Effects of Resolution of Satellite-based Rainfall Estimates on Hydrologic
Modeling Skill at Different Scales

Humberto Vergara
Department of Civil Engineering and Environmental Science, and Advanced Radar
Research Center, University of Oklahoma, and National Severe Storms Laboratory,
Norman, Oklahoma

Yang Hong
Department of Civil Engineering and Environmental Science, and Advanced Radar
Research Center, University of Oklahoma, Norman, Oklahoma

Jonathan J. Gourley
National Severe Storms Laboratory, Norman, Oklahoma

Emmanouil N. Anagnostou, Viviana Maggioni and Dimitrios Stampoulis
Department of Civil Engineering and Environmental Science, University of Connecticut,
Storrs, Connecticut

Pierre-Emmanuel Kirstetter
Department of Civil Engineering and Environmental Science, and Advanced Radar
Research Center, University of Oklahoma, and National Severe Storms Laboratory,
Norman, Oklahoma

Revised Draft

Submitted electronically as an article to: Journal of Hydrometeorology

08 August 2013

1 Corresponding Author: Yang Hong, National Weather Center ARR Suite 4610, 120 David L. Boren Blvd, Norman, OK 73072-7303 Tel: 405 325-3644; Email: yanghong@ou.edu
Abstract

Uncertainty due to resolution of current satellite-based rainfall products is believed to be an important source of error in applications of hydrologic modeling and forecasting systems. A method to account for input’s resolution and to accurately evaluate the hydrologic utility of satellite rainfall estimates is devised and analyzed herein. A radar-based Multi-sensor Precipitation Estimator (MPE) rainfall product (4-km, 1-hr) was utilized to assess the impact of resolution of precipitation products on the estimation of rainfall and subsequent simulation of streamflow on a cascade of basins ranging from approximately 500 km\(^2\) to 5000 km\(^2\). MPE data were resampled to match TRMM’s 3B42RT satellite rainfall product resolution (25-km, 3-hr) and compared with its native resolution data to estimate errors in rainfall fields. It was found that resolution degradation considerably modifies the spatial structure of rainfall fields. Additionally, a sensitivity analysis was designed to effectively isolate the error on hydrologic simulations due to rainfall resolution using a distributed hydrologic model. These analyses revealed that resolution degradation introduces a significant amount of error in rainfall fields, which propagated to the streamflow simulations as magnified bias and dampened aggregated error (Root Mean Square Errors - RMSE). Furthermore, the scale dependency of errors due to resolution degradation was found to intensify with increasing streamflow magnitudes. The hydrologic model was calibrated with satellite- and original-resolution MPE using a multi-scale approach. The resulting simulations had virtually the same skill, suggesting that the effects of rainfall resolution can be accounted for during calibration of hydrologic models, which was further demonstrated with 3B42RT.

**Keywords:** QPE resolution impacts, satellite QPE, hydrologic model calibration, basin scales
1. Introduction

Satellite-based rainfall products have become an important resource for a broad variety of hydrological applications over the globe (e.g., flood forecasting, assessing water resources, water management; Hong et al. 2006; Hong et al. 2007; Su et al. 2008; Pereira Filho et al. 2010; see also https://servirglobal.net/Global.aspx), particularly for regions lacking in-situ observational systems such as rain gauge networks. The success of current satellite-based rainfall estimates lies in their capability to resolve spatial patterns at scales largely unachievable on a global or continental scale by rain gauge networks. However, these estimates are accompanied by a host of uncertainties associated with the indirectness of distant radiance measurements and the platforms’ limitations to capture the high spatiotemporal variability of the precipitation-related observations (McCollum et al. 2002; Sapiano and Arkin 2009; Gourley et al. 2010). Although progress has been made on characterizing the statistical properties of satellite retrieval error (e.g. Kidd et al. 2003; Hossain and Anagnostou 2006a, 2006b; Dinku et al. 2010; Scheel et al. 2011) and their interaction with hydrologic processes and basin characteristics represented in hydrological models (e.g. Hossain and Lettenmaier 2006; Hong et al. 2007; Hossain and Huffman 2008; Nikolopoulos et al. 2010; Gourley et al. 2011; Maggioni et al. 2011), much work is still necessary for an optimal use of satellite-based precipitation in hydrology (Hossain and Katiyar 2008).

Since rainfall is the main forcing for hydrologic models in the generation of runoff, it is critical to understand how uncertainties associated to its estimation will propagate through the non-linear dynamics embedded in these models. Arguably, the effects of error in rainfall volume on the simulation of hydrologic variables, such as
streamflow, are rather intuitive. The spatial and temporal patterns inherent to rainfall, on the other hand, have been the interest of many hydrological studies (e.g. Ogden and Julien 1993; Obled et al. 1994; Koren et al. 1999; Arnaud et al. 2002; Smith et al. 2004; Segond et al. 2007; Younger et al. 2009). Moreover, the association of errors in volume and spatiotemporal representation of rainfall estimates with limitations in the precipitation sampling process has been extensively explored in relation to the effect on hydrologic applications (e.g. Kouwen and Garland 1989; Krajewski et al. 1991; Obled et al. 1994; Ogden and Julien 1994; Faurès et al. 1995; Dinku et al. 2002; Syed et al. 2003; Wang et al. 2009; Kirstetter et al. 2010; Mohamoud and Prieto 2012). This uncertainty stemming from sampling errors is an important drawback of the use of satellite-based rainfall estimates. Several studies have already pointed out that satellite rainfall products exhibit significant and complex uncertainty at high spatial and temporal resolutions (e.g., Kidd et al. 2003; Hossain and Anagnostou 2004; Nijssen and Lettenmaier 2004; Hong et al. 2006; Scheel et al. 2011). Thus, a natural consequence is that although the error variability in coarse resolution rainfall products may be smoothed out, their resolution may not be acceptable at the scales required for some hydrologic modeling applications (e.g., flash flood forecasting). It has been shown that the spatial scale of the application can impact the propagation of this uncertainty in flood modeling (e.g., Nikolopoulos et al. 2010). Therefore, understanding how satellite rainfall retrieval uncertainty manifests in hydrologic prediction requires studying the effects of basin scale and spatio-temporal resolution on the rainfall error. Accounting for the error due to rainfall input resolution in an explicit manner has proven to be fundamental for evaluating satellite rainfall inputs’ hydrologic potential (Gourley et al. 2011).
In this work, we expand the work of Gourley et al. (2011) by presenting a methodology to explicitly account for the effects of rainfall estimates’ resolution in hydrologic modeling, whose basis is a parameter estimation procedure that considers basin scale. As the effect of satellite rainfall resolution on hydrologic simulation is a significant one, this study promotes its consideration when evaluating satellite rainfall inputs in a hydrologic context. A systematic analysis of the effects of current satellite rainfall products’ resolutions on the hydrologic modeling is performed for a cascade of basins of small to medium size (500-5000 km²). It must be recognized that the errors in the hydrologic simulation result from a complex interaction between the uncertainty in the input (i.e., rainfall), the model structure and approximations, the estimation of model parameters and other variables (e.g., potential evapotranspiration), and observations (e.g., gauged streamflow). In the present study, the effect of satellite rainfall products’ resolution was effectively isolated from other sources of uncertainty through a synthetic experiment with a hydrologic model used in operational settings. Studying the errors in hydrologic simulation due to rainfall resolution alone can serve the purpose of indicating the appropriateness of satellite product resolutions regarding basin scale and streamflow magnitude (e.g., water resources management vs. high-impact flooding events). Likewise, it can help understand limitations and indicate possible improvements in the next generation of satellite rainfall products from the Global Precipitation Measurement (GPM) mission. In a similar work, Nijssen and Lettenmaier (2004) explored the impact of sampling error on the precipitation volume and its subsequent effect on the hydrologic modeling of a cascade of basins using gridded gauge-based estimates and a macroscale hydrologic model. In the present study, the impact of satellite product resolution on the
spatial variability and structure of rainfall estimates is additionally analyzed.

