USING NEXRAD AND RAIN GAUGE PRECIPITATION DATA FOR HYDROLOGIC CALIBRATION OF SWAT IN A NORTHEASTERN WATERSHED

A. M. Sexton, A. M. Sadeghi, X. Zhang, R. Srinivasan, A. Shirmohammadi

ABSTRACT. *The value of watershed-scale, hydrologic and water quality models to ecosystem management is increasingly evident as more programs adopt these tools to evaluate the effectiveness of different management scenarios and their impact on the environment. Quality of precipitation data is critical for appropriate application of watershed models. In small watersheds, where no dense rain gauge network is available, modelers are faced with a dilemma to choose between different data sets. In this study, we used the German Branch (GB) watershed (~50 km^2), which is included in the USDA Conservation Effects Assessment Project (CEAP), to examine the implications of using surface rain gauge and next-generation radar (NEXRAD) precipitation data sets on the performance of the Soil and Water Assessment Tool (SWAT). The GB watershed is located in the Coastal Plain of Maryland on the eastern shore of Chesapeake Bay. Stream flow estimation results using surface rain gauge data seem to indicate the importance of using rain gauges within the same direction as the storm pattern with respect to the watershed. In the absence of a spatially representative network of rain gauges within the watershed, NEXRAD data produced good estimates of stream flow at the outlet of the watershed. Three NEXRAD datasets, including (1) non-corrected (NC), (2) bias-corrected (BC), and (3) inverse distance weighted (IDW) corrected NEXRAD data, were produced. Nash-Sutcliffe efficiency coefficients for daily stream flow simulation using these three NEXRAD data ranged from 0.46 to 0.58 during calibration and from 0.68 to 0.76 during validation. Overall, correcting NEXRAD with rain gauge data is promising to produce better hydrologic modeling results. Given the multiple precipitation datasets and corresponding simulations, we explored the combination of the multiple simulations using Bayesian model averaging. The results show that this Bayesian scheme can produce better deterministic prediction than any single simulation and can provide reasonable uncertainty estimation. The optimal water balance obtained in this study is an essential precursor to acquiring realistic estimates of sediment and nutrient loads in future GB modeling efforts. The results presented in this study are expected to provide insights into selecting precipitation data for watershed modeling in small Coastal Plain catchments.*

Keywords. *Hydrologic modeling, Model calibration, MPE, NEXRAD, Rain gauge, SWAT.*

Model simulation of hydrologic processes is only as good as the input used to drive the model. Precipitation is one of the most important inputs to any hydrologic model. Except for several experimental watersheds, the National Climatic Data Center (NCDC) rain gauge data (approximately one gauge per 800 km^2) is the major source of observed precipitation data for most watersheds in the U.S. Another important source of precipitation data is next-generation radar (NEXRAD), which provides spatially continuous estimations at approximately 4 × 4 km^2 resolution. In small watersheds, where no dense rain gauge network is available, we are faced with a dilemma to choose between different data sets.

Site-specific precipitation data are generally scarce due to lack of a sufficient number of rain gauges and/or due to measurement errors. These issues have been major concerns in watershed model calibration (Groisman and Legates, 1994; Neff, 1977; Skinner et al., 2009; Huebner et al., 2003) and have led researchers to explore other sources of rainfall data, such as NEXRAD. Along with rain gauges, NEXRAD data contain measurement and algorithm errors (Young et al., 2000; Jayakrishnan et al., 2004; Hunter, 1996). Attempts to validate NEXRAD data using rain gauge data as ground truth have also encountered some difficulties because of the scarcity of gauge data and the difference in sampling area (Young et al., 2000; Jayakrishnan et al., 2004; Skinner et al., 2009). Although these errors and difficulties exist, radar estimations are still a viable source of rainfall data in hydrologic modeling, especially as ra-

Submitted for review in October 2009 as manuscript number SW 8273; approved for publication by the Soil & Water Division of ASABE in March 2010.

The authors are **Aisha M. Sexton, ASABE Member Engineer,** Postdoctoral Research Associate, Fischell Department of Bioengineering, University of Maryland, College Park, and USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland; **Ali M. Sadeghi,** Soil Scientist, USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland; **Xuesong Zhang, ASABE Member Engineer,** Research Scientist, Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, Maryland; **Raghavan Srinivasan,** Professor, Department of Ecosystem Sciences and Management, Spatial Sciences Laboratory, Texas A&M University, College Station, Texas; and **Adel Shirmohammadi, ASABE Fellow,** Associate Dean and Professor, College of Agriculture and Natural Resources, University of Maryland, College Park, Maryland. **Corresponding author:** Aisha M. Sexton, USDA-ARS Hydrology and Remote Sensing Laboratory, 10300 Baltimore Ave., BARC-West Bldg. 007, Beltsville, MD 20705; phone: 301-504-8554; fax: 301-504-8931; e-mail: aisha.sexton@gmail.com.

dar algorithms are improved and denser gauge networks are created for radar validation (Habib et al., 2009).

NEXRAD rainfall data have been developed and improved through four different processing steps. In stage I, the Hourly Digital Precipitation (HDP) network was developed by entering reflectivity measurements into the Precipitation Processing System (PPS) algorithm. These data have been organized into 4×4 km^2 grids in the Hydrologic Rainfall Analysis Project (HRAP) coordinate system. The stage II product is a combination of stage I radar data corrected for bias using hourly rain gauge observations for a single radar site. Collective stage II radars covering River Forecasting Center (RFC) regions are mosaicked and corrected for bias to produce stage III data. Finally, stage IV data were created by combining stage III data (RFC regions) to cover the entire U.S. Several references describe the details of the PPS algorithm and its updates (Fulton et al., 1998; Ahnert et al., 1983; Ahnert et al., 1984).

To date, stage III data have been the most widely used of all NEXRAD rainfall data in the area of hydrologic and water quality modeling (Starks and Moriasi, 2009; Tobin and Bennett, 2009; Jayakrishnan et al., 2005; Moon et al., 2004; Neary et al., 2004; Di Luzio and Arnold, 2004). However, improvements to the PPS algorithm have led to another product, the Multisensor Precipitation Estimator (MPE). MPE enhancements over the stage III algorithm include: delineation of effective areal coverage of radar, mosaicking based on radar sampling geometry, service area-wide precipitation analysis, improved mean-field bias correction, and local bias correction (Seo and Breidenbach, 2002; Nelson et al., 2006; Wang et al., 2008). The MPE product has proven to be superior to stage III data, which is why RFCs have replaced the production of stage III data with MPE data in recent years. Although the MPE product has been validated in a number of studies, the impact of MPE data in hydrologic modeling is not well known. Neary et al. (2004) mentioned the need for more studies to evaluate the recently adopted MPE products in hydrologic modeling.

