This is an Accepted Manuscript of an article published in the Journal of Hydrologic Engineering in 2013 available online: http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000687

Estimating Watershed-Scale Precipitation by Combining Gauge and Radar Derived Observations

Mehmet B. Ercan [1] Jonathan L. Goodall [2]

ABSTRACT

Watershed modeling requires accurate estimates of precipitation, however in some cases it is necessary to simulate streamflow in a watershed for which there is no precipitation gauge records within close proximity to the watershed. For such cases, we propose an approach for estimating watershed-scale precipitation by combining (or fusing) gauge-based precipitation time series with radar-based precipitation time series in a way that seeks to match input precipitation for the watershed model with observed streamflow at the watershed outlet. We test the proposed data fusion approach through a case study where the Soil and Water Assessment Tool (SWAT) model is used to simulate streamflow for a portion of the Eno River Watershed located in Orange County, North Carolina. Results of this case study show that the proposed approach improved model accuracy (E = 0.60; R^2 = 0.74; PB = -10.2) when compared to a model driven by gauge data only (E = 0.50; R^2 = 0.54; PB = -25.5) or radar data only (E = 0.33; R^2 = 0.61; PB = -13.7). While this result is limited to a single watershed case study, it suggests that the proposed approach could be a useful tool for hydrologic engineers in need of retrospective precipitation estimates for watersheds that suffer from inadequate gauge coverage.

Keywords: Precipitation, Watershed Modeling, Data Fusion

---

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, College of Engineering and Computing, University of South Carolina, 300 Main Street, Columbia, SC 29208

[2] Assistant Professor, Department of Civil and Environmental Engineering, College of Engineering and Computing, University of South Carolina, 300 Main Street, Columbia, SC 29208

INTRODUCTION

There are many challenges associated with applying watershed models to water quantity and quality problems (Singh and Woolhiser, 2002). One of the most basic challenges is obtaining accurate input data for running the model, and one of the most important input datasets required to run a watershed model is precipitation (Biemans et al., 2009). Two common approaches for estimating precipitation for use in watershed modeling are (1) as observations made at gauging stations that typically use a tipping bucket instrument to capture rainfall intensity and (2) as estimates derived from radar which, in general, relate a reflectivity factor obtained from backscattered power of the echo returns to precipitation intensity (e.g., Hitschfeld and Bordan, 1954; Lakshmanan et al., 2007). When watershed models are used in engineering practice, it is sometimes the case that there is a lack of representative precipitation gauges for a watershed, and low confidence in the accuracy of radar-based precipitation estimates (Wilson and Brandes, 1979; Droegemeier et al., 2000; Young et al., 2000; Krajewski et al., 2010). The goal of this research is to build from well established idea that watershed-scale precipitation can be better estimated by combining gauge and radar-based precipitation estimates (Hildebrand et al., 1979; Smith and Krajewski, 1991; Legates, 2000) by testing an approach for fusing gauge and radar-based precipitation time series based on informational content within the watershed stream discharge record.

Our motivation for the proposed approach is that both gauge and radar-based methods for estimating precipitation have strengths and weaknesses. For example, there are gauge measurement errors associated with wind effects, wetting losses when emptying the collector, and evaporation and splashing during the storm events (Legates and DeLiberty, 1993; Groisman and Legates, 1995). Likewise, there are radar measurement errors including reflectivity errors from beam blockage, ground clutter, beam broadening with range, and bright-band contamination (Droegemeier et al., 2000). After taking these measurement errors into account, precipitation measurements at gauging locations tend to be more accurate than radar based estimates. This is because gauges directly measure precipitation using a tipping bucket or similar instrument, whereas radar-based systems indirectly estimate precipitation rates based on reflectivity of hydro-meteors (e.g. rain and hail). However, precipitation estimates from radar have the advantage that they capture the spatial variability and provide a more complete spatial coverage of the storm event.

In watershed management applications, available gauge data may be physically located miles away from the watershed being modeled, and this decreases the likelihood that the gauges capture the actual precipitation falling within the watershed. Therefore while gagued data may be preferred, it is not always available to support a watershed model. Even if nearby gagued data is available, there is evidence that different storm events (e.g., convective, frontal, etc.) may be best captured by different measuring approaches. Many studies have shown this to be the case, with some of the earliest studies being Huff (1970) who showed that warm season storm events in Illinois required additional rain gauges in order to capture spatial variability during such events, and Hildebrand et al. (1979) who showed that for low gauge densities, gauge-corrected radar precipitation estimates may be more accurate than gage-only measurements for convective storm events. More recently, Olivera et al. (2008) showed through an analysis of Areal Reduction Factors (ARF) in Texas that storms have different orientations during different seasons, and the authors attributed this finding to different mechanisms for generating precipitation (fronts vs convection) that are prevalent during different seasons.

Previous studies in radar based precipitation estimates have focused primarily on increasing the accuracy of radar generated precipitation estimates by using observed precipitation and streamflow data. Recent work has included an approach by Tuppad et al. (2010) that used a Soil and Water Assessment Tool (SWAT) model to adjust Next Generation Weather Radar (NEXRAD) Stage III data for the Smoky Hill River/Kanopoilis Lake watershed based on observed streamflow. The results show that NEXRAD Stage III data overestimated precipitation during warm months and underestimated precipitation during cold months compared to gauge estimated precipitation data. Smith and Krajewski (1991) developed a procedure to estimate the mean field bias of radar precipitation estimates based on precipitation gauge and NEXRAD data. They applied their method to an area in Norman, Oklahoma for a storm on May 27, 1987 that caused flood damage and found that the correlation between radar and gage data ranged from 0.71 to 0.96.

Legates (2000) similarly introduced a procedure to calibrate NEXRAD estimations in real-time using gauge precipitation observations and illustrated the procedure with a storm event on the Southern Great Plains. The study show that the NEXRAD precipitation estimates represented spatial distribution much better than spatially interpolated gauge precipitation. Jayakrishnan et al. (2004) compared the multi-sensor estimated hourly NEXRAD precipitation data with 545 rain gauges for a five year time period over the Texas-Gulf basin. The study showed that radar data generally underestimated the precipitation compare to the rain gauges, and the performance of radar data varied greatly both spatially and temporary. However, this study was conducted over the period 1995-1999 and a correction of the NEXRAD precipitation processing reported by Fulton et al. (2003) was incorporated in 2003, which is before our study period.

Our study builds on this prior work, but is different in that we assume the precipitation falling over a watershed during the course of a year will be better captured for some events by a gauge and for other events by a radar. This approach is justified by the uncertainty of both methods for measuring precipitation, particularly when precipitation gauges are not located within the watershed being modeled. Therefore, we are not performing a gauge-correction of the radar data, but instead proposing an algorithm for fusing the gauge and radar-based precipitation time series by selecting from one of the two sources for any given day to better capture watershed-scale precipitation. The radar data used in this study is, in fact, a radar product where the radar-based estimates are adjusted to match precipitation observations with gauge precipitation estimates because they considered to be the "ground truth" (Lawrence et al., 2003).

In this study we test an approach for combining gauge and radar-derived precipitation time series. The approach makes use of the idea that the true precipitation for a given day is reflected in the observed streamflow for that day. Therefore, the approach assumes that the time of concentration for the watershed is less than a day and there is higher confidence placed on the streamflow record than on the precipitation estimates. We test if combined dataset improves precipitation estimates by capturing storm events that may have been missed by one of the observational approaches, but captured by the other observation approach.

METHODS AND MATERIALS

Our general methodology was to apply the 2005 version of the Soil Water & Assessment Tool (SWAT) watershed model to simulate daily streamflow over a six year period for the Eno Watershed in North Carolina driven by three different precipitation input datasets: gauge, radar, and combined. The predicted streamflow obtained for each of these three precipitation cases was compared to observed streamflow at the watershed outlet in order to quantify the effectiveness of each precipitation dataset for estimating streamflow. The following subsections provide details of the methods and materials used in this study.

Study Area

The Eno Watershed is near the city of Hillsborough in Orange County, North Carolina (Figure 1) and has a drainage area of 171 km^2 with gently rolling topography and a mild, four-season climate. The Eno Watershed is a typical rural watershed that is large enough to take advantage of radar precipitation estimates, but small enough not to introduce significant computational challenges, particularly with model calibration.

