

**Improving satellite-based rainfall
accumulation estimates using spaceborne
surface soil moisture retrievals**

WADE T. CROW*

USDA ARS HRSL, BELTSVILLE, MD, USA

GEORGE J. HUFFMAN

SSAI/NASA GSFC, GREENBELT, MD, USA

RAJAT BINDLISH

SSAI/USDA ARS HRSL, BELTSVILLE, MD, USA

THOMAS J. JACKSON

USDA ARS HRSL, BELTSVILLE, MD, USA

* *Corresponding author address:* Wade Crow, USDA ARS Hydrology and Remote Sensing Laboratory,
Rm. 104, Blg. 007, BARC-W, Beltsville, MD 20705, USA.

E-mail: wade.crow@ars.usda.gov

ABSTRACT

Over land, remotely-sensed surface soil moisture and rainfall accumulation retrievals contain complementary information that can be exploited for the mutual benefit of both product types. Here a Kalman filtering-based tool is developed that utilizes a time series of spaceborne surface soil moisture retrievals to enhance short-term (2- to 10-day) satellite-based rainfall accumulation products. Using the National Center for Environmental Prediction's Climate Prediction Center rain gauge data as a validation source, and a soil moisture product derived from the Advanced Microwave Scanning Radiometer aboard the NASA Aqua satellite, the approach is evaluated over the contiguous United States. Results demonstrate that, for areas of low to moderate vegetation cover density, the procedure is capable of improving short-term (2- to 10-day) rainfall accumulation estimates extracted from a variety of satellite-based rainfall products. The approach is especially effective for correcting rainfall accumulation estimates derived without aid of ground-based rain gauge observations. Special emphasis is placed on demonstrating that the approach can be applied in continental areas lacking ground-based observations and/or long-term satellite data records.

1. Introduction

For the majority of global land areas, satellite-based rainfall estimates offer the only possible source of near real-time precipitation accumulation information for operational hydrologic applications. However, providing such information at required accuracy levels has proven difficult (Hossain et al., 2004; Hossain and Anagnostou, 2004). The expected (next decade) deployment of the Global Precipitation Mission (GPM) satellite constellation represents a critical advance in these efforts (Hou, 2006). In addition to GPM, a currently under-explored possibility for improving land rainfall retrieval lies in the development of efficient techniques for leveraging complementary spaceborne water cycle observations (Crow et al., 2006; McCabe et al., 2006). A promising class of such techniques for rainfall correction are based on interpreting variations in soil water storage realized upon the assimilation of satellite-based surface soil moisture retrievals into a water balance model (Pan et al., 2006). Despite the limited heritage of soil moisture remote sensing methods and products relative to their rainfall equivalents, surface soil moisture retrievals - unlike instantaneous rainfall rate measurements - possess memory reflecting antecedent rainfall amounts and can therefore be sampled at relatively low temporal frequencies (e.g. once every 1 to 3 days) and still provide useful rainfall accumulation information (Crow, 2003).

Recently, Crow (2007) demonstrated that spaceborne estimates of surface soil moisture can be processed (via their Kalman filter-based assimilation into a water balance model) to reveal useful information concerning the sign and magnitude of antecedent rainfall accumulation errors. Their approach is based on the assimilation of remotely-sensed surface soil moisture retrievals into a simple soil water balance model using a Kalman filter. The

model is assumed to be forced primarily by a satellite-based rainfall product. Net additions or subtractions of soil water suggested by the filter upon assimilation of a soil moisture retrieval (commonly referred to as “analysis increments”) contain useful information about recent rainfall errors. That is, the underestimation (or overestimation) of antecedent rainfall accumulations by a satellite-based rainfall product should result in the subsequent addition (or removal) of water by the filter upon assimilation of a post-event surface soil moisture retrieval. Correlation between rainfall errors and subsequent filter analysis increments implies that analysis increments realized during soil moisture data assimilation are adequately compensating water balance predictions for the impact of antecedent precipitation errors. Based on this reasoning, Crow and Zhan (2007) use the magnitude of the correlation between rainfall error and analysis increment correlations as a metric to evaluate the continental-scale performance of various remotely-sensed surface soil moisture products. A separate possibility examined here is that the presence of such correlation can also be exploited to dynamically filter errors existing in satellite-based rainfall accumulation estimates. That is, analysis increments could be combined with satellite-based rainfall retrievals in such a way that error in rainfall accumulation estimates is minimized. The successful application of variance-minimizing data assimilation techniques to this problem would ensure that the soil moisture-based correction of a rainfall time series would never increase the root-mean-square (RMS) error of rainfall estimates - even in areas where soil moisture retrievals are of poor quality. While past work (see e.g. Crow and Bolten (2007)) has demonstrated the potential for passively validating satellite-based rainfall products using soil moisture observations, the active correction of rainfall products with remotely-sensed soil moisture retrievals has not yet been attempted.

Here, we present an approach for correcting short-term (2- to 10-day) satellite-based rainfall accumulation products over land using analysis increments calculated during the sequential assimilation of a remotely-sensed soil moisture product into a simple water balance model. Surface soil moisture products are based on the application of the Jackson (1993) single-channel retrieval algorithm to H-polarized 10.6 GHz (X-band) brightness temperature (T_B) observations acquired from the Advanced Microwave Scanning Radiometer (AMSR-E) aboard the NASA Aqua satellite (Jackson et al., 2007). The correction scheme is applied to a number of satellite-based precipitation products, including rainfall accumulation products generated by the Tropical Rainfall Measurement Mission (TRMM) Precipitation Analysis (TMPA). As an initial validation exercise, the procedure is evaluated over the contiguous United States (CONUS) based on comparisons with the Climate Prediction Center (CPC) Unified rain gauge dataset (Higgins et al., 2002). However, the ultimate value of the approach is likely to be the greatest for continental areas possessing relatively limited ground-based observing capabilities. As required by such areas, the procedure is designed to operate on satellite-based data alone.

2. Kalman Filtering

Our approach is based on using a satellite-based rainfall product (P') to drive a spatially distributed, daily antecedent precipitation index (API) model

$$\text{API}_{i,j} = \gamma_i \text{API}_{i-1,j} + P'_{i,j} \quad (1)$$

where i and j are time and space indices (respectively) and γ in (1) is varied according to day-of-year (d) as

$$\gamma_i = \alpha + \beta \cos(2\pi d_i/365). \quad (2)$$

Following Crow and Zhan (2007), α and β here are held constant (in both space and time) at 0.85 and 0.10. P' in (1) represents the accumulation depth of rainfall on day i and, consequently, API is expressed in dimensions of water depth.

Remotely-sensed surface soil moisture estimates $\theta_{i,j}$ are used to update (1) via a Kalman filter

$$\text{API}_{i,j}^+ = \text{API}_{i,j}^- + K_{i,j}(\theta_{i,j} - \text{API}_{i,j}^-). \quad (3)$$

Here “-” and “+” denote API values before and after Kalman filter updating. Following Reichle and Koster (2005), daily $\theta_{i,j}$ estimates (in water depth dimensions) for a particular pixel are obtained by linearly rescaling a time series of raw surface soil moisture retrievals θ° (in volumetric soil moisture dimensions) such that retrievals long-term mean (μ) and standard deviation (σ) match those derived from a multi-year integration of API calculated for the same pixel

$$\theta_{i,j} = (\theta_{i,j}^\circ - \mu_j^\theta) \frac{\sigma_j^{\text{API}}}{\sigma_j^\theta} + \mu_j^{\text{API}}. \quad (4)$$

Implicit in this transformation is an assumption that API estimates and the satellite soil moisture retrievals possess the same vertical support within the soil column. While the purely linear API models lack an explicit representation of a soil layer depth, its inherent memory of past rainfall (and therefore its storage capacity) is specified via γ . The sensitivity of subsequent results to our particular parameterization of γ in (2) is examined in Section 5.e. Also, note that API mean and standard deviation statistics for (4) are sampled from a time

series generated using (1) and no Kalman filter updating. The required length of the data heritage required to obtain stable estimates of these statistics are also examined later (see Section 5.f).

The Kalman gain K in (3) is given by

$$K_{i,j} = T_{i,j}^- / (T_{i,j}^- + S_j), \quad (5)$$

where T is the error variance for API forecasts and S is the error variance for θ retrievals. At measurement times, T is updated via

$$T_{i,j}^+ = (1 - K_{i,j})T_{i,j}^-. \quad (6)$$

Between soil moisture retrievals, and the adjustment of API and T via (3) and (6), API is forecasted in time using observed P' and (1). In parallel, forecast error T is updated using

$$T_{i,j}^- = \gamma_i^2 T_{i-1,j}^+ + Q_j \quad (7)$$

where Q relates the forecast uncertainty added to an API estimate between time $i - 1$ and i .

The goal of the Kalman update in (3) is optimally updating API estimates given the availability of (potentially noisy) satellite soil moisture retrievals. For example, note how greater uncertainty in these retrievals (i.e. larger S) leads to smaller K and a reduction in the weighting applied to observations by (3). A critical aspect then is the estimation of the error parameter S in (5) (describing soil moisture retrieval error) and Q in (7) (determining uncertainty in API forecasts). Here S and Q are assumed to constant, scalar quantities which are calibrated on a pixel-by-pixel basis until a time series of filtering innovations

$$\nu_{i,j} = (\theta_{i,j} - \text{API}_{i,j}^-) / (T_{i,j}^- + S_j)^{0.5} \quad (8)$$

is obtained which is temporally uncorrelated and has a variance of one. In particular, the lack of temporal correlation in ν ensures that the filter is accurately partitioning total error between observation and modeling sources (Gelb, 1974). Note that such calibration requires no outside information other than time series variables already used in the filtering procedure (e.g. θ° and P'). Despite the potential for non-Gaussian rainfall errors, such an approach has been successfully applied to estimate modeling errors in a similar soil moisture data assimilation system (Crow and Bolten, 2007).