The Soil Water Assessment Tool (SWAT) is a widely used hydrologic and water quality model that was developed to simulate the effects of changing land uses and climate on watershed water quality (Arnold et al. 1998). SWAT is the major modeling tool used in the CEAP program, which aims to quantify the environmental benefits of conservation practices implemented under USDA conservation programs. The utility of NEXRAD rainfall data in SWAT has the potential to improve flow estimates by accounting for the spatial variability of rainfall. The majority of studies that have implemented NEXRAD data in the SWAT model have used stage III data (Moon et al., 2004; Di Luzio and Arnold, 2004; Jayakrishnan et al., 2005; Tobin and Bennett, 2009). Moon et al. (2004) evaluated the use of NEXRAD rainfall data on stream flow estimation of the SWAT model. NEXRAD rainfall inputs provided a better flow estimate than gauge data. Another study found that SWAT simulations with NEXRAD provided better stream flow results on a monthly time scale compared to Tropical Rainfall Measurement Mission (TRMM) and rain gauge data (Tobin and Bennett, 2009). Jayakrishnan at el. (2005) found stream flow estimates using NEXRAD stage III data better than rain gauge estimates without calibrating the model. They pointed out the potential of improving radar data and subsequent model simulations by calibrating radar data using gauge data. NEXRAD data, however, have not provided better flow simulation results in all cases. Kalin and Hantush (2006) compared gauge and NEXRAD MPE driven simulations of SWAT on an eastern Pennsylvania watershed. There was no significant difference in model flow prediction during the calibration period on the monthly or daily time scale. Validation on a monthly basis using NEXRAD data resulted in higher model efficiencies than using gauge data, while validation on a daily basis using gauge data resulted in higher model efficiencies than NEXRAD data. More research is needed in this area to determine the circumstances in which NEXRAD will provide better stream flow estimates than surface rain gauge data in hydrologic modeling.

The goal of this study was to evaluate the ability of SWAT to estimate stream flow in a watershed containing no rain gauges using proximal rain gauge and NEXRAD MPE precipitation data. A new GIS tool for incorporating NEXRAD data into ArcSWAT, NEXRAD_SWAT (Zhang and Srinivasan, 2009), was utilized to provide MPE data as well as rain gauge calibrated radar data. In addition to comparing gauge and NEXRAD flow estimates, NEXRAD_SWAT enabled us to compare flow estimates derived with original MPE data as well as MPE data that received additional calibration using rain gauge data. Given the multiple precipitation datasets and corresponding simulations, we explored the combination of the multiple simulations using Bayesian model averaging. Special attention was also given to the location of surface rain gauges with respect to the direction of storm flow patterns in capturing rainfall amounts representative of that in the watershed.

Materials and Methods

Site Description

The German Branch (GB) watershed (~50 km^2) is a tributary within the non-tidal zone of the larger Choptank River basin located in the Coastal Plain of Maryland on the eastern shore of Chesapeake Bay (fig. 1). Upland soils of the watershed are mostly composed of Ingleside sandy loam on 2% to 5% slopes. Baseflow contributes about 65% of the total flow in the watershed (Bachman et al., 1998). The major land uses are agriculture (~61%) and forest (~33%), followed by developed land (~5%) and water (~1%). The agricultural landscape in the region is dominated by the poultry industry. Corn and soybeans are grown to supply feed to those operations, and poultry litter is used to fertilize the crops. This watershed is being evaluated because it was initially selected as one of the CEAP special emphasis watersheds in which several tributaries have been identified as "impaired waters" under Section 303(d) of the Clean Water Act due to high levels of nutrients and sediments. It is to be noted that the CEAP program managers have recently moved this watershed from the "special emphasis watersheds" into the list of permanent watersheds called "CEAP core watersheds" (NRCS, 2009).

Figure 2 shows the four National Climatic Data Center (NCDC) surface rain gauges in closest proximity to the GB watershed outlet. The Chestertown and Royal Oak gauges were chosen to be included in this study because of their closeness to the GB outlet as well as their directional location with respect to the watershed and the direction of storms on the eastern shore of Maryland. Both gauges are located west of the watershed, which is the direction that storms generally travel from in temperate latitudes. Therefore, their storm pattern was more likely to resemble storm occurrences in the GB watershed.

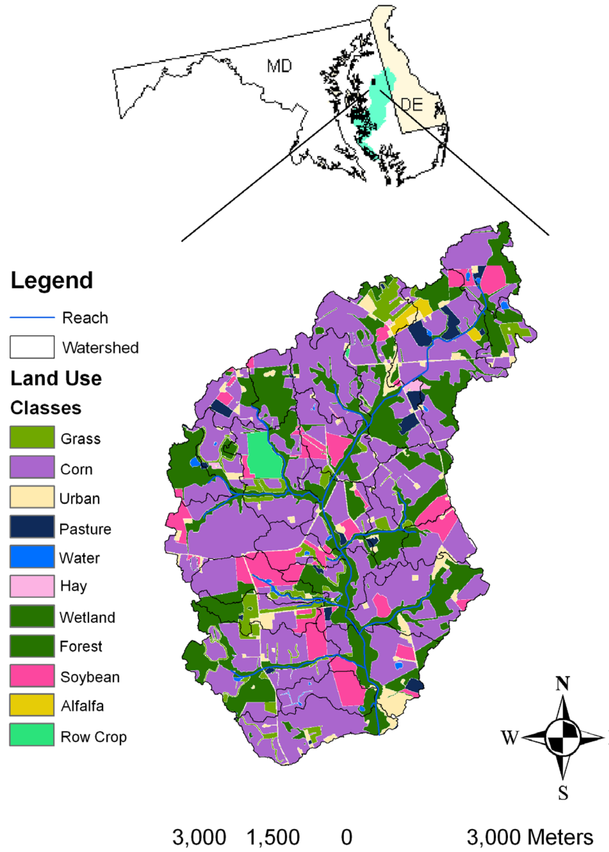

Figure 1. Location of German Branch watershed in Maryland and land use map showing delineated subbasins.

Model Selection and Data Acquisition

SWAT Model

SWAT is a complex, physically based, semi-distributed model that operates in continuous time on a daily time step. The main components of SWAT include: climate, hydrology, land cover/plant growth, erosion, nutrients, pesticides, land management, channel routing, and reservoir routing. Algorithms from the QUAL2E model were incorporated into SWAT to provide in-stream water quality modeling capabilities (Ramanarayanan et al., 1996).

