

——, O. P. Prat, D.-J. Seo, and E. Habib, 2016: Assessment and implications of Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Okamoto, K., T. Ushio, T. Iguchi, N. Takahashi, and K. Iwanami, 2005: The Global Satellite Mapping of Precipitation (GSMaP) project. *2005 Int. Geoscience and Remote Sensing Symp.*, Seoul, South Korea, IEEE, 3414–3416, https://doi.org/10.1109/IGARSS.2005.1526575.

Prat, O. P., and B. R. Nelson, 2013: Precipitation contribution of tropical cyclones in the southeastern United States from 1998 to 2009 using TRMM satellite data. *J. Climate*, **26**, 1047–1062, https://doi.org/10.1175/JCLI-D-11-00736.1.

——, ——, S. E. Stevens, and D.-J. Seo, 2014: Long-term large-scale bias-adjusted precipitation estimates at high spatial and temporal resolution derived from the National Mosaic and Multi-sensor QPE (NMQ/Q2) precipitation reanalysis over CONUS. *2014 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract A11A-3007.

——, ——, ——, E. Nickl, D.-J. Seo, B. Kim, J. Zhang, and Y. Qi, 2015: Merging radar quantitative precipitation estimates (QPEs) from the high-resolution NEXRAD reanalysis over CONUS with rain-gauge observations. *2015 Fall Meeting*, San Francisco, CA, Amer. Geophys. Union, Abstract H24E-04.

Rafieeinasab, A., A. Norouzi, D.-J. Seo, and B. Nelson, 2015: Improving high-resolution quantitative precipitation estimation via fusion of multiple radar-based precipitation products. *J. Hydrol.*, **531**, 320–336, https://doi.org/10.1016/j.jhydrol.2015.04.066.

Schweppe, F. C., 1973: *Uncertain Dynamic Systems*. Prentice-Hall, 563 pp.

Seber, G. A. F., and C. J. Wild, 1989: *Nonlinear Regression*. John Wiley and Sons, 768 pp.

Seo, D.-J., 1996: Nonlinear estimation of spatial distribution of rainfall – An indicator cokriging approach. *Stoch. Hydrol. Hydraul.*, **10**, 127, https://doi.org/10.1007/BF01581763.

——, 1998a: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol.*, **208**, 37–52, https://doi.org/10.1016/S0022-1694(98)00141-3.

——, 1998b: Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. *J. Hydrol.*, **208**, 25–36, https://doi.org/10.1016/S0022-1694(98)00140-1.

——, 2013: Conditional bias-penalized kriging (CBPK). *Stochastic Environ. Res. Risk Assess.*, **27**, 43–58, https://doi.org/10.1007/s00477-012-0567-z.

——, and J. A. Smith, 1996: On the relationship between catchment scale and climatological variability of surface-runoff volume. *Water Resour. Res.*, **32**, 633–643, https://doi.org/10.1029/95WR03641.

——, 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, https://doi.org/10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2.

——, ——, and E. R. Johnson, 1999: Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol.*, **223**, 131–147, https://doi.org/10.1016/S0022-1694(99)00106-7.

——, S. Perica, E. Welles, and J. C. Schaake, 2000: Simulation of precipitation fields from probabilistic quantitative precipitation forecast. *J. Hydrol.*, **239**, 203–229, https://doi.org/10.1016/S0022-1694(00)00345-0.

——, A. Seed, and G. Delrieu, 2010: Radar and multisensor rainfall estimation for hydrologic applications. *Rainfall: State of the Science, Geophys. Monogr.*, Vol. 191, Amer. Geophys. Union, 79–104, https://doi.org/10.1029/2010GM000952.

——, R. Siddique, Y. Zhang, and D. Kim, 2014: Improving real-time estimation of heavy-to-extreme precipitation using rain gauge data via conditional bias-penalized optimal estimation. *J. Hydrol.*, **519**, 1824–1835, https://doi.org/10.1016/j.jhydrol.2014.09.055.

——, M. M. Saifuddin, and H. Lee, 2018: Conditional bias-penalized Kalman filter for improved estimation and prediction of extremes. *Stochastic Environ. Res. Risk Assess.*, **32**, 183–201, https://doi.org/10.1007/s00477-017-1442-8; Corrigendum, **32**, 3561, https://doi.org/10.1007/s00477-018-1626-x.

Shen, H., D.-J. Seo, and H. Lee, H., 2019: Adaptive conditional bias-penalized Kalman filter for improved estimation of extremes and its approximation for reduced computation. arXiv, https://arxiv.org/abs/1908.00482.

Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteor.*, **30**, 397–412, https://doi.org/10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2.

——, D.-J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035–2045, https://doi.org/10.1029/96WR00270.

Smith, T. M., P. A. Arkin, J. J. Bates, and G. J. Huffman, 2006: Estimating bias of satellite-based precipitation estimates. *J. Hydrometeor.*, **7**, 841–856, https://doi.org/10.1175/JHM524.1.

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, https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Tang, L., D.-J. Seo, M. Nabatian, J. Zhang, K. W. Howard, and D. Kitzmiller, 2019: Comparative evaluation of merging and local bias correction for radar-gauge QPE. *Ninth Conf. on Transition of Research to Operations*, Phoenix, AZ, Amer. Meteor. Soc., 11.2, https://ams.confex.com/ams/2019Annual/webprogram/Paper353793.html.

Turk, F. J., and S. D. Miller, 2005: Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.*, **43**, 1059–1069, https://doi.org/10.1109/TGRS.2004.841627.

Vanmarcke, E., 1983: *Random Fields: Analysis and Synthesis*. MIT Press, 382 pp.

Vasiloff, S. V., and Coauthors, 2007: Improving QPE and very short term QPF: An initiative for a community-wide integrated approach. *Bull. Amer. Meteor. Soc.*, **88**, 1899–1911, https://doi.org/10.1175/BAMS-88-12-1899.

Vicente, G. A., R. A. Scofield, and W. P. Menzel, 1998: The operational GOES infrared rainfall estimation technique. *Bull. Amer. Meteor. Soc.*, **79**, 1883–1898, https://doi.org/10.1175/1520-0477(1998)079<1883:TOGIRE>2.0.CO;2.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. International Geophysics Series, Vol. 100, Academic Press, 648 pp.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, https://doi.org/10.1175/2011BAMS-D-11-00047.1.

——, and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, https://doi.org/10.1175/BAMS-D-14-00174.1.