3. Rainfall Correction

Of particular importance for this analysis are updates to API dictated by (3)

$$\delta_{i,j} = \text{API}_{i,j}^+ - \text{API}_{i,j}^- = K_i(\theta_{i,j} - \text{API}_{i,j}^-). \quad (9)$$

If θ retrievals are (at least) minimally skilled and the Kalman filter properly parameterized, δ values obtained from (9) should correlate with antecedent errors in P' values used to force (1). In regions where ground-based rain gauge observations are dense enough to be considered a validation data product P , error in P' can be explicitly calculated as $P' - P$. To examine the relationship between δ and $P' - P$, Crow (2007) temporally aggregated both quantities within a series of non-overlapping windows of length m

$$[\delta]_{k,j} = \sum_{i=(km)+n}^{i=[(k+1)m]+n} \delta_{i,j} \quad (10)$$

$$[P' - P]_{k,j} = \sum_{i=km}^{i=(k+1)m} (P'_{i,j} - P_{i,j}) \quad (11)$$

and calculated the long-term correlation coefficient between $[\delta]$ and $[P' - P]$. The new index $k = 0, 1, 2, \dots$ counts non-overlapping m -day time periods and a lag of n days is introduced in (10) to account for the causal relationship between past rainfall and current soil moisture.

Using $m = 7$, $n = 1$ and an AMSR-E soil moisture product, Crow and Zhan (2007) found a statistically significant correlation between $[\delta]$ and $[P' - P]$ (at 95% confidence) for 84% of the CONUS land surface. This result suggests that $[P' - P]$ rainfall accumulation errors can be estimated, and systematically reduced, based on $[\delta]$ values realized during the sequential assimilation of remotely-sensed surface soil moisture into (1). However, because $[\delta]$ reflects the correction of all errors sources in API predictions (e.g. the poor parameterization of soil water loss) and not just those associated with rainfall, the linear relationship between δ and $[P' - P]$ is not generally one-to-one. In response, we propose to correct m -day remotely-sensed rainfall accumulations ($[P']$) using an additive correction of the form

$$[\widetilde{P}']_{k,j} = [P']_{k,j} + \lambda_j [\delta]_{k,j}. \quad (12)$$

and incorporating a time constant scaling factor λ . While consistent with the application of a Kalman filter, it is possible that an additive (as opposed to multiplicative) error model for precipitation accumulations has shortcomings at the space-times scales under study (1° lat/long, 2-10 days), so results obtained from (12) will be scrutinized for evidence concerning the appropriateness of this assumption. One immediate consequence of an additive correction is that (12) can yield negative $[\widetilde{P}']$ values. Here, all such estimates are replaced with zeros. In addition, correction via (12) is attempted only for m -day periods containing more than $m/2$ distinct soil moisture retrievals.

The eventual accuracy of the corrected rainfall product $[\widetilde{P}']$ in (12) is also dependent on

deriving a feasible estimation strategy for individually estimating the time-constant scaling factor λ for each pixel in which rainfall is corrected. Unfortunately, the optimal choice for λ demonstrates a theoretical dependence on a number of unknown factors. As noted above, the magnitude of λ is sensitive to the relative partitioning of total modeling error in (1) between external rainfall forcing and shortcomings in internal model structure. For instance, if error in (1) is dominated by the poor treatment of soil water loss, as opposed to error in P' , analysis increments will mostly reflect corrections due to highly inaccurate soil loss prediction and overestimate the volume of water required to compensate for (more modest) rainfall errors. This, in turn, implies that λ values less than one are required in (12) to obtain optimally accurate $[\widetilde{P}']$ estimates. Conversely, where rainfall uncertainty is a relatively greater source of overall model error, optimal values for λ will be higher as updates to API made via (3) will more strongly reflect the impact of rainfall uncertainty.

A further complication in applying (1) is the role of surface runoff and quick drainage of the near-surface (1-3 cm) soil moisture layer after the end of rainfall events, but prior to the subsequent acquisition of a surface soil moisture retrieval. Both processes reduce the volume of the rainfall error that manifests itself in the satellite-estimated surface soil moisture anomaly. Consequently, their impact necessitates an increase in the value of optimized λ in order to compensate for the undetected rainfall volume that either runs off or infiltrates beyond the shallow sensor measurement depth.

These factors suggest that the estimation of λ will pose a significant obstacle for the operational implementation of (12). Here, we pursue and evaluate two potential strategies for dealing with this issue. First, a naive “default” strategy of simply assigning λ to a fixed value of either 0.5 or 1 for all grid cells at all times. Second, a non-parametric “estimated

λ ” strategy of utilizing a second, independently-acquired satellite rainfall data set (P'') and tuning time-constant values of λ on a pixel-by-pixel basis until the root-mean-square (RMS) difference between $[\widetilde{P}']$ and $[P'']$ is minimized. If errors in P' and P'' are independent, λ values that minimize the mean-squared average of $[\widetilde{P}' - P'']$ will also minimize actual RMS error in $[\widetilde{P}']$ (calculated here relative to the benchmark CPC dataset). This “estimated λ ” approach can be evaluated based on comparisons with an “optimized λ ” strategy in which a time-constant λ values are explicitly tuned (on a pixel-by-pixel basis) to minimize the RMS difference between $[\widetilde{P}']$ and the gauge-based CPC accumulation product ($[P]$). This final approach is, of course, not feasible in a global setting as it requires long-term access to P values acquired from dense, ground-based rain gauge networks.

In summary, it is important to stress that not all differences between API predictions in (1) and re-scaled AMSR-E soil moisture retrievals obtained from (4) can be attributed to the impact of errors in P on API predictions. As discussed above, a portion of such differences arise from both AMSR-E soil moisture retrievals errors and non-rainfall sources of error in API predictions. In this context, our entire methodology can be viewed as a two-step filtering procedure designed to isolate these non-rainfall-based error sources before they can be misattributed to rainfall. The first step in this filtering process is the application of a Kalman filter in (9) to estimate API analysis increments δ . Here, the total difference between re-scaled AMSR-E retrievals (θ) and background API forecasts (API^-) is reduced by a fractional Kalman gain ($K < 1$) calculated in (5) based on consideration of the relative magnitude of soil moisture retrievals and API errors. As such, it quantifies the fraction of the difference between θ and API^- attributable to modeling, as opposed to AMSR-E retrieval, error. Subsequently, we apply a secondary correction via a temporally constant

λ parameter in (13) to extract, from a temporal aggregation of δ ($[\delta]$), the fraction of the total modeling error directly attributable to rainfall and discard the complementary portion due to non-rainfall sources of error in API forecasts. Both of these two filtering steps are critical before resulting water depths can be correctively applied to a satellite-based rainfall accumulation estimate.

4. Data and Approach

Remotely-sensed surface (1- to 3-cm) soil moisture retrievals θ° are obtained from application of the single-polarization Jackson (1993) algorithm to X-band AMSR-E T_B data (Jackson et al., 2007). Climatological normalized difference vegetation index (NDVI) composite products derived from the Advanced Very High Resolution Radiometer (AVHRR) and the vegetation water content (VWC)/NDVI regression relationship of Jackson et al. (1999) are used to estimate VWC. Surface soil moisture retrievals are acquired with a spatial resolution of about 40^2 km² and measurement frequency of 1 to 2 days at mid-latitudes. Screening is performed to mask areas with snow cover and/or experiencing active rainfall. After screening, retrievals obtained from both ascending and descending overpasses between 2002 July 1 and 2006 December 31 are combined and aggregated to form a (near) daily, 1° latitude/longitude product. Prior to their assimilation into (1), θ° retrievals are linearly rescaled (on a pixel-by-pixel basis) using (4) to form a new product (θ) with the same temporal mean and standard deviation as API products derived from (1).

During this same 2002 to 2006 time period, a number of different remotely-sensed rainfall data sets are used to estimate the daily rainfall accumulation P' in (1). The TRMM Multi-

satellite Precipitation Analysis (TMPA) is computed retrospectively as the Version 6 3B42 product. It combines multiple passive microwave estimates, microwave-calibrated infrared (IR) estimates, and monthly gauge data (Huffman et al. 2007). The real-time, combined passive microwave portion of the TMPA is computed experimentally as the 3B40RT product (Huffman et al. 2007). It differs from the Version 6 3B42 product in only using microwave data, being real-time, and having a heterogeneous computational record. In particular, the inventory of microwave data approximately doubled in February 2005. The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) employs a neural network to calibrate IR estimates with passive microwave data (Sorooshian et al. 2000). Finally, the Hydroestimator (HE) product is based on radar-calibrated IR estimates and uses numerical weather model data to adjust for moisture availability, height of the convective equilibrium level, and orographic influences (Scofield and Kuligowski 2003). Use of the HE product is delayed until after July 2003 due to temporal gaps in its coverage. Benchmark rainfall magnitudes P for 2002 to 2005 are obtained from the National Center for Environmental Prediction’s (NCEP) Climate Prediction Center (CPC) retrospective rain gauge analysis product within the contiguous United States (Higgins et al., 2002). For 2006 only, P is based on the real-time CPC product - derived from slightly fewer rain gauges (see <http://www.cpc.noaa.gov/products/precip/realtime/index.shtml>). Although individual rain gauges provide sampling that is quite different from that of satellite-based rainfall estimates, it is conventional to use interpolated analyses of “sufficiently dense” gauge networks as validation for satellite estimates (e.g., Ebert et al, 2007).

