MULTI-SCALE COMPARISON OF STAGE IV NEXRAD (MPE) AND GAUGE PRECIPITATION DATA FOR WATERSHED MODELING

K. Price, S.T. Purucker, and S.R. Kraemer

AUTHORS: U.S. EPA Office of Research and Development, Ecosystems Research Division, 960 College Station Rd, Athens, Georgia, 30605.
REFERENCE: *Proceedings of the 2011 Georgia Water Resources Conference*, held April 11–13, 2011, at the University of Georgia.

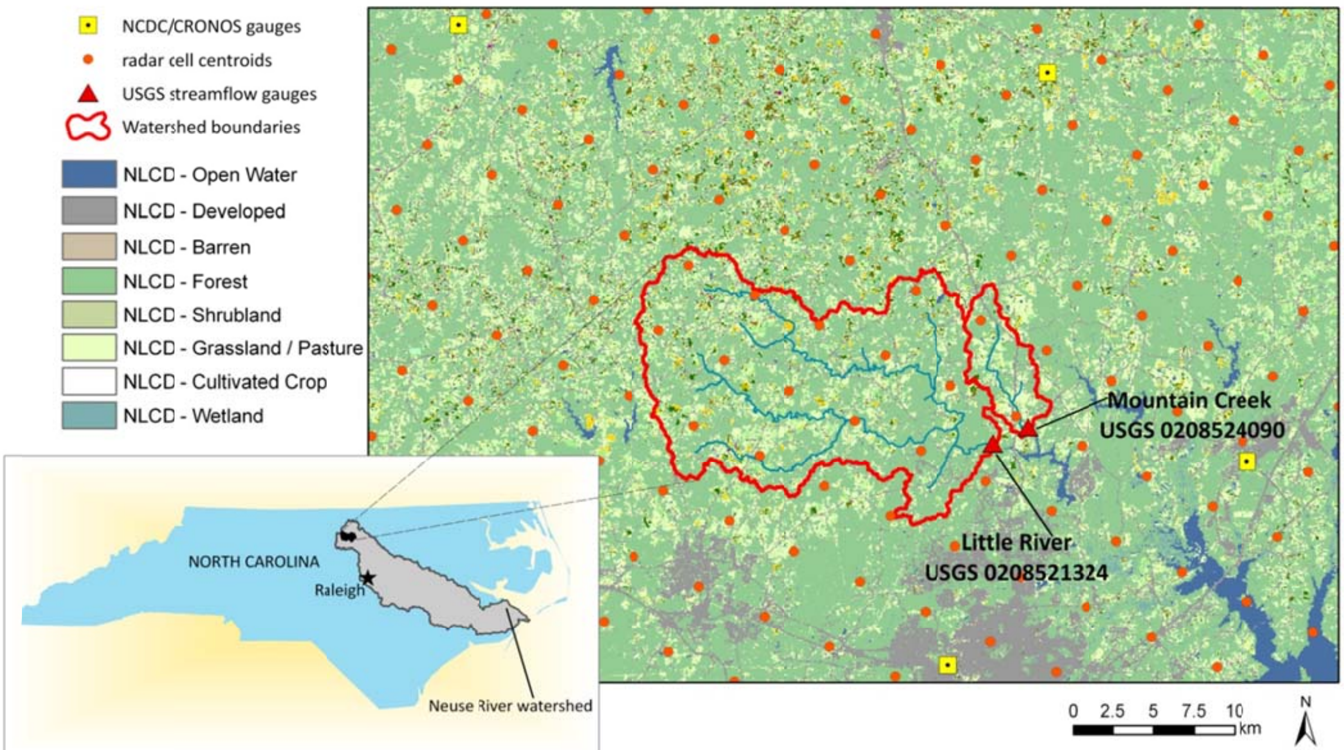
Abstract. Watershed hydrologic and fate-and-transport models are widely used to forecast water quantity and quality responses to alternative land use and climate change scenarios. The ability of such tools to forecast changes in ecosystem services with reasonable accuracy depends on calibrating reliable simulations of streamflow, which in turn require accurate climatic forcing data. Precipitation is widely acknowledged to be the largest source of uncertainty in watershed modeling. Most watershed models are designed to easily incorporate publicly-available precipitation data from rain gauges (e.g., data provided by the National Climatic Data Center), but several additional data products from ground-based radar and satellite-based sensors are now available and can potentially be used to generate more precise, spatially-explicit precipitation estimates. Here, we investigate whether the use of higher-resolution Multisensor Precipitation Estimator (MPE, also known as Stage IV NEXRAD) data can improve the accuracy of daily streamflow simulations using the Soil and Water Assessment Tool (SWAT) watershed hydrology model. Simulated vs. observed streamflow and model calibrations are compared for two Piedmont sub-basins of the Neuse River in North Carolina (21 and 203 km^2 watershed area) for an 8 year simulation period (January 1, 2002 to August 31, 2010). MPE simulations led to more accurate simulations of daily streamflow magnitude and frequency measures than gauge data, and differences were more pronounced in the smaller watershed. Compared with USGS-observed flows, MPE simulations produced R^2 values of 0.64 and 0.54 for the larger and smaller watershed, respectively, while gauge data produced R^2 values of 0.19 in both watersheds. Nash-Sutcliffe Efficiency and other goodness-of-fit indices also showed much better simulations associated with MPE data. Additionally, the temporal structure of MPE-simulated streamflows more closely approximated that of the observed streamflows. These results are likely extendable to the Piedmont of the broader southeastern U.S. Ongoing research on this topic investigates additional spatial and temporal scales, as well as additional precipitation data types.

