AWRA 2010 SUMMER SPECIALTY CONFERENCE
San Juan, Puerto Rico
August 30– September 1, 2010

EVALUATION OF UPSCALING PARAMETERS AND THEIR INFLUENCE ON HYDROLOGIC PREDICTABILITY IN UPLAND TROPICAL AREAS

Alejandra M. Rojas González, Eric W. Harmsen, Sandra Cruz Pol, and Yaritza De Jesús Arce *

ABSTRACT: A hydrologic model will be used to perform real-time flood forecasting within the Mayaguez Bay drainage basin (847 km^2), located in western Puerto Rico. Minimizing run times, for this relatively large model, is important, therefore, the grid size should be optimized (minimizing runtime), while still producing accurate results. In this study we used a numerical distributed hydrologic model to simulate the hydrologic response in a small Test Bed Subwatershed (TBSW) over a three months period (October 17–December 31, 2007) using gridded rainfall derived from a dense rain gauge network (28 rain gauges within 16 km^2). The TBSW is located within the Mayaguez Bay drainage basin and is characterized by high rainfall variability over short distances (< 200 m), has an area of 3.55 square km, 29 percent average slope, and soils classified predominantly as Hydrologic Group B (60% of TBSW). Stream stage elevation was measured continuously at the TBSW outlet for use in the model calibration. An evaluation of the interrelation between different upscaling parameters is being conducted to quantify their influence on hydrologic predictability. In this paper we present the results for the baseline grid size (10 m) ensemble with 100 m rainfall resolution for an event that occurred on October 22, 2007. This study also evaluates potential evapotranspiration (PET) calculations using different methods with a daily time step for 2 years, at 2 weather stations and at the TBSW.

KEY TERMS: distributed model; hydrologic predictions; gridded rainfall; ensemble members, potential evapotranspiration.

INTRODUCTION

The overall goal of this research is to enhance flood prediction in real time and to alert the population of any flood potential, using radar information and a distributed hydrologic model. Knowledge of hydrologic prediction sensitivity within a basin is important when calibrating a model and for flood prediction accuracy. The quantification of the probability distribution function (pdf) is an important measure for understanding the basin response to changes in parameters and inputs. An up-scaling experiment will be developed by increasing the grid size to produce incrementally coarser resolution maps of each parameter and terrain input to be evaluated by an ensemble approach and generalized likelihood uncertainty estimation (GLUE) methodology. Each combination of terrain (topography) and precipitation map is called an "ensemble", and each simulation with a different parameter variation is called an "ensemble member". In this paper we present the results for one ensemble, configured with the terrain resolution of 10 m and rainfall interpolation of 100 meters for a rainfall event that occurred on October 22, 2007. Hydraulic conductivity and the roughness (Manning's n) were perturbed at 5 levels within their physically realistic limits. In total, 100 simulations were performed to create the ensemble and a best fit probability distribution function (pdf) was developed. Another specific objective presented here was to compare the Hargreaves-Samani (Hargreaves and Samani, 1985) and Penman-Monteith (Allen et al., 1998) reference or potential evapotranspiration (PET) methods with observed data at two Natural Resource Conservation Service (NRCS) – Soil Climate Analysis Network (SCAN) stations, and to derive a correction factor for the study area.

Technical Approach

To evaluate the predictability limits in the hydrologic simulation and study the interrelation between rainfall scales and terrain scales, a rain gauge network (28 tipping buckets rain gauges) was installed in a testbed subwatershed (TBSW) in western Puerto Rico (Figure 1). The rain gauges are located within a single GOES Satellite Hydro-Estimator (HE) pixel (4 km x 4 km) and 64% of the rain gauges are within in the TBSW. The rain gauge network will provide a high resolution rainfall data set (temporally and spatially) to evaluate the Collaborative Adaptive Sensing of the Atmosphere (CASA) radars,