Data Preparation

Terrain data for the Eno Watershed were obtained from the National Elevation Dataset (NED) (Figure 2). The NED provides Digital Elevation Models (DEMs) at three spatial scales: 1 arc second ($\approx$ 30 m), 1/3 arc second ($\approx$ 10 m), and 1/9 arc second ($\approx$ 3 m). The 1/9 arc second DEM provided an incomplete coverage of the entire watershed area, therefore the 1/3 arc second DEM was used for this study. According to the NED, the elevation within the Eno Watershed ranges from 149 m to 261 m with an average elevation of 200 m above sea level. The terrain slope in the watershed ranges from 0 to 153% with an average slope of 5.9%.

Land cover data for the Eno Watershed were obtained from the National Land Cover Dataset (NLCD) (Figure 2). The NLCD is available for three different years: 1992, 2001, and 2006. Each of these land use maps has a 30 m spatial resolution. We elected to use the NLCD 2006 land use map because it was nearest to the study period. This dataset shows that forest, pasture lands, and developed area (mostly open space) dominate the watershed covering 55.5%, 24.5%, and 11.6% of the watershed, respectively. Open water, scrub, grassland, and cultivated crops each cover about 2% of the watershed.

Soil data for the Eno Watershed were obtained from the State Soil Geographic (STATSGO) dataset. Although there is the higher resolution SSURGO data available for much of the United States, it was not available in a spatial data format for Orange County, NC at the time of the study. Soil types in the study area are named with Map Unit Identifier (MUID) as NC061, NC062, NC082 and NC083, and these types represent 67.2, 8.4, 20, and 4.4% of the watershed, respectively. These data indicate that the first 30 cm of soil consists of either silt loam or sandy loam over the watershed area, while deeper layers also contain clay. The NC061, NC062 and NC082 soil groups are hydrologic group B soils and the NC083 soil is a hydrologic group C soil.

Weather observations including temperature, wind speed, humidity, and precipitation were obtained from the National Climatic Data Center (NCDC). Figure 1 provides the location of the five nearest weather stations and shows that none of the stations are located within the watershed boundary. Again, the fact that none of the gauges are within the watershed boundary is one of the challenges addressed by this study. Based on these data, we found that the average daily maximum and minimum air temperature were 22.4 °C and 9.7 °C, respectively, while the humidity and wind speed were 62% and 1.60 m/s, respectively, over the period 2005-2010. Based on the gauged precipitation data, the daily average precipitation was 3.14 mm.

Radar-based precipitation estimates from the NEXRAD program were downloaded from the National Weather Service (NWS) website (`http://water.weather.gov/precip/download.php`). The radar daily precipitation data were available in a shapefile format as the National Hydrologic Rainfall Analysis Project (HRAP) grid cells (Figure 1). The spatial resolution of the dataset is 4 km and the data are available from 2005 to present. This radar precipitation data produced by the NEXRAD program uses the Multi-sensor Precipitation Estimator (MPE) (Lawrence et al., 2003). Based on radar data, the daily average precipitation was 2.69 mm, 0.45 mm less than the gauged data, over the period of analysis.

Finally, the USGS has maintained a streamflow gauge on the Eno River near Hillsborough, NC since October of 1972 and we obtained this streamflow time series for use in the modeling activities. Based on these records, daily average streamflow at the outlet of the watershed is 1.08 m^3/s over the period of study.

Model Setup

To create the SWAT model, we used the ArcSWAT extension to subdivide the Eno Watershed into 15 subbasins, which were then subdivided further into 130 Hydrologic Response Units (HRUs). ArcSWAT uses terrain processing tools in GIS such as flow direction and flow accumulation to determine subbasin areas from the Digital Elevation Model (DEM). Subbasin outlet points for the Eno Watershed were selected based on the size and heterogeneity of the land surface within the watershed. The resulting size of subbasins ranged from 8.02 to 15.8 km^2 with an average size of 11.4 km^2.

Each subbasin was further subdivided into HRUs that represent homogeneous areas within a subbasin in terms of the land use, soil type, and slope within that subbasin. While HRUs are not spatially defined within the SWAT model, they provide a means for capturing subbasin variability of soil type, terrain slope and land use types. The process of defining HRUs was done by defining thresholds for each land use, soil, and slope so that areas within those threshold values can be lumped into a single HRU within a given subbasin. The SWAT documentation recommends between 1 and 10 HRUs be used per subbasin, so to be within the recommended range, we set threshold values of 10% for land use and soil type, and a value of 20% for slope.

The resulting simplified land use map used in the model included only the dominated land use types within the watershed: deciduous forest, hay, evergreen forest, low density and medium density residential area. The resulting soil map was not altered by the threshold value, likely because we used the coarser STATSGO soil map. While we could have selected alternative thresholds for selecting HRUs, previous SWAT studies have shown that the number of HRUs does not have a significant effect on hydrologic predictions but could impact water quality predictions (e.g., Jha et al. 2004; Arabi et al. 2006; Migliaccio and Chaubey 2008). We therefore do not expect our particular HRU classification scheme to significantly impact the results of this study.

SWAT provides two methods for estimating surface runoff: the Natural Resource Conservation Service (NRCS) Curve Number (CN) method (Kenneth, 1972) and the Green & Ampt infiltration method (Green and Ampt, 1911). The NRCS CN method was chosen for this study because we judged it to be an acceptable approach for simulating a watershed of this size and type on a daily time step, and because most of large-scale models still use NRCS CN method (Arnold et al., 2010). SWAT also provides three methods to calculate potential evapotranspiration (PET): the Penman-Monteith method (Allen, 1986; Allen et al., 1989; Neitsch et al., 2005), the Priestley-Taylor method (Priestley and Taylor, 1972) and the Hargreaves method (Hargreaves and Samani, 1985). The Penman-Monteith method was chosen because it considers factors such as land cover and wind speed, which the other two methods ignore. SWAT allows for two channel routing approaches: Muskingum method or the variable storage method. The variable storage routing method was used in this model because we judged it to be an acceptable approach for the size and complexity of the watershed.

Model Simulations

The Eno watershed model was used to simulate daily averaged streamflow on a daily simulation time step using three different precipitation input datasets: combined, gauge, and radar-based estimates. The SWAT Weather Generator was used to spin-up the model during the period 2002-2004 in order to establish initial conditions such as antecedent soil moisture conditions. We calibrated the model individually for each of the three input precipitation datasets during the period 2005-2007 and then used the streamflow record from 2008-2010 to validate each model. All input datasets for the watershed model were held constant so that only the precipitation input and resulting calibration parameters were allowed to vary. Additional detail for the three model simulations, including a description of how the three input precipitation time series were created, follows.

Gauged Precipitation Case: The five nearest gauges to the watershed (shown in Figure 1), all within 18 km of the watershed boundary, were included in the analysis. Ordinary Kriging (OK) spatial interpolation was used to estimate subbasin precipitation from the gauged observations. We selected OK as the spatial interpolation method based on work reported by Goovaerts (2000) that concluded OK is more robust for precipitation estimation compared to other interpolation methods including Inverse Distance Weighting and Thiessen polygons.

Radar Precipitation Case: The precipitation data in 4 km radar grid cells were rescaled to subbasin averages using an Areal Weighting (AW) spatial interpolation. This method keeps the radar grid cells as areal averages and performs a weighted average of precipitation values for subbasins based on the proportion of the radar grid cells that intersect the subbasin area.

Combined Precipitation Case: The gauged and radar precipitation case time series were combined into a new time series using the following algorithm. The combined precipitation value for day i and subbasin j ($P_{c,i,j}$) was selected using the condition

$$P_{c,i,j} = \begin{cases} P_{g,i,j} & if\ |q_i - p_{g,i,j}| \leq |q_i - p_{r,i,j}| \\ P_{r,i,j} & else \end{cases} \tag{1}$$

where $P_{g,i,j}$ and $P_{r,i,j}$ are the gauge and radar precipitations, respectively, for day i and interpolated (using OK and AW methods, respectively) to subbasin j. The terms q_i, $p_{g,i,j}$, and $p_{r,i,j}$ represent a percent difference between an observed value on day i and an average term. These terms are calculated as

$$q_i = \frac{Q_{m,i} - \overline{Q_{m,i}}}{Q_{m,i}} \tag{2}$$

$$p_{g,i,j} = \frac{P_{g,i,j} - \overline{P_{g,i,j}}}{P_{g,i,j}} \tag{3}$$

$$p_{r,i,j} = \frac{P_{r,i,j} - \overline{P_{r,i,j}}}{P_{r,i,j}} \tag{4}$$

where $Q_{m,i}$ is the measured streamflow at the outlet of the watershed for day i. The terms $\overline{Q_{m,i}}$, $\overline{P_{g,i,j}}$, and $\overline{P_{r,i,j}}$ are average terms that take into account three time windows around the the observation recorded on day i: the average of all observations taken within the same month and year, observations taken within the same year, and all observations within the study period. These three time window averages are then averaged themselves as