INTRODUCTION AND BACKGROUND

Simulation of streamflow, sediment, and dissolved constituents requires climatic forcing data. Temperature can be reasonably estimated for hydrologic modeling from a sparse network of stations within and surrounding the study watershed (Attorre et al., 2007). However, accurate representation of precipitation spatial and temporal variability from available resources has proven to be a challenge for hydrologic modeling. Failure to incorporate such variability potentially introduces large amounts of uncertainty to hydrologic and fate-and-transport modeling efforts (Jordan, 2000; Andréassian et al., 2001; Schuurmans and Bierkens, 2007; Villarni et al., 2008).

There are two predominant approaches to evaluating precipitation data sources. In the first approach, interpolations or area averages from a network of precipitation gauges is treated as the “actual” precipitation as a basis of comparison for other sources of precipitation data (e.g., sparser gauge networks, radar, satellite). In the absence of a very dense network of rain gauges, it is inappropriate to treat any rainfall data source as the actual precipitation, because of the known uncertainty of all available data types (Schuurmans and Bierkens, 2007; Villarni et al., 2009; Habib et al., 2009; Mandapaka et al., 2009). Because of this, many studies have adopted an alternative approach using streamflow simulations from a watershed model as an independent assessment of the precipitation data accuracy (e.g., Borga, 2002; Su et al., 2008; Schuurmans and Bierkens, 2007; Tobin and Bennett, 2007; Starks and Moriasi, 2009). In this approach, the disparities between streamflow simulations are attributed to differences in precipitation data accuracy. This technique is especially appropriate for studies whose objective is to evaluate the potential of various precipitation data sources as forcing data for hydrologic modeling and is employed in this study.

Recent studies evaluating and comparing precipitation data sources have shown mixed results, suggesting that there is no universally optimal precipitation data source for hydrologic modeling. Hossain et al. (2004)



showed that the uncertainty of MPE and gauge data was similar for storm events in a 116 km^2 mountainous catchment. Similarly, rain gauge and satellite rainfall data were shown to give very similar storm totals for an extreme event in Vietnam (Valeriano et al., 2009). MPE data were shown to have a conditional bias when compared against a dense network of precipitation gauges, where low-intensity events were overestimated and high-intensity events were underestimated in southern Louisiana (Habib et al., 2009). In that study, the relationship between MPE and gauge data demonstrated significant scatter, particularly during low-intensity events.

Based on the results of these studies, there is reason to believe that there may be spatial scale dependencies for the accuracy of these data types. While several studies have demonstrated scale dependencies of precipitation data accuracy (e.g., Villarni et al., 2008), no known studies explicitly address these scale issues in a comparative framework incorporating multiple precipitation data types. There may be thresholds of optimal applicability of these various data types, and knowledge of such thresholds is critical for reducing uncertainty in hydrologic and fate-and-transport modeling efforts. As various landscape attributes are very important to the partitioning of overland flow, infiltration, and recharge, inaccurate spatial distribution of rainfall may lead to erroneous streamflow simulations. This is a particular concern in fate-and-transport modeling (Das et al., 2008).