The version of SWAT used in this study was SWAT-2005, which operates in the ArcGIS interface (Winchell et al., 2007). The three basic GIS maps required to run SWAT include a digital elevation model (DEM), land cover/land use, and soils data. A 2 m resolution DEM based on LIDAR data collected by the State of Maryland was produced at USDA-ARS. A 3 × 3 pixel low-pass filter was used to eliminate "no data" values. Additional "no data" values were removed by hand and replaced by local averages. A high-resolution land use map was developed through on-screen digitizing in ArcMap 9.1 using 1998 National Aerial Photography Program (NAPP) digital orthophoto quad imagery (1:12,000 scale). Soil Survey Geographic (SSURGO) data downloaded from the USDA-NRCS Soil Data Mart server were used as soils input into the model.

The delineated watershed was separated into 26 subbasins based on tributary drainage areas (fig. 1). Within each subbasin, the superimposing of similar land uses, soil types, and slopes created 233 hydrologic response units (HRUs) in the GB watershed. Threshold area values of $\geq$10%, $\geq$15%, and $\geq$15% were used to include land use, soils, and slope types, respectively, in the HRU definition process. Surface rain gauge and NEXRAD data were obtained from the National Weather Service (NWS). Surface rain gauge and temperature data from NCDC were obtained for two Maryland sites, one in Chestertown and the other in Royal Oak. NEXRAD data missing from the NWS website were obtained from the Mid-Atlantic River Forecasting Center (MARFC). Daily solar radiation, wind speed, relative humidity, and missing precipitation and temperature data were generated using SWAT's weather generator (Neitsch et al., 2005).

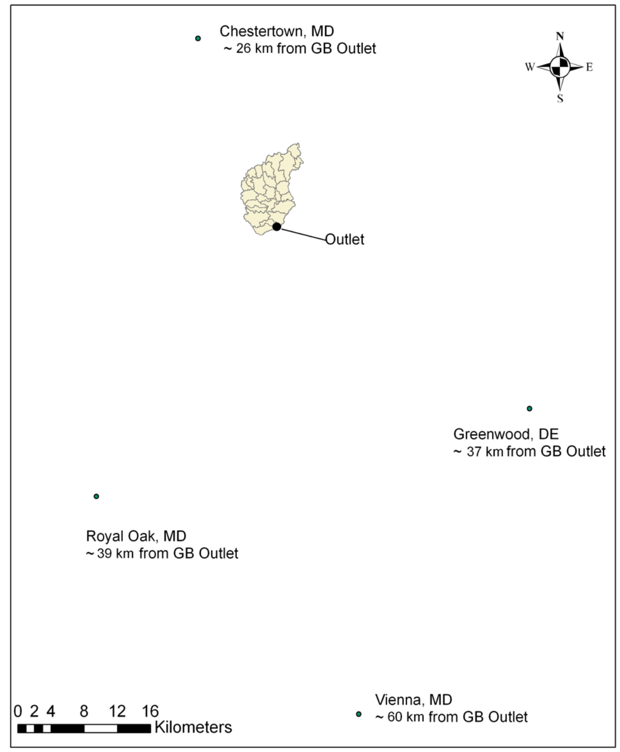

Figure 2. Location of nearby rain gauges with respect to GB watershed outlet.

Using NEXRAD in SWAT

The ArcGIS interface of SWAT can automatically select the rain gauge closest to each subbasin and read in precipitation records stored in text and database format for each rain gauge. In order to use NEXRAD in SWAT, a new GIS program was developed for SWAT (NEXRAD-SWAT) (Zhang and Srinivasan, 2009) to automatically read in the binary NEXRAD MPE data and estimate the spatial average precipitation for each subbasin. NEXRAD-SWAT can evaluate and correct NEXRAD data using rain gauge data. Several geostatistical methods can be employed by the tool to produce a spatial precipitation map (in grid format) for use in hydrologic modeling. The methods include: nearest bias correction (BC), simple kriging (SK), ordinary kriging (OK), inverse distance weighted (IDW), simple kriging with varying local means (SKlm), and kriging with external drift (KED). Kriging methods only perform better than IDW when there is a dense rain gauge network. In this study area, data from only four surface rain gauges were available near the watershed. Therefore, NEXRAD-SWAT was used to supply non-corrected (NC), bias-corrected (BC), and inverse distance weighted (IDW) corrected MPE data to ArcSWAT.

NEXRAD_SWAT makes bias corrections to MPE data using the following equations:

$$R_{adj} = B \cdot R_{ori} \tag{1}$$

$$B = \frac{\sum_{i=1}^{n} Z(\mathbf{x}_i)/n}{\sum_{i=1}^{n} R(\mathbf{x}_i)/n} \tag{2}$$

where R_{adj} is the bias-adjusted NEXRAD data, R_{ori} is the NEXRAD-estimated precipitation, B is the bias adjustment factor, $Z(\mathbf{x}_i)$ and $R(\mathbf{x}_i)$ are, respectively, the rain gauge observed and NEXRAD-estimated precipitation at sampled locations $\mathbf{x}_i$ (i = 1, 2, ..., n), and n is the number of data sampled using rain gauges.

In the implementation of IDW to correct NEXRAD data, the precipitation amount (Z) within a spatial domain is decomposed into a trend (m) and a residual (ε), where $Z(\mathbf{x}) = m(\mathbf{x}) + \varepsilon(\mathbf{x})$. The drift trend ($m$) is fitted using linear regression analysis. The general form of $m(\mathbf{x})$ is:

$$m(\mathbf{x}) = \sum_{k=0}^{K} \beta_k y_k(\mathbf{x})$$

where $y_1(\mathbf{x})$, $y_2(\mathbf{x})$, ..., $y_K(\mathbf{x})$ are external explanatory variables, the β_k are unknown drift model coefficients to be determined, and K is the number of predictors. In this study, the trend surface used by IDW was obtained using $m(\mathbf{x}) = \beta_1 + \beta_2 R(\mathbf{x})$. The unknown residual at the unsampled location ($\mathbf{u}$) is a linear combination of neighboring observed residuals, that is:

$$\varepsilon(\mathbf{u}) = \sum_{i=1}^{n} \lambda_{ui}[\varepsilon(\mathbf{x}_i)]$$

The IDW method interpolates precipitation by weighting the points closer to the prediction location greater than those farther away. Equation 3 denotes the procedure (Zhang and Srinivasan, 2009):

$$\varepsilon(\mathbf{u}) = \frac{1}{\sum_{i=1}^{n} \lambda_{ui}} \sum_{i=1}^{n} \lambda_{ui} Z(\mathbf{x}_i), \text{ where } \lambda_{ui} = \frac{1}{h_{ui}{}^{p}} \tag{3}$$

where $\varepsilon(\mathbf{u})$ is the interpolated residual value, λ_{ui} is the weight of the sampled data at location $\mathbf{x}_i$, h_{ui} denotes the distance between unsampled location $\mathbf{u}$ and sampled location $\mathbf{x}_i$, and p is the power of h_{ui}.