$$\overline{Q_{m,i}} = (1/3)\left(\left(\overline{Q_m}\right)_{\text{Year-Month}(i)} + \left(\overline{Q_m}\right)_{\text{Year}(i)} + \overline{Q_m}\right) \tag{5}$$

$$\overline{P_{g,i,j}} = (1/3)\left(\left(\overline{P_{g,j}}\right)_{\text{Year-Month}(i)} + \left(\overline{P_{g,j}}\right)_{\text{Year}(i)} + \overline{P_{g,j}}\right) \tag{6}$$

$$\overline{P_{r,i,j}} = (1/3)\left(\left(\overline{P_{r,j}}\right)_{\text{Year-Month}(i)} + \left(\overline{P_{r,j}}\right)_{\text{Year}(i)} + \overline{P_{r,j}}\right) \tag{7}$$

where the terms Year-Month(i) and Year(i) represent the month of the year and the year for day i, respectively. Using Equation 1, the combined precipitation time series for subbasin j is calculated as $P_{c,j} = [P_{c,i=1,j}, P_{c,i=2,j}, ..., P_{c,i=n,j}]$ where n is the total number of values in the time series. This procedure is then repeated for all subbasins in the watershed $(j = 1, 2, ...m)$

where m is the total number of subbasins.

The approach can be explained by the following example. Suppose that the precipitation is observed using radar on March 14, 2005 and interpolated to one of the watershed subbasins. This observation would be represented by $P_{r,i,j}$ in our nomenclature where r stands for radar, i is March 14, 2005, and j is the subbasin identifier. First the term $\overline{P_{r,i,j}}$ would be calculated using Equation 7. In Equation 7, the term $\left(\overline{P_{r,j}}\right)_{\text{Year-Month}(i)}$ would be the average of all precipitation observations taken by radar and interpolated to subbasin j during the month of March, 2005; the term $\left(\overline{P_{r,j}}\right)_{\text{Year}(i)}$ would be the average of all precipitation observations taken by radar and interpolated to subbasin j during the year 2005; lastly, the term $\overline{P_{r,j}}$ would be the average of all precipitation observations taken by radar and interpolated to subbasin j over the period of the study. Next Equation 4 would be used to quantify a percent difference between the observation on March 14, 2005 and the average term $(\overline{P_{r,i,j}})$ that takes into account monthly, annual, and long term averages. These calculations are repeated for the gauge precipitation observations and the streamflow observations. Finally Equation 1 is used to select either the gauge or the radar precipitation observation for March 14, 2005 based on whether the gauge or radar percent difference term is closer to the streamflow ranking for that day.

After completing this analysis for the Eno Watershed, the resulting combined precipitation dataset for all subbasins and all time steps (P_c) included 32,865 values. Of these values, 10,190 (or 31%) came from the gauge precipitation time series (P_g) and 5,513 (or 17%) came from the radar precipitation time series (P_r) and the remaining gauge and radar precipitation values are equal. Although the algorithm took most of the precipitation values from the gauge estimates, the resulting time series, when averaged over all time steps for each subbasin, is closer to radar precipitation estimates (Figure 3). Figure 3 also shows that radar precipitation estimates tend to be lower than gauge precipitation estimates, as we saw with the daily averaged calculations reported earlier in this paper.

It should be noted that this method does not account for solid precipitation (e.g. snow, hail, sleet) because it was not required for this particular study watershed, which has a mild climate. Examination of the precipitation record showed only 6 recordings of snow not melting on the same day that it fell across all 5 gauges in the study region, which suggests that solid precipitation is not significant in this watershed. We believe that incorporating solid precipitation into this method is possible, however, by accounting for snowmelt processes and therefore lags between precipitation and streamflow.

Model Calibration

Each model scenario (gauge, radar and combined precipitation) was separately calibrated using two algorithms: the Shuffled Complex Evolution algorithm (SEA-UA) (Sahu and Gu, 2009) and the Dynamically Dimensioned Search (DDS) calibration method (Tolson and Shoemaker, 2007). The SEA-UA algorithm is capable of efficiently and effectively identifying the optimal values for the model parameters (Duan et al., 1992) and has been successfully applied for estimating SWAT model parameters (Eckhardt and Arnold 2001; van Griensven et al. 2002). The Dynamically Dimensioned Search (DDS) calibration method (Tolson and Shoemaker, 2007) was also used to confirm the calibrated parameter values. The parameters used in the calibration were the initial NRCS runoff curve number for moisture condition II (Cn2), the soil evaporation compensation factor (Esco), the available water capacity of the soil layers (SolAwc) and the surface runoff lag coefficient (Surlag). These parameters are the most commonly used calibration parameters for SWAT modeling applications (Arnold et al. 2010; Gassman et al. 2007).

Model Evaluation

There are a variety of approaches for quantifying the effectiveness of a watershed model. A widely used approach (McCuen et al., 2006) commonly used in SWAT applications (Gassman et al., 2007) is the Nash-Sutcliffe Coefficient (E) (Nash and Sutcliffe, 1970). According to Nash and Sutcliffe (1970), model efficiency can be calculated as

$$E = 1 - \frac{\sum_{i=1}^{n}(Q_{m,i} - Q_{p,i})^2}{\sum_{i=1}^{n}(Q_{m,i} - \overline{Q_m})^2} \tag{8}$$

where $Q_{m,i}$ is the measured streamflow at the outlet of the watershed for day i, $Q_{p,i}$ is the predicted streamflow at the outlet of the watershed for day i, and $\overline{Q_m}$ is the average of the measured streamflows. E values range from negative infinity to unity and, as the value of E approaches unity, the model efficiency increases such that when $E = 1$, the predicted streamflow perfectly matches the measured streamflow.

A second approach is to use a coefficient of determination (R^2), which measures the amount of variation of the simulated streamflow that is explained by variation in the observed streamflow (Santhi et al., 2006). The coefficient of determination (R^2) is calculated as

$$R^2 = \left(\frac{n \sum Q_p Q_m - (\sum Q_p)(\sum Q_m)}{\sqrt{n \sum Q_p^2 - (\sum Q_p)^2} \sqrt{n \sum Q_m^2 - (\sum Q_m)^2}} \right)^2 \tag{9}$$

where n is the number of days and the summations are over all observations in the time series. R^2 values range from zero to unity and, as the value of R^2 approaches unity, the model is able to explain more of the variability present within the observed streamflow dataset.

A third approach is the Percent Bias (PB) calculated as

$$PB = \frac{\overline{Q_m} - \overline{Q_p}}{\overline{Q_m}} \tag{10}$$

where $\overline{Q_p}$ is the predicted average streamflow and $\overline{Q_m}$ is the measured average streamflow, as indicated before. As the value of PB approaches zero, the model becomes less biased in terms of either over or under predicting streamflow. A negative PB value indicates that the predicted streamflow overestimates the measured streamflow, while a positive PB value indicates that the predicted streamflow underestimates the measured streamflow. We used each of these statistics as means for evaluating the model results driven by the different precipitation input datasets.

RESULTS AND DISCUSSION

Model Calibration and Evaluation Results

The changes in model parameter values resulting from calibration are given in Table 1. These parameters were assigned initial values based on input terrain, soil, and land use datasets and the two calibration routines described in the Model Calibration section were used to identify optimal model parameters within an acceptable range of values in order to best match observed streamflow. Changes in the Cn2 and SolAwc parameters from their original estimates are expressed in absolute percent differences, while changes in the Esco and Surlag parameters are expressed in absolute values. The Range column in Table 1 indicates the constraints placed on the parameters during the calibration process. These constraints limit the resulting parameter values to a range that is physically meaningful.

The calibration process resulted in an increase in the Cn2 parameter from its initial value for all three models. The Cn2 parameter controls the partitioning of precipitation between runoff and infiltration, therefore an increase in this parameter results in an increase in runoff. The higher Cn2 value for the radar case may be due to the fact that radar precipitation is generally lower than gauge precipitation estimates. The Esco parameter, which controls soil water evaporation, was set to a low value for all the three cases. The SolAwc parameter, which controls the available water capacity in the soil for use by plants, showed the greatest difference between radar and the other two models. Calibration of the combined case resulted in a slightly lower SolAwc value compare to gauge case. We investigated if the evaporation estimates resulted in unrealistic values by comparing the estimates with those derived from remote sensing imagery (Mu et al., 2007, 2011), but concluded that the evaporation estimates from all models were within a reasonable range. Lastly, calibration of the Surlag parameter, which controls storage within the watershed, resulted in nearly identical values for all three models.