The key aspects that this work seeks to explore can be put into context as follows:

Given that satellite precipitation products are available in regions where other sources of rainfall observations are sparse or non-existent, it is critical to establish their usefulness for hydrologic applications. A common approach in determining the hydrologic utility of satellite rainfall consists of calibrating a hydrologic model to the product subject to evaluation (see for example Yilmaz et al. 2005; Artan et al. 2007; Yong et al. 2010). While this method can show the performance of the satellite product within the particular modeling setup, the utility of satellite rainfall is obscured by the calibration strategy and the model’s ability to accommodate the observational data (e.g., streamflow) during estimation of its parameters. An alternative method consists of calibration of the hydrologic model to what is considered the “true” rainfall, followed by an evaluation of the hydrologic simulation’s performance forced by the satellite product (see for example Su et al. 2008; Li et al. 2009; Gourley et al. 2011). In this case, the differences in hydrologic skill stem from inherent abilities to estimate rainfall of each product and their corresponding spatiotemporal resolutions. Therefore, eliminating the latter conditioning factor in the evaluation process is required to accurately reveal the hydrologic utility of satellite precipitation products. In other words, an accurate evaluation of the hydrologic utility of satellite precipitation products as a function of basin scale and streamflow magnitude is herein defined as one that can account for the effects of the precipitation product’s resolution. In order to address this issue, this study aims to establish how important the effects of rainfall estimates’ resolution are for hydrologic applications, whether they can be accounted for or not, and whether this can help to determine the
hydrologic utility of satellite rainfall estimation. Several scientific questions are in order:
1) What is the impact of sampling error on rainfall fields in terms of magnitude, spatial variability and spatial structure? 2) How are the corresponding hydrologic simulations affected? And given the non-linear nature of hydrologic processes, 3) are these effects magnified or dampened, and does the size of the watershed play an important role? Since model calibration is usually employed to compensate for errors in the input data along with other sources of uncertainty in the modeling system, 4) can the effects analyzed in 2) and 3) be effectively mitigated through adjustment of model parameters? Finally, and most importantly, 5) can model calibration serve as a method to explicitly account for input’s resolution and accurately reveal the utility of satellite rainfall products?

In order to answer the aforementioned questions, a high-resolution, radar-based rainfall product was resampled to match the resolution of 3B42RT product from the Tropical Rainfall Measurement Mission (TRMM) (Huffman et al. 2007). Additionally, the dependency of the product’s resolution effect on rainfall or streamflow relative magnitudes (i.e., mean values vs. extreme high values) is studied. Finally, we present a strategy through the estimation of the hydrologic model parameters to account for the impact of resolution in the evaluation of satellite rainfall products. The method is put to the test using TRMM’s 3B42RT estimates to force the hydrologic simulations from the resulting model setup. The remainder of the paper is organized as follows. Section 2 presents the study area and the data employed for the experiments. In section 3, questions 1), 2) and 3) are addressed through the synthetic experiment for isolating the effect of rainfall product’s resolution. Section 4 reports the analysis of questions 4) and 5) using a
general-purpose optimization algorithm, and section 5 summarizes the work and outlines the conclusions.

2. Study area and data

This study region focuses on the catchment of the US Geological Survey (USGS) station located at Tarboro (Streamgage No 02083500), a sub-catchment of the Tar-Pamlico River basin located in coastal North Carolina (Fig. 1). The basin is periodically affected by heavy rainfall from tropical storms and hurricanes, at which times major flood events occur. Because of these coastal impacts, the basin has been a subject of research for the National Oceanic and Atmospheric Administration (NOAA)’s Coastal and Inland Flooding Observation and Warning (CI-FLOW) project (Van Cooten et al. 2011). Tarboro’s catchment consists of the upper Tar River and Fishing Creek sub-basins for a total drainage area of 5,709 km$^2$. Located on the coastal plain, the elevation for this catchment ranges from 2.8 - 225.7 m above mean sea level (AMSL). The Tar is a perennial river with a minimum recorded daily flow of 0.79 m$^3$/s, a mean of 62 m$^3$/s and a maximum of 1,996 m$^3$/s at Tarboro. Sub-hourly streamflow observations are available at eight interior locations of Tarboro spanning from October 1985 to present at the USGS Instantaneous Data Archive (IDA; USGS 2011). Five of the measurement sites located on the main stem were selected to derive hourly streamflow observations for the experiments conducted herein. The distribution of the sites forms a cascade of sub-catchments ranging from 529 to 5,709 km$^2$.

The quantitative precipitation estimation (QPE) product used herein as ground reference was the US National Weather Service (NWS) Multisensor Precipitation
Estimation (MPE) product. MPE data is derived employing a blend of automated and interactive procedures to combine information from satellite, radar, and rain gauges (Briedenbach and Bradberry 2001; Fulton 2002; Seo et al. 2010). Hourly MPE data are available for the Southeast River Forecast Center (SERFC) region at the Hydrologic Rainfall Analysis Project (HRAP; Schaake 1989) resolution from 2002 to present. The size of an HRAP grid cell is approximately 4x4 km$^2$.

The satellite product used in this work was the near real-time infrared and microwave merged TRMM precipitation product 3B42RT, version 6. 3B42RT estimates are available at 0.25 degree (~25 km) and 3-hourly time intervals (Huffman et al. 2007). A total of eight years of data (2002 – 2009) overlapping MPE and 3B42RT archives were employed and analyzed in this study. 3B42RT estimates were resampled to MPE’s HRAP resolution for comparison and subsequent use in the hydrologic model. Figure 2 presents an evaluation of 3B42RT performance as compared to MPE for the period of data. Mean annual totals (Fig. 2a,b) are shown along with spatially distributed relative bias (Eq. 1; Fig. 2c) and relative RMSE (Eq. 2; Fig. 2d).

$$\text{Bias(\%)} = \frac{\sum_{i=1}^{N} (X_{i}^{\text{sim}} - X_{i}^{\text{ref}})}{\sum_{i=1}^{N} X_{i}^{\text{ref}}} \times 100 \quad (1)$$

$$\text{RMSE(\%)} = \frac{\sqrt{\sum_{i=1}^{N} (X_{i}^{\text{sim}} - X_{i}^{\text{ref}})^2}}{\bar{X}^{\text{ref}}} \times 100 \quad (2)$$

where $X_{i}^{\text{sim}}$ is the variable under evaluation (i.e. 3B42RT) value at time $i$, $X_{i}^{\text{ref}}$ is the reference value (i.e. MPE) at time $i$, and $\bar{X}^{\text{ref}}$ is the average reference value over the
simulation interval. The clear southeast – northwest precipitation gradient over the basin shown in Fig. 2a is generally well captured by 3B42RT (Fig. 2b). However, 3B42RT significantly underestimates at the upper part of the basin where the lowest amounts of rainfall occur. This can also be seen in Fig. 2c where values of negative bias range from about -5% to -15% for the upper part, and areas of overestimation on the mid and lower parts of the basin. The relative RMSE is significantly high with values ranging from about 550% to 750%, where the highest values generally occur at the lower part of the basin. These errors result from the combination of discrepancies between 3B42RT and MPE in terms of both spatiotemporal resolution and retrieval capabilities. The following section presents an analysis to describe the contribution of input’s resolution to these errors and their impact on hydrologic predictions.