Presented here are the prelimary results of research that evaluates streamflow simulations using precipitation data from multiple sources, at two spatial scales of watershed hydrologic model operation. The objective of these comparisons was to determine whether higher resolution precipitation data produces superior streamflow simulations in the Piedmont, and, if so, if these improvements in streamflow resolution are scale-dependent. Two types of precipitation data are compared in this study: 1) Gauge data from the National Climate Data Center (NCDC) and North Carolina network, and 2) Multisensor Precipitation Estimate (MPE). Gauge data are measurements of precipitation depth collected in a rain gauge operated and maintained by the National Weather Service or by cooperative institutions and agencies. Rainfall totals measured by these gauges represent the near-actual amount of rainfall collected at the scale of a localized point, but the ground-based gauge networks tend to be relatively sparse. MPE data are doppler radar precipitation data (known as both NEXRAD and WSR-88D) that have been adjusted based on gauge data from the highest accuracy precipitation stations. These data are available in grid format, with total precipitation reported for each

NEXRAD pixel (4 x 4 km), resulting in much higher spatial resolution than available for gauges (Figure 1).

METHODS

This study includes two sub basins of the Neuse River watershed in North Carolina, Mountain Creek (21 km^2) and Little River (203 km^2). Both watersheds are located in the Piedmont physiographic province, and are characterized by relatively low relief and moderately disturbed land use (Figure 1).

The Soil and Water Assessment Tool (SWAT; (Gassman et al., 2007) was used to simulate daily streamflow using both precipitation inputs for Mountain Creek and Little River (Figure 1), for which streamflow is gauged by the U.S. Geological Survey (USGS) during the entire study period. Standard data requirements for SWAT include temperature and precipitation time series data, as well as spatial coverages of topography, soils and land cover. In order to compare streamflow simulations using the two separate precipitation data types, all other SWAT inputs were identical during model runs. A 10 m digital elevation model (DEM), 2009 Cropland Data Layer (CDL) 30 m land cover data, and SSURGO digital county soil data were obtained from the Geospatial Data Gateway (http://datagateway.nrcs.usda.gov/). Daily maximum and minimum temperature data were obtained from the National Climatic Data Center (NCDC, http://www.ncdc.noaa.gov/oa/ncdc.html) for all stations active in the upper Neuse during the study period. MPE data were obtained from the Earth Observing Laboratory (http://data.eol.ucar.edu/). Additional meteorological parameters (wind speed and relative humidity) were simulated using SWAT's weather generator.

For both temperature and precipitation data, gauging stations missing > 10% of observations during the study period were not included. For the stations that were missing < 10% of observations, missing values were filled from the nearest station with available data. SWAT was designed to incorporate rainfall data in time series form, and spatially associated with point stations (as in standard rain gauge data). It was thus necessary to manipulate the MPE data format for incorporation into the model: After obtaining daily MPE grids in GRIB format, the files were converted from GRIB to NetCDF format using the free "degrib" program available from NOAA (http://www.nws.noaa.gov/mdl/degrib/). A program was written in R statistical software to clip the MPE grid by the watershed boundary (including a 5 km buffer) and create a time series of daily precipitation data for each individual grid cell in the study area. The coordinates of the centroid of each cell were treated as the point associated with each time series, meaning that SWAT incorporated the MPE data as though there were a gauging station in the center of each 4 x 4 km radar pixel.

ArcSWAT 2009 was used for watershed delineation and hydrologic response unit (HRU) definition. In SWAT, the nearest station or radar cell centroid is assigned to a given sub-basin, and each daily precipitation total is assumed to fall uniformly over the sub-basin. In this study, watershed delineation was performed to achieve an average sub-basin size of approximately 4 km^2. This ensured that information from each radar pixel and available rain gauge were incorporated into the streamflow simulation. This resulted in four and 50 sub-basins for Mountain Creek and Little River, respectively. HRU definition included 5% thresholds for minimum coverage of soil, slope, and land cover class, with all classes of wetland and urban land use retained. Calibration was performed on the period 2002-2006, with the best calibration parameters determined by the coefficient of determination (R^2), modified coefficient of determination (bR^2), Nash-Sutcliffe Efficiency (NSE), modified Nash-Sutcliffe Efficiency (mNSE), mean error, root mean squared error, and visual correspondence of simulated low and high flows compared with the observed hydrograph (Krause et al., 2005). Validation indices and precipitation data comparisons were based on simulations for the entire study period. Calibrations targeted the most sensitive parameters for each combination of data type and watershed, which included indices related to curve number (CN2), baseflow recession (Alpha_Bf), groundwater depth required for return flow (GWQMN), and soil available water (Sol_AWC).