Statistical summaries of the streamflow predictions compared to observations using the approaches described in the Model Evaluation section (Table 2) provide a quantitative means for judging the accuracy of the models. The statistics between daily observed and simulated streamflow for the three models are shown in Table 2 for the calibration and evaluation periods. During the calibration period, the combined case produced the highest E and R^2 values, suggesting that this model was best able to predict streamflow. The combined method, however, had only the second best PB value during the calibration period. The PB suggested that the model tended to underpredict observed streamflow on average during this time period. During the evaluation period, the combined method performed the best as judged by the all three statistics.

Based on evaluations of watershed models presented in Moriasi et al. (2007), the combined model would be classified as "good" during the calibration period and "satisfactory" during the evaluation period. However, this classification scheme was designed for E values based on a watershed model calibrated to estimate monthly streamflows. Our E values are based on daily predictions with a model calibrated for daily streamflow. Past SWAT studies show that estimated daily statistics are lower than monthly statistics (Gassman et al. 2007), therefore the model classification would likely improve if we used monthly E values generated by a monthly rather than a daily calibration.

We tested if differences in the predicted streamflow between the combined case and the gauge and radar cases were statistically significant using a two-tailed t-test. We found that differences between gauge vs. combined and radar vs. combined were significant with a 95% confidence interval for the calibration period. For the evaluation period, the differences between gauge vs. combined were significant, but the differences between combined vs. radar were not significant at a 95% confidence interval. Because the majority of values in the combined dataset came from the gauge time series, this result may be the result of the combined method selecting radar precipitation for larger streamflow events during the evaluation period.

We tested the impact of wet and dry periods on the model calibration by reversing the calibration and evaluation periods (calibrating on 2008-2010 and validating on 2005-2007). We did this because the period 2005-2007 was drier than the period 2008-2010 and we wanted to test if the combined method still performed best when calibrated over a wet instead of a dry period. The results of this analysis were that the combined method still performed the best as judged by the evaluation statistics. In fact, there were no significant changes in the evaluation statistics across all three models after making this change. This result supports our finding that the combined method performs best at estimating precipitation, even if the model is calibrated over a wet rather than a dry period.

We also tested the sensitivity of our findings to the particular model parameters chosen for the model runs. This was done by using constant parameter sets across all three model simulations. We did this using two different parameter sets: (1) the uncalibrated parameter set and (2) the calibrated parameter set obtained from the model driven by only gauge observed precipitation values. Both cases resulted again in the same ordering of goodness-of-fit for the three model scenarios. This test suggests that study findings are independent of the particular model parameters chosen for the model scenarios.

Streamflow Predictions

Comparing daily observed and modeled streamflow for two different years, one during the model calibration period (Figure 4) and the other during the model evaluation period (Figure 5), provides a visual means for judging model accuracy. Figure 4 shows that the model was generally able to reproduce observed streamflow for all three input precipitation datasets during the calibration period. Notable difference between the gauge and radar predictions include summertime precipitation events that produced streamflow and were observed by the radar but not by the gauge. These events suggest that radar may be better able to capture summertime convective storms. During fall months, the gauge case seemed to overestimate streamflow peaks when compared to the radar case. During winter months, the radar case produced streamflow estimates that matched well with observed streamflow but the gauge case overestimated the observed baseflow conditions. Finaly, during spring months, the gauge case matched well with observed streamflow but the radar case underestimated the observed baseflow conditions. These results suggest seasonal trends in the accuracy of the gauge and radar-based precipitation estimates.

The combined case improved on both the gauge and radar cases by selecting the optimal precipitation from these two datasets to best match observed streamflow. For example, the summertime storms observed by the radar but not by the precipitation gauges were correctly identified and incorporated into the combined dataset. In some cases (e.g., the storm in February, 2007) the combined case resulted in streamflow estimates that improved on both the gauge and radar based estimates. Also, during winter and spring months, the combined case resulted in streamflow estimates both for peaks and for baseflow conditions that generally improved on estimates derived from the gauge or radar datasets alone. During the evaluation period (Figure 5), similar seasonal characteristics were observed, despite the overall assessment that model predictions were, as expected, less accurate during the evaluation period compared to the calibration period.

A scatter plot of the predicted vs. observed streamflow values is another means for judging model accuracy (Figure 6). Focusing on the evaluation time period, low flow events (less than 5 mm) appear to be best modeled by the gauge or combined precipitation estimates, while the radar case often overestimated streamflow for low flow events. For medium flows (5 mm - 20 mm), it is clear that the gauge case most often underpredicted streamflow while the radar and combined cases had greater error, but less bias. For the two high flow events (greater than 20 mm), the gauge case significantly underpredicted streamflow while the combined case did the best at estimating these high flow events.

Monthly accumulation of the daily average streamflows reveal interesting characteristics of the different precipitation cases over both the calibration and evaluation time periods (Figure 7). The combined case performed well overall, despite poor performance in certain years. During the calibration period in late 2005 and early 2006, for example, the combined method was less accurate than the gauge case and more closely matched the radar case, which underpredicted observed streamflow. However, in late 2006 the combined case performed best at matching observed streamflow, while the gauge case overpredicted streamflow and the radar case underpredicted streamflow. During the evaluation period in late 2009 and early 2010, the combined case did not perform as well as the radar case at predicting the high flows, although it did perform better than the gauge case.

The annual aggregation results (Figure 8) show that, during the calibration period (2005-2007), the combined case produced an annual water balance that fell between the gauge and radar water balances in the years 2005 and 2007. In both of these years, the gauge was the most accurate annual water balances. In 2006, the combined case produced an annual water balance slightly lower than the radar case, but none of the three cases produced a very accurate annual water balance for this year. During the evaluation time period (2008-2009), the combined case did well in 2009 at capturing the water balance, but did poorly in 2008 relative to the radar case. In 2010, all three models over estimated the streamflow by a similar amount.

Given these annual accumulation results, we can say that the combined case most often falls between the gauge and radar cases, as expected. However, both the gauge and radar cases frequently either under or over predict observed streamflow, and therefore the combined case also under or over predicts the annual accumulated flow for these years. Yet when all data is accumulated over the period of analysis, the combined case does produce the most accurate total water balance by a slight margin over the radar case. It is also clear from the total water balance summations that the gauge case produces streamflow predictions that overestimate the total water balance observed in the streamflow record.

Consideration of Alternative Combination Approaches

The approach presented for combining gauge and radar-based precipitation estimates in this study was one of four methods tested as part of our research. We focused the discussion on this single method because it performed best of the methods tested. Two of the other methods were attempts to identify convective from frontal storms, and to use radar for convective storms and gauge data for frontal storms. The hypothesis was that, because convective storms are more heterogeneous than frontal storms, radar would best capture those storm events. However, because frontal storms are more homogeneous, and because gauge observations of precipitation are generally more accurate than radar-based observations, gauge-observations would be optimal for frontal storms. The other method we tested was also based on streamflow matching, like the method described in this paper, but attempted to match modeled streamflow using radar and gauge-based precipitation estimates to observed streamflow, rather than matching the gauge and radar precipitation records themselves to observed streamflow.

For the convective vs. frontal storm hypothesis, we devised and tested two combination approaches. First, we simply selected radar precipitation for summer months and gauged precipitation for other months. This simple combination approach did not perform well, and in fact did not perform as well as simply using the gauged precipitation time series to drive the model. In the second approach we took the five gauged precipitation values for each day, and if there were high variance between them, we assumed that a convective storm event occurred and used radar precipitation estimates for those days. This second approach did perform better than using either the gauge or radar estimates alone, but did not improve on the combined method described in this paper. The alternative streamflow-matching approach also performed better than using either the gauge or radar estimates alone, but again did not improve on the combined method described in this paper. This was surprising because we expected this approach to better handle the spatial distribution of subbasin precipitation selection.

CONCLUSIONS

We tested a method for combining precipitation observations from gauging stations and from radar-based estimates in order to improve streamflow estimates using a watershed model. The method is based on the concept of selecting the precipitation estimates from one of these two datasets for each day and for each subbasin based on a matching of observed precipitation with observed streamflow. We compared streamflow estimates generated with SWAT models calibrated to three different input precipitation datasets (gauge, radar, combined) to observed streamflow to test if the combined method produced better streamflow estimates.