3. Isolating uncertainty due to rainfall product resolution

This work concentrated on the error contributed by the rainfall products’ spatiotemporal resolution. Specifically, this is the resolution error associated with current satellite-based QPEs as compared to ground reference products, which are generally considered as the ground truth. In order to create a scenario where all other sources of uncertainty can be virtually neglected, the reference high-resolution QPE product (i.e., MPE: 4-km/hourly) was resampled to the coarser resolution of TRMM satellite products (i.e., 25-km/3-hourly). The resampling method consisted of a simple computation of the unconditional mean of all the 4-km pixels within a given 25-km pixel for the spatial aggregation (Eq. 3), and a centered weighted average computation involving the four closest time steps for the temporal aggregation (Eq. 4):
\[ R_{i,j}^{25\text{km}} = \frac{\sum_{k=1}^{n} \left( \sum_{l=1}^{m} R_{k,l}^{4\text{km}} \right)}{n \times m} \]  (3)

\[ R_{i,j,t}^{25\text{km}} = w_1 \left( R_{i,j,t-1}^{25\text{km}} + R_{i,j,t+1}^{25\text{km}} \right) + w_2 \left( R_{i,j,t-2}^{25\text{km}} + R_{i,j,t+2}^{25\text{km}} \right) \]  (4)

where \( R_{i,j}^{25\text{km}} \) is the rain rate at the resampled \( i^{th} \), \( j^{th} \) 25-km grid point resulting from the average of the 4-km grid points in the region delimited by the \( n \times m \) (rows, columns) box, \( w_1 \) and \( w_2 \) are weights for the temporal averaging of rain rates in the interval \( t - 2 \) to \( t + 2 \) with values 0.15 and 0.35 respectively, and \( R_{i,j,t}^{25\text{km}} \) is the rain rate degraded in both spatial and temporal resolution at time \( t \). In this section, results and their corresponding analysis are based on comparisons between the high- and the coarse-resolution versions of MPE, referred to hereafter as HR-MPE and CR-MPE respectively.

a. Effects on rainfall estimates

The first step in this analysis was an examination of the overall impact of resolution on rainfall estimates. Since this particular aspect has been well studied and documented in previous studies (e.g. Obled et al. 1994; Ogden and Julien 1994), some of the results described herein are not included in the main document but can be found in the appendix. The difference in resolution introduces considerably large errors in the estimates with values of RMSE of about 150% to 350% for annual grids, and about 650% – 750% for the entire period of data (appendix A - Fig. A1). Although the values of bias are in general not significant, there is considerable variability in some of the years with biases ranging from more than 20% overestimation to about 10% underestimation. Moreover, a clear inter-annual variability was observed in terms of these error signatures. This is most probably related to the rainfall regime and characteristics of individual
storms of each year. Years dominated by small-scale convection are more likely to be associated with relatively high RMSE due to a smoothing effect when aggregating up to the coarser 25-km pixel scale. This issue was explored through an analysis of the effect of QPE’s resolution degradation on the spatial structure of rainfall for 100 representative storm events from the entire period of record. A relatively simple method based on variograms (Journel and Huijbregts 1978) proposed by Lebel et al. (1987) and followed by subsequent studies (e.g. Berne et al. 2004; Kirstetter et al. 2010; Kirstetter et al. 2012) was used herein to characterize the differences in spatial structure between HR-MPE and CR-MPE. Assuming the rainfall field to be a realization of a random function, the variogram represents the spatial correlation of the rain field and is defined as half the expectation of the quadratic increments between estimates of rainfall as a function of interdistance between the points of estimation. In the multi-realization case (e.g. rainfall fields in successive hourly accumulations for a given event), it is convenient to take into account information from all the realizations to infer a single and more robust empirical variogram, assuming the fields to have similar statistical characteristics except for a constant factor. Times with zero rainfall are not considered in computing the variogram. In this study, we used the approach proposed by Lebel et al. (1987) based on a normalized variogram established by scaling each squared increment by the field variance. A standard model is fitted to the empirical mean variogram. The exponential model of the following general form was found suitable:

\[ \gamma(h) = C_0 + (C - C_0) \left(1 - e^{-h/d} \right) \]  

(5)

where \( \gamma \) is the variogram function ranging from 0 to 1, \( h \) is the interdistance in km, \( C_0 \), and \( d \) are the variogram model parameters respectively called the nugget (dimensionless)
and the range (in kilometers). The nugget is the value of the variogram at the origin, which can be related either to variability poorly resolved by the product or to measurement error (Kirstetter et al. 2010). The sill, $C$, is the value at which the variogram asymptotically approaches when $h$ tends to infinity. In the case of normalized variograms, the sum of the nugget and sill equals 1.0. The range is the interdistance at which the variogram reaches the sill and is defined as the mean decorrelation distance of the measurements, which can be used as an indicator of the characteristic spatial scale of a storm: large systems such as tropical storms and hurricanes can be characterized by large range values, while small range values can be associated to small scale convection. Because this analysis requires a large enough spatial domain to correctly characterize the spatial structure of rainfall, it is only applied at the largest scale (i.e. B5 basin). The nugget and range were used herein to describe the spatial structure of the selected storms.

Figure 3 shows scatter plots of selected variogram parameters and their association to the effect of resolution degradation for the 100 events. Panels a) and b) directly demonstrate the modification of rainfall spatial structure features (i.e. range and nugget respectively). The effect is clear and directly proportional to both the nuggets and the ranges of the storms as observed by the high-resolution product. The trend of the nugget indicates that for events with values less than about 0.1 the effect of resolution is higher. Furthermore, panels c) and d) of Fig. 3 present the association of the modification of these spatial structure characteristics of rainfall to error due to resolution degradation. For the case of the variogram range, there is a direct relationship with the basin averaged relative RMSE. This indicates that introduction of spurious spatial correlation due to smoothing effects promotes the introduction of aggregated error, which increases with
decreasing storm scale as shown in panel a). The inverse relationship between nugget ratio and RMSE displayed in panel d) clearly shows that increasing inabilities to resolve variability at small scales results in error as well. These results illustrate that coarser resolution QPE results in error as quantified through an analysis of the spatial characteristics of rainfall fields.

b. Propagation into hydrologic modeling

1) HYDROLOGIC MODEL

The National Weather Service (NWS) Hydrology Laboratory – Research Distributed Hydrologic Model (HL-RDHM), whose precursor was the NWS HL-RMS (Research Modeling System; Koren et al. 2004), was employed here to evaluate the effect of rainfall forcing resolution on the simulation of streamflow. The modeling system consists of the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash et al. 1973) for the water balance/runoff generation and a kinematic wave model for the hillslope and channel routing. These two components are applied to each cell in a rectangular grid at the HRAP resolution in a polar stereographic projection, which directly corresponds to the MPE product.