All rainfall accumulation products (TRMM 3B42, TRMM 3B40RT, PERSIANN, HE and NCEP CPC) are re-sampled to a daily, 1° latitude/longitude grid overlying the CONUS

area. For these products, we define daily accumulation as the total depth of rainfall occurring between 12 and 12 UTC. However, for the soil moisture product the same day is defined by a period shifted 12 hours into the future (0 to 24 UTC). Here, this 12-hour shift is assumed to effectively capture the necessary delay between rainfall and resulting soil moisture, and n in (10) is set to zero. Results will focus primarily on a choice of 3 days (i.e. $m = 3$) for the time scale of corrected accumulation products. The spatial domain of interest is the entire CONUS land area with a special focus on a lightly-vegetated Southern Great Plains (SGP) sub-domain between 33 to 40°N and 100 to 105°W that is generally considered to be well-suited for soil moisture remote sensing (Jackson et al., 1999). Future lower-frequency (1.4 GHz) satellite sensors (Kerr et al., 2001) should yield higher accuracy soil moisture products over moderately- to heavily-vegetated surfaces. Consequently, AMSR-E SGP results, derived using higher frequency X-band retrievals over a lightly vegetated area, are likely representative of future 1.4 GHz L-band spaceborne results for a geographic domain extending beyond the SGP to more densely vegetated regions.

5. Results

For the single lightly-vegetated 1° box centered at 35° N and 100° W, Figure 1 plots time series of a) daily API values derived from (1) using the TRMM 3B40RT rainfall product for P' , b) the raw AMSR-E surface soil moisture product (θ°), and c) the analysis increments (δ) realized when assimilating rescaled θ° into (1). For the same grid box over a shorter time period during summer 2004, Figure 2 shows the impact of using (12) and the δ time series in Figure 1c to correct 3-day TRMM 3B40RT accumulation products. Periods of time when the

TRMM 3B40RT product overestimates (or underestimates) benchmark CPC accumulation products are labeled. Note that, for each instance, the corrected products obtained from (12) are able to revise the TRMM product in the correct direction (i.e. remove accumulated rainfall in overestimated cases and add rainfall volume in underestimated cases). The source of this skill is the time series of δ values (Figure 1c) derived from θ° (Figure 1b) using (4) and (9).

a. Estimation of λ

Figure 2 plots both “optimized λ ” and “estimated λ ” rainfall correction results. As discussed in Section 3, optimized results are based on tuning a temporally constant λ factor on a pixel-by-pixel basis to minimize the RMS difference between corrected 3-day accumulation products $[\widetilde{P}']$ and 3-day benchmark CPC rainfall accumulations $[P]$. Such explicit tuning, however, is not possible for the application of (12) outside of limited areas with extensive ground-based observations. A globally-feasible alternative is to calibrate λ in (12) to minimize the RMS difference between $[\widetilde{P}']$ and a second, independently-acquired, satellite rainfall product $[P'']$ (in this case, the HE product). For the single 1° box examined in Figure 2, relatively little difference is observed between the two approaches. Over the entire CONUS area, Figure 3 explicitly tests the accuracy of this alternative approach by plotting time-constant estimated λ values - obtained by minimizing the RMS difference between the corrected 3-day TRMM 3B40RT rainfall product ($[\widetilde{P}']$) and a 3-day Hydroestimator (HE) rainfall product ($[P'']$) - against optimized λ results which explicitly minimize the RMS difference between $[\widetilde{P}']$ and $[P]$. Unless otherwise noted, all estimated λ values are based on HE data acquired

between July 2003 and December 2006.

Values of optimized λ plotted on the ordinate in Figure 3 typically fall between zero and one. The relative lack of optimized λ values greater than one suggests that, while $[\delta]$ and $[P' - P]$ tend to be positively correlated, the entire volume of $[\delta]$ should not necessarily be attributed directly to rainfall error. More importantly, the relatively strong ($R^2 = 0.56$) correlation in Figure 3 implies that optimized time-constant λ values can be effectively estimated without the aid of extensive ground-based rainfall observations. However, Figure 3 also demonstrates a tendency for estimated λ values to slightly under predict their optimized counterparts. One consequence of this bias is that negative estimated λ values are calculated for a small number of 1° pixels (representing about 5% of the total CONUS area). Since no physical basis exists for the positive correlation between $[\delta]$ and $[P' - P]$ (required by a negative value for λ), negative λ values are reset to zero. Due to the form of (12), this prevents the estimated λ approach from modifying rainfall in these pixels.

b. SGP correction of TRMM 3B40RT

For all 1° boxes in the SGP sub-domain, Figure 4 plots original and corrected 3-day TRMM 3B40RT accumulations against their benchmark CPC equivalents. Specifically, Figure 4a plots benchmark CPC 3-day accumulations versus original (uncorrected) TRMM 3B40RT accumulations. As suggested by earlier results in Figure 2, this relationship is enhanced through application of the optimized λ correction procedure (Figure 4b). Furthermore, an approximately equivalent correction is obtained when utilizing estimated (as opposed to optimized) λ values in (12) (Figure 4c). Relatively small differences between

Figures 4b and 4c suggest that the relationship in Figure 3 produces sufficiently accurate λ estimates to form the basis of a robust, and operationally feasible, correction procedure.

An alternative, less encouraging, explanation for the lack of difference between Figures 4b and 4c is that HE daily accumulations (used as the independent rainfall data source [P''] to define λ) are significantly more accurate (relative to TRMM 3B40RT) and can therefore be substituted for benchmark CPC rainfall estimates with little subsequent impact on λ estimates. However, if this effect was truly responsible for the small differences seen between Figures 4b and 4c, than our estimation approach should yield much poorer results when the roles of HE and TRMM 3B40RT are reversed (i.e. when TRMM 3B40RT data are used as [P''] to define λ and correct HE accumulations). In fact, HE correction results for this reverse case demonstrate only very small differences between optimized and estimated λ results. Specifically, relative to an R^2 of 0.48 for uncorrected HE 3-day accumulations, duplication of Figure 4 for this reverse case (not shown) leads to an R^2 of 0.64 for optimized λ versus 0.63 for estimated λ . The lack of difference between estimated and optimized results for both the original and reverse cases implies that the success of our estimation approach is not contingent on a significant difference in accuracy between the two satellite-based products (HE and 3B40RT) but rather on the mutual independence of their errors.

Results in Table 1 summarize the accuracy of original and corrected rainfall products (relative to the CPC benchmark) using a variety of accuracy metrics. Root-mean-square-error (RMSE) and square of the correlation coefficient (R^2) results are based on the comparison of corrected 3-day accumulation products with the benchmark CPC product. The false alarm ratio (FAR) is defined as the fraction of estimated events that are actually non-events in the CPC product. The probability of detection (POD) relates that fraction of all actual events

that are correctly estimated. Here, an event is defined as a 3-day rainfall accumulation that exceeds the 95th quantile for all 3-day CPC accumulations intervals in each 1° box. Since our selection of the 95th quantile is essentially arbitrary, results for other quantile choices will be discussed later (Section 5.e). Table 1 is also broken down according to both geographical domain (i.e. the entire CONUS area and the smaller SGP sub-domain) and the method utilized for obtaining λ (i.e. the “estimated”, “optimized” and “default” approaches introduced in Section 3 and discussed above).

Note that the correction procedure improves all five performance metrics relative to the uncorrected, original TRMM 3B40RT product within the SGP sub-domain (Table 1). As in Figure 4, relatively small differences exist between corrected results obtained via the optimized λ and estimated λ approaches. Both approaches also slightly outperform the $\lambda = 1$ default case. However, the fixed default choice of $\lambda = 0.5$ for all pixels comes close to matching results for the estimated λ correction. This lack of a significant difference emphasizes that, despite the need to quantify λ in a reasonable way, SGP accuracy improvements demonstrated in Table 1 are ultimately due to skill imbedded in temporally variable Kalman filter analysis increments (Figure 1c) and not the pixel-by-pixel tuning of a time-constant λ factor performed by both the estimated and optimized λ strategies. Improvement in all rainfall accumulation metrics is possible even for the simplistic treatment of λ as a temporally and spatially fixed variable.

Returning to the additive form for (12) and the occasional negative values of $[\widetilde{P}']$ that it allows, recall that such physically-unrealistic values are removed by automatically resetting them to zero (see Section 3). For the SGP-domain, removing this post-processing step and allowing negative $[\widetilde{P}']$ values degrades estimated λ results in Table 1, but only modestly (e.g.