Streamflow simulations generated by gauge and MPE data were compared separately for each watershed. Summary statistics and indices of low and high flows were compared between observed flows and simulations using both precipitation data types. Paired t-tests were performed to evaluate significance of differences between streamflows simulated with each data type, and between simulated and observed flows. The temporal structures of simulated and observed flows were examined, as streamflow temporal variability is related to the environmental flows concept and bears important implications for water supply, water quality, and habitat availability.

RESULTS AND DISCUSSION

MPE precipitation data produced more accurate streamflow simulations than gauge data, by all standard goodness of fit metrics in both watersheds (Table 1, Figure 2). Most notably, the MPE data produced strikingly good fits with uncalibrated flows (R^2 0.63 and 0.53, Nash-Sutcliffe Efficiency 0.47 and 0.49). The gauge data tended to produce overestimates of streamflows. Improvements from using higher-resolution precipitation data were evident at

Table 1. Goodness of fit and summary statistics

Mountain Creek	USGS observed	MPE uncalibrated	NCDC uncalibrated	MPE calibrated	NCDC calibrated	MPE entire pd.	NCDC entire pd.
goodness of fit statistics							
R^2	-	0.63	0.11	0.64	0.19	0.49	0.18
bR^2	-	0.22	0.01	0.32	0.06	0.31	0.07
Nash-Sutcliffe Efficiency	-	0.47	-0.38	0.59	0.14	0.46	0.08
modified Nash-Sutcliffe	-	0.37	0.08	0.40	0.27	0.34	0.15
mean error (bias)	-	-0.04	0.11	-0.03	0.02	0.01	0.05
RMSE	-	0.51	0.67	0.45	0.65	0.48	0.62
flow descriptive statistics							
mean flow	0.19	0.15	0.24	0.16	0.21	0.19	0.23
median flow	0.05	0.09	0.14	0.07	0.06	0.08	0.12
5th percentile flow	<0.01	0.01	0.01	<0.01	<0.01	<0.01	<0.01
99th percentile flow	2.13	1.06	1.80	1.86	1.52	2.21	2.16
flow standard deviation	0.70	0.28	0.50	0.42	0.46	0.56	0.48
Little River	USGS observed	MPE uncalibrated	NCDC uncalibrated	MPE calibrated	NCDC calibrated	MPE entire pd.	NCDC entire pd.
goodness of fit statistics							
R^2	-	0.53	0.18	0.54	0.19	0.54	0.20
bR^2	-	0.24	0.09	0.28	0.08	0.29	0.09
Nash-Sutcliffe Efficiency	-	0.49	-0.14	0.53	0.01	0.54	0.00
modified Nash-Sutcliffe	-	0.38	0.20	0.39	0.31	0.38	0.19
mean error (bias)	-	-0.59	0.26	-0.55	-0.06	-0.21	0.36
RMSE	-	4.07	6.07	3.91	5.66	3.81	5.60
flow descriptive statistics							
mean flow	2.16	1.57	2.41	1.61	2.09	1.79	2.30
median flow	0.85	0.83	1.32	0.72	0.99	0.98	1.02
5th percentile flow	<0.01	0.04	0.04	0.03	0.01	0.03	0.03
99th percentile flow	24.2	24.1	20.8	16.6	17.8	16.8	21.5
flow standard deviation	5.69	3.17	5.62	3.80	4.83	3.76	4.87

Goodness of fit statistics calculated using the HydroGOF package in R (from Krause et al., 2005, *Adv. Geosciences*)

All flow descriptive statistics are in m^3s^{-1}; "calibrated" period is 2002-2006, "entire" period is 2002-2010