The SAC-SMA is a conceptual rainfall-runoff (CRR) watershed model widely used within the NWS River Forecast System (Sorooshian et al. 1993; Boyle et al. 2000; Koren et al. 2004). The model structure and other details are well described in Koren et al. (2000; 2003; 2004) and Yilmaz et al. (2008), and so we focus on the presentation of the model parameters and their estimation here. The model simulates runoff generation using 17 conceptual parameters (appendix B – Table B1). Koren et al. (2000) developed
an approach to estimate values for 11 of the SAC-SMA parameters based on the State Soil Geographic soil data (STATSGO) (Soil Survey Staff 1994, 1996). The remaining six parameters use lumped values established by the NWS from previous experience on different basins (Pokhrel et al. 2008; Yilmaz et al. 2008). Potential Evaporation (PE) data is another input in SAC-SMA (i.e., in addition to gridded precipitation data). Twelve climatological mean monthly potential evaporation grids, available at the HRAP resolution for the Continental US, were used in this study. HL-RDHM’s user manual (NWS et al. 2008) describes how these grids were derived.

The flow routing component in HL-RDHM is divided into hillslope routing (overland flow) and channel routing. Values for the parameters controlling overland flow were derived from a high-resolution digital elevation model (DEM), land use data, and results reported from initial tests on the model (NWS et al. 2008). The channel routing parameters, $q_m$ and $q_0$, were specified using a discharge to cross-sectional area relationship through a rating curve method (Koren et al. 2004):

$$Q = q_0 A^{\alpha_m}$$

(6)

where $q_m$ and $q_0$ were determined using measurements of $Q$ and $A$ data at eight stations within Tarboro’s catchment (USGS stations 02081500, 02081747, 02082506, 02082585, 02082770, 02082950, 02083000, 02083500). Using HL-RDHM’s built-in tools, spatially distributed grids for the two channel routing parameters were created.

The aforementioned setup yielded a model referred to hereafter as the baseline model. Koren et al. (2003) report that the a priori estimates give reasonable initial values for SAC-SMA parameters and reduce uncertainties in their ranges. Figure 4 presents a 1-year hourly simulation of streamflow with the baseline model forced by HR-MPE at the
smallest and largest basins (Fig. 4a and b, respectively) to evaluate the validity of the
setup. The simulation was compared to the observed streamflow and metrics of goodness
of fit, commonly employed to assess hydrologic model performance, were computed to
illustrate the model’s skill in describing the observed overall response of the basin to
rainfall. Visual inspection of the hydrographs reveals that even with no calibration
involved in the estimation of model parameters, the skill of the simulation in capturing
the direct response to rainfall (i.e., occurrence of peak flows) as well as the baseflow in
both basins is high. The values of the error metrics (i.e., relative Bias and relative RMSE)
are only slightly higher at the smallest basin due to underestimation of the magnitude of
some of the peaks. The value of the Nash-Sutcliffe Coefficient of Efficiency (NSCE), a
measure commonly used in hydrology to summarize simulation skill, is relatively high at
the smallest basin (NSCE = 0.65) and relatively low at the largest basin (NSCE = 0.23).
The lower value of NSCE for the simulation at the largest scale is primarily a result of
low covariance between observed and simulated streamflow time-series due to timing
offsets. This timing issue is caused by limitations in the estimation of the routing
parameters (i.e. $q_m$ and $q_o$ in Eq. 6), which have a larger impact on medium to large
basins given that the lag in the response to rainfall is directly correlated to catchment size.
Dramatic improvement in simulation timing can be achieved by only adjusting the two
routing parameters (NSCE above 0.8; not shown here).
2) SENSITIVITY OF STREAMFLOW SIMULATIONS TO RAINFALL RESOLUTION ALONE

We use the baseline model to isolate error in the hydrologic simulations associated to rainfall forcing’s resolution alone, because the parameters are based on soil and land use data that are independent from the rainfall algorithms. A calibrated model, on the other hand, will be conditioned on the input and reference data (e.g., measured streamflow). To neglect the other sources of uncertainty, including model structure and approximations, the simulation forced by HR-MPE was used as the reference to evaluate precipitation resolution errors in the hydrologic simulation forced by CR-MPE. The use of hydrologic simulations as reference also enables us to compare simulations at each grid point across scales. The basic assumption of this synthetic experiment is that the hydrologic model represents basin response to rainfall, and that the HR-MPE dataset represents the actual surface rainfall; loss in skill in the hydrologic simulations using the CR-MPE forcing is from precipitation estimate’s resolution alone. The analysis presented herein was fixed to the hourly modeling temporal scale since, first, this is the reference resolution of HR-MPE and, secondly, it is the scale at which resampled rain estimates exhibited the greatest uncertainty (appendix A - Fig. A2).

Figure 5 presents a grid-based comparison of simulated streamflow forced by CR-MPE and the synthetic reference for the 2002 – 2009 period. Overall, we notice that streamflow is underestimated across the catchment, although there is some overestimation in Fishing Creek’s sub-catchment (upper portion of Tarboro’s catchment). Additionally, the propagation of aggregated errors from the CR-MPE rainfall to streamflow presents a clear pattern. The relatively low RMSE and high NSCE values are
collocated in large-scale streams and rivers. Furthermore, it can be seen how the model performance degrades moving in the upstream direction towards headwaters. This is consistent with results from previous studies in relation to the inverse association between resolution impact and drainage area (e.g. Nijssen and Lettenmaier 2004). A more detailed investigation on basin-scale dependence follows.

A summary of the magnitude of errors in rainfall and streamflow for the 2002-2009 period and the five basins considered in this work are presented in Table 1. The relative RMSE and NSCE of streamflow generally improve with increasing basin scale. The relative bias, on the other hand, does not display a clear trend with basin scale. The values of bias seem to worsen from B1 to B4, but it improves when going to B5. This behavior is consistent with the mean relative bias of rainfall and the gradient of mean rainfall over the basin shown in Fig. 2a. Moreover, it can be seen that the global bias in streamflow simulations stemming from QPE resolution difference is relatively low, while the RMSE is considerably high. Figure 6 shows the same metrics in Fig. 5 but expressed as ratios to corresponding rainfall error and plotted as a function of basin catchment area for different streamflow thresholds. These results illustrate the error propagation and possible dampening or magnification of resolution-related errors by the non-linear transformation of rainfall to runoff in integrated basins. Nikolopoulous et al. (2010) examined rainfall error propagation in a complex terrain basin in NE Italy and found that it depended on the metric of reference. For example, the relative error in rainfall translated to a lower relative error in runoff volume, while the same rainfall error resulted in a magnified relative error in peak discharge (in most ensembles). In our case, the global bias in rainfall significantly magnified into higher global bias in streamflow.
Moreover, this magnification increased with decreasing basin scale with an abrupt jump from a factor of about 6.0 to a factor of a little over 20.0 when going from B2 to B1 for the case that considers all flow values (i.e. the thick black line - Q(100%)). On the other hand, the aggregated errors in rainfall were greatly dampened in the streamflow simulations. Furthermore, the reduction in error directly correlated with basin scale with a factor close to 6.0 (Q to R RMSE ratio of ~0.17) for B1 and over 20.0 (Q to R RMSE ratio of ~0.048) for B5 for the case that considers all flow values.

