Hydrology and Earth System Sciences Discussions

A multitemporal remote sensing approach to parsimonious streamflow modeling in a southcentral Texas watershed, USA

B. P. Weissling[1,2], H. Xie[1], and K. E. Murray[1]

[1]Department of Earth and Environmental Sciences, University of Texas San Antonio, 6900 N. Loop 1604 West, San Antonio, Texas, 78249, USA
[2]SWCA Environmental Consultants, 6200 UTSA Boulevard, Suite 102, San Antonio, Texas, 78249, USA

Correspondence to: B. P. Weissling (bweissling@swca.com)

1

Abstract

Soil moisture condition plays a vital role in a watershed's hydrologic response to a precipitation event and is thus parameterized in most, if not all, rainfall-runoff models. Yet the soil moisture condition antecedent to an event has proven difficult to quan-
5   tify both spatially and temporally. This study assesses the potential to parameterize a parsimonious streamflow prediction model solely utilizing precipitation records and multi-temporal remotely sensed biophysical variables (i.e. from Moderate Resolution Imaging Spectroradiometer (MODIS)/Terra satellite). This study is conducted on a 1420 km^2 rural watershed in the Guadalupe River basin of southcentral Texas, a basin
10   prone to catastrophic flooding from convective precipitation events. A multiple regression model, accounting for 78% of the variance of observed streamflow for calendar year 2004, was developed based on gauged precipitation, land surface temperature, and enhanced vegetation Index (EVI), on an 8-day interval. These results compared favorably with streamflow estimations utilizing the Natural Resources Conservation Ser-
15   vice (NRCS) curve number method and the 5-day antecedent moisture model. This approach has great potential for developing near real-time predictive models for flood forecasting and can be used as a tool for flood management in any region for which similar remotely sensed data are available.

1   Introduction

20   A watershed's hydrologic regime, in terms of water balance models, is critical to understanding numerous environmental and biophysical processes operating at varying spatial and temporal scales. Quantifying the spatial and temporal response of runoff to a precipitation event is critical to understanding processes of erosion, sediment transport, contaminant loadings, flood hydrology, and ecological impact. Soil types, texture,
25   and condition, surface storage, and soil moisture are primary parameters that influence both infiltration and subsequent runoff. Other primary and secondary parameters

that influence infiltration and runoff are vegetation type, density and condition, topography, impervious surface, and various artificial barriers to natural flow such as roads, culverts, and berms.

Effective hydrologic modeling of a watershed's response to a precipitation event is partially dependent on the assessment and parameterization of the soil moisture condition within that watershed in the days or weeks preceding the precipitation event. Commonly referred to as the antecedent moisture condition (AMC), the soil moisture condition can account for large variations in runoff from similar precipitation events within a watershed or from the same storm event in otherwise identical watersheds (Heggen, 2001).

Both physically- and empirically-based hydrologic models depend on the characterization and/or quantification of soil moisture antecedent to the precipitation events and the subsequent runoff being modeled. Modeling approaches based on physical parameterization, such as that employed in the Green and Ampt and Holtan models for infiltration, depend on the quantification of soil moisture as available storage or effective saturation. These data must either be collected from soil moisture probe studies or inferred from empirical analysis of the precipitation history of the watershed or catchment being studied. The spatial variability of soil properties and condition and their respective infiltration characteristics suggest that physical approaches quickly lead to massive data requirements if point estimations of soil moisture and infiltration dynamics are required across an entire watershed.

The question as how to capture the physical dynamics of infiltration and runoff in minimally parameterized hydrologic models has traditionally been answered by the application of empirical parameterization of rainfall-runoff relationships. The widely used rainfall-runoff model, the Natural Resources Conservation Service (NRCS) curve number (CN) method utilizes a "triad" of antecedent moisture conditions (dry, normal, and wet) based on 5-day antecedent precipitation for both growing and dormant vegetative season (USDA-SCS, 1985). While this antecedent soil moisture and precipitation approach was developed based on midwestern U.S. agricultural watersheds of less

than $250\,\text{km}^2$ (Ponce, 1989), it has been applied in widely varying environmental, geographic, hydrologic and climatologic settings (Mustafa, 2005; Melesse et al., 2004).