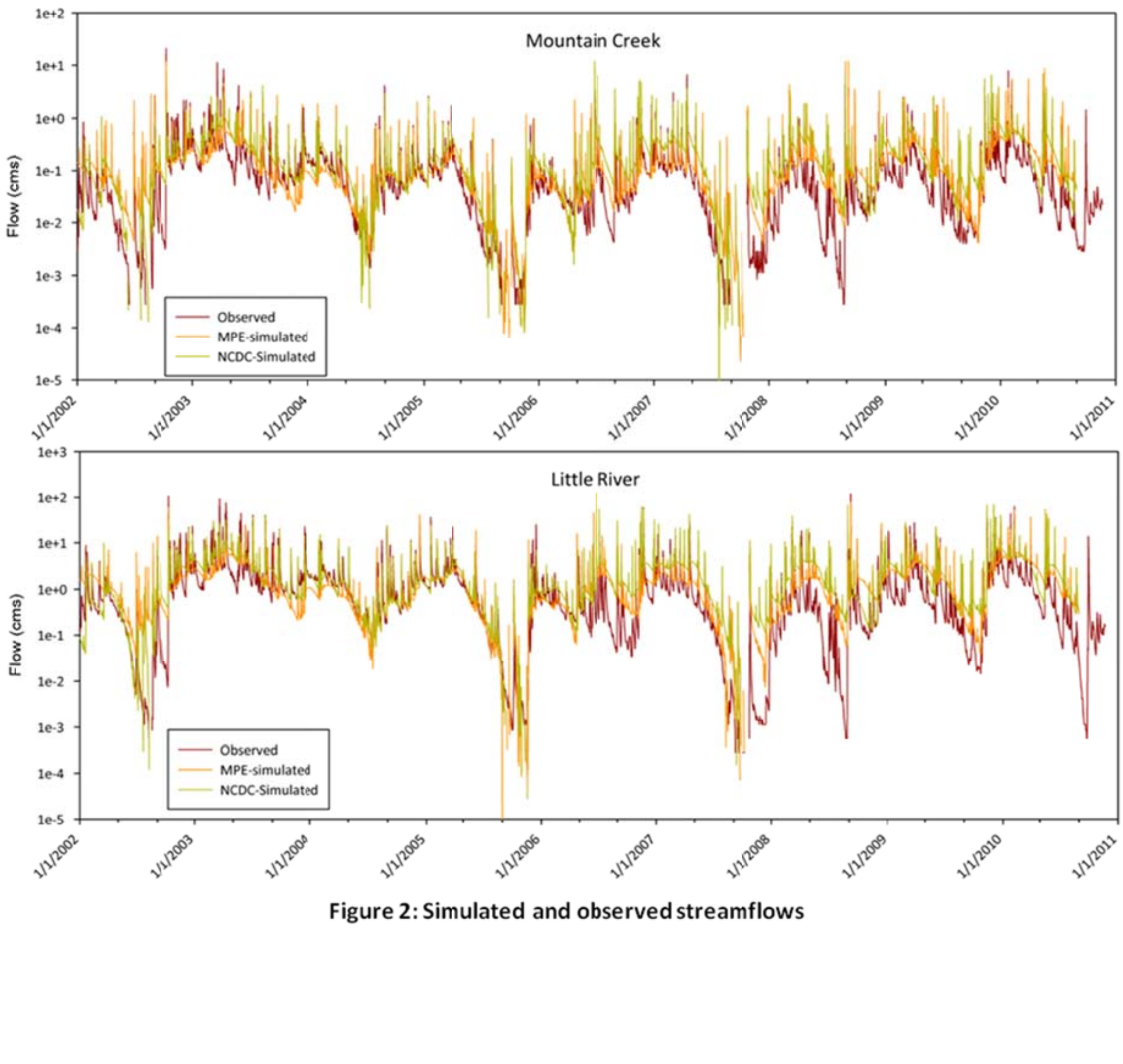
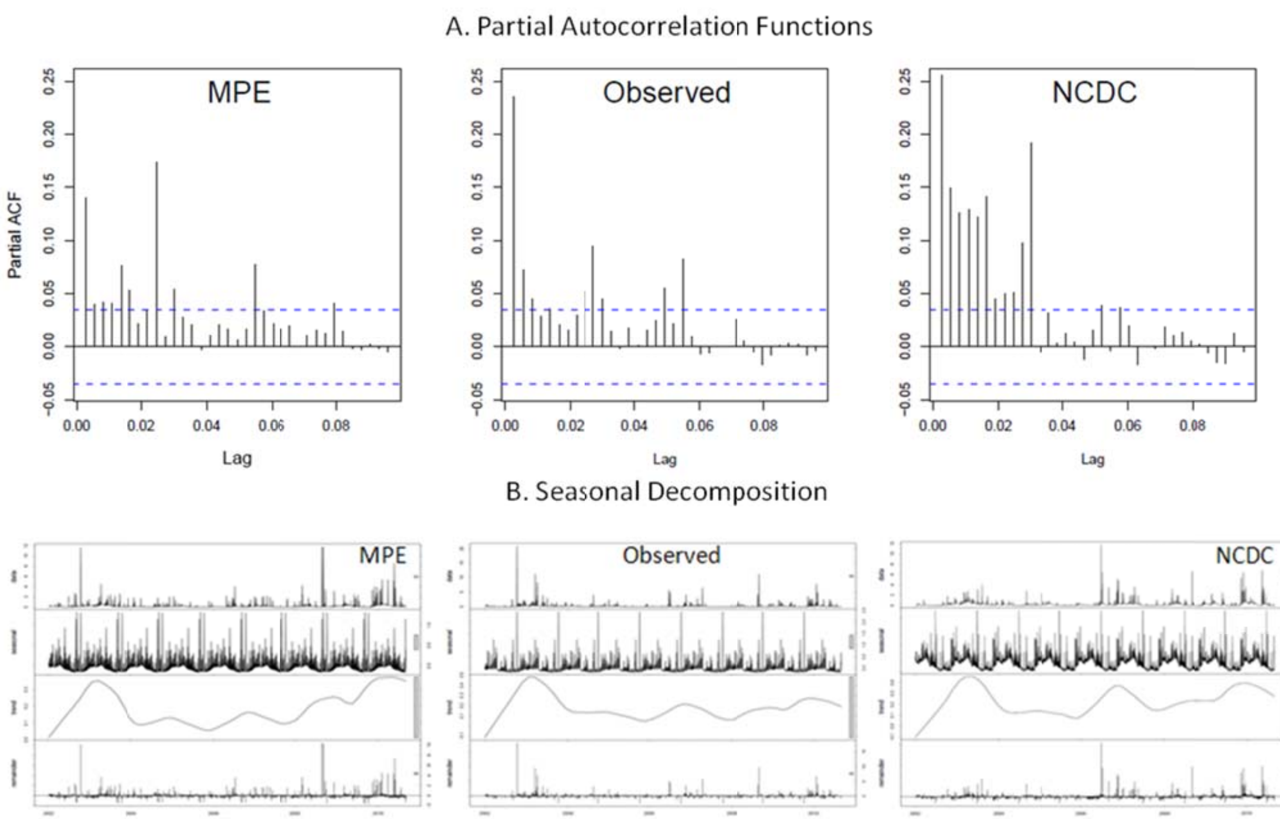