Figure 6 also presents metric values for different streamflow exceedance thresholds. It can be seen that the impact of QPE resolution difference increases non-linearly as the probability of occurrence of streamflow decreases, which negatively affects the skill of the modeling system to simulate extreme flooding events. The difference in relative RMSE (inferred from the differences in Q to R ratio) and NSCE between Q(100%) and Q(10%) is greatest at the smallest basin scales. Also, the QPE resolution impact appears to be greatest for the highest flow thresholds and at the smallest basin scales, which have implications on hydrologic applications such as flash flood modeling. Furthermore, the dampening of random error decreases with increasing streamflow threshold. On the other hand, the global bias and its magnification are not significantly affected. A reduction on the magnification, however, can be noticed when considering the highest flow thresholds (i.e. Q(10%)).

Lastly, the effect of resolution is assessed at the scale of specific events, which allows for a more detailed examination of the rainfall error propagation to streamflow simulations. Figure 7 presents a summary of the error features of rainfall estimates (i.e., magnitude, spatial variability and spatial structure) introduced by resolution degradation.
and their propagation to hydrologic simulations. The sample of 100 events used in the analysis in section 3.a was employed for this purpose. The error magnitude was defined as the basin averaged mean differences between CR-MPE and HR-MPE over the event interval. Similarly, the error spatial variability was defined as the mean basin standard deviation of differences between CR-MPE and HR-MPE over the event interval. Finally, the error spatial structure was defined computing the variogram of rainfall differences between CR-MPE and HR-MPE for each event and comparing their ranges to those of HR-MPE’s variograms. These rainfall error features were contrasted to the corresponding streamflow error across the basin (y-axis for error magnitude and error variability panels) and at the basin outlet described by the relative bias (upper panels) and relative RMSE (lower panels).

In terms of error magnitude (Fig. 7a,d), there is a definitive connection between basin-averaged rainfall and streamflow errors. Moreover, the relative bias in streamflow simulations at the outlet is clearly associated to the aforementioned connection of error magnitudes. The relative RMSE displays a similar association with the basin-averaged differences. Furthermore, it can be observed that most of events with relatively low RMSE (i.e. below ~20% - blue colored events) are associated to events with positive bias (around 10 – 20%). This indicates that the effect of resolution for these particular events is mostly reflected in relative bias. In terms of error spatial variability (Fig. 7b,e), rainfall differences also correlate to streamflow differences across the basin, although not as strongly as the error magnitude does. However, there is not a clear association with either bias or RMSE of streamflow at the basin outlet. Finally, the graphs of error spatial structure (Fig. 7c,f) show a linear and direct relationship between the size and/or level of
organization of the storms and the spatial correlation structure of the rainfall error, which is consistent with results from section 3.a. In terms of the association with error in streamflow simulations at the outlet, there are mainly two aspects that can be observed: 1) even though there is not a very clear trend of the error as depicted by the color gradients, it is possible to distinguish subgroups of events associated with sectors in the error range (i.e., events with RMSE values greater than ~60% – red colored points – are grouped towards the lower left side of the graph, which corresponds to the lower values of variogram’s range), which is seen in both the relative bias and the relative RMSE; and 2) events with the highest values of streamflow error, both bias and RMSE, are associated to lower characteristic spatial scale. Aspect 1) is most probably related to specific conditions of each event that were not accounted for, such as the antecedent soil conditions or the proximity of the storm to the basin outlet. This might also explain why it is not possible to observe a clear relationship between error in the spatial structure of rainfall and error in streamflow at the basin outlet. However, aspect 2) indicates that error due to resolution in the estimation of rainfall from small-scale convection will have higher impact on streamflow simulations.

4. Mitigating input’s resolution impacts through model calibration

a. Evaluation of differences due to rainfall resolution

A standard procedure when setting up an operational hydrologic model is the calibration of its parameters using observations of rainfall and streamflow. This has become a common practice in studies that assess the hydrologic utility of satellite-based QPE, where the hydrologic model is calibrated at the scale corresponding to the
algorithm’s pixel resolution or to the resolution of a reference rainfall dataset (e.g., from rain gauges). The calibrated parameter set is highly dependent on the characteristics of the QPE used to force the model during the calibration period. Other variables such as the length of period during calibration and the variability of flow conditions and events in the calibration period also play an important role during the adjustment of the parameters. However, the discussion in this section is focused on the differences in model performance caused by QPE resolution. More specifically, the objective is to assess the extent to which model calibration can compensate for errors due to QPE resolution. Other issues in the satellite QPE, such as biases and random errors inherent to the retrieval process, are not explicitly accounted for in the model parameter estimation.

The hydrologic model was calibrated using a general-purpose optimization algorithm entitled Differential Evolution Adaptive Metropolis (DREAM) (Vrugt et al. 2009). DREAM is an adaptation of the widely used Shuffled Complex Evolution algorithm (SCE-UA; Duan et al. 1993). The algorithm employs Markov Chain Monte Carlo (MCMC) sampling to estimate the posterior probability density function of parameters using a formal likelihood function, which ensures a collective evolution of the model parameters (i.e., all parameters are optimized simultaneously to account for interdependencies among them). The algorithm is designed to operate in complex, high-dimensional sampling problems (Vrugt et al. 2008). The reader is encouraged to review the work by Vrugt et al. (2009) for more details regarding DREAM. A period of three years (2004 – 2006) of USGS gauge streamflow observations was used as the reference in the calibration process to compute the Sum of Squared Residuals (SSR; i.e. the objective function). In addition to SSR, a constraint was set in DREAM’s algorithm for
selecting parameter sets that produced relatively unbiased simulations (i.e. relative bias < 10%). The starting point or benchmark of this calibration was the baseline model described in 3.b.1. Since the baseline model has spatially distributed parameters estimated from soil and stream observations, the calibration approach was used to generate scaling factors and so the spatial variability dictated by the observations are maintained (i.e., the relative differences among grid cells is preserved). An independent validation period consisting of the data not used for calibration (i.e. 2002 – 2003 and 2007 – 2009) was used to assess the suitability of the resulting parameter sets in terms of relative performance of model skill, conditioned on the different parameter sets. Note that the validation of the method to account for QPE product’s resolution is presented in section 4.b.