Although many studies utilizing the original or adjusted NRCS AMC curve number approach report favorable results, other studies have questioned their validity. Miserocchi and Savi (2005), Mishra et al. (2004), Centolani and Savi (2003), and Heggen (2001) have concluded that the curve number itself is far from being a static parameter applied to a designated association of land use land cover (LULC), and soil type or hydrologic response, and that the variability of CN on a storm by storm basis cannot be adequately explained by the generalized NRCS method for assessing soil moisture antecedent to a storm event. This growing recognition of the stochastic nature of the CN has not lessened the enthusiasm of CN practitioners. Papers continue to be published in peer-reviewed journals describing CN applications to watershed hydrology, with both conventional and modified AMC models (Jacobs et al., 2003; Nachebe, 2006, Melesse and Graham, 2004; Fennessey et al., 2001; Moglen, 2000; Grove et al., 1998, Shirmohammadi et al., 1997). A trend, however, of many of these recent studies is to augment and extend the common CN methodology, as well as other empirical and physical models, with incorporation of remotely sensed hydrologic parameters.

The advent of remote sensing (RS) with its synoptic view, multiple sensors, and repetitive coverage coupled with Geographic Information System (GIS) tools have significantly facilitated the provision of land use and hydrologic information as well as the processing, management and interpretation of hydrologic data. While many researchers have recognized the importance of remote sensing in hydrologic studies in general, others have just recently begun to evaluate the role of RS in the characterization and quantification of soil moisture as it relates to hydrologic models (Cashion et al., 2005; Jacobs et al., 2003; Schmugge et al., 2002). RS passive microwave and radar products have been incorporated into traditional curve number models in lieu of the NRCS AMC adjustments and have achieved improved performance of those models in estimating and predicting runoff from precipitation events (Jacobs et al., 2003). Land surface temperature, as a measure of the thermodynamic state of the surface

and near-surface landscape, is influenced by soil and vegetation canopy moisture, by surface roughness, and by albedo (Schmugge et al., 2002). Durre and Wallace (2000) derived a dependency relation between summertime daily maximum temperature and antecedent soil moisture in the Central and Eastern U.S., whereby a temperature signal retains a memory of a soil moisture anomaly for up to several weeks. In another study (Adegoke and Carleton, 2002), remote sensing vegetation indices of midwestern U.S. cropland and forest were found to correlate to field-acquired soil moisture. The authors concluded that a long-term memory of soil water content, of several weeks, existed in the vegetation canopy.

Despite many improvements of empirical runoff models with the consideration of antecedent soil moisture, whether by generalization or by the inclusion of remotely-sensed data, these models have limited application in many parts of the world. Water managers for regions that do not have or have only incomplete LULC and soil data will not be able to apply any variation of the NRCS CN method because these parameters are essential inputs for the model. By exploiting the documented "memory" of soil moisture in land surface temperature and vegetation cover, this study tests a hypothesis that a rainfall-runoff model can be built solely from these remote sensing parameters that indicate the antecedent soil moisture condition of a watershed and the resultant runoff or streamflow response to that moisture condition. For comparison purposes, a precipitation-runoff model was also developed for the studied watershed utilizing both the curve number and the 5-day AMC adjustments.