Results of the study show that fusing the two precipitation data sources using the combined methodology improved model streamflow estimates. The increase of model accuracy (measured by E and R^2) was expected because the method is allowed to select the best precipitation data sources from the gauge and radar options judged by how well each corresponds to observed streamflow. Our justification for this approach is that a given storm event may be better captured by one observational approach (e.g., gauging stations) compared to another approach (e.g., radar). Therefore, the goal in the selection process is to reconstruct the true precipitation with the assumption that both precipitation observing approaches are uncertain.

The result of this study can aid watershed modelers and decision makers in creating input precipitation datasets for watershed models where precipitation gauges are inadequate either because the gauges are not in close proximity to the watershed of interest, or because there is insufficient spatial coverage of gauges for the watershed area. By considering precipitation time series datasets from different sources as imperfect records of the true precipitation that fell over the watershed, it becomes reasonable to attempt to merge the two datasets in order to reconstruct the true precipitation that fell over that watershed.

There are certainly other data approaches for the data fusion algorithm that could be tested besides the ones described in this paper. For example, our approach ignores valuable information such as watershed conditions including vegetative cover or antecedent moisture conditions, which could prove valuable in the algorithm. We believe that the primary value of this work, therefore, is an argument that imperfect datasets of precipitation can be combined into a new dataset using algorithms that attempt to maximum informational content extraction.

Despite the success of the combined methodology presented here, we caution that the results of this study may be dependent on conditions specific to the region studied (e.g., climate, ecology, and geology). Therefore the methodology we followed for testing the combined precipitation datasets and the alterable approaches for combining the two time series briefly described in this paper should be applied when using this approach for other watersheds.

REFERENCES

Allen, R. (1986). "A Penman for all seasons." *Journal of Irrigation and Drainage Engineering-ASCE*, 112(4), 348–368.

Allen, R., Jensen, M., Wright, J., and Burman, R. (1989). "Operational estimates of reference evapotranspiration." *Agronomy Journal*, 81(4), 650–662.

Arabi, M., Govindaraju, R. S., Hantush, M. M., and Engel, B. A. (2006). "Role of watershed subdivision on modeling the effectiveness of best management practices with SWAT." *Journal of the American Water Resources Association*, 42(2), 513–528.

Arnold, J. G., Allen, P. M., Volk, M., Williams, J. R., and Bosch, D. (2010). "Assessment of different representations of spatial variability on SWAT model performance." *Transactions of the ASABE*, 53(5), 1433–1443.

Biemans, H., Hutjes, R., Kabat, P., Strengers, B., Gerten, D., and Rost, S. (2009). "Effects of precipitation uncertainty on discharge calculations for main river basins." *Journal of Hydrometeorology*, 10(4), 1011–1025.

Droegemeier, K., Smith, J., Businger, S., Doswell III, C., Doyle, J., Duffy, C., Foufoula-Georgiou, E., Graziano, T., James, L., Krajewski, V., et al. (2000). "Hydrological aspects of weather prediction and flood warnings: Report of the ninth prospectus development team of the U.S. Weather Research Program." *Bulletin of the American Meteorological Society*, 81(11), 2665–2680.

Duan, Q., Sorooshian, S., and Gupta, V. (1992). "Effective and efficient global optimization for conceptual rainfall-runoff models." *Water Resources Research*, 28(4), PP. 1015–1031.

Eckhardt, K. and Arnold, J. G. (2001). "Automatic calibration of a distributed catchment model." *Journal of Hydrology*, 251(1-2), 103–109.

Fulton, R., Ding, F., and Miller, D. (2003). "Truncation errors in historical WSR-88D rainfall products." *Proceedings of the 31st Conference on Radar Meteorology*, Vol. 6. 12.

Gassman, P., Reyes, M. R., Green, C., and Arnold, J. (2007). "The soil and water assessment tool: Historical development, applications, and future research directions." *Transactions of the ASABE*, 50(4), 1211–1250.

Goovaerts, P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." *Journal of hydrology*, 228(1), 113–129.

Green, W. and Ampt, G. (1911). "Studies on soil physics part I - the flow of air and water through soils." *Journal of Agricultural Science*, 4, 1–24.

Groisman, P. and Legates, D. (1995). "Documenting and detecting long-term precipitation trends: where we are and what should be done." *Climatic Change*, 31(2), 601–622.

Hargreaves, G. H. and Samani, Z. A. (1985). "Reference crop evapotranspiration from temperature." *Applied Engineering in Agriculture*, 1(2), 96–99.

Hildebrand, P. H., Towery, N., and Snell, M. R. (1979). "Measurement of convective mean rainfall over small areas using High-Density raingages and radar." *Journal of Applied Meteorology*, 18, 1316–1326.

Hitschfeld, W. and Bordan, J. (1954). "Errors inherent in the radar measurement of rainfall at attenuating wavelengths." *Journal of Meteorology*, 11, 58–67.

Huff, F. A. (1970). "Sampling errors in measurement of mean precipitation." *Journal of Applied Meteorology*, 9, 35–44.

Jayakrishnan, R., Srinivasan, R., and Arnold, J. (2004). "Comparison of raingage and WSR-88D stage III precipitation data over the Texas-Gulf basin." *Journal of Hydrology*, 292(1-4), 135–152.

Jha, M., Gassman, P. W., Secchi, S., Gu, R., and Arnold, J. (2004). "Effect of watershed subdivision on SWAT flow, sediment, and nutrient predictions." *Journal of the American Water Resources Association*, 40(3), 811–825.

Kenneth, M. (1972). "Part 630 Hydrology, National Engineering Handbook." *Chapter 15*, United States Department of Agriculture, Natural Resources Conservation Service, Washington D.C.

Krajewski, W., Villarini, G., and Smith, J. (2010). "Radar-rainfall uncertainties." *Bulletin of the American Meteorological Society*, 91(1), 87–94.

Lakshmanan, V., Fritz, A., Smith, T., Hondl, K., and Stumpf, G. (2007). "An automated technique to quality control radar reflectivity data." *Journal of applied meteorology and climatology*, 46(3), 288–305.

Lawrence, B. A., Shebsovich, M. I., Glaudemans, M. J., and Tilles, P. S. (2003). "Enhancing precipitation estimation capabilities at national weather service field offices using multi-sensor precipitation data mosaics." *Chapter 15.1*, Office of Hydrologic Development, National Weather Service, NOAA, Silver Spring, Maryland.

Legates, D. (2000). "Real-time calibration of radar precipitation estimates." *Professional Geographer*, 52(2), 235–246.

Legates, D. and DeLiberty, T. (1993). "Precipitation measurement biases in the united states1." *JAWRA Journal of the American Water Resources Association*, 29(5), 855–861.

McCuen, R. H., Knight, Z., and Cutter, A. G. (2006). "Evaluation of the NashSutcliffe efficiency index." *Journal of Hydrologic Engineering*, 11(6).

Migliaccio, K. and Chaubey, I. (2008). “Spatial distributions and stochastic parameter influences on SWAT flow and sediment predictions.” *Journal of Hydrologic Engineering*, 13(4), 258–269.

Moriasi, D., Arnold, J., Liew, M. V., Bingner, R., Harmel, R., and Veith, T. (2007). “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations.” *Transactions of the ASAE*, 50(3), 885–900.

Mu, Q., Heinsch, F., Zhao, M., and Running, S. (2007). “Development of a global evapotranspiration algorithm based on modis and global meteorology data.” *Remote Sensing of Environment*, 111(4), 519–536.

Mu, Q., Zhao, M., and Running, S. (2011). “Improvements to a modis global terrestrial evapotranspiration algorithm.” *Remote Sensing of Environment*.

Nash, J. and Sutcliffe, J. (1970). “River flow forecasting through conceptual models part I-A discussion of principles.” *Journal of Hydrology*, 10(3), pp. 282–290.

Neitsch, S., Arnold, J., Kiniry, J., and Williams, J. (2005). “Soil and water assessment tool theoretical documentation.” *Version 2005*, Grassland Soil and Water Research Laboratory, Agricultural Research Service, Temple, Texas.

Olivera, F., Choi, J., Kim, D., Li, M., et al. (2008). “Estimation of average rainfall areal reduction factors in texas using NEXRAD data.” *Journal of Hydrologic Engineering*, 13, 438.

Priestley, C. H. B. and Taylor, R. J. (1972). “On the assessment of surface heat flux and evaporation using Large-Scale parameters.” *Monthly Weather Review*, 100(2), 81–92.

Sahu, M. and Gu, R. R. (2009). “Modeling the effects of riparian buffer zone and contour strips on stream water quality.” *Ecological Engineering*, 35(8), 1167–1177.

Santhi, C., Srinivasan, R., Arnold, J., and Williams, J. (2006). "A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in texas." *Environmental Modelling & Software*, 21(8), 1141–1157.