RMSE increases from 9.15 to 9.80 mm and R^2 decreases from 0.53 to 0.49). In addition, because our approach is based on re-scaling raw θ° retrievals into a potentially biased API climatology prior to correction, it cannot correct for long-term accumulation bias. All positive impacts noted in Table 1 are therefore based on the correction of random and/or slowly-varying accumulation error components. Correction of long-term bias in rainfall products will require the implementation of other techniques (e.g., Smith et al. (2006)).

c. CONUS correction of TRMM 3B40RT

Results presented up to this point have been based on the application of our approach to a lightly vegetated area (the SGP) known to be well-suited for soil moisture remote sensing (Jackson et al., 1999). Relative to the SGP sub-domain, application of the estimated λ correction procedure to the entire CONUS area yields substantially smaller relative improvements in all accuracy metrics except FAR (Table 1). This reduction likely reflects the wider range of land surface conditions outside of the SGP sub-domain, some of which are not well-suited to X-band soil moisture remote sensing (Njoku et al., 2003). As with the SGP region, only small CONUS-wide differences are noted between corrected results acquired using optimized and estimated λ values in (12). In contrast, all four accuracy metrics for the $\lambda = 1$ default correction case are degraded relative to the original (uncorrected) TRMM 3B40RT product. The failure of this naive approach underscores the importance of estimating λ in an appropriate manner. However, as in the SGP, relatively good results are obtained for the better default choice of $\lambda = 0.5$ (Table 1).

Figure 5a shows CONUS-wide imagery of the relative RMSE improvement observed

between original and corrected TRMM 3B40RT accumulation products ($(\text{RMSE}_{\text{corrected}} - \text{RMSE}_{\text{original}})/\text{RMSE}_{\text{original}}$). Despite the use of estimated, rather than explicitly optimized, λ values in (12), relative RMSE improvement is observed over almost the entire domain, and, within a large area of the central United States, relative RMSE reductions greater than 0.30 are found. Increased RMSE in corrected rainfall is limited to a small number (~ 30) of 1° cells in heavily forested areas (e.g. New England). The mischaracterization of non-rainfall error sources in our approach would almost certainly lead to areas of increased rainfall accumulation RMSE. The relative absence of such degraded areas in Figure 5a demonstrates that our particular rainfall correction approach is adequately filtering out the corrupting effect of soil moisture retrieval uncertainty and non-rainfall sources of error in API estimates.

While the approach requires the specification of several parameters using retrospective data (e.g. Kalman filtering parameters in Q and S in (5) and (7) and the scaling factor λ in (12)), all results in Figure 5 are based on the approximation of these parameters using only remotely-sensed data products. At no point during the correction process is access to the benchmark CPC rain gauge data required (or assumed) to tune a particular parameter. In particular, λ can be adequately estimated from satellite-based precipitation products (Figure 4). Even a simple default case of $\lambda = 0.5$ leads to substantial correction (Table 1). Consequently, RMSE correction results displayed in Figure 5a can be considered representative of results obtainable in an operational setting lacking access to ground data.

However, improved RMSE results do not guarantee the enhancement of other rainfall accuracy metrics. Figure 5b plots CONUS results for the absolute difference in R^2 ($R_{\text{corrected}}^2 - R_{\text{original}}^2$) results calculated between $[\widetilde{P}']$ and $[P]$. Here clusters of slightly degraded (i.e. reduced) R^2 are found in areas of high vegetation cover (e.g. the Pacific Northwest, the

Ozarks and along the Appalachian Mountains). Nevertheless, outside of these clusters, areas of more substantial enhancement can be seen - particularly along a broad north-south swath of the Western United States. Enhanced POD is also observed throughout relatively lightly vegetated and non-mountainous areas of the Western United States (Figure 5c). Figure 5d shows that while FAR skill is generally enhanced throughout the entire CONUS domain, the pattern of improvement is somewhat erratic in that it does not conform to any known spatial variation in vegetation cover or climatological rainfall characteristics .

d. Correction of TRMM 3B42, HE and PERSIANN

To this point, all results have been based on the using TRMM 3B40RT product for P' in (1). Table 2 also summarizes correction results based on application of our estimated λ correction procedure to three other rainfall products (TRMM 3B42, HE and PERSIANN). Note that the TRMM 3B40RT product is used as the independent P'' data source during the correction of HE and PERSIANN accumulation products. All results in Table 2 are given in terms of absolute metric differences (corrected - original). Within the SGP sub-domain, all four products are enhanced with respect to all four accuracy metrics (i.e. reduced RMSE and FAR and increased R^2 and POD). Corrections made to the satellite-only TRMM 3B40RT, PERSIANN and HE products capture the potential range of improvement obtainable for satellite-base rainfall estimates acquired using a variety of algorithms and satellite-based observations. For instance, slightly smaller corrections made to the PERSIANN product (relative to TRMM 3B40RT) may be attributable to its incorporation of more frequent, although less precise, thermal infrared remote-sensing observations not considered in the

microwave-only TRMM 3B40RT product. A larger difference can be noted between results for the purely satellite-derived products and the gauge-corrected TRMM 3B42 product (Huffman et al., 2007). For all metrics except FAR, the retrospective gauge-based correction of the TRMM 3B42 appears to limit the added utility associated with our surface soil moisture-based correction procedure. This implies that the primary value of remotely-sensed soil moisture for rainfall correction will be in areas lacking adequate rain gauge coverage for use in either real-time or retrospective rainfall analyses. Finally, as in Table 1, enhancements to RMSE, R^2 and POD skill for all four products are reduced when expanding the domain of interest from the lightly-vegetated SGP sub-domain to the wider range of land coverage types found within the entire CONUS region.

e. Parameter sensitivity and correction robustness

Our procedure contains four separate parameters that could conceivably be modified to impact results: the choice of $\alpha = 0.85$ and $\beta = 0.10$ in (2), the use of a 3-day accumulation window ($m = 3$) in (10) and (11), and the choice of a 95th quantile (of 3-day rainfall accumulations) as the event threshold for POD and FAR calculations. The summary below examines potential sensitivity associated with these parameters for correction of the TRMM 3B40RT product within the SGP sub-domain.

Values of α and β used here (0.85 and 0.10) are based on default values used previously in Crow and Zhan (2007). Modest sensitivity is observed to variations in these parameters (not shown). The most significant trend is a tendency for obtaining slightly better correction results for lower values of α (e.g. between 0.75 and 0.80). In addition, correction results

degrade for choices of α greater than 0.90 and β values less than 0.05.

With regards to variations in the length of the accumulation window, Figure 6 plots original and corrected TRMM 3B40RT RMSE and R^2 results over the SGP sub-domain as a function of accumulation period length. The relatively small differences between results based on optimized and estimated λ in Figure 6 underscores the robustness of our estimation procedure for λ over a range of accumulation time scales. In addition, the relative magnitude of corrections for a 3-day accumulation period ($m = 3$) are generally representative of any accumulation length between 2 and 10 days. However, correction skill declines significantly at accumulation time scales finer than the 1-2 day retrieval frequency of the assimilated AMSR-E soil moisture product. At present, this reduces the effectiveness of our approach at daily and sub-daily times scales.

FAR correction results are also degraded when specifying lower event thresholds (not shown). In fact, SGP FAR results for corrected rainfall products slightly increase (relative to the original, uncorrected TRMM 3B40RT product) when defining an event threshold below the 85th quantile for 3-day CPC accumulations. Since λ is restricted to positive values, the additive form of (12) implies a positive $[\widetilde{P}']$ value for any positive filter analysis increment $[\delta]$. Positive $[\delta]$ are commonly associated with underestimated rainfall but can also arise during periods of no precipitation and excessively rapid dry-down dynamics in API predictions (Section 3). Despite our attempts to separate out the effect of such non-rainfall error sources, it is possible that the additive form of (12) may erroneously add small accumulation depths to P to compensate for what is, in reality, a model evaporative or drainage parameterization problem. This, in turn, would lead to the false detection of low-to moderate-intensity rainfall events. Difficulties with FAR correction are also seen in Figure

2, where (12) is typically able to only partially correct accumulation periods in which the TRMM 3B40RT product substantially overestimates rainfall accumulations. Consequently, the corrected accumulation time series remains prone to false-alarm errors.

f. Training period requirements

A potential obstacle to the operational implementation of this approach lies in the need to sample long-term mean and variance statistics in order to parameterize both the scaling relationship described in (4) and obtain RMS statistics on which to base our estimated λ correction procedure (see Section 3). Up to this point, both calculations have been performed assuming that data is available for the entire AMSR-E data period (July 2002 to December 2006). In reality, an operational implementation of this approach using a new soil moisture or rainfall product would need to acquire these statistics within a shorter time frame in order to commence real-time product generation.

Within the SGP sub-domain, Figure 7 examines this issue by plotting the R^2 performance metric for the estimated λ correction procedure as a function of training period length. Data available within these finite training periods are sampled to obtain both the rescaling statistics used in (4) and the RMS error statistics needed to estimate λ . Here, training period length refers to the number of days with good data that are sampled to obtain such statistics. As in Table 1, results are based on the correction of TRMM 3B40RT accumulations within the SGP sub-domain. However, to maximize the length of the data time series, the longer PERSIANN times series data was used for P'' in place of the HE product. Multiple grey lines on Figure 7 are generated by starting finite training periods at various times along the

entire July 2002 to December 2006 period. These starting points are spaced 100 days apart in order to sample across the seasonal cycle.

Results in Figure 7 demonstrate that improved accumulation estimates can generally be obtained after sampling for as little as 20 days, and results essentially consistent with the use of the entire period as a training period (dashed black line in Figure 7) are possible after a training period of about 200 days. These results suggest that it may prove possible to introduce a new data source with preliminary coefficients after one month of training and subsequently optimize them using about one year of data.