2 Background of study site

The Guadalupe and San Antonio river basins extend from their headwaters on the Edwards plateau in central Texas to the Texas gulf coast estuaries and bays (Fig. 1). The middle regions of these basins, as selected for this study, are bounded to the northwest by the Balcones escarpment, and to the southeast by the gently sloping coastal plains. Topographically, these mid-basin watersheds are characterized by low rolling hills and

gently sloped stream valleys. Sandies Creek, a tributary of the Guadalupe River, drains approximately 2000 square kilometers of predominately rural land in Guadalupe, Gonzales, Wilson, Karnes, and DeWitt counties. Land cover types in Sandies Creek watershed are mainly oak and hickory woodlands with relict stands of long-leaf pine, interspersed with shrub and grasslands. Agricultural use is primarily grazing and hay production. Soil types in the watershed vary hydrologically from clay-rich, high runoff potential soils of the hydrologic groups C and D predominating in the inter-stream areas of the watershed and sandy, low runoff potential soils of the hydrologic groups A and B predominating in floodplains and along stream courses. Hydrologic group A soils are also associated with the occurrence of the pine stands.

Precipitation in the region is primarily convective, typically associated with the convergence of frontal systems and moisture masses moving inland from the Gulf of Mexico, as well as from the Gulf of California and Pacific Ocean. Mean annual precipitation across the respective river basins ranges from 750 to 1050 mm, along a west to east gradient. The majority of significant rainfall events occur in late Spring (May–June) and early Autumn (September–October). The region is also known for locally intense precipitation events that are attributed to the orographic influence of the Edwards Plateau and Balcones Escarpment, tropical moisture streams from the Gulf of Mexico and/or the Pacific, and the destabilizing influence of late Spring and early Autumn frontal systems. This region of central Texas holds numerous world records for precipitation intensities, from 2 h events to 2 day events (Larkin and Bowmar, 1983; Smith et al., 2000). Likewise, the region is also home to catastrophic flood events.

Sandies Creek watershed has an extensive history of agricultural use and the stream course itself has recently been listed, along with neighboring Elm Creek, as an impaired stream by the Texas Commission on Environmental Quality (TCEQ, 2006). The watershed also has a significant history of flooding. An understanding of the watershed's runoff response to precipitation events is therefore critical to developing a water quality assessment and flood mitigation program.

3 Methodology

3.1 Acquisition of precipitation and streamflow data

Two National Weather Service (NWS) Cooperative weather stations, Gonzales 10 SW and Gonzales 1 N, were selected as data sources for gauged precipitation. Gonzales 10 SW is located a few kilometers just inside the north-eastern boundary of the watershed and Gonzales 1N is located 13 km north of the watershed. With one gauge station within the watershed boundary, the assumption that a single station data record would reliably represent spatially varying precipitation within the watershed was questionable. Monthly precipitation totals for both gauges from 1997 to 2005 were analyzed for collinearity to address this issue (Table 1). Precipitation events at the two stations showed the strongest correlations in 1998 and 2004, indicating a greater likelihood for spatially uniform precipitation across the region in these two years. On the basis of the availability of multi-temporal remote sensing data, year 2004 was selected for the model development of this study.

The sole gauging point for Sandies Creek, situated several kilometers upstream from the confluence of Sandies Creek and the Guadalupe River, is a United States Geologic Survey (USGS) gauge station (ID 8175000) (29.2153 N, 97.4494 W). With a 75 year data record, this station was selected as the outlet for a 1420 km^2 watershed as delineated with ArcGIS hydrology tools (ESRI Inc., 2006). The overall mean flow rate for the year 2004 at this gauge station was 3.54 m^3 s^{-1} with the mode occurring at 0.40 m^3 s^{-1}, a rate essentially equivalent to the baseflow component of the total flow. The maximum daily flow rate, corresponding to a 7-day precipitation event in the third week of November 2004, was 121.12 m^3 s^{-1} and the minimum rate was 0.017 m^3 s^{-1}, occurring on 17 August 2004 after a dry period of 20 days and record high temperatures.

Daily mean precipitation and streamflow events for Sandies Creek watershed (Fig. 2) for 2004 are representative of the seasonal precipitation trends in the region with late Spring and mid Autumn events responsible for most of the runoff from the watershed.