To explore the effectiveness of model calibration for mitigating rainfall resolution errors as a function of spatial scale and to account for scale dependencies in the hydrologic model, each QPE product (i.e., HR-MPE and CR-MPE) was used to automatically calibrate the model parameters at each basin, independently; that is, calibration was performed over the entire area associated to each basin (i.e. B1 through B5) ignoring observations at interior locations (e.g., B3 is calibrated over its entire catchment ignoring the fact that B2 and B1 exist within it). This array of multiple independent parameter sets is referred to hereafter as “Local” calibration, as each parameter set was derived specifically to the basin and to the QPE input. The parameter set that was generated by calibrating the model at the largest scale (i.e., B5) was used to simulate streamflow at interior gauged points to be compared to simulations produced by
the “Local” calibration. Therefore, the B5 calibration is referred to hereafter as “Regional”.

Metrics of overall performance for the resulting calibrated model realizations during both calibration and validation periods are presented in Fig. 8 as a function of basin area. When comparing the results from CR-MPE and HR-MPE in the “Local” calibration runs, we see their skill metrics are very similar with the only exception in the smallest scale basin where differences exist mainly in terms of the direction of bias (i.e., HR-MPE yields a bias of about 10%, while CR-MPE yields a bias of about -10%). Although these similarities in skill (i.e., the closeness of performance metrics across basins) are somewhat reduced during the validation period, particularly for drainage areas under 2,000 km², these results indicate that errors due to QPE resolution can be mitigated or, in other words, accounted for with adjustment of the hydrologic model parameters. Figure 9 presents a comparison of the resulting adjusted parameters after calibration runs forced by HR-MPE and CR-MPE. The values are presented as ratios of CR-MPE-based parameters to HR-MPE-based parameters. These ratios were plotted against basin area to check for correction of basin scale effects observed in the differences between CR-MPE and HR-MPE in section 3.b.2. Some parameters display a clear trend with respect to basin area, which is an indication of the impact of basin scale and rainfall resolution differences effects on the calibration results. Such is the case for parameters like UZTWM, UZK, PCTIM, RIVA, ZPERC, LZTWM, LZFPM and LZPK. Other parameters do not display any trend, which is an indication of low sensitivity to basin scale effects. Moreover, these random differences in the adjustment of some model parameters between CR-MPE and HR-MPE demonstrates the compensatory effects of
model calibration given the random nature of errors introduced by resolution degradation and the stochastic features of the optimization algorithm. The global bias (i.e., the negative bias of CR-MPE streamflow simulations as compared to HR-MPE simulations), on the other hand, was consistently corrected after adjustment of some parameters towards an increase in runoff generation (i.e., the upper zone depletion rate UZK or the impervious area fraction PCTIM, which directly control runoff generation, and the ratio of percolation rates ZPERC, which controls how much water is lost to the soil). These results might indicate that some parameters could yield reduction of systematic errors when perturbed in a pre-defined fashion, which would have implications on the issue of parameter transferability to other basins. However, this aspect was not further explored in the present study.

The comparison between “Local” and “Regional” simulations reveals much improved skill during the calibration period at all basin scales and QPE resolutions using the locally optimized parameter settings (Fig. 8). These local-scale improvements, however, diminish and are essentially indistinguishable from the skill attained with the Regional model during the validation period. Note these levels of skill are only achievable by accounting for the QPE-resolution effect in the calibration process. This result has significant implications for forecasting streamflow at ungauged, interior points using a model calibrated with observations at the basin outlet. A natural extension of this result is an analysis of the limiting catchment size for which the Regional calibration is effective over different regions, which should be investigated in future works.

Lastly, we notice that the relative RMSE is the only metric that clearly displays the impact of basin scale even during the calibration period, which is most evident for
basins under 2,000 km². The impact of basin scale on relative RMSE observed here is consistent with what was observed in the synthetic experiment (section 3.b.2) and previous studies such as the ones by Nijssen and Lettenmaier (2004) and Nikolopoulos (2010), although the latter found a rather counterintuitive behavior: the values of relative RMSE increased with increasing basin size. Nevertheless, these results suggest that the effects of basin scale on aggregated errors might be resistive to model calibration.

b. Validation of method with 3B42RT: Revealing the utility of satellite estimates

Up to this point, it has been shown that the use of rainfall estimates matching satellite’s coarse resolution has a significant impact on the hydrologic modeling of a cascade of basins as compared to the use of a high-resolution rainfall product. Furthermore, it has been demonstrated that this impact can be reduced through adjustment of the parameters of the hydrologic model. The last step of this study was to evaluate whether the information gained in the calibration process in terms of input’s resolution could be transferred to simulations forced by 3B42RT estimates and more accurately reveal its hydrologic skill. This has important implications in relation to the use of satellite rainfall estimates for hydrologic applications in ungauged basins. Note that the parameter sets resulting from the calibration process in section 4.a are assumed to be representative of the true and best characterization of basin physical properties and processes. Therefore, the resulting model configuration may be used for an evaluation of alternative input data such as satellite rainfall products.

Streamflow simulations for the period of data (i.e. 2002 – 2009) were thus generated forcing the hydrologic model with 3B42RT using both locally optimized sets
of parameters (i.e., “Local” CR-MPE and HR-MPE associated parameter sets). Figure 10 presents metrics of skill for the overall performance of both models at the selected basin scales and for different streamflow thresholds. In terms of relative RMSE, there are little differences between the simulations, which might indicate that aggregated errors from other sources of uncertainty in the generation of the 3B42RT product overshadow the resolution effect. In terms of relative bias, however, it can be seen that the model calibrated with CR-MPE performed better than the model calibrated with HR-MPE, although the differences reduce with decreasing basin scale and streamflow probability of occurrence. In terms of the general performance of 3B42RT, the significant negative bias, which increases toward the headwater basin, are consistent with the gradient of rainfall bias observed in Fig. 2c. These results clearly indicate that accounting for satellite-based QPE’s resolution in an explicit manner can help in the evaluation of 3B42RT’s hydrologic utility.

This overall improvement in bias on the simulations forced by 3B42RT relates back to the compensation from the adjustment of parameters analyzed in section 4.a. These results support the idea that model calibration can be used to account for inputs’ resolution. The basis of the approach lies in the fact that it is possible to force model calibration to place emphasis on a desired correction. We have purposely introduced “extra” error in the input data (i.e. HR-MPE) by altering its resolution (i.e. CR-MPE), which results in adjustments of model parameters to reduce the overall error in the hydrologic simulations including that from the resolution degradation. This method can be used to establish a benchmark model for the evaluation of satellite rainfall products. Establishing the hydrologic utility of GPM-era QPE products, whose spatial resolution is
anticipated to be upgraded to 0.1 degrees (~10 km; Smith et al. 2007; Huffman et al. 2011), will be a subject of many future studies. The results from this work should, therefore, be considered in the design of such studies. Additionally, accounting for input’s resolution in the model calibration process helps preventing the necessity of recalibration every time an upgrade on the algorithm for QPE generation is done (e.g., to date, the 3B42 algorithm has been upgraded seven times).