Singh, V. P. and Woolhiser, D. A. (2002). "Mathematical modeling of watershed hydrology." *Journal of Hydrologic Engineering*, 7(4), 270–292.

Smith, J. A. and Krajewski, W. F. (1991). "Estimation of the mean field bias of radar rainfall estimates." *Journal of Applied Meteorology*, 30(4), 397–412.

Tolson, B. and Shoemaker, C. (2007). "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration." *Water Resources Research*, 43(1), W01413.

Tuppad, P., Douglas-Mankin, K. R., Koelliker, J. K., Hutchinson, J. M. S., and Knapp, M. C. (2010). "NEXRAD Stage III precipitation local bias adjustment for streamflow prediction." *Transactions of the ASABE*, 53(5), 1511–1520.

van Griensven, A., Francos, A., and Bauwens, W. (2002). "Sensitivity analysis and auto-calibration of an integral dynamic model for river water quality." *Water Science & Technology*, 45(9), pp 325–332.

Wilson, J. and Brandes, E. (1979). "Radar measurement of rainfall -A summary." *Bulletin of the American Meteorological Society*, 60, 1048–1060.

Young, C., Bradley, A., Krajewski, W., Kruger, A., and Morrissey, M. (2000). "Evaluating nexrad multisensor precipitation estimates for operational hydrologic forecasting." *Journal of Hydrometeorology*, 1(3), 241–254.

List of Tables

TABLE 1. Resulting changes in parameter values from model calibrations

Parameter	Precipitation Scenario			Range	Operation
	Gauge	Radar	Combined		
Cn2	8.0	23.1	21.0	±25%	% Added
Esco	0.04	0.13	0.10	0.01-1.00	Replaced
SolAwc	24.5	-1.6	22.8	±25%	% Added
Surlag	0.78	0.75	0.89	0-10	Replaced

TABLE 2. Model statistics for the calibration and evaluation periods

Time Period	Statistic	Gauge	Radar	Combined
	E	0.58	0.59	0.75
2005-2007[a]	R^2	0.59	0.62	0.77
	PB (%)	-6.2	41.0	32.1
	E	0.50	0.33	0.60
2008-2010[b]	R^2	0.54	0.61	0.74
	PB (%)	-25.5	-13.7	-10.2

a Calibration period.
b Evaluation period.

List of Figures

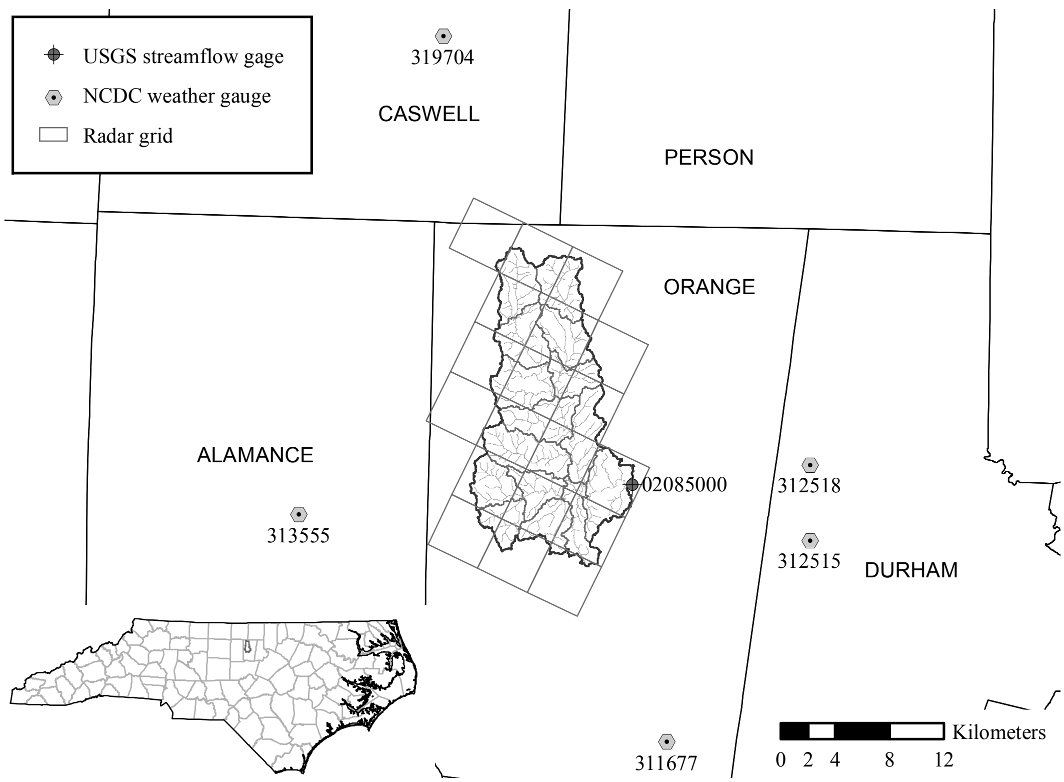

FIG. 1. The Eno Watershed with radar precipitation grids and precipitation gauges

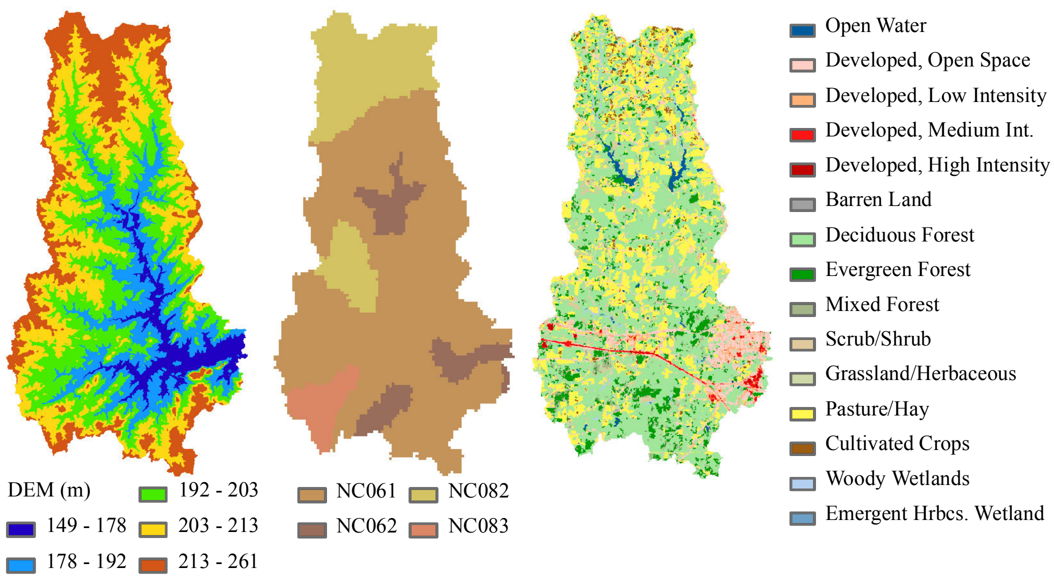

FIG. 2. Input geospatial datasets for the Eno Watershed. From left to right: Elevation, soil and land cover