The data requirements for NEXRAD-SWAT include hourly NEXRAD-MPE data in XMRG format (~4 km resolution), a rain gauge shape file, daily precipitation records for each rain gauge, and a subbasin shape file. The map projection of the rain gauge and subbasin shape files must be known and have the same projection. NEXRAD-SWAT will automatically convert the projection used by the modeler to the HRAP projection used by NEXRAD. In SWAT, precipitation is modeled on a subbasin basis. Therefore, the subbasin map is discretized into a grid (fig. 3a). NEXRAD-SWAT interpolation methods are then implemented at each grid to estimate precipitation. Precipitation estimates from all grids contained in each subbasin are then averaged to represent precipitation at the subbasin center (fig. 3b).

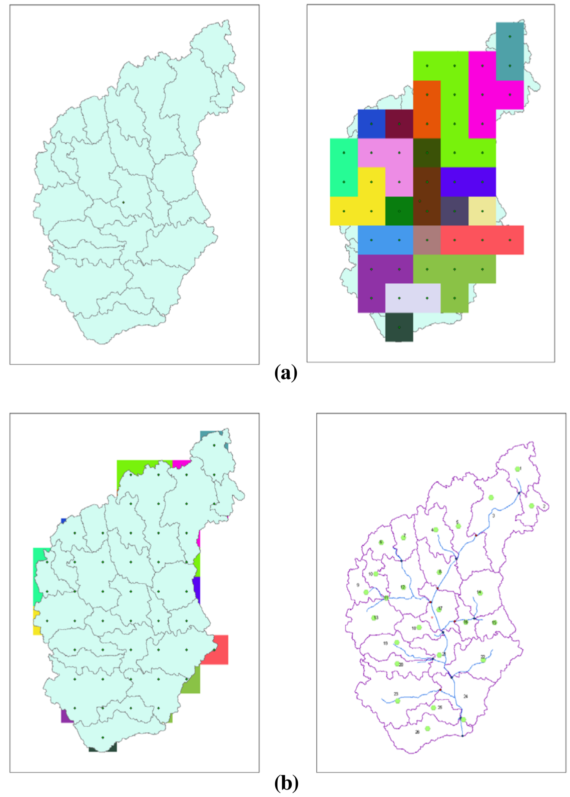

Figure 3. Maps of German Branch subbasins showing (a) NEXRAD-SWAT discretized watershed grid, and (b) grid centroids over subbasins and subbasin centroids.

Sensitivity Analysis

Sensitivity analysis was conducted using the Latin hypercube one-at-a-time (LH-OAT) method (van Griensven et al., 2006), which is part of the slate of evaluation tools built into the SWAT model. A sensitivity index (S_i) was calculated by averaging the sensitivity indices for each interval of each parameter, denoted by:

$$S_{i,j} = \left| \frac{100 * \left(\frac{M(e_1,\ldots,e_i*(1+f_i),\ldots,e_p) - M(e_1,\ldots,e_i,\ldots,e_p)}{[M(e_1,\ldots,e_i*(1+f_i),\ldots,e_p) + M(e_1,\ldots,e_i,\ldots,e_p)]/2} \right)}{f_i} \right| \tag{4}$$

where $M(\ )$ is the model function, f_i is the fraction by which parameter e_i is changed, i is the number of parameters , and j is the LHS point or interval number. Based on the sensitivity analysis, 13 parameters were chosen to be sensitive (table 1) and were therefore included in model calibration. Parameters Cn2, Rchrg_Dp, Esco, Alpha_Bf, and Sol_Awc were ranked

Table 1. Sensitivity index (S_i) and ranking of parameters used in sensitivity analysis.[a]

Rank	Parameter	Description	S_i
1	**Cn2**	Initial SCS runoff curve number for moisture condition II	1.55
2	**Rchrg_Dp**	Deep aquifer percolation factor	0.88
3	**Esco**	Soil evaporation compensation factor	0.65
4	Gwqmn	Threshold depth of water in the shallow aquifer required for return flow to occur	0.54
5	**Alpha_Bf**	Baseflow alpha factor	0.25
6	Sol_Z	Depth from soil surface to bottom of layer	0.22
7	**Sol_Awc**	Available water capacity	0.21
8	Blai	Maximum potential leaf area index	0.08
9	**Timp**	Snow pack temperature lag factor	0.07
10	**Ch_K2**	Effective hydraulic conductivity in main channel alluvium	0.07
11	Canmx	Maximum canopy storage	0.06
12	**GW_Revap**	Groundwater "revap" coefficient	0.02
13	Slope	Average slope steepness	0.02
14	**Sol_K**	Saturated hydraulic conductivity	0.02
15	**Surlag**	Surface runoff lag coefficient	0.01
16	Revapmn	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur	0.01
17	**GW_Delay**	Groundwater delay time	0.01
18	**Ch_N2**	Manning's roughness coefficient for the main channel	0.01
19	**Smtmp**	Snow melt base temperature	0.01
20	Biomix	Biological mixing efficiency	0.01

[a] **Bold** type indicates the 13 parameters chosen to be sensitive.

highest. Other parameters affecting hydrograph timing, such as Timp, GW_Revap, and Surlag, were ranked lower but still considered important. Some parameters, such as Gwqmn and Sol_Z, were ranked high in sensitivity but not included in the calibration because their values were not well known or the default values were a best estimate.

Model Calibration and Validation

Calibration Algorithm

The optimization method used to calibrate the model was parameter solutions (Parasol) (van Griensven and Meixner, 2007). It uses shuffled complex evolution (SCE, a global search algorithm) to minimize a single objective function or multiple objective functions. Objective functions include sum of the squares of the residuals (SSQ) and SSQ after ranking. The equation for SSQ is:

$$\mathrm{SSQ} = \sum_{i=1,n} \left[x_{i,measured} - x_{i,simulated}\right]^2 \tag{5}$$

where $x_{i,measured}$ are measured data, $x_{i,simulated}$ are simulated data, and n represents the number of observations. Up to 16 parameters can be adjusted in one optimization run. The model was calibrated on a daily basis using years 2005 and 2006 with one year of spin-up (2004). Validation was conducted using the 1 January to 15 April 2007 (1/1/07 to 4/15/07) time period. Five different calibration scenarios were run using three sources of rainfall data. Two scenarios utilized rainfall data from each of the two surface rain gauges (Chestertown and Royal Oak). Another scenario included MPE data with no correction (NC). The remaining two scenarios utilized data from both rain gauges to correct MPE data for bias (BC) and with the inverse distance weighted (IDW) method.