* Respectively, Graduate Research Assistant, Dept. of Civil Engineering and Surveying, University of Puerto Rico, Mayagüez. P.O. Box 9000, Dept. of Civil Engineering and Surveying, University of Puerto Rico, Mayagüez, PR 00681-9000; phone: 1-787-464-5834; email: alejandra.rojas@upr.edu; Professor, Dept. of Agricultural and Biosystems Engineering, University of Puerto Rico–Mayagüez Campus. P.O. Box 9030; phone: 1-787-834-2575; fax: 1-787-265-3853; e-mail: eric.harmsen@.upr.edu; Professor, Dept. of Computer and Electrical Engineering, University of Puerto Rico, Mayagüez; email: SandraCruzPol@ieee.org; Undergraduate Student. Dept. of Computer and Electrical Engineering; email: yaritza.dejesus@upr.edu

estimate biases of NEXRAD products (Rojas et al., 2009), determine sub-pixel satellite rainfall variability (Harmsen et al., 2008), and understand the hydrologic response and predictability limits as well at small-scale spatial variations.

In this study we used the numerical hydrologic distributed model V*flo* (Vieux and Associates, Inc., 2004) to simulate the hydrologic response in the TBSW, with rainfall derived from the dense rain gauge network. Most of the parameter data for the distributed model were prepared in ArcGIS using a digital elevation model (DEM) with 10 m resolution, land use and SSURGO soil maps. Overland slope, flow direction, flow accumulation and stream locations were determined using the Arc Hydro Tools extension used commonly in water resources.

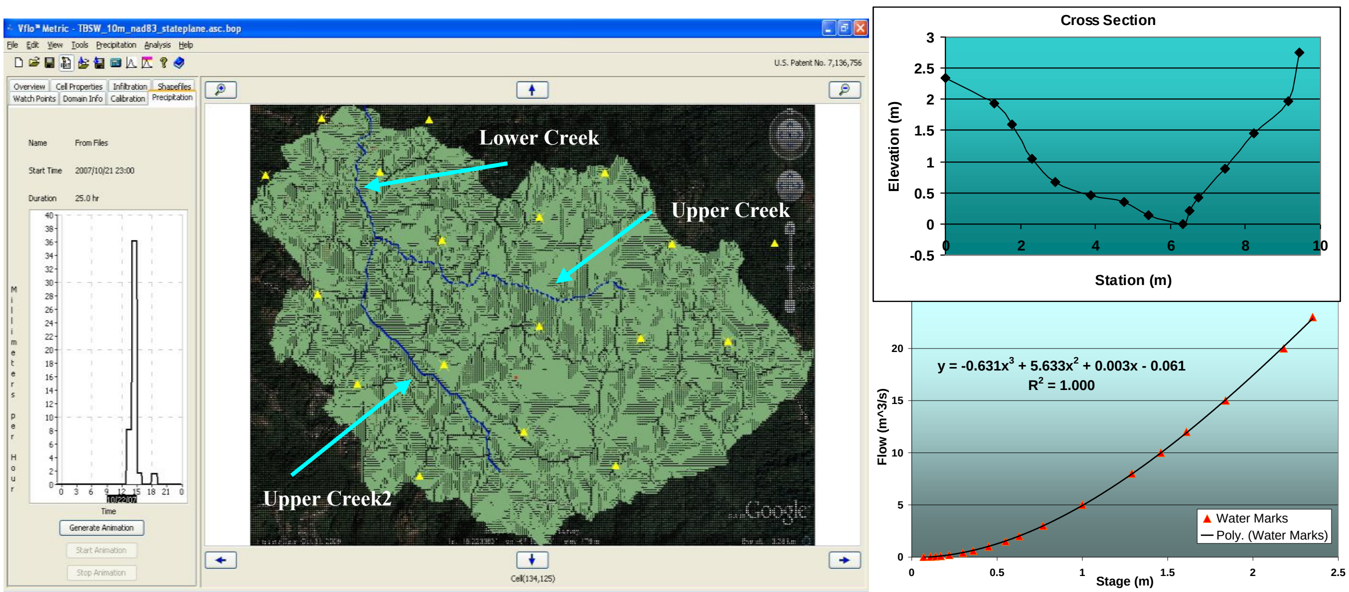

Figure 1. TBSW and rain gauge network viewed in V*flo*™ (left). Cross section measured at TBSW outlet and flow-stage curve with a third order polynomial equation (right).