7

3.2 Pre-processing remote sensing data

Whereas many land surface remote sensing studies demand high spatial and spectral resolution data, a rainfall-runoff hydrologic model to be based on remote sensing data is more dependent on high temporal resolution. The National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra satellite, launched in 1999, images the entire earth's surface every 1 to 2 days in 36 spectral bands from the visible to the middle infrared (http://modis.gsfc.nasa.gov). MODIS raw data (Level 1A and 1B) and processed products (Level 2, 3 and 4) at spatial resolutions ranging from 250 m to 1000 m are processed and delivered at no cost to the public by the Earth Resources Observation Systems (EROS) data center of the USGS . Data products known to be sensitive to or indicative of land surface moisture condition assessed in the study were: the Land Surface Temperature/Emissivity (MOD11A2) 8-day (1 km) product, the Vegetation Indices (MOD13A2)16-day (1 km) product, and the Albedo (MOD43B3) 16-day (1 km) product.

The Land Surface Temperature/Emissivity (MOD11A2) product provides a per-pixel average surface temperature estimation of the composited vegetation canopy and soil, with an accuracy of 1 degree Kelvin (Wan et al., 2004). The Vegetation Indices (MOD13A2) product provides both the standard Normalized Difference Vegetation Index (NDVI), and an Enhanced Vegetation Index (EVI). EVI has improved sensitivity to high biomass regions and is de-sensitized to canopy background signal and atmospheric influences (Gao et al., 2000; Miura et al., 2001). The Albedo (MOD43B3) product provides surface albedo, defined as the ratio of upwelling or reflected radiative flux to downwelling flux, corrected for both direct and diffuse components. All MODIS products are corrected for atmospheric scattering and absorption.

An additional remote sensing product examined in this study was a daily microwave emission or brightness temperature product from the Advanced Microwave Scanning Radiometer for the EOS (AMSR-E) sensor launched in 2002 on board NASA's Aqua satellite. The global coverage AMSR-E sensor assesses the microwave temperature

8

of the earth's surface twice daily, on both ascending (night) and descending (day) orbits. This level 3 processed product (AE_Land3) includes daily surface soil moisture and vegetation water content and roughness data, both of which are key inputs to hydrologic and climate models. AMSR-E products are available free of charge at the National Snow and Ice Data Center (NSIDC), online at http://nsidc.org/data/amsre/.

All original MODIS datafiles, in HDF format, were re-projected from the original Equal-Area Sinusoidal projection to a UTM Zone 14 WGS84 datum projection and then spatially subsetted to the bounding coordinates of the study area in a new GeoTIFF image format. This was accomplished using a software tool called the MODIS Re-Project Tool, available from the USGS web portal online at http://edcdaac.usgs.gov/datatools. asp. An ArcInfo Arc Macro Language (AML) script (Xie et al., 2005; Zhou et al., 2005) was employed to convert the subsetted MODIS TIFF images to an ArcGIS grid format, to clip each grid by the polygon coverage of the watershed boundary, and to create ASCII format output of all pixel values. These pixel values, as 8 and 16 day forward-in-time event means, were lumped or averaged for the entire watershed.

The original AMSR-E datafiles were subsetted to the bounding coordinates of the watershed prior to download from the NSIDC. The processing routine consisted of manually regridding and geolocating each AMSR file (in an original Equal-Area Scalable Earth Grid) utilizing an ENVI georeferencing routine to map all image pixels to known X and Y coordinates using a Geographic Lookup Table (GLT). This GLT is built in ENVI using the geolocation geometry information downloaded with the AMSR product. A second manual processing step involved converting all georeferenced files to TIFF format. A similar ArcInfo AML script, as described above, was employed to resample the 25×25 km AMSR pixel to a finer resolution of 2.5×2.5 km in order to achieve a more accurate fit to the polygonal boundary of the watershed, which at the 25 km pixel resolution the watershed would encompass less than 4 pixels. Once resampled, the new TIFF formatted file was clipped to the watershed boundary, and summary pixel values were output to a text file. In particular, the AMSR-E daily soil moisture product for daytime descending orbits was examined in this study.