Model Performance Measures

The performance measures used to evaluate model calibration and validation were the Nash-Sutcliffe efficiency coefficient (NSE), coefficient of determination (r^2), root mean squared error (RMSE), and percent bias (PBIAS). They are defined as follows:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2} \tag{6}$$

$$r^2 = \left(\frac{\sum_{i=1}^{n} (O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n} (O_i - \overline{O})^2} \sqrt{\sum_{i=1}^{n} (P_i - \overline{P})^2}} \right)^2 \tag{7}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (O_i - P_i)^2}{n}} \tag{8}$$

$$\mathrm{PBIAS} = \left[\frac{\sum_{i=1}^{n} (O_i - P_i) * 100}{\sum_{i=1}^{n} O_i} \right] (\%) \tag{9}$$

where O_i are observed data, P_i are predicted data, $\overline{O}$ and $\overline{P}$ are observed and predicted mean values, respectively, and n is the number of observations. Time series plots were also used to evaluate model performance.

Ensemble Model Prediction

Given the several precipitation datasets described above, we can produce several streamflow simulations. Instead of selecting one simulation with the best performance, multiple simulations can be combined to provide ensemble model prediction and uncertainty analysis. Bayesian model averaging (BMA) is a standard approach to inference in the presence of multiple competing models (Raftery et al., 2005). The BMA algorithm described by Zhang et al. (2009), which was derived based on Raftery et al. (2005) and Duan et al. (2007), is applied in this study. In BMA, the probabilistic distribution of a hydrologic prediction (y) is the weighted average of the posterior distribution of each model under consideration:

$$p(y \mid f_1, f_2, ..., f_K) = \sum_{k=1} w_k g(y \mid f_k) \tag{10}$$

where K is the number of competing models, k is the index of each model, f_k denotes the bias-corrected prediction of a candidate model M_k, and w_k is $p(f_k \mid D)$, the posterior probability of model prediction f_k, also known as the likelihood of model prediction f_k being the correct prediction given the observational data D. The value of w_k is nonnegative and with a sum $\left(\sum_{k=1}^{K} w_k\right)$ of 1. Finally, $g(y\,|\,f_k)$ represents the conditional probability distribution function (PDF) of y conditional on f_k. The conditional distribution, $g(y \mid f_k)$, can be represented as a normal distribution, $N(a_k + b_k f_k, \sigma_k^2$, where a_k and b_k are regression coefficients obtained through least square linear regression. Given equation 10, it is straightforward to derive the expected mean prediction and uncertainty interval. Detailed information on calculating terms in equation 10 is provided by Zhang et al. (2009).

The four evaluation coefficients described in the previous section were used to evaluate the performance of the deterministic expected mean prediction produced by BMA. Another evaluation coefficient, percentage of coverage (POC) of observations in the uncertainty interval, was used to evaluate the uncertainty intervals obtained by BMA. A smaller difference between POC and the expected coverage percentage of an uncertainty interval indicates better performance of the estimated uncertainty interval. The 95% uncertainty interval, which is expected to include 95% of the observations, was derived from equation 10 and evaluated using POC.

Results and Discussion

Stream Flow Estimates Using Surface Rain Gauges

Table 2 shows model performance measures for daily stream flow estimation during calibration and validation periods using precipitation data from two separate rain gauges (Royal Oak and Chestertown) located outside the GB watershed. Results indicated that stream flow was estimated best using precipitation data collected at the Chestertown gauge. Chestertown simulations produced higher NSE values of 0.49 and 0.73, compared to 0.42 and 0.54 for Royal Oak simulations, during the calibration and validation periods, respectively (table 2). RMSE values during calibration and validation were lower for Chestertown (0.87 and 1.07) than for Royal Oak (0.93 and 1.41) simulations, also indicating better estimation of stream flow using Chestertown rainfall data. Negative PBIAS values in table 2 show that stream flow was generally underestimated using both Royal Oak and Chestertown rainfall measurements. Systematic over- or under-estimations of streamflow using Royal Oak data were negligible during the calibration period.

These results can be explained first by the closeness of the Chestertown gauge to the watershed outlet (~26 km) compared to the Royal Oak gauge (~39 km), as shown in figure 2. This difference of 13 km (8 miles) may be enough to have significance; however, the directional location of the gauges with respect to the GB watershed and the general direction of storms in this region may have also played a role in capturing the most accurate rainfall patterns. Previous studies have shown that storm direction and velocity have pronounced effects on runoff hydrographs (Foroud et al., 1984; Singh, 1998; Tsanis et al., 2002; de Lima et al., 2003). Given that fact, it follows that rain gauges that are in a position to capture those effects will provide the most representative rainfall measurements.

Storms generally flow from west to east in temperate latitudes. However, according to the National Climatic Data Center (NCDC), prevailing winds in Maryland flow from the northwest quadrant during approximately nine months of the year (NESDIS, 2009). For this reason, precipitation collected at the Chestertown rain gauge (located northwest of GB) is more likely to resemble rainfall patterns in the German Branch watershed than Royal Oak measurements (located southwest of GB) during most of the year. This is illustrated by the monthly SSQs (eq. 5) between observed and simulated stream flow for Royal Oak and Chestertown in 2005 (table 3). During that year, SSQs compiled on a monthly basis for Royal Oak exceeded Chestertown monthly SSQs in 8 out of 12 months of the year, indicating that flow simulated using Chestertown data was 67% more accurate than Royal Oak flow simulations in 2005.

NEXRAD Stream Flow Estimates

Model performance measures for daily stream flow simulation using NEXRAD rainfall data are shown in table 4. Calibration results showed that, in most cases, SWAT estimated stream flow more accurately using NEXRAD precipitation data than rain gauge data (tables 2 and 4). This is likely due to the fact that the rain gauges were located outside of the watershed. The Chestertown gauge data provided better daily stream flow estimates than bias-corrected (BC) MPE (multi-sensor precipitation estimator) data during calibration as well as non-corrected (NC) MPE and inverse distance weighted (IDW) MPE during validation. This further demonstrated the representativeness of the Chestertown gauge data over the Royal Oak gauge data, which did not outperform any of the MPE datasets in estimating daily stream flow at the GB watershed outlet.

Table 2. Model performance measures for daily stream flow estimation for Royal Oak and Chestertown surface rain gauges. Values in parentheses are performance measures for non-calibrated models.[a]

	Rain Gauge	NSE	r^2	RMSE (cms)	PBIAS (%)
Calibration (2005-2006)	Royal Oak	0.42 (-4.44)	0.42 (0.24)	0.93 (2.85)	0.52 (59.20)
	Chestertown	0.49 (-0.71)	0.50 (0.36)	0.87 (1.60)	-15.33 (-1.39)
Validation (1 Jan. to 15 April 2007)	Royal Oak	0.54 (-1.04)	0.60 (0.34)	1.41 (2.96)	-11.94 (19.86)
	Chestertown	0.73 (-0.21)	0.75 (0.51)	1.07 (2.28)	-19.82 (-18.46)

[a] NSE = Nash-Sutcliffe coefficient of efficiency, RMSE = root mean square error, and PBIAS = percent bias.