6. Conclusions

Remotely-sensed rainfall and surface soil moisture retrievals contain complementary information that can be exploited to enhance both types of hydrologic observations. Past work has demonstrated how the root-mean-square accuracy of daily rainfall accumulations can be estimated through the assimilation of surface soil moisture retrievals into an API model (Crow and Bolten, 2007). Here, we expand on this by moving past the passive evaluation of rainfall products to demonstrate how remotely-sensed soil moisture retrievals can be used to actively enhance the accuracy of short-term (2- to 10-day) rainfall estimates derived from satellites. Due to the availability of high-quality rain gauge datasets for validation purposes, this initial study is limited to the well-instrumented CONUS region. However, the procedure is explicitly designed to require only satellite-based inputs and can therefore be applied over continental areas lacking extensive ground instrumentation.

Results demonstrate that our additive correction model (12) can be adequately parame-

terized using two quasi-independent remotely-sensed rainfall datasets (Figure 2). Application of this approach leads to large-scale improvements in the accuracy of the TRMM 3B40RT rainfall product for a range of rainfall accuracy metrics (Table 1, Figures 4 and 5). While qualitatively similar results are found for other satellite-only products (HE and PERSIANN) (Table 2), reduced improvements noted for the TRMM 3B42 product (which incorporates a retrospective correction based on ground-based rain gauge observations) suggest that the highest utility for the procedure lies in enhancing rainfall measurements in continental regions lacking adequate ground-based rain radar or rain gauge coverage.

Because of its basis in Kalman filtering, the approach can be widely applied in regions where soil moisture retrievals suffer poor accuracy (due to the presence of dense vegetation) without reducing the RMS accuracy of the rainfall product (Figure 5a). However, improvements in other rainfall accuracy metrics (e.g. R^2 and POD) are generally limited to areas of light and moderate vegetation cover (Figures 5b and 5c). This limitation is consistent with the performance of current-generation soil moisture retrievals derived from AMSR-E X-band (10.7 GHz) T_B observations. However, the future availability of lower-frequency L-band (1.4 GHz) T_B observations from the European Space Agency (ESA) Soil Moisture and Ocean Salinity Mission (SMOS; Kerr et al., 2001) should significantly enhance results in areas of moderate and dense vegetation. Relative to X-band, L-band sensors will also reduce the impact of atmospheric hydrometeors on soil moisture retrievals and should therefore improve our ability to retrieve soil moisture during and immediately after storm events.

Despite such generally encouraging results, a close examination of our approach reveals limitations that may be attributable to our implicit choice of an additive error model for rainfall accumulations (via our decision to implement a Kalman filter and adopt an additive

form for (12)). In particular, FAR correction results at low event thresholds appear to pose a particular challenge for (12) (Section 5e). Error associated with an inappropriate rainfall error model may also underlie the erratic pattern of FAR correction results in Figure 5d. These problems suggest that it may prove advantageous to reformulate the approach using an Ensemble Kalman or particle filtering approach capable of better representing the (potentially) multiplicative structure of rainfall errors. Additionally, improved FAR results may also require explicitly considering saturation effects whereby land surface signals become insensitive to additional amounts of antecedent rainfall once runoff generation occurs. Finally, limitations in the temporal frequency and spatial resolution of satellite-based soil moisture retrievals may limit the time/space scales (i.e. 2-10 day and 1°) at which our procedure can successfully correct rainfall, and, consequently, its value for certain hydrologic applications. Further research is required to fully address these issues.

Finally, future study is required to clarify the potential benefit of more complex modeling and data assimilation approaches. In particular, the Constrained Ensemble Kalman filtering (CEnKF) concept proposed by Pan et al. (2006) provides an optimal framework for explicitly decomposing water balance analysis increments into runoff, rainfall, evapotranspiration and soil moisture storage error components. Consequently, it forms a natural framework for the correction of rainfall accumulation products via soil moisture remote retrievals, and may enhance the efficiency of our approach when a more complex land surface model is utilized. It should also be noted that the basic structure of our correction (in which near-past rainfall accumulations are modified using current soil moisture retrievals) is more suggestive of a smoothing re-analysis - rather than filtering - problem. Consequently, it may prove beneficial to reformulate the approach using a fixed-lag smoother to update rainfall accumulation

estimates (Dunne and Entekhabi, 2004). Future work in this direction could further enhance encouraging results noted here.

Acknowledgments.

This research was partially supported by NASA EOS grant NNH04AC301. Data support from Robert Kuligowski (NOAA NESDIS) and the NOAA Climate Prediction Center is gratefully acknowledged.

REFERENCES

- Crow, W.T., 2003: Correcting land surface model predictions for the impact of temporally sparse rainfall rate measurements using an Ensemble Kalman filter and surface brightness temperature observations. *J. Hydrometeor.*, **4**, 960–973.
- Crow, W.T., D. Entekhabi, R.D. Koster, and R.H. Reichle, 2006: Multiple spaceborne water cycle observations would aid modeling. *EOS*, **87**(15), 11 April 2006.
- Crow, W.T., 2007: A novel method for quantifying value in spaceborne soil moisture retrievals. *J. Hydrometeor.*, **8**(1), 56–67.
- Crow, W.T., and J.D. Bolten, 2007: Estimating precipitation errors using spaceborne surface soil moisture retrievals. *Geophys. Res. Lett.*, **34**, L08403, doi:10.1029/2007GL029450.
- Crow, W.T., and X. Zhan, 2007: Continental-scale evaluation of spaceborne soil moisture products. *IEEE Geosci. Rem. Sens. Lett.*, **4**(3), 451–455.
- Dunne, S., and D. Entekhabi, 2005: An ensemble-based reanalysis approach to land data assimilation. *Wat. Resour. Res.*, **41**, W02013, doi:10.1029/2004WR003449.
- Ebert, E.E., J.E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**(1), 47–64.
- Gelb, A., 1974: *Applied Optimal Estimation*. 374 pp., The MIT Press, Cambridge, MA.

- Higgins, R.W., W. Shi and E. Yarosh, 2000: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center ATLAS*, **7**, 40 pp.
- Hossain, F., E. N. Anagnostou, and T. Dinku, 2004: Sensitivity analysis of satellite rainfall retrieval and sampling error on flood prediction uncertainty. *IEEE Trans. Geosci. Rem. Sens.*, **42**(1), 130-139.
- Hossain, F., and E. N. Anagnostou, 2004: Assessment of current passive-microwave- and infrared-based satellite rainfall remote sensing for flood prediction. *J. Geophys. Res.*, **109**, D07102, doi:10.1029/2003JD003986.
- Hou, A., 2006: The Global Precipitation Mission (GPM): An overview. *Proceedings of the 2006 EUMETSTAT Meteorological Satellite Conference*, Helsinki, Finland 12-16, 2006.
- Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, and D.B. Wolff, 2007: The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. *J. Hydrometeor.*, **8**, 28–55.
- Jackson, T.J., 1993: Measuring surface soil moisture using passive microwave remote sensing. *Hydrol. Process.*, **7**, 139–152.
- Jackson, T.J., D.M. Le Vine, A.Y. Hsu, A. Oldak, P.J. Starks, C.T. Swift, J. Isham, and M. Haken, 1999: Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment. *IEEE Trans. Geosci. Rem. Sens.*, **37**, 2136–2151.

- Jackson, T.J., M.H. Cosh, R. Bindlish, and J. Du, 2007: Validation of AMSR-E soil moisture algorithms with ground based networks. *Proceedings of the International Geoscience and Remote Sensing Symposium 2007*, Barcelona, Spain, July 23-27, 2007, 1181–1184.
- Kerr, Y.H., P. Waldteufel, J.-P. Wigneron, J.-M. Martinuzzi, J. Font, and M. Berger, 2001: Soil moisture retrieval from space: the soil moisture and ocean salinity mission (SMOS). *IEEE Trans. Geosci. Rem. Sens.*, **39**(8), 1729–1735.
- McCabe, M., E.F. Wood, R. Wojcik, M. Pan, J. Sheffield, H. Gao and H. Su, 2007: Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies. In press, *Rem. Sens. Environ.*
- Njoku, E.G., T.J. Jackson, V. Lakshmi, T. Chan, and S.V. Nghiem, 2003: Soil moisture retrieval from AMSRE. *IEEE Trans. Geosci. Rem. Sens.*, **41**(2), 215–229.
- Pan, M., and E.F. Wood, 2007: Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *J. Hydrometeor.*, **7**, 534–547.
- Reichle, R.H. and R.D. Koster, 2005: Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophys. Res. Lett.*, **32**, L02404, doi:10.1029/2004GL021700.
- Scofield, R.A. and R.J. Kuligowski, 2003: Status and outlook of operational satellite precipitation algorithms for extreme precipitation events. *Mon. Wea. Rev.*, **124**, 2126–2151.
- Smith, M.S., P.A. Arkin, J.J. Bates and G.J. Huffman, 2006: Estimating bias of satellite-based precipitation estimates. *J. Hydrometeor.*, **7**, 841–856.

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.