The stream map generated was utilized to define the channel cells in V*flo;* channel side slopes were assumed to be 1:1; and bed channel width was set to 5 m. In most of the river sections the bed width is about 4 or 8 m. Stream flow and flow volume are not sensitive to bed width, however, stream stage is sensitive to bed width. For good flood estimation, stream cross section and rating curve information were assigned to selected model cells corresponding to the principal stream channel. The bed channel slopes were assigned by segments using the average longitudinal slope between cross sections digitized from the DEM (10 m) and corroborated with field measures (Figure 1). The Lower Creek has a longitudinal average slope of 1.25%, the Upper Creek has 2.22% and the Upper Creek2 was divided in two segments, the upstream segment shows a slope of 11.27% and the downstream segment, 3.27%. Flow stage measurements have been taken from October 2007 to October 2009. The instrument is located at 18.232667° latitude, -67.119533° longitude and 25 m elevation. Daily minimum barometric pressures were used to correct the stage measurements at the Miradero KPRMAYAG1 weather station (Lat. 18.2º and Lon. -67.13º and 22.86 m elevation), available at www.weatherunderground.com. The average adjusted stage value using the minimum pressure was 0.847 m with 0.0225 m standard deviation.

Figure 1 shows the stream cross section at the TBSW outlet, and the stream rating curve that was fitted to a third order polynomial equation with a regression coefficient of 1. The abstractions in the distributed model are calculated with the Green Ampt equation, where the principal parameters are: saturated hydraulic conductivity calculated using the percent of clay and sand, and bulk density of each applicable soil in the Rosetta Lite version 1.1 computer program (Schaap, 2001); effective porosity, soil depth and wetting front. All the parameters were assigned to a grid model resolution of 10 m. The dominant soil is Consumo soil, accounting for 59.85 % of the area. Although a clay texture soil, its B Hydrologic Group rating enhances the infiltration in the zone. Average infiltration parameters for the TBSW are tabulated in Table 1. Bouwer (1966) suggested multiplying the hydraulic conductivity by 0.5 for the saturated hydraulic conductivity in Green-Ampt model, therefore the average saturated hydraulic conductivity for the TBSW is 0.69 cm/hr. The predominant land use is forest low density with 39.36 % of the area, brush rangeland with 38.17 % of the area and 14.51 % urban land use, respectively (Table 2). The area weighted average roughness value is 0.12 for the TBSW.

Table 1. Soils type classification (SSURGO), hydrologic group and infiltration parameters in the TBSW.

Soil Name	Texture	Hydrologic Group	Area (%)	Wetting Front (cm)	K sat (cm/hr)	Depth (cm)	Effective Porosity
Consumo	Clay	B	59.85	31.63	1.273	300	0.415
Daguey	Clay	C	15.11	31.63	1.266	300	0.451
Humatas	Clay	C	25.03	31.63	1.736	300	0.454
Serpentinite	Rock Serpentine	D	0.01	3.00	5.7	300	0.26
Toa	Silty Clay Loam	B	0.01	27.30	0.294	300	0.377
Average				31.62	1.38		0.43

Table 2. Land use classification, n Manning values and K_c for Evapotranspiration quantification in the TBSW