Table 3. Sum of squares of residuals (SSQ) of the differences compiled on a monthly basis for 2005.

Rain Gauge	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Royal Oak	10.16	2.41	98.16	227.59	51.82	2.07	8.58	3.42	1.97	15.74	0.82	8.62
Chestertown	7.42	2.26	76.62	233.35	71.15	5.54	1.67	3.32	1.89	13.90	1.91	4.85

Table 4. Model performance measures for daily stream flow estimation using NEXRAD precipitation data. Values in parentheses are performance measures for non-calibrated models.[a]

	NEXRAD Data[b]	NSE	r^2	RMSE (cms)	PBIAS (%)
Calibration (2005-2006)	NC	0.58 (-0.90)	0.60 (0.40)	0.79 (1.68)	-28.44 (9.40)
	BC	0.46 (-2.36)	0.47 (0.36)	0.90 (2.24)	-11.80 (33.05)
	IDW	0.54 (-0.50)	0.56 (0.40)	0.83 (1.50)	-25.32 (3.65)
Validation (1 Jan. to 15 April 2007)	NC	0.73 (-2.71)	0.75 (0.49)	1.07 (3.99)	1.14 (41.11)
	BC	0.76 (-0.58)	0.76 (0.53)	1.02 (2.61)	-10.33 (10.74)
	IDW	0.68 (-0.13)	0.70 (0.50)	1.18 (2.21)	-18.91 (-4.51)

[a] NSE = Nash-Sutcliffe coefficient of efficiency, RMSE = root mean square error, and PBIAS = percent bias.
[b] NC = NEXRAD precipitation with no correction, BC = NEXRAD precipitation with bias correction, and IDW = NEXRAD precipitation with inverse distance weighted interpolation.

Results among MPE data were mixed. The model performed best using NC MPE data (NSE = 0.58) during the calibration period and using BC MPE data (NSE = 0.76) during the validation period (table 4). With such a small number of rain gauges available to correct NEXRAD data, there was not much improvement over NC MPE data. In addition, correction of NEXRAD data using distant gauge data can decrease the accuracy of radar data in local areas (Hunter, 1996). This can be seen in the daily stream flow time series plots for NC and IDW during the validation period (fig. 4). Events A, B, and C show significant degradation of flow predictions when using IDW corrected MPE data compared to NC MPE data. NEXRAD data generally produced underestimations of flow, especially during the calibration period. However, daily flow was mostly overpredicted during validation using NC MPE data. This may be explained by the fact that 2007 was a dry year (fig. 5) and the period used for validation was mostly dry (fig. 4), leading to overestimations during dry periods using NC MPE data.

If more than five gauges were available near the watershed, then the kriging interpolation methods of NEXRAD_SWAT could have been employed. Those methods may potentially provide better rainfall estimates due to their sophistication. However, since their relative benefit mainly increases with rainfall network density, they may render estimates similar to simple methods (e.g., BC and IDW) using low-density networks (Goudenhoofdt and Delobbe, 2009). Furthermore, having surface rain gauges located within the watershed would provide better estimates of stream flow due to more accurate measurements of rainfall.

Model performance measures for the different sources of rainfall without model calibration are shown in tables 2 and 4. There was no major improvement of baseline model behavior by using NEXRAD data. Although IDW had the best model performance without calibration, there was not much of an improvement over non-calibrated Chestertown gauge results. Hence, the use of NEXRAD data, in this case, did not eliminate the need for model calibration, and neither did these data reduce calibration efforts. Calibration improved model performance in every case.

The decision to use one source of rainfall data over the other depends on the availability of accurate rain gauge data, the variability of rainfall in the watershed, and the ability of NEXRAD data to accurately account for that variability. Overall results in this study indicate that, in the absence of properly designed rain gauge network data in a given watershed, the best set of rainfall data for use in watershed hydrologic assessment is NEXRAD data. The second option for sources of data in the absence of the watershed-based rain gauge network data is data obtained from the closest rain

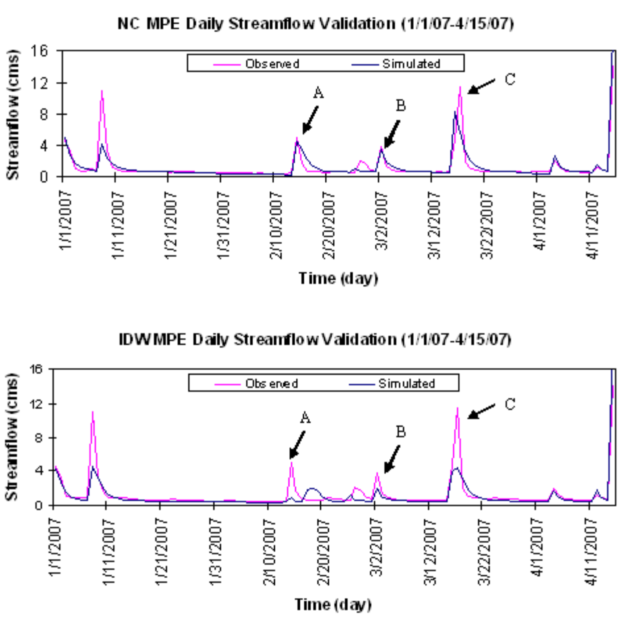

Figure 4. Times series plots showing NC MPE flow prediction (top) and IDW MPE flow prediction (bottom) for events A (14 Feb. 2007), B (2 March 2007), and C (17 March 2007).

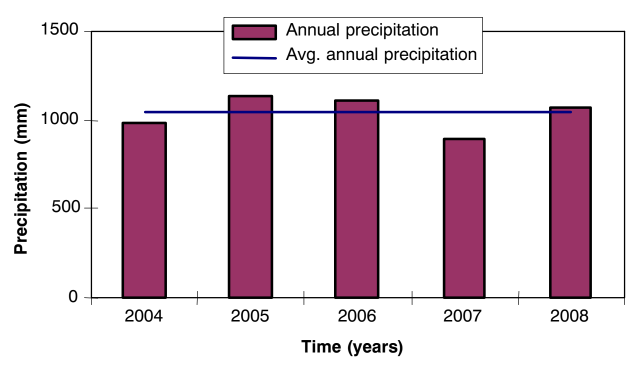

Figure 5. Observed annual precipitation at the Chestertown, Maryland, surface rain gauge (2004-2008).

Table 5. BMA performance measures for daily stream flow estimation.