List of Figures

1	For the 1° box centered at 35° N and 100° W, time series of a) API forced by the TRMM 3B40RT product, b) raw AMSR-E soil moisture retrievals (θ°) and c) analysis increments (δ) realized when assimilating rescaled θ° into (1) using (4) and (9).	34
2	Time series of original (TRMM 3B40RT), benchmark (CPC) and corrected 3-day TRMM rainfall accumulations for the 1° box centered at 35° N and 100° W. Overestimation (underestimation) labels refer to time periods in which TRMM 3B40RT rainfall accumulations are greater than (less than) the CPC benchmark.	35
3	For 3-day TRMM 3B40RT precipitation accumulations, comparison between optimized λ (derived by minimizing RMS differences with CPC rainfall) and estimated λ (derived by minimizing RMS differences with HE rainfall).	36
4	For all 1° boxes in the SGP sub-domain, scatterplots of 3-day CPC accumulations versus a) TRMM 3B40RT, b) optimized λ corrected TRMM 3B40RT and c) estimated λ corrected TRMM 3B40RT 3-days accumulations.	37
5	For 3-day TRMM 3B40RT precipitation accumulations, relative ((corrected - original)/original) changes in RMSE and absolute (corrected - original) changes in R^2 , POD and FAR realized upon implementation of the estimated λ correction procedure. Blue shading represents an improvement in accuracy and white areas inside the CONUS region denote 1° pixels where application of the procedure led to no change.	38

6	The impact of accumulation period length on R^2 and RMSE (calculated versus the benchmark CPC rain gauge product) for original and corrected TRMM 3B40RT accumulation products within the SGP sub-domain.	39
7	Impact of training period length on R^2 correction results within the SGP sub-domain. Multiple lines are generated by choosing various starting points for a training sample used to generate statistics required for (4) and the estimated λ correction procedure (see Section 3).	40

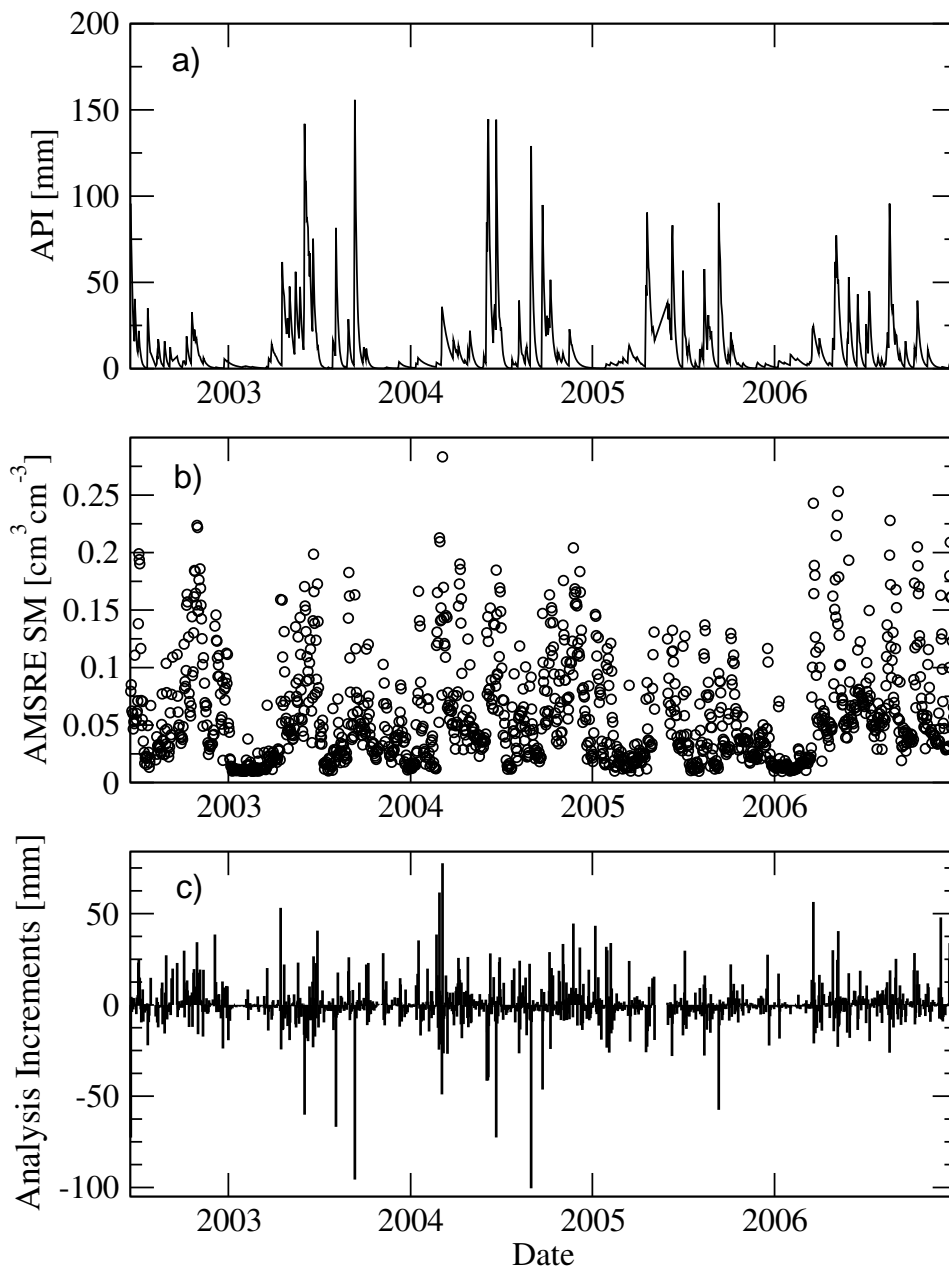


FIG. 1. For the 1° box centered at 35° N and 100° W, time series of a) API forced by the TRMM 3B40RT product, b) raw AMSR-E soil moisture retrievals (θ°) and c) analysis increments (δ) realized when assimilating rescaled θ° into (1) using (4) and (9).

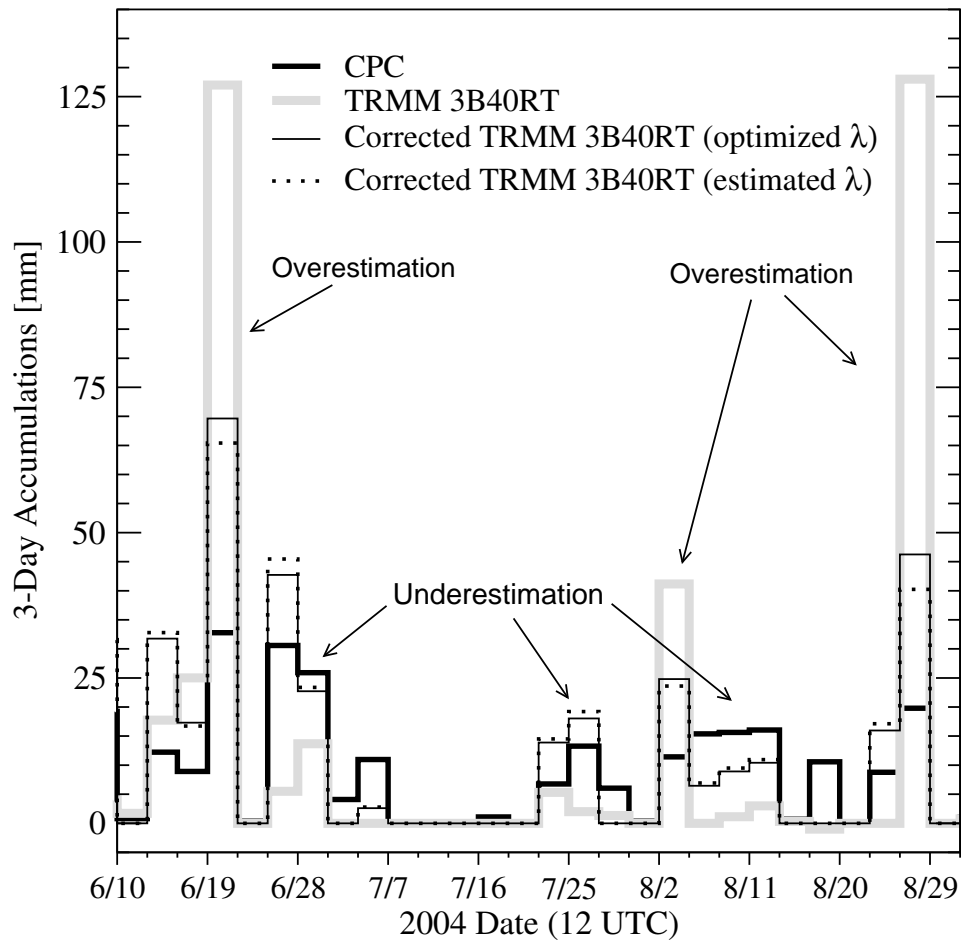


FIG. 2. Time series of original (TRMM 3B40RT), benchmark (CPC) and corrected 3-day TRMM rainfall accumulations for the 1° box centered at 35° N and 100° W. Overestimation (underestimation) labels refer to time periods in which TRMM 3B40RT rainfall accumulations are greater than (less than) the CPC benchmark.