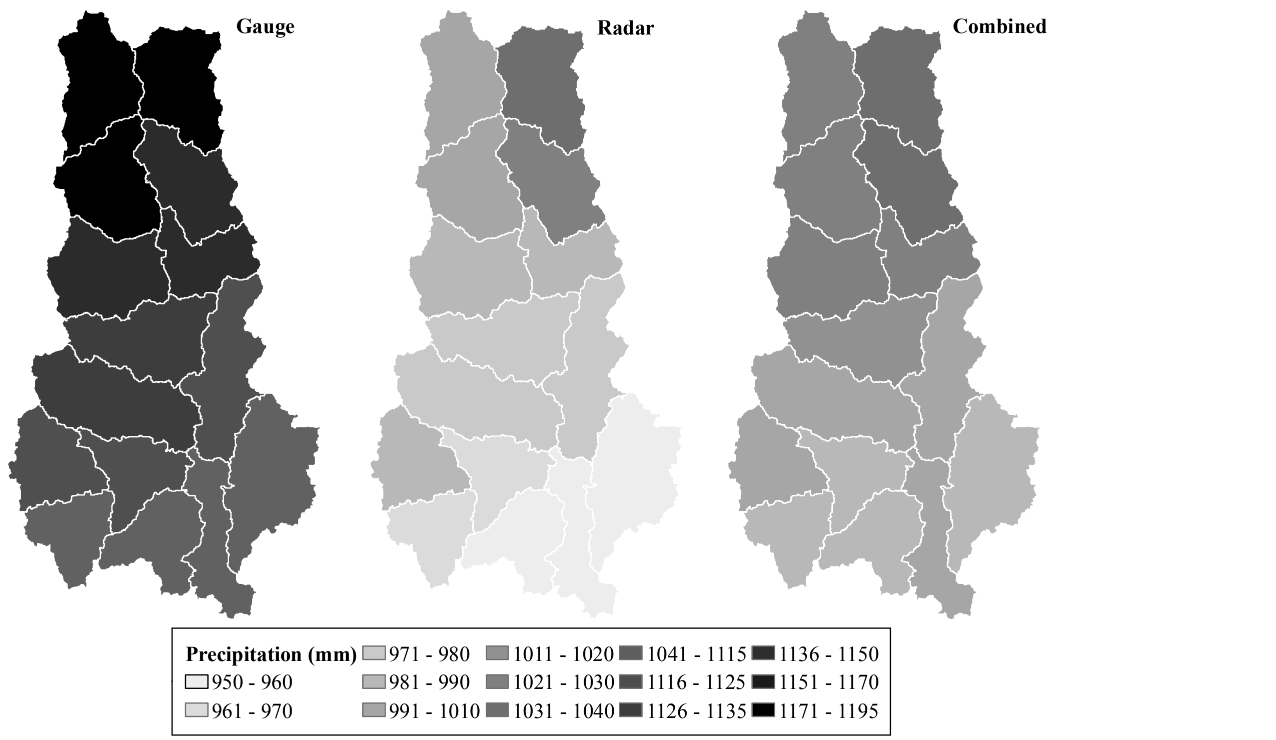

FIG. 3. Average annual precipitation during the period 2005-2010 for gauge, radar and combined precipitation cases

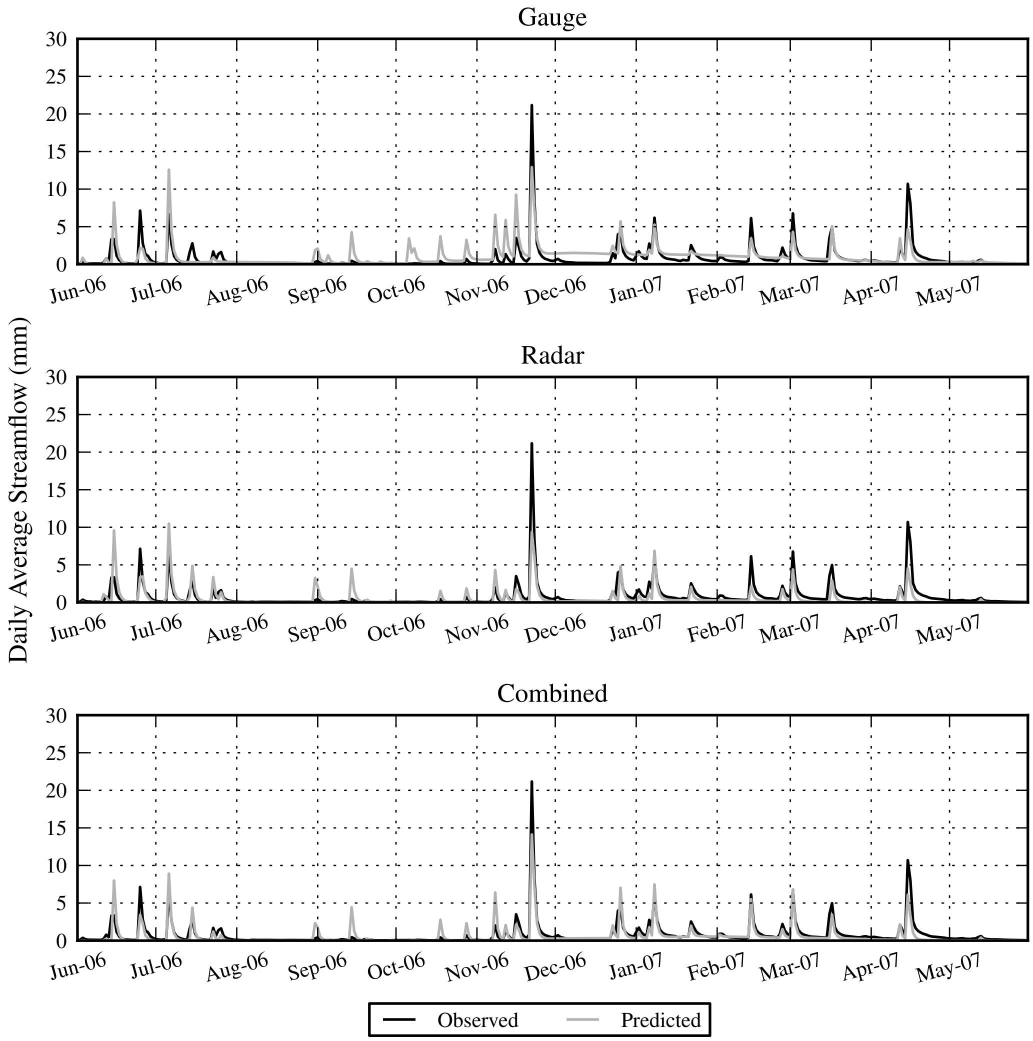

FIG. 4. Comparison of observed daily streamflow with modeled daily streamflow for a year during the calibration period.

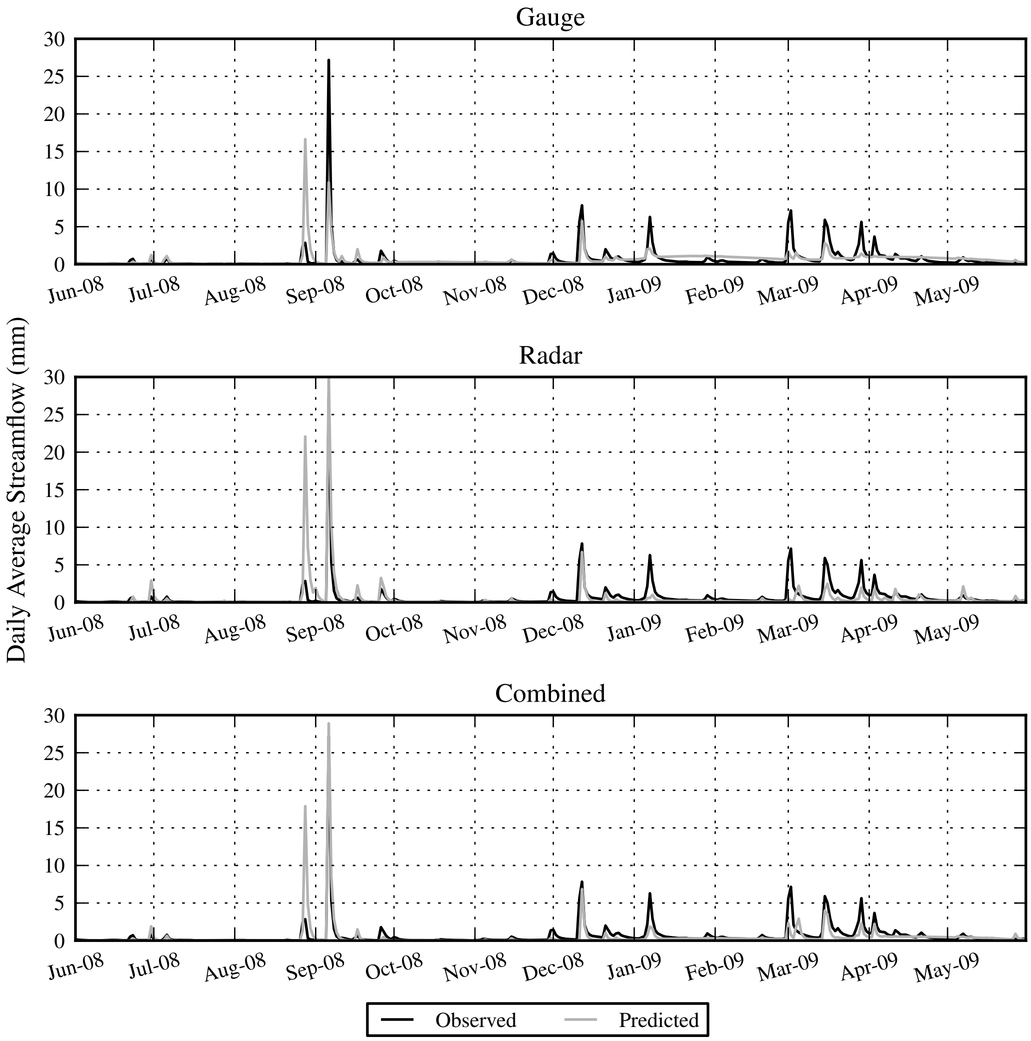

FIG. 5. Comparison of observed daily streamflow with modeled daily streamflow for a year during the evaluation period.