Figure 2: Simulated and observed streamflows

both watershed scales, but were more pronounced in the ller watershed (Mountain Creek), than in the larger watershed (Little River). Further comparisons at additional watershed scales will show whether there is consistent scale-dependency for improved streamflow simulations using higher-resolution precipitation data. To address this, the precipitation comparison will be extended to two larger watersheds in the Neuse River system: Neuse River near Clayton (USGS 02087500, 2,979 km^2) and Neuse River near Fort Barnwell (02091814, 10,100 km^2).

Paired t-tests indicated that differences between all combinations of simulated and observed flows, as well as all combinations of simulations from various data sources, were statistically significant at the $p<0.001$ level. The most striking aspect of these results is the difference between the observed flows and the uncalibrated simulations with the two data types (Table 1). In all cases, the uncalibrated MPE flows produced much higher R^2 and NSE than even the calibrated gauge simulations. As more complex calibration strategies are explored, it may prove that gauge data can achieve similarly successful simulations to the MPE data.

Spectral analysis showed that both data types produced realistic periodic variation in streamflows, as compared with observed flows (Figure 3). Streamflow periodicity is of prime importance in environmental flows management for ecosystem service protection (O'Keefe, 2009). These results suggest that capturing general patterns of streamflow magnitude may be possible with gauge data, to a greater extent than capturing the exact flow timing necessary for model calibration. The partial autocorrelation function (pacf) was used to determine the extent of the lag within the Mountain Creek time series (Figure 3a). The significant lag for the observed data is strong for three days, and persists for up to five days, whereas the significant lags for the gauge-simulated data are rather uniformly strong for up to 11 days, and the lags for the MPE-simulated data show significant lags through one week. However, the calibration has been performed using goodness of fit statistics for the flow magnitude and not the temporal structure. Therefore, such differences between the selected model and the observed data are possible. Overall pacf behavior suggests a smoothing effect of the model when compared to the observed flow time series, longer periods of self-similar flow regimes occur in the simulated flows than in the observed data, but this effect is less noticeable in the MPE-simulated streamflows than in the gauge-simulated flows. It may be possible to design calibration diagnostic statistics that explicitly account for temporal structure to improve assessment endpoints consistent with the environmental flows concept. We also performed a seasonal decomposition of a locally