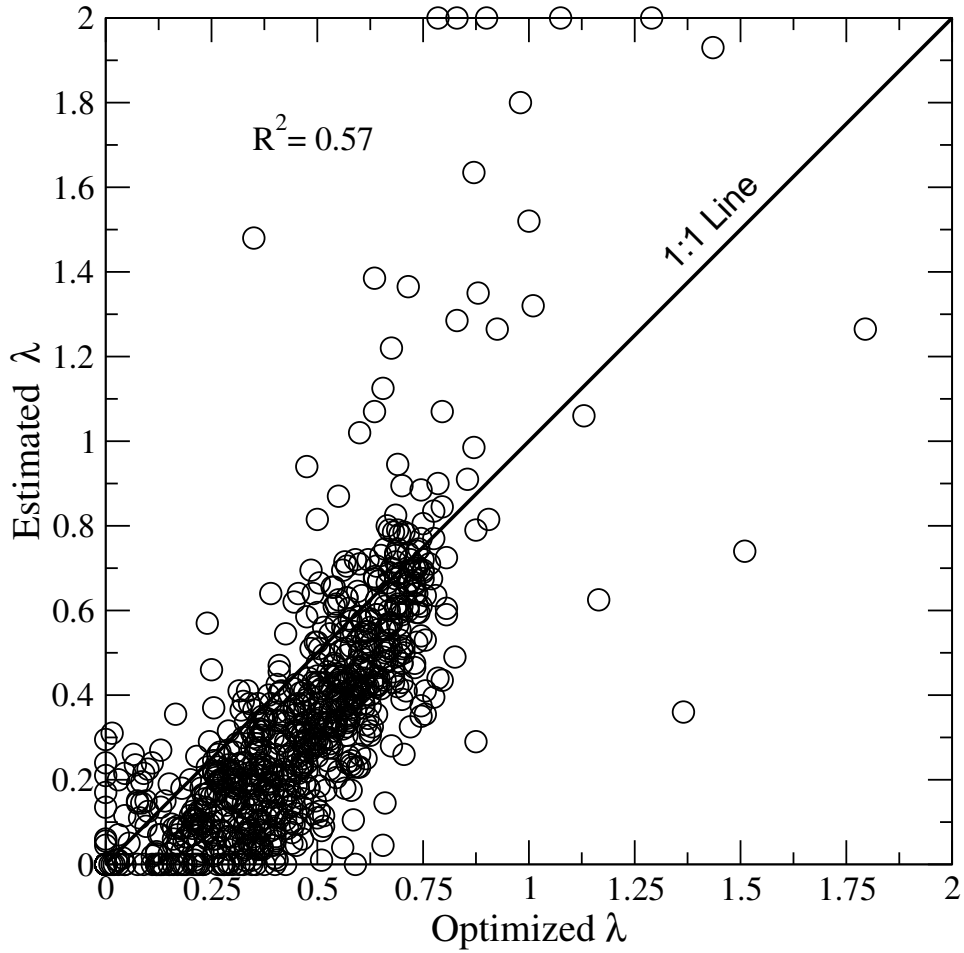


FIG. 3. For 3-day TRMM 3B40RT precipitation accumulations, comparison between optimized λ (derived by minimizing RMS differences with CPC rainfall) and estimated λ (derived by minimizing RMS differences with HE rainfall).

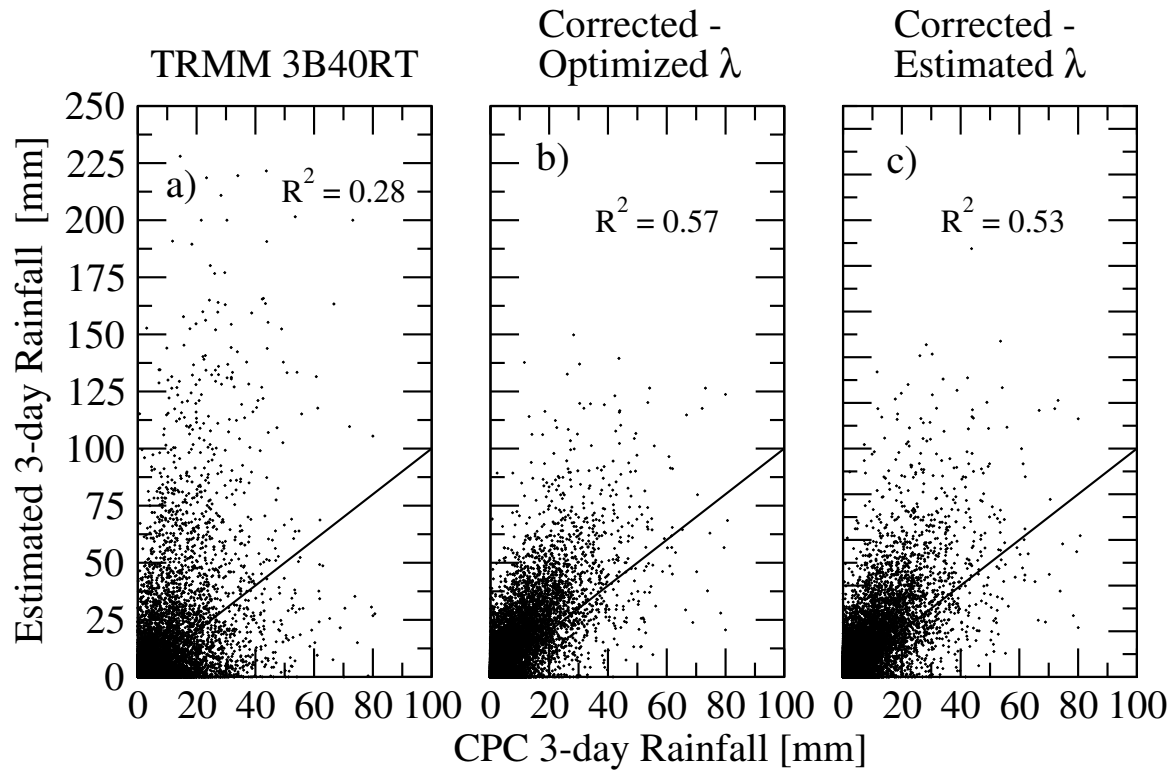


FIG. 4. For all 1° boxes in the SGP sub-domain, scatterplots of 3-day CPC accumulations versus a) TRMM 3B40RT, b) optimized λ corrected TRMM 3B40RT and c) estimated λ corrected TRMM 3B40RT 3-days accumulations.

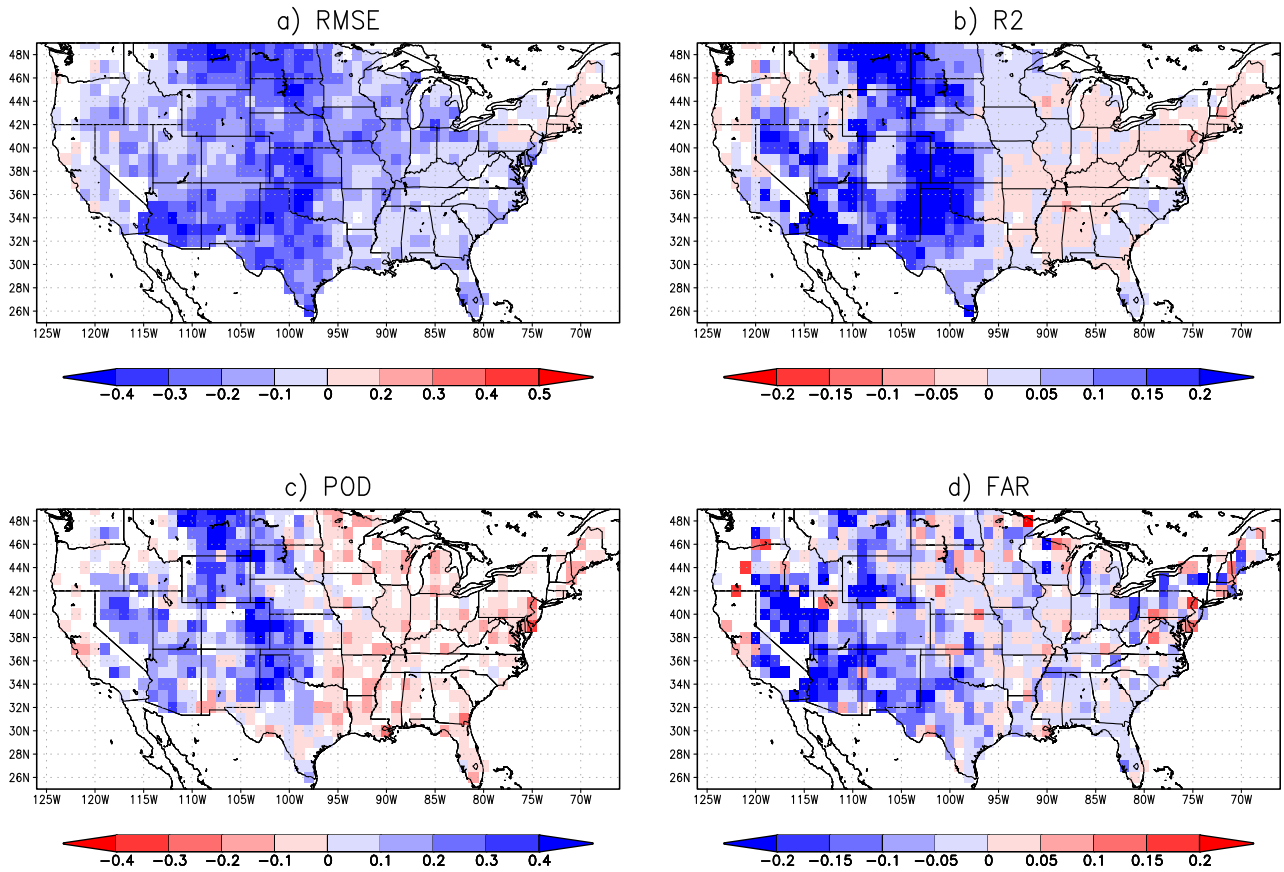


FIG. 5. For 3-day TRMM 3B40RT precipitation accumulations, relative ((corrected - original)/original) changes in RMSE and absolute (corrected - original) changes in R^2 , POD and FAR realized upon implementation of the estimated λ correction procedure. Blue shading represents an improvement in accuracy and white areas inside the CONUS region denote 1° pixels where application of the procedure led to no change.