Land use Classification	n Manning	K_c	Area (m^2)	Area Percent
Forest low density	0.1500	1.100	1399376	39.36
Shrub and brush rangeland	0.1300	1.000	1357012	38.17
Urban or built-up land	0.0150	0.300	515711	14.51
Forest high density	0.1500	1.200	208260	5.86
Baren land	0.0150	0.300	37800	1.06
Transition area	0.0500	0.300	21600	0.61
Transportation, communication	0.0150	0.300	8341	0.23
Streams and canals	0.0300	1.050	4500	0.13
Gravel pit	0.0150	0.100	1800	0.05
Native pastures	0.0450	0.850	900	0.03

The rain gauge network data was interpolated using the Inverse Distance weighted method at 100 m resolution and incorporated into the Vflo model for calibration proposes. Once the calculated and measured values are sufficiently close, an up-scaling procedure will be developed for calibrating the TBSW and the regional scale model (i.e., Mayaguez Bay drainage basin). A forecast ensemble provides the capability to estimate the forecast probability distribution function (pdf). The ensemble member will be generated using the hydrologic simulations of the Vflo model for each terrain, rainfall input and parameter perturbations. Vieux and Moreda (2003) used scalars to uniformly change the parameter maps in the calibration process with a distributed model preserving the spatial variability. The scalars used to multiply the parameters sets are calculated with the following equation:

$$N_i = \frac{1}{8}(2+3i)\Big|_{i=0,1,2,3,4} \quad (1)$$

where N_i is the adjustment factor. From a previous sensitivity analysis study (Rojas, 2008) in a larger basin in western Puerto Rico, the parameters that affected principally the peak flow are roughness, initial soil moisture and hydraulic conductivity. The discharge volume was affected by initial soil moisture, followed by overland hydraulic conductivity and soil depth. Five perturbations for each parameter mentioned above were made with the adjustment factors of 0.25, 0.625, 1, 1.375 and 1.75; except for initial soil moisture saturation that was set to 40%, 60%, 80% and 100%. During the previous month (September, 2007), prior to the hydrologic simulation, a total of 247.4 mm of rainfall by the Multisensor Precipitation Estimator (MPE) was reported and 241.45 mm of rainfall by the rain gauge network, with a monthly bias between them of 0.98 (Rojas, 2009). During the first 21 days of October, 2007, a spatial rainfall average of 57.98 mm was reported by the rain gauges and on October 22 45.0 mm of rainfall was registered.

The Vflo model requires potential or reference evapotranspiration as input. Reference evapotranspiration calculated by the Penman-Monteith method (Allen et al., 1998) at the NRCS SCAN weather stations (USDA Tropical Agricultural Research Station (TARS) at Mayaguez and Maricao Forest, PR) were compared with the Hargreaves Samani method using equation 3 below with a daily time step (October, 2007 through October 2009).

$$PET = 0.0135 * R_s * (Tave + 17.8) \quad (2)$$

where R_s is solar radiation in units of mm per day and Tave is average air temperature in degrees Celsius. The coefficient of determination (R^2) between methods is 0.9375 and the bias is 0.956 for this period, indicating that the Hargreaves Samani constant (0.0135) could be corrected by the factor 0.956 for the study area using a more simplistic formula than Penman-Monteith. Goyal et al. (1988) developed monthly linear regression equations for air temperature (mean (Tave), maximum (Tmax) and minimum (Tmin)) in Puerto Rico, which depend on the surface elevation. PET can be calculated using these linear regressions (Goyal et al., 1988) along with equ (2) and equ (3).