5. Summary and Conclusions

This work was designed to explore the effects that resolution of QPE has on hydrologic simulations in the context of current and next generation satellite-based QPE products. The study focused on the Tar River basin in coastal North Carolina, where a cascade of five catchments with drainage areas ranging from 529 to 5,709 km$^2$ were used to assess the scale dependence, regionalization of parameter sets, and effect of QPE resolution on hydrologic simulations. A high-quality, high-resolution QPE product, used operationally in the US National Weather Service, was employed as the rainfall reference product. The reference QPE product called MPE (4km/1hr) was resampled to current satellite-based QPE’s resolution in space and time (25km/3hr). Differences in rainfall estimates between the high-resolution MPE (HR-MPE) and the coarse-resolution MPE (CR-MPE) were evaluated considering different error features, including spatial structure and variability through an analysis of 100 events. A sensitivity analysis with a synthetic experiment using the uncalibrated model was then performed to effectively isolate QPE’s resolution effect on hydrologic simulations. Finally, the hydrologic model was calibrated using measurements from the gauged sites for the two QPE resolutions (i.e., CR-MPE
and HR-MPE). The resulting calibrated models were then forced with TRMM’s 3B42RT rainfall estimates. Conclusions derived from this work are summarized as follows:

- Degradation of QPE resolution has a clear impact on rainfall fields in terms of global bias and aggregated errors. Moreover, it was observed that resolution degradation alters the spatial structure of rainfall introducing considerable amounts of aggregated error. This alteration and its resulting error are directly proportional to the characteristic spatial structure and scale of each storm event.

- The sensitivity analysis on the overall impacts of resolution degradation on hydrologic simulations shows that the global bias of rainfall is magnified while the aggregated errors are dampened. These error propagation behaviors are correlated with the size of the basin. The magnification of the bias increases with decreasing basin area reaching magnification factors of about 20. The dampening of aggregated errors increases with increasing basin size also reaching dampening factors of about 20. Furthermore, the dampening effect of aggregated errors significantly decreases with increasing streamflow values, while the magnification of bias is not significantly affected.

- The analysis on the three error features of rainfall (namely magnitude, spatial variability, and spatial structure) indicates that errors in streamflow are mostly related to the effects that resolution has on the magnitude of rainfall. Even though the spatial variability of rainfall error is well correlated with the spatial variability of streamflow error, the hydrologic filtering effect of the basin obscures its impact on the errors at integration locations (i.e., basin outlets). Similarly, the effect on the spatial structure of rainfall could not be well established for the same reason.
and the fact that specific conditions of each event were not accounted for. However, this analysis indicated that events of significant streamflow error were associated to error in the spatial structure of rainfall from small-scale convection.

- The results from the calibration procedure clearly demonstrate that adjustment of model parameters can be used as a mean to account for the effects of input’s resolution over hydrologic simulations. The validation with 3B42RT demonstrated that significant reduction in relative bias is attained when using the parameter set that accounted for input’s resolution.

Overall, these results illustrate the importance of resolution and its close relation with scale for satellite-based applications, particularly when determining the hydrologic utility of satellite rainfall. Accounting for the rainfall resolution effect is essential when evaluating satellite rainfall algorithms using a hydrologic model. Likewise, this work indicates the significant improvements that could potentially be expected from the next generation satellite rainfall products associated with GPM, considering the anticipated upgrade in resolution and improvements in retrieval capabilities. Results from this study indicate improvements due to resolution are most needed in the smallest scale basins for high-magnitude events, such as flash floods. Furthermore, the analysis on calibration results suggests that model calibration needs to consider scale. Future studies need to explore the issue on the persistent impact of basin scale on aggregated errors. Likewise, it would be helpful to extend some of the analysis performed herein to different regions and a wider range of basin sizes to improve our understanding on issues such as 1) the transferability of information from basin outlet to interior locations; 2) transferability of systematic perturbations of
particular model parameters to account for input resolution; and 3) characterization of
the impact of resolution degradation on particular storm types.
References


Burnash, R. J. C., R. L. Ferral, and R. A. McGuire, 1973: A general streamflow simulation system - Conceptual modeling for digital computers. In Report by the Joint Federal State River Forecasts Center, Sacramento, California,


Sapiano, M., and P. Arkin, 2009: An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *Journal of Hydrometeorology*, 10, 149-166


Su, F., Y. Hong, and D. P. Lettenmaier, 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


List of figures

Figure 1: Study area showing the selected sub-basins (B1 – B5) and associated gauge locations within the catchment of Tarboro’s USGS streamflow gauge station in Tar river basin, North Carolina - US. The hydrography of the basin is presented highlighting main streams.

Figure 2: Performance evaluation of 3B42RT rainfall estimates for 2002 - 2009 period at 4-km and 1-hour resolutions. Panels a), and b) present the spatial distribution of mean annual rainfall from MPE, and 3B42RT, respectively. Panels c) and d) present the spatial distribution of the hourly-based mean relative bias and percent RMSE, respectively.

Figure 3: Effect of resolution degradation on the spatial structure of rainfall for 100 events: a) HR-MPE range vs. CR-MPE range (solid line represents 1:1 relation); b) HR-MPE nugget vs. CR-MPE nugget (solid line represents 1:1 relation); c) Basin averaged relative RMSE (%) as a function of the ratio of ranges; and d) Basin averaged relative RMSE (%) as a function of the ratio of nuggets (log scale).

Figure 4: Simulation of streamflow during 2006 with the baseline model at a) smallest scale basin - B1, and b) largest scale basin - B5. Measures of performance for the simulation are presented in the legend.

Figure 5: Grid-based comparison of streamflow forced by CR-MPE and HR-MPE for the synthetic experiment: a) Relative Bias, b) Relative RMSE, and c) NSCE.

Figure 6: Overall propagation of mean rainfall error to streamflow simulations. Metrics of model performance as functions of basin area and streamflow threshold are presented: a) ratio of streamflow (Q) relative bias to rainfall (R) relative bias; b) ratio of streamflow (Q) relative RMSE to rainfall (R) relative RMSE; and c) NSCE of streamflow
simulations. The values between parentheses indicate the probability of occurrence of the corresponding threshold.

Figure 7: Summary of error features and their propagation to streamflow modeling at basin outlet for the 100 events. Left panels (a and d): error magnitude; center panels (b and e): error variability; right panels (c and f): error spatial structure. The relative bias (upper panels) and relative RMSE (lower panels) were computed for streamflow simulations forced by CR-MPE as compared to streamflow simulations forced by HR-MPE at the basin outlet.

Figure 8: Summary of calibration (left panels) and validation (right panels) results. Metrics of model performance are plotted against basin area for the five sub-basins. The "R." and "L." refer to "Regional" and "Local" calibration runs respectively.

Figure 9: Comparison of adjusted hydrologic model parameters after calibrations forced by HR-MPE and CR-MPE. The graphs present CR-MPE to HR-MPE ratio of parameter basin average values (y-axis) versus basin area (x-axis). The horizontal thick line in each panel indicates where the ratio is 1.0.