9

3.3 Curve number model development

3.3.1 Soils and land use land cover data

The hydrologic group designation for any soil type can be either A, B, C, or D, where the runoff potential increases from A to D. The source for specific soil hydrologic group type (SSURGO) is distributed by the NRCS of the USDA. This SSURGO database includes both tabular soils data and vector maps of soils as mapped by county agencies. Land cover information, as 21 separate land cover classes, is available in raster format from the National Land Cover Dataset (NLCD) distributed by the USGS. These data are generated from unsupervised classifications of Landsat 5 imagery collected in 1992. The landscape bounded by the extent of the Sandies Creek watershed was represented by 16 individual land cover classes in 8 categories. Four categories – forested land, shrubland, grassland, and pasture/cultivated lands – represented 98.9% of the total land area of the watershed, the remainder being wetlands, urban lands, barren lands, and open water. Soil hydrologic group types were predominately D and C (soil types with low infiltration and high runoff potential), representing 57% and 21% of the total watershed area, respectively. Soil types B and A (soils with higher infiltration but lower runoff potential), represented 13% and 9% of watershed area, respectively.

Geographic Information Systems (GIS) allows for merging vector-based soil map units with the raster-based land use and land cover (LULC). The resultant product is a raster image of each soil map unit and LULC classification, commonly referred to as a hydrologic response unit (HRU). Curve numbers for each LULC and soil type combination are readily available in lookup tables found in various hydrology handbooks and manuals, such as the TR-55 handbook (USDA, 1986). The specific area of each HRU can be used to calculate a spatially weighted composite curve number (CN_{comp}) for the watershed according to the following equation

$$CN_{comp} = \frac{\sum A_i CN_i}{\sum A_i} \qquad (1)$$

10

with A_i and CN_i being the area and CN for each HRU within the watershed.

3.3.2 Antecedent moisture condition model

The actual composite curve number for the watershed was adjusted for antecedent moisture condition on an event by event basis according to the following empirical equations.

$$CN(I) = \frac{4.2CN(II)}{10 - 0.058CN(II)} \tag{2}$$

$$CN(III) = \frac{23CN(II)}{10 + 0.13CN(II)} \tag{3}$$

Where CN(II) represents the composite mean curve number for the watershed, CN(I) the adjusted curve number for dry conditions, and CN(III) the adjusted curve number for wet conditions (Chow et al, 1985) based on the National Engineering Handbook (NEH-4) (USDA-SCS, 1985). Table 4.2 of NEH-4 provides guidance for this adjustment based on cumulative precipitation for the preceding 5 days, with adjustments for dormant and growing season conditions (Table 2). For this application, the growing season was defined as being from 1 May to 30 September.

In recognition that a watershed's moisture condition must vary dynamically and not statically across just three states "dry, normal, and wet" as suggested by the AMC model above, an exponential smoothing filter (with a damping factor of 0.7) was run across the calculated curve number values for the data period record. This resulted in a more hydrologically realistic gradational change of CN (Fig. 3).

3.3.3 Calculation of uniform depth of runoff

For a given precipitation event, the CN method partitions a given uniform depth of precipitation into a runoff component and an infiltration component through the following

11

two equations (USDA, 1986).

$$S = \frac{25400}{CN_{comp}} - 254 \tag{4}$$

$$Q = \frac{(P - I_a)^2}{(P - I_a + S)} \tag{5}$$

Where S represents the maximum potential retention of the watershed in mm. I_a, the initial abstraction, represents that portion of retention associated with interception, ponding, and wetting of soil and vegetation surfaces. I_a is usually approximated as 0.2S. For a given precipitation event (P), Q, the uniform depth of runoff