$$PET = 0.0023 * R_a * (Tave + 17.8) * (T\max - T\min)^{0.5} \tag{3}$$

where R_a is the extraterrestrial radiation in units of mm per day and Tmax and Tmin are the average daily maximum and minimum air temperature, respectively. Solar radiation is highly variable in space (Harmsen, 2009), therefore, the effectiveness of equ (2) and (3) to estimate PET using the temperature versus elevation relationships developed by Goyal at short time scales (daily) was evaluated in this study. Constants in Goyal's monthly linear regressions were interpolated to daily constants. All input parameters needed in the Penman-Monteith and Hargreaves-Samani methods (equ. 2) are measured by the SCAN stations. The elevation in the TARS is 13.72 m with an average temperature (Tave) of 23.9 C for the period of analysis (October, 2007-October, 2009); and in Maricao Forest the elevation is 747 m with Tave 19.7 C. The results show that the Goyal regressions at a daily time step predict the Tave with a coefficient of determination R^2 = 0.46 for TARS and 0.62 for Maricao. However, if PET is calculated with the solar radiation measured at the stations along with the Tave derived from the Goyal regressions, R^2 values of 0.987 and 0.992 are obtained at TARS and Maricao Forest, respectively (Figure 2). Values of R^2 of 0.2145 for TARS and 0.0013 for Maricao were obtained using Goyal's elevation model and equ (3). The R^2 is increased to 0.2254 in for the Maricao station if the PET is calculated using the Tave from the Goyal equations and the solar radiation is assumed to be equal to the TARS solar radiation (Figure 2). These results show that solar radiation is a spatially sensitive parameter in the PET calculation and that solar radiation cannot be assumed equal at locations distant from each other. Remotely sensed satellite measurements are suggested for a better spatially distributed solar radiation dataset, such as the method used by Harmsen et al. (2009). In this study, the PET for the TBSW was calculated using equ (2) and assuming that the solar radiation is the same as TARS, due to its relatively close proximity, around 2.5 km, compared to 16.3 km between TARS and Maricao Forest stations.

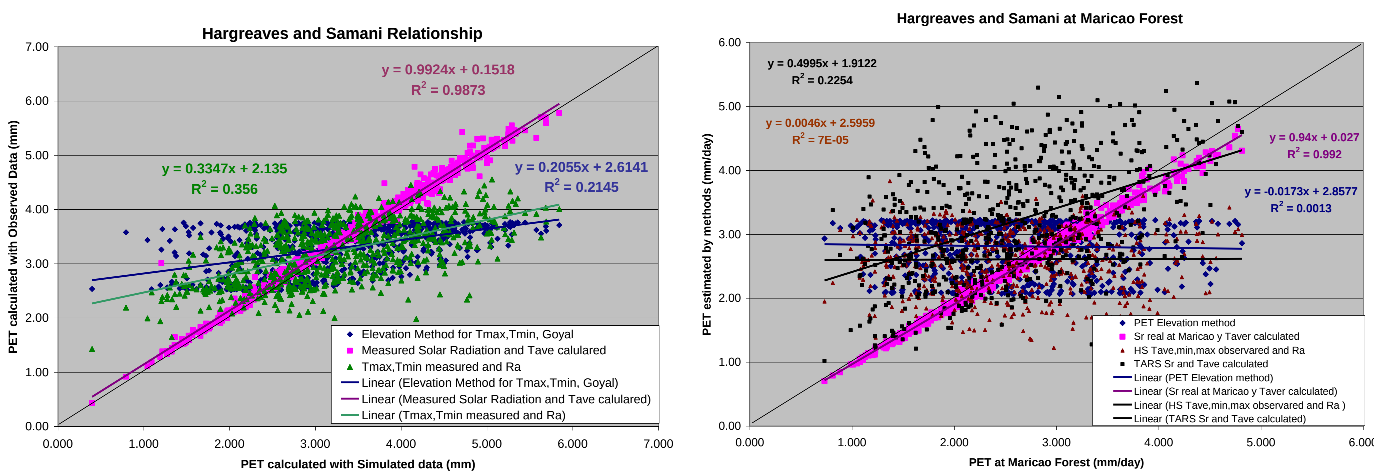

Figure 2. PET with Hargreaves-Samani relationship for observed Tmax, Tmin, Tave, solar radiation, extraterrestrial radiation; and temperatures predicted by Goyal relationships at SCAN-NRCS stations (left, TARS and right, Maricao Forest).