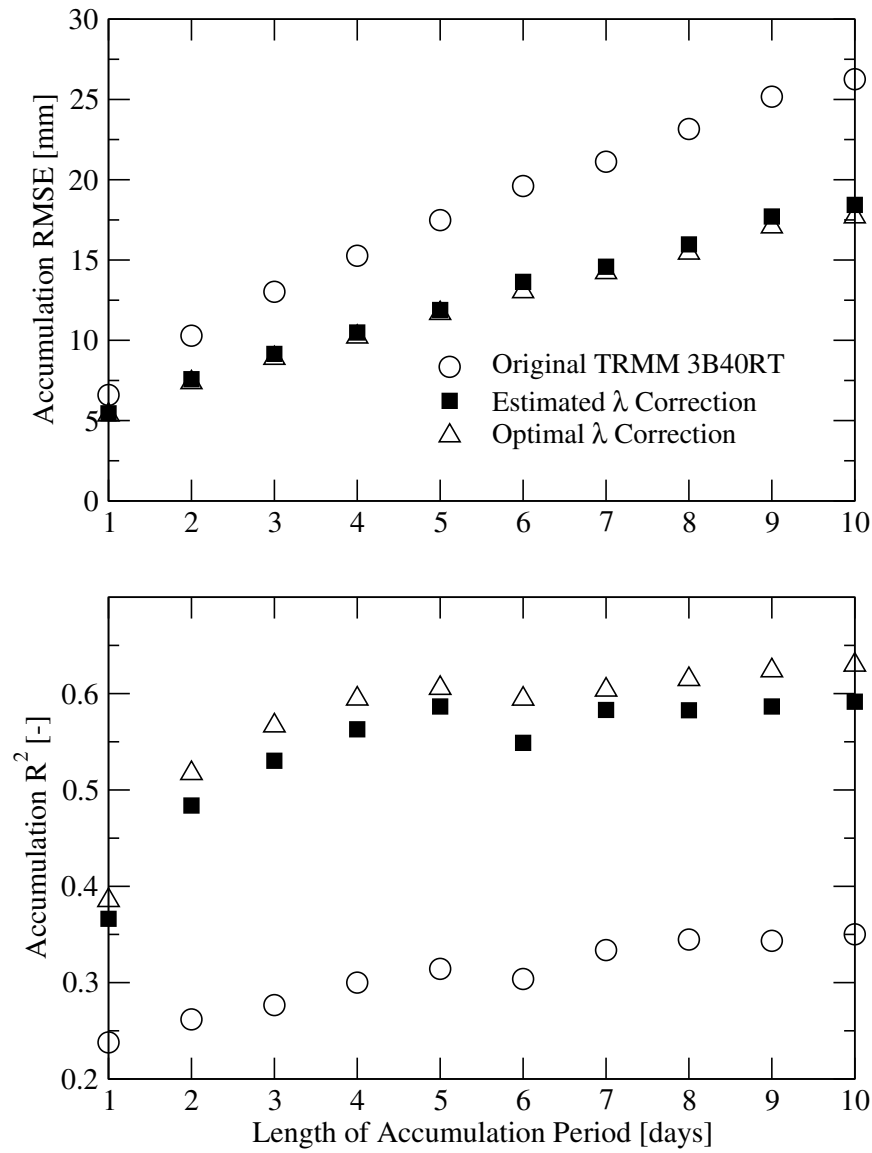


FIG. 6. The impact of accumulation period length on R^2 and RMSE (calculated versus the benchmark CPC rain gauge product) for original and corrected TRMM 3B40RT accumulation products within the SGP sub-domain.

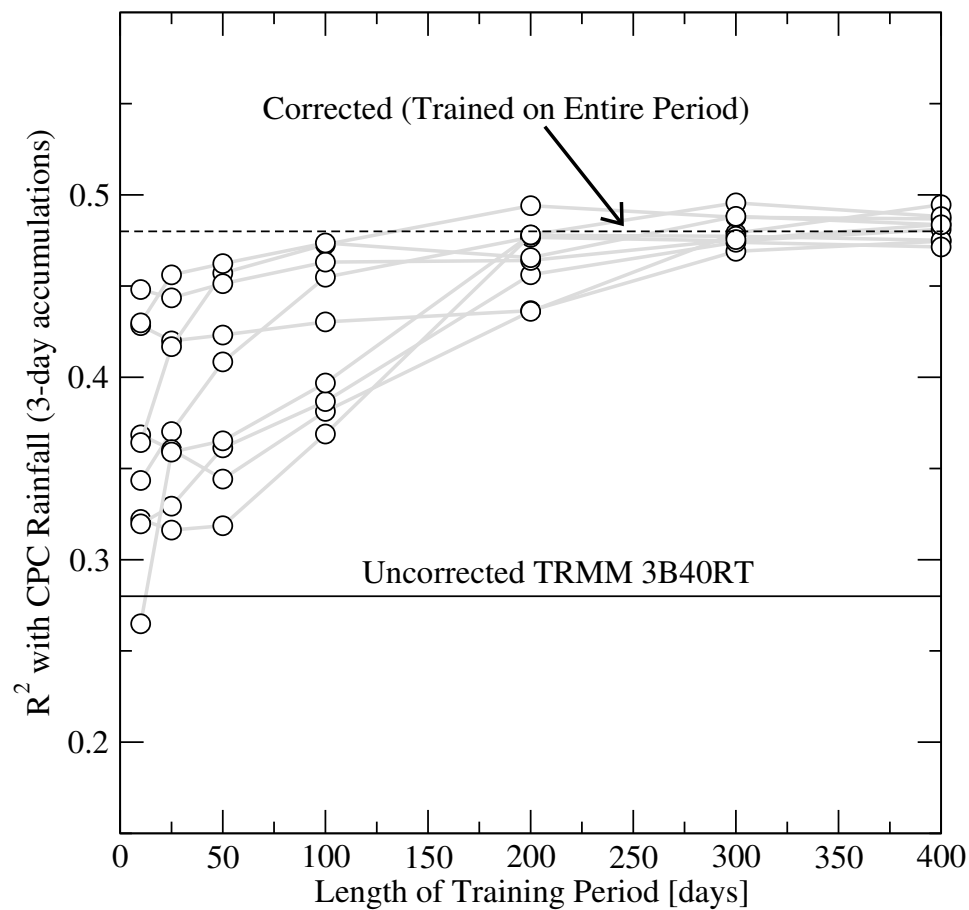


FIG. 7. Impact of training period length on R^2 correction results within the SGP sub-domain. Multiple lines are generated by choosing various starting points for a training sample used to generate statistics required for (4) and the estimated λ correction procedure (see Section 3).

List of Tables

- 1 Spatially-averaged root-mean square error (RMSE), R^2 , false alarm ratio (FAR) and probability of detection (POD) statistics for 3-day accumulation estimates derived from both the original and corrected versions of the TRMM 3B42RT dataset. FAR and POD statistics are based on exceeding the 95th quantile for 3-day CPC accumulations. 42
- 2 For estimated λ corrections derived from various precipitation products, spatially-averaged absolute change (corrected - original) in RMSE, R^2 , FAR, and POD for 3-day rainfall accumulation estimates. FAR and POD statistics are based on exceeding the 95th quantile for 3-day CPC accumulations. 43

TABLE 1. Spatially-averaged root-mean square error (RMSE), R^2 , false alarm ratio (FAR) and probability of detection (POD) statistics for 3-day accumulation estimates derived from both the original and corrected versions of the TRMM 3B42RT dataset. FAR and POD statistics are based on exceeding the 95th quantile for 3-day CPC accumulations.

Version	Domain	RMSE [mm]	R^2 [-]	FAR [-]	POD [-]
Original ($\lambda = 0$)	SGP	13.02	0.28	0.67	0.49
Estimated λ		9.15	0.53	0.60	0.77
Optimized λ		8.88	0.57	0.62	0.80
Default ($\lambda = 1$)		10.09	0.55	0.67	0.83
Default ($\lambda = 0.5$)		9.26	0.52	0.59	0.75
Original ($\lambda = 0$)	CONUS	11.78	0.30	0.62	0.37
Estimated λ		10.07	0.37	0.58	0.41
Optimized λ		9.66	0.37	0.57	0.41
Default ($\lambda = 1$)		13.07	0.19	0.72	0.32
Default ($\lambda = 0.5$)		10.17	0.32	0.57	0.37

TABLE 2. For estimated λ corrections derived from various precipitation products, spatially-averaged absolute change (corrected - original) in RMSE, R^2 , FAR, and POD for 3-day rainfall accumulation estimates. FAR and POD statistics are based on exceeding the 95th quantile for 3-day CPC accumulations.

Product	Domain	Δ RMSE	ΔR^2	Δ FAR	Δ POD
		[mm]	[-]	[-]	[-]
TRMM 3B40RT	SGP	-2.87	0.25	-0.06	0.28
HE		-0.94	0.16	-0.03	0.22
PERSIANN		-1.92	0.15	-0.07	0.19
TRMM 3B42		-0.60	0.07	-0.07	0.03
TRMM 3B40RT	CONUS	-1.71	0.07	-0.05	0.05
HE		-1.20	0.02	-0.03	0.02
PERSIANN		-1.07	0.03	-0.03	0.02
TRMM 3B42		-0.21	0.01	-0.02	0.00