weighted scatterplot smoothing of the observed and modeled time series. (Figure 3b). This yields a seasonal component, a moving-average determined trend, and residuals from the seasonal plus trend fit that represent irregular components. These figures demonstrate reasonable similarity between the time series signal of the observed time series compared to both sets of the simulated data. The seasonal decomposition shows stronger periodicity for the observed data, however, the modeled time series do reproduce the main features of the trend rather well. The residuals are much more variable for the observed time series loess when compared to the model, demonstrating the "smoothing" effect of the model compared to the observed flow data.

CONCLUSIONS

The prelimary results of this multi-scale comparison of precipitation data sources in watershed modeling show that MPE data generate more accurate streamflows than gauge data. We interpret this difference to be due to the much higher spatial resolution of the MPE data, as opposed to error in the gauge data. Spectral characteristics were reasonably represented by both data types, but indicated an overall smoothing of temporal variability in modeled streamflow, irrespective of precipitation data type. Ongoing research will incorporate additional watershed sizes and precipitation data types, such as PRISM (Daly et al., 2002) and the emergent, very high resolution (1 x1 km) Q2 data (http://nmq.ou.edu/). Further research will also investigate the role of data resolution and analysis of all data types at varied simulation timesteps.

ACKNOWLEDGMENTS

Valuable advice and assistance with this project were provided by several scientists within the U.S. EPA's Office of Research and Development, especially Robin Dennis and Ellen Cooter of the Atmospheric Modeling and Analysis Division, Chris Knightes, John Johnston, and Mike Cyterski of the Ecosystems Research Division, and Heather Golden of the Ecological Exposure Research Division. Ken Howard and Heather Moser (NOAA- National Severe Storms Laboratory) also provided assistance. Feedback on drafts of this manuscript was provided by Mike Cyterski and two anonymous reviewers. This paper has been reviewed in accordance with the U.S. EPA's peer and administrative review policies and has been approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the agency, nor does the mention of trade names or commercial products constitute endorsement or recommendation for use.

REFERENCES

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I. and Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. Journal of Hydrology, 250: 206-223.

Attorre, F., Alfo, M., De Sanctis, M., Francesconi, F. and Bruno, F., 2007. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. International Journal of Climatology, 27: 1825-1843.

Borga, M., 2002. Accuracy of radar rainfall estimates for streamflow simulation. Journal of Hydrology, 267(1-2): 26-39.

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L. and Pasteris, P., 2002. A knowledge-based approach to the statistical mapping of climate. Climate Research, 22: 99-113.

Das, T., Bardossy, A., Zehe, E. and He, Y., 2008. Comparison of conceptual model performance using different representations of spatial variability. Journal of Hydrology, 356(1-2): 106-118.

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

Habib, E., Larson, B.F. and Graschel, J., 2009. Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in south Lousiana. Journal of Hydrology, 373: 463-478.

Hossain, F., Anagnostou, E.N., Dinku, T. and Borga, M., 2004. Hydrological model sensitivity to radar rainfall estimation uncertainty. Hydrological Processes, 18: 3277-3291.

Jordan, P., 2000. Sampling errors in radar estimates of rainfall. Journal of Geophysical Research - Atmospheres, 105(D2): 2247-2257.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences 5:89-97.

Mandapaka, P.V., Krajewski, W.F., Ciach, G.J., Villarini, G. and Smith, J.A., 2009. Estimation of radar-rainfall error spatial correlation. Advances in Water Resources, 32(7): 1020-1030.

O'Keefe, J.O. 2009. Sustaining river ecosystems: balancing use and protection. Progress in Physical Geography 33(3): 339-357.

Schuurmans, J.M. and Bierkens, M.F.P., 2007. Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. Hydrology and Earth System Sciences, 11(2): 677-693.

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

Su, F., Hong, Y. and Lettenmaier, D.P., 2008. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. Journal of Hydrometeorology, 9: 622-640.

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

Valeriano, O.C.S. et al., 2009. Flood simulation using different sources of rainfall in the Huong River, Vietnam. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 54(5): 909-917.

Villarni, G., Mandapaka, P.V., Krajewski, W.F. and Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research, 113(D11102).