The event of October 22, 2007 was selected to generate one ensemble, because it was one of the largest flows measured at the flow gauge, with a discharge volume of 16.64 mm and 10.1 cubic meters per second (cms) flow, and a runoff-rainfall ratio of 0.37. A total of 125 runs were made, for each initial saturation value, where peak flow, volume discharge and peak time were compared with observed data. In this sub-watershed, simulated changes in peak discharge and volume were not apparent with changes in soil depth, therefore, those runs were eliminated from the ensemble forecast. A total of 100 runs remained in the forecast ensemble that describes the dispersion of hydrologic predictions at a 10 m terrain resolution and 100 m rainfall interpolation. Goodness of fit statistics were calculated to compare different probability distributions to the hydrologic simulation output. The Pearson correlation coefficient measures the strength of the linear relationship between the X and Y variables on a probability plot (a value close to 1 indicates that the relationship is highly linear). The Anderson-Darling statistic (AD) is a measure of how far the plotted points fall from the fitted line in a probability plot. The statistic is a weighted squared distance from the plot points to the fitted line with larger weights in the tails of the distribution, values near zero are desirables. The simulated data are plotted using the principal probabilistic distributions in Figure 3, to quantify the best fit for the ensemble. The extreme values in the runoff depth tend not to have a good fit with the pdf. The soil moisture saturation of 100% produced a very high and constant values of runoff depth, and high values of peak flow but with marked variation to changes in roughness. The best fit probability distribution was the Weibull and 3-Parameter Weibull with low AD (Table 3) and Pearson correlation coefficients of 0.98 for runoff volume and 0.989 for peak discharge (Table 3).

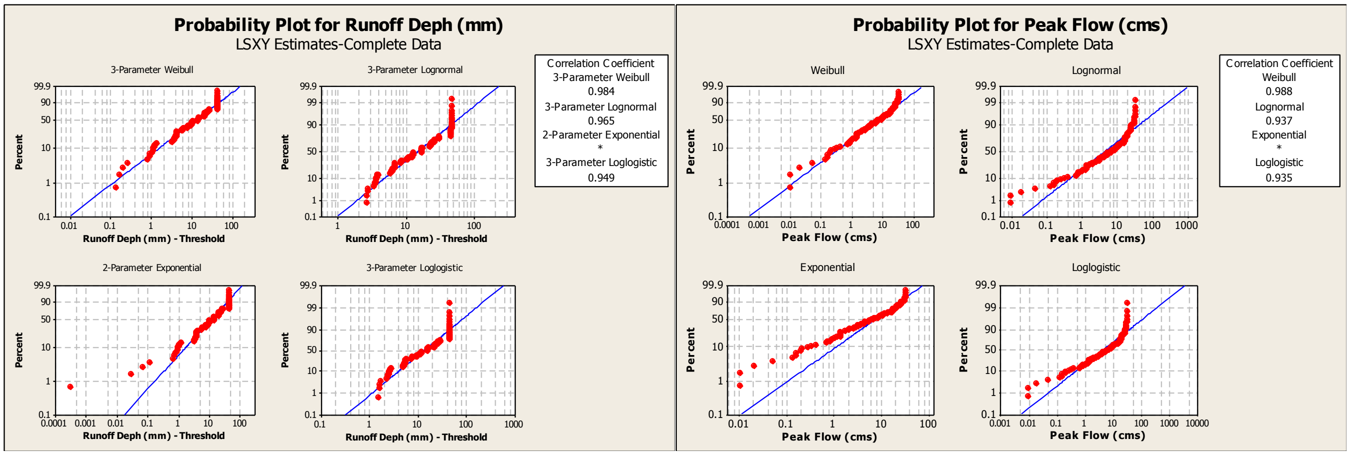

Figure 3. Probability distribution functions pair to ensemble forecast simulated data for runoff depth and peak flow.