Figure 10: Relative Bias and RMSE of 3B42RT-forced simulations from the two resulting locally calibrated parameter sets over the 2002 – 2009 period. Error metrics are computed using actual streamflow observations from gauged locations. The model using HR-MPE based parameter set (dashed lines) is compared to the CR-MPE based parameter set (solid lines) at different streamflow threshold values.
Figure 1: Study area showing the selected sub-basins (B1 – B5) and associated gauge locations within the catchment of Tarboro’s USGS streamflow gauge station in Tar river basin, North Carolina - US. The hydrography of the basin is presented highlighting main streams.
Figure 2: Performance evaluation of 3B42RT rainfall estimates for 2002 - 2009 period at 4-km and 1-hour resolutions. Panels a), and b) present the spatial distribution of mean annual rainfall from MPE, and 3B42RT, respectively. Panels c) and d) present the spatial distribution of the hourly-based mean relative bias and percent RMSE, respectively.
Figure 3: Effect of resolution degradation on the spatial structure of rainfall for 100 events: a) HR-MPE range vs. CR-MPE range (solid line represents 1:1 relation); b) HR-MPE nugget vs. CR-MPE nugget (solid line represents 1:1 relation); c) Basin averaged relative RMSE (%) as a function of the ratio of ranges; and d) Basin averaged relative RMSE (%) as a function of the ratio of nuggets (log scale).
Figure 4: Simulation of streamflow during 2006 with the baseline model at a) smallest scale basin - B1, and b) largest scale basin - B5. Measures of performance for the simulation are presented in the legend.
Figure 5: Grid-based comparison of streamflow forced by CR-MPE and HR-MPE for the synthetic experiment: a) Relative Bias, b) Relative RMSE, and c) NSCE.
Figure 6: Overall propagation of mean rainfall error to streamflow simulations. Metrics of model performance as functions of basin area and streamflow threshold are presented: a) ratio of streamflow (Q) relative bias to rainfall (R) relative bias; b) ratio of streamflow (Q) relative RMSE to rainfall (R) relative RMSE; and c) NSCE of streamflow simulations. The values between parentheses indicate the probability of occurrence of the corresponding threshold.
Figure 7: Summary of error features and their propagation to streamflow modeling at basin outlet for the 100 events. Left panels (a and d): error magnitude; center panels (b and e): error variability; right panels (c and f): error spatial structure. The relative bias (upper panels) and relative RMSE (lower panels) were computed for streamflow simulations forced by CR-MPE as compared to streamflow simulations forced by HR-MPE at the basin outlet.
Figure 8: Summary of calibration (left panels) and validation (right panels) results. Metrics of model performance are plotted against basin area for the five sub-basins. The "R." and "L." refer to "Regional" and "Local" calibration runs respectively.
Figure 9: Comparison of adjusted hydrologic model parameters after calibrations forced by HR-MPE and CR-MPE. The graphs present CR-MPE to HR-MPE ratio of parameter basin average values (y-axis) versus basin area (x-axis). The horizontal thick line in each panel indicates where the ratio is 1.0.
Figure 10: Relative Bias and RMSE of 3B42RT-forced simulations from the two resulting locally calibrated parameter sets over the 2002–2009 period. Error metrics are computed using actual streamflow observations from gauged locations. The model using HR-MPE based parameter set (dashed lines) is compared to the CR-MPE based parameter set (solid lines) at different streamflow threshold values.
APPENDIX A

Error Evaluation of Rainfall Estimates after Resolution Degradation

Figure A1: Comparison of rain fields between HR-MPE and CR-MPE for the study area from 2002-2009: a) Mean annual rainfall over the study area for HR-MPE and CR-MPE; b) Relative bias (%) for each year; c) Relative Root Mean Squared Error (RMSE; %) for each year; d) Relative bias for the whole period; and e) Relative RMSE for the whole period. Histograms summarizing the spatial variability of the errors are presented for each panel. Computations are based on hourly rainfall estimates.
Figure A2: Error evaluation of basin-averaged rainfall time series as a function of rainfall threshold at different temporal scales (h – hour, D – daily, M – monthly, S – seasonal, Y – annual): a) Relative Bias for basin B1, b) same as panel a but for basin B5, c) Relative RMSE for basin B1, and d) same as panel c but for basin B5. The 1-h scale refers to instantaneous rain rates (disaggregated from 3h) while the remaining scales refer to values aggregated from 3-h rain rates. The value between parentheses in the legend indicates the frequency associated to the threshold rainfall value (i.e. percentage of time a given value occurred) based on non-zero data points.
Table 1: Summary of overall error metrics for basin averaged rainfall and streamflow at basin outlet for the 2002 - 2009 period. Values are presented for all sub-basins.

<table>
<thead>
<tr>
<th>Sub-Basin</th>
<th>Rainfall Error</th>
<th>Streamflow Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias (%)</td>
<td>RMSE (%)</td>
</tr>
<tr>
<td>B1 - 529 km²</td>
<td>-0.25</td>
<td>674</td>
</tr>
<tr>
<td>B2 - 1089 km²</td>
<td>-1.10</td>
<td>675</td>
</tr>
<tr>
<td>B3 - 1982 km²</td>
<td>-2.67</td>
<td>675</td>
</tr>
<tr>
<td>B4 - 2365 km²</td>
<td>-2.95</td>
<td>682</td>
</tr>
<tr>
<td>B5 - 5709 km²</td>
<td>-1.96</td>
<td>691</td>
</tr>
</tbody>
</table>
## APPENDIX B

List and description of Sacramento Soli Moisture Accounting Model (SAC-SMA)

Table B1: Description of SAC-SMA model parameter

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SPATIALLY DISTRIBUTED</strong></td>
<td></td>
</tr>
<tr>
<td>UZTWM</td>
<td>The upper layer tension water capacity, mm</td>
</tr>
<tr>
<td>UZFWM</td>
<td>The upper layer free water capacity, mm</td>
</tr>
<tr>
<td>UZK</td>
<td>Interflow depletion rate from the upper layer free water storage, day$^{-1}$</td>
</tr>
<tr>
<td>ZPERC</td>
<td>Ratio of maximum and minimum percolation rates</td>
</tr>
<tr>
<td>REXP</td>
<td>Shape parameter of the percolation curve</td>
</tr>
<tr>
<td>LZTWM</td>
<td>The lower layer tension water capacity, mm</td>
</tr>
<tr>
<td>LZFSM</td>
<td>The lower layer supplemental free water capacity, mm</td>
</tr>
<tr>
<td>LZFPM</td>
<td>The lower layer primary free water capacity, mm</td>
</tr>
<tr>
<td>LZSK</td>
<td>Depletion rate of the lower layer supplemental free water storage, day$^{-1}$</td>
</tr>
<tr>
<td>LZPK</td>
<td>Depletion rate of the lower layer primary free water storage, day$^{-1}$</td>
</tr>
<tr>
<td>PFREE</td>
<td>Percolation fraction that goes directly to the lower layer free water storages</td>
</tr>
<tr>
<td><strong>LUMPED</strong></td>
<td></td>
</tr>
<tr>
<td>PCTIM</td>
<td>Permanent impervious area fraction</td>
</tr>
<tr>
<td>ADIMP</td>
<td>Maximum fraction of an additional impervious area due to saturation</td>
</tr>
<tr>
<td>RIVA</td>
<td>Riparian vegetation area fraction</td>
</tr>
<tr>
<td>SIDE</td>
<td>Ratio of deep percolation from lower layer free water storages</td>
</tr>
<tr>
<td>RSERV</td>
<td>Fraction of lower layer free water not transferable to lower layer tension water</td>
</tr>
<tr>
<td>EFC</td>
<td>Effective forest fraction</td>
</tr>
</tbody>
</table>