	NSE	r^2	RMSE (cms)	PBIAS (%)	POC (%)
Calibration (2005-2006)	0.63	0.65	0.74	0	96.9
Validation (1 Jan. to 15 April 2007)	0.77	0.77	0.99	0.09	96.2

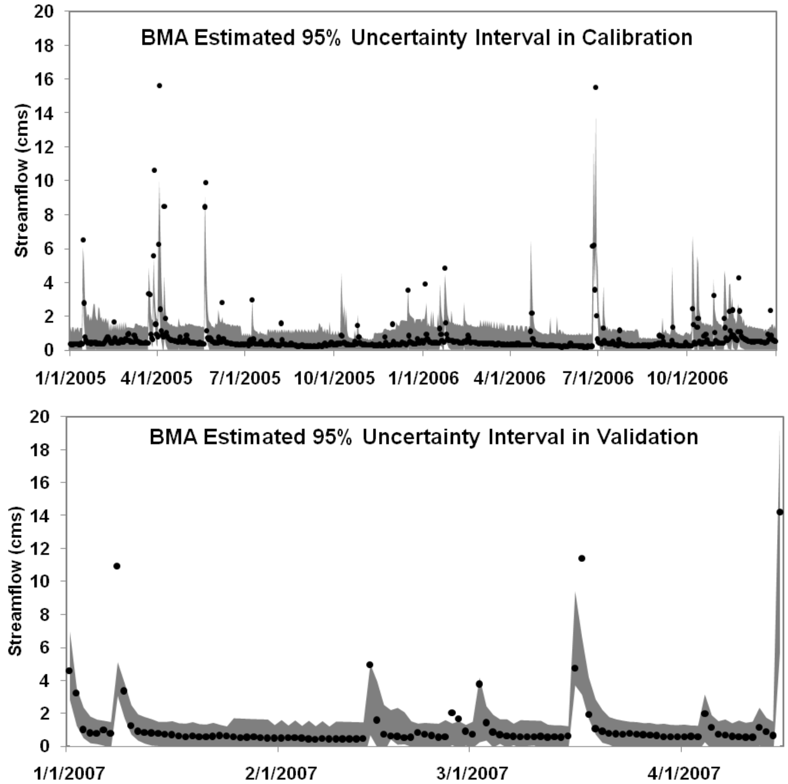

Figure 6. 95% uncertainty intervals estimated by BMA for the calibration and validation periods.

gauge to the watershed. In addition, one should consider the location of such a rain gauge with respect to the direction of storm travel in the region.

Bayesian Model Averaging Estimates

Based on tables 2 and 4, the model performed better during the validation period compared to the calibration period. It is risky to select one model for prediction. The BMA algorithm was used to combine the five simulations to derive expected mean prediction and 95% uncertainty intervals. The evaluation coefficients for BMA mean prediction are listed in table 5. NSE and r^2 values obtained by BMA are consistently larger than other predictions with one set of precipitation data in both calibration and validation periods. Meanwhile, BMA also obtained the smallest RMSE and PBIAS values compared to the statistics listed in tables 2 and 4. For uncertainty analysis, the 95% uncertainty intervals derived using equation 10 are shown in figure 6. Visually, the uncertainty intervals in figure 6 contain most of the observations. POC values are 96.9% and 96.2% in the calibration and validation periods, respectively (table 5). The difference between POC and expected coverage percentage is less than 2% for both periods. In general, the combination of multiple SWAT model simulations using different precipitation datasets produced better deterministic prediction than any simulation with one set of precipitation input data. The BMA algorithm implemented in this study also provided reasonable uncertainty estimation results, which is valuable for water resources related investigations and decision making processes.

Conclusions

In most cases of watershed hydrologic and water quality assessments, measured rainfall data are either sparse or nonexistent. For watersheds within which there is no well designed surface rain gauge network, this study demonstrated the importance of considering rain gauge proximity to the watershed and potentially the direction of storm patterns in the region when choosing the most representative rain gauge. In Maryland, storms travel from the northwest quadrant during most of the year. This factor along with the closeness of the gauge to the watershed may explain why the gauge located northwest of the watershed (Chestertown) provided better precipitation input than the gauge located southwest of the watershed (Royal Oak). However, further study of this notion should be explored in watersheds with denser rain gauge networks.

This study also showed that NEXRAD rainfall data can be a viable alternative to using rainfall data collected from surface rain gauges located outside of the watershed. NEXRAD MPE data produced comparable and, in most cases, better estimates of flow than rain gauge data. This is likely due to having a better network of radar grid cells located within the watershed boundaries, allowing NEXRAD to better account for spatial variability of rainfall. The surface rain gauges used in this study were located outside the watershed and thus were not able to represent the watershed rainfall distribution as would a well designed rain gauge network or properly kriged NEXRAD data. NEXRAD rainfall data can be a good alternative when rain gauge data are not available or where gauges are not located within the storm path of the watershed. As the quality of NEXRAD data is further improved, NEXRAD data will become increasingly suitable for use in hydrologic modeling.

Given the multiple precipitation datasets and corresponding simulations, we explored the combination of the multiple simulations using Bayesian model averaging. The results show that this Bayesian scheme can produce better deterministic prediction than any single simulation and can provide reasonable uncertainty estimation results. Overall, when there are several available precipitation datasets, it is suggested to combine the simulations.

Acknowledgements

Special recognition goes to Greg McCarty and Laura McConnell (CEAP, PIs). We also appreciate Megan Lang, Dean Hively, and other affiliates at the USDA-ARS Hydrology and Remote Sensing Laboratory (HRSL) in Beltsville, Maryland, for providing the 2 m resolution DEM, detailed land use data, and observed stream flow measurements. Joseph Ostrowski (NOAA, NWS) is also acknowledged for providing supplementary NEXRAD data.

References

Ahnert, P., M. Hudlow, E. Johnson, D. Greene, and M. Rosa Dias. 1983. Proposed "on-site" precipitation processing system for NEXRAD. In *Preprints of 21st Conf. on Radar Meteorology*, 378-385. Washington, D.C.: American Meteorological Society.

Ahnert, P., M. Hudlow, and E. Johnson. 1984. Validation of the "on-site" precipitation processing system for NEXRAD. In *Preprints of 22nd Conf. on Radar Meteorology*, 192-201. Washington, D.C.: American Meteorological Society.

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large-area hydrologic modeling and assessment: Part I. Model development. *J. American Water Resources Assoc.* 34(1): 73-89.

Bachman, L. J., B. Lindsey, J. Brakebill, and D. S. Powars. 1998. Ground-water discharge and baseflow nitrate loads of nontidal streams, and their relation to a hydrogeomorphic classification of the Chesapeake Bay watershed, Middle Atlantic Coast. *Water Resources Investigations Report* 98-4059. Washington, D.C.: U.S. Geological Survey.

de Lima, J. L. M. P., V. P. Singh, and M. I. P. de Lima. 2003. The influence of storm movement on water erosion: Storm direction and velocity effects. *Catena* 52(1): 39-56.