Table 3. Best fit statistic for probability distributions

	Runoff Depth		Peak Discharge	
Distribution	**Anderson-Darling (AD)**	**Correlation Coefficient**	**Anderson-Darling (AD)**	**Correlation Coefficient**
Weibull	**1.759**	**0.981**	**1.243**	**0.988**
Lognormal	3.714	0.925	3.358	0.937
3-Parameter Weibull	**1.704**	**0.984**	**1.18**	**0.989**
3-Parameter Lognormal	2.469	0.965	2.248	0.97
Normal	6.927	0.923	5.409	0.936

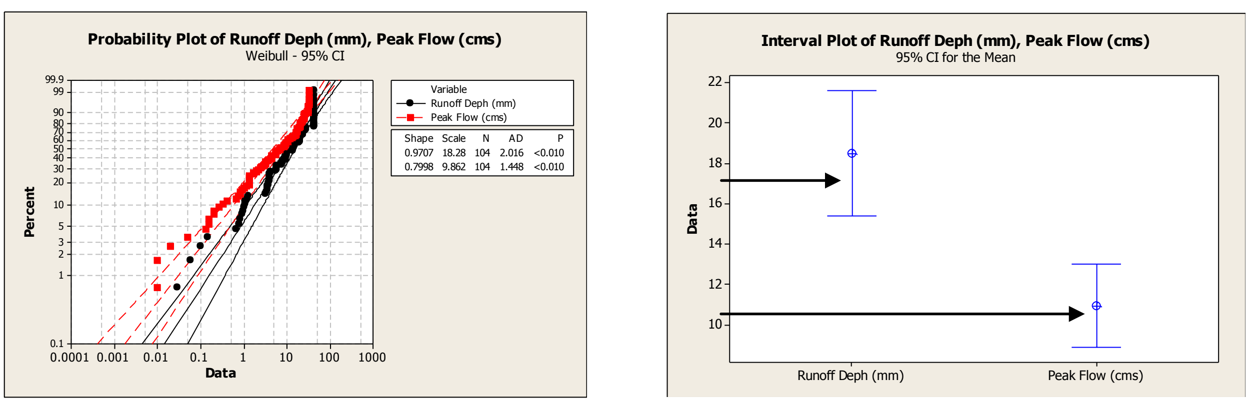

Figure 4. Probability and interval plot of runoff depth and peak flow with a 95% confidence interval and Weibull distribution.

Figure 4 shows the 95% confidence intervals for two variables for a Weibull distribution (left) and the 50% quartile or median with the lower (5%) and upper (95%) quartiles, (right). From Figure 4 it can be observed that the "true" runoff depth and peak flow have a 55% and 60% chance of being selected, respectively. Statistics of behavior of this ensemble (10 m terrain and 100 m rainfall resolution) for one event (October 22, 2007) shows a mean absolute error (MAE) of 8.9 cms for peak flow, 1.04 minutes for time to peak and 13.5 mm for runoff volume. Calculated root mean square errors (RMSE) were as follows respectively: 10.56 cms, and 16.07 mm for runoff volume. A bias is defined here as the ratio between the average of simulated and observed data, computing 1.14 for peak flow, 1.12 for time to peak and 1.0 for runoff volume.

CONCLUSIONS

In this paper we presented the results for the baseline grid size (10 m) ensemble with 100 m rainfall resolution for an event that occurred on October 22, 2007, in the TBSW located in western PR. The pdf for runoff depth and peak flow, and various other statistics were presented based on 100 simulations. The best fit pdf for the simulated data were the Weibull and the 3 parameter Weibull distributions for peak flow and runoff volume. Assumption of normal distribution is not recommended for these data. Estimated biases were 1.14 for peak flow, 1.12 for time to peak and 1.0 for runoff volume. In the future, statistics of the others ensembles will be developed and compared between them to measure the behavior of flood prediction with changes in parameters and inputs. A better understanding is needed of the watershed to improve the calibration and flood prediction; and this can be achieved by evaluation of additional ensembles using wet and dry season events. Future efforts will also include recommendations for optimal radar spatial resolution needed to achieve accurate hydrologic predictions.