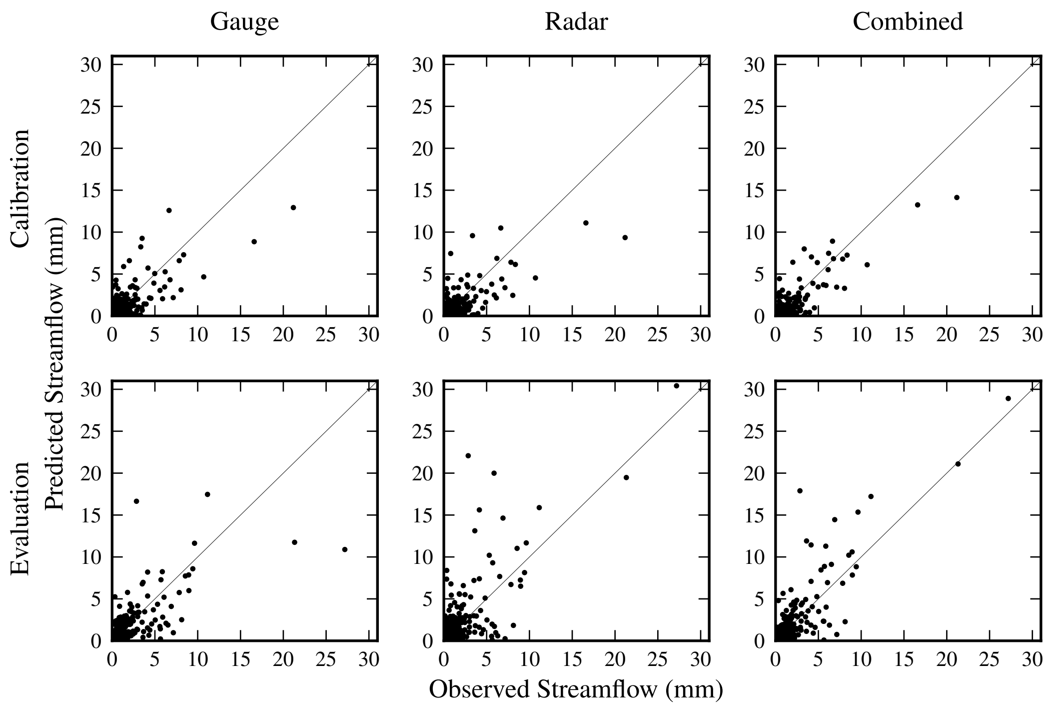

FIG. 6. Scatter plots of observed and modeled daily streamflow during the model calibration (2005-2007) and evaluation (2008-2010) periods

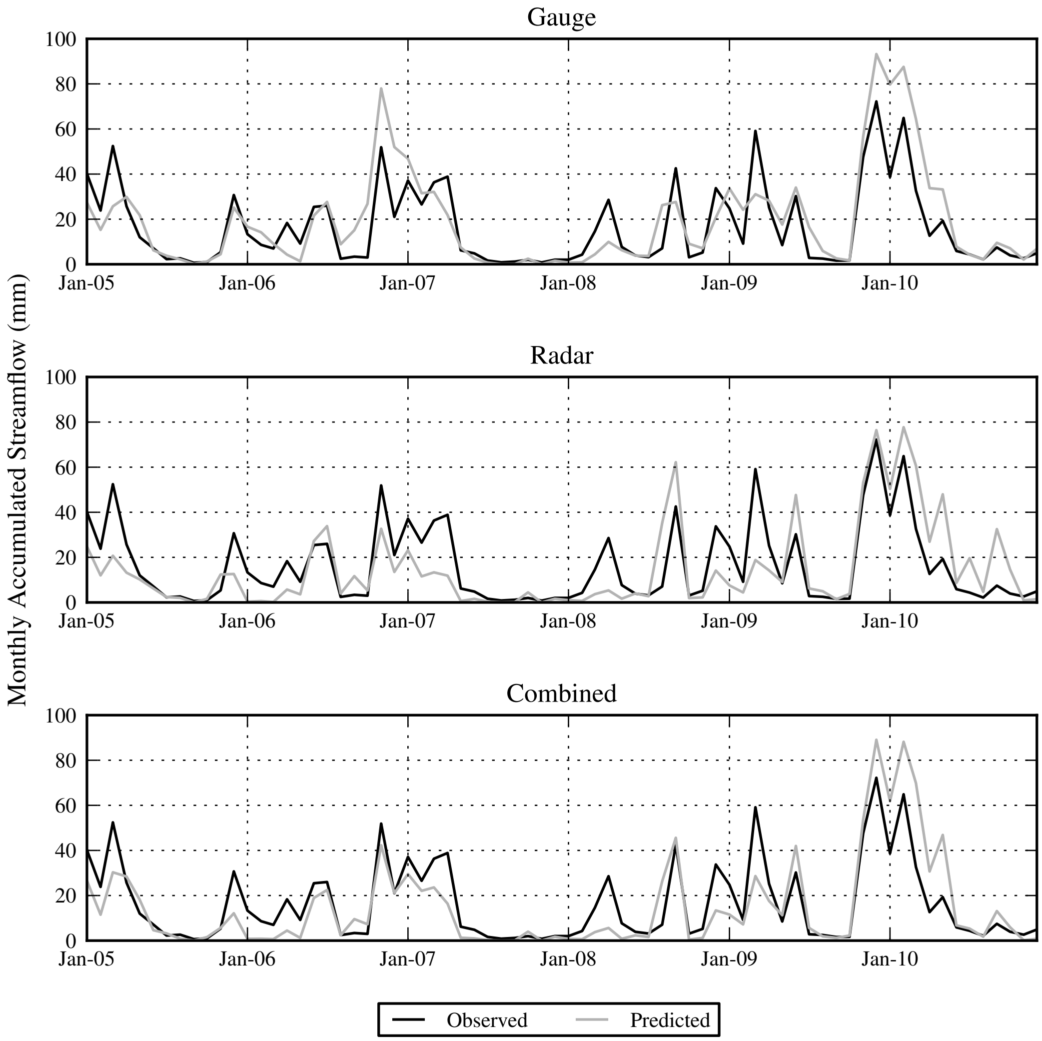

FIG. 7. Comparison of observed and modeled daily streamflows aggregated to monthly summations for the calibration (2005-2007) and evaluation (2008-2010) periods.

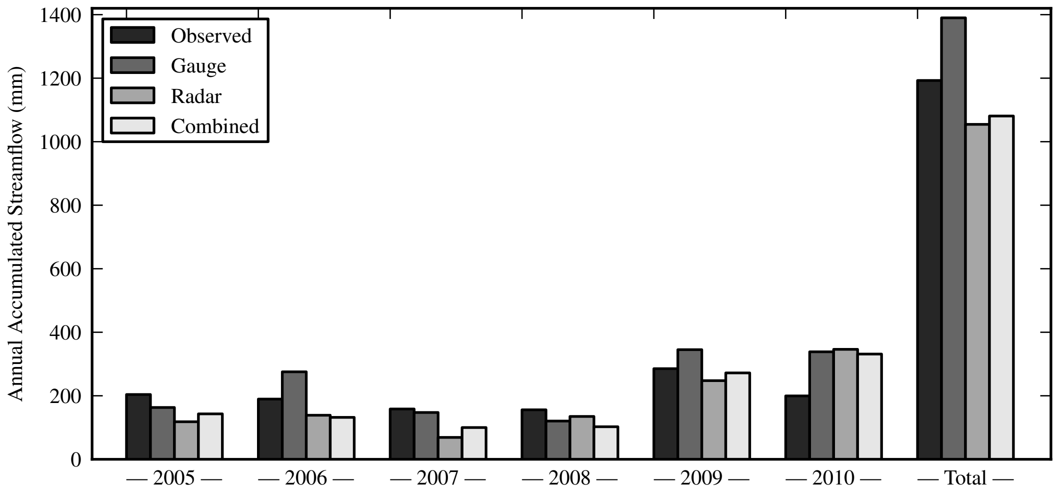

FIG. 8. Comparison of observed and modeled daily streamflows aggregated to annual summations for the calibration (2005-2007) and evaluation (2008-2010) periods.