Di Luzio, M., and J. G. Arnold. 2004. Formulation of a hybrid calibration approach for a physically based distributed model with NEXRAD data input. *J. Hydrol.* 298: 136-154.

Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian. 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources* 30(5): 1371-1386.

Foroud, N., R. S. Broughton, and G. L. Austin. 1984. The effect of a moving rainstorm on direct runoff properties. *J. American Water Resources Assoc.* 20(1): 87-91.

Fulton, R. A, J. P. Breidenbach, D. J. Seo, D. A. Miller, and T. O'Bannon. 1998. The WSR-88D rainfall algorithm. *Weather and Forecasting* 13(2): 377-395.

Goudenhoofdt, E., and L. Delobbe. 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol. and Earth System Sci.* 13(2): 195-203.

Groisman, P. Y., and D. R. Legates. 1994. The accuracy of United States precipitation data. *Bull. American Meteorol. Soc.* 75(3): 215-227.

Habib, E., B. F. Larson, and J. Graschel. 2009. Validation of NEXRAD multisensory precipitation estimates using an experimental dense rain gauge network in south Louisiana. *J. Hydrol.* 373(3-4): 463-478.

Huebner, S., C. S. Pathak, and B. C. Hoblit. 2003. Development and use of a NEXRAD database for water management in south Florida. In *Proc. ASCE EWRI World Water and Environmental Resources Congress.* Reston, Va.: ASCE.

Hunter, S. M. 1996. WSR-88D radar rainfall estimation: Capabilities, limitations, and potential improvements. *NWA Digest* 20(4): 26-36.

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

Jayakrishnan, R., R. Srinivasan, C. Santhi, and J. G. Arnold. 2005. Advances on the application of the SWAT model for water resources management. *Hydrol. Proc.* 19(3): 749-762.

Kalin, L., and M. M. Hantush. 2006. Hydrologic modeling of an eastern Pennsylvania watershed with NEXRAD and rain gauge data. *J. Hydrol. Eng.* 11(6): 555-569.

Moon, J., R. Srinivasan, and J. H. Jacobs. 2004. Stream flow estimation using spatially distributed rainfall in the Trinity River basin, Texas. *Trans. ASAE* 47(5): 1445-1451.

Neary, V. S., E. Habib, and M. Flemming. 2004. Hydrologic modeling with NEXRAD precipitation in Middle Tennessee. *J. Hydrol. Eng.* 9(5): 339-349.

Neff, E. L. 1977. How much rain does a rain gage gage? *J. Hydrol.* 35(3-4): 213-220.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2005. *Soil and Water Assessment Tool Theoretical Documentation*. Version 2005. Temple, Tex.: USDA-ARS Grassland, Soil and Water Research Laboratory and Blackland Research Center.

Nelson, B. R, D. J. Seo, D. Kim, and J. J. Bates. 2006. Multisensor precipitation reanalysis. In *Proc. AMS 2006 Annual Meeting*, Paper No. 4.6. Washington, D.C.: American Meteorological Society.

NESDIS. 2009. Climatography of the United States No. 60: Climate of the states. Silver Spring, Md.: NOAA National Environmental Satellite, Data, and Information Service (NESDIS). Available at: http://cdo.ncdc.noaa.gov/climatenormals/clim60/states/Clim_MD_01.pdf. Accessed 28 July 2009.

NRCS. 2009. CEAP fact sheet. Washington, D.C.: USDA-NRCS. Available at: www.nrcs.usda.gov/technical/NRI/ceap/ceapgeneralfact.pdf. Accessed 28 October 2009.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski. 2005. Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* 133(5): 1155-1174.

Ramanarayanan, T. S., R. Srinivasan, and J. G. Arnold. 1996. Modeling Wister Lake watershed using a GIS-linked basin-scale hydrologic/water quality model. Santa Barbara, Cal.: National Center for Geographic Information and Analysis (NCGIA). Available at: www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/ramanarayanan_tharacad/my_paper.html. Accessed 3 June 2005.

Seo, D. J., and J. P. Breidenbach. 2002. Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeorology* 3(2): 93-111.

Singh, V. P. 1998. Effect of the direction of storm movement on planar flow. *Hydrol. Proc.* 12(1): 147-170

Skinner, C., F. Bloetscher, and C. S. Pathak. 2009. Comparison of NEXRAD and rain gauge precipitation measurements in south Florida. *J. Hydrol. Eng.* 14(3): 248-260.

Starks, P. J., and D. N. Moriasi. 2009. Spatial resolution effect of precipitation data on SWAT calibration and performance: Implications for CEAP. *Trans. ASABE* 52(4): 1171-1180.

Tobin, K. J., and M. E. Bennett. 2009. Using SWAT to model streamflow in two river basins with ground and satellite precipitation data. *J. American Water Resources Assoc.* 45(1): 253-271.

Tsanis, I. K., M. A. Gad, and N. T. Donaldson. 2002. A comparative analysis of rain-gauge and radar techniques for storm kinematics. *Advances in Water Resources* 25(3): 305-316.

van Griensven, A., and T. Meixner. 2007. A global and efficient multi-objective auto-calibration and uncertainty estimation method for water quality catchment models. *J. Hydroinformatics* 9(4): 277-291.

van Griensven, A., T. Meixner, S. Grunwald, T. Bishop, M. Di Luzio, and R. Srinivasan. 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J. Hydrol.* 324(1-4): 10-23.

Wang, X., H. Xie, H. Sharif, and J. Zeitler. 2008. Validating NEXRAD MPE and stage III precipitation products for uniform rainfall on the Upper Guadalupe River basin of the Texas hill country. *J. Hydrol.* 348(1-2): 73-86.

Winchell, M., R. Srinivasn, M. Di Luzio, and J. G. Arnold. 2007. *ArcSWAT Interface for SWAT2005-User's Guide.* Temple, Tex.: USDA-ARS Grassland, Soil and Water Research Laboratory and Blackland Research Center.

Young, C. B., A. Bradley, W. Krajewski, A. Kruger, and M. Morrissey. 2000. Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. *J. Hydrometeorology* 1(3): 241-254.

Zhang, X., and R. Srinivasan. 2009. GIS-based spatial precipitation estimation: A comparison of geostatistical approaches. *J. American Water Resources Assoc.* 45(4): 894-906.

Zhang, X., R. Srinivasan, and D. Bosch. 2009. Calibration and uncertainty analysis of the SWAT model using genetic algorithms and Bayesian model averaging. *J. Hydrol.* 374(3-4): 307-317.