The temperature/elevation linear regression equations of Goyal et al. (1988) were evaluated to calculate the PET at a daily time step using the Hargreaves-Samani equation and the results show similar regression coefficients between observed and calculated Tmax, Tmin and Tave. Therefore the most sensitive parameter is the solar radiation, because the elevation model (Goyal et al., 1988) cannot represent the spatial variability of this parameter using the daily interpolation for extraterrestrial radiation and the Tmax and Tmin calculated with the elevation model. In this study, PET was estimated using the Hargreaves-Samani equation with the solar radiation for the TBSW assumed to be equal to that at the TARS, and average daily temperature estimated from the Goyal et al. (1988) model.

ACKNOWLEDGMENTS

This work is supported primarily by the Engineering Research Centers Program of the National Science Foundation under NSF (grant 031374) and NOAA-CREST (grant NA06OAR4810162).

REFERENCES

Allen, R. G., L. S. Pereira, Dirk Raes and M. Smith, 1998. Crop Evapotranspiration Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, Rome.

Bouwer, H., Rapid Field Measurements of Air-Entry Value and Hydraulic Conductivity of Soil as Significant Parameters in Flow System Analysis. Water Resour. Res., v. 2, n. 4, pp. 729-738, 1966.

Hargreaves, G. H. and Z. A. Samani, 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric., ASAE. 1(2). pp.96-99.

Harmsen, E. W., S. E. Gomez Mesa, E. Cabassa, N. D. Ramirez Beltran, S. C. Pol, R. J. Kuligowski ,R. Vasquez, 2008. Satellite Sub-Pixel Rainfall Variability. WSEAS Transactions on Signal Processing. Issue 8, Volume 7, (504-513)

Harmsen, E. W., J. Mecikalski, M. J. Cardona-Soto, A. Rojas González and R. Vasquez, 2009. Estimating daily evapotranspiration in Puerto Rico using satellite remote sensing. WSEAS Transactions on Environment and Development. Vol. 6(5):456-465.

Gourley, J. J., Vieux, B. E. (2005). A method for evaluating the accuracy of Quantitative Precipitation Estimates from a Hydrologic modeling perspective. American Meteorological Society. April, 2005. 115-133.

Goyal, M.R., E.A. González and C. Chao de Báez, 1988. Temperature versus elevation relationships for Puerto Rico. J. Agric. Univ. P.R. Vol. 72(3): 449-467. July, 1988

Rojas-González, A.M., E. W. Harmsen, S. Cruz-Pol, 2010. Performance and Evaluation of Multi-sensor Precipitation Estimation Algorithm using a High Density Rain Gauge Network and Hydrologic Simulations. 21st Century Watershed Technology: Improving Water Quality and the Environment, Earth University, Costa Rica, February 21-24, 2010.

Rojas-González, A. M., E. W. Harmsen and S. Cruz Pol, 2009. Performance evaluation of MPE rainfall product at hourly and daily temporal resolution within a Hydro-Estimator pixel. WSEAS Transactions on Environment and Development. Vol. 7(5):478-487.

Schaap, M.G., F.J. Leij, and M. Th. van Genuchten. 2001. Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. Journal of Hydrology. 251:163-176.

Vieux, B.E., Moreda, F.G. (2003). Ordered physics-based parameter adjustment of a distributed model. In: Duan, Q., Sorooshian, S., Grupta, H.V., Rousseau, A.N., Turcotte, R. (Eds.), Water Science and Application Series, vol. 6. American Geophysical Union, pp. 267-281, ISBN 0-87590-355-X (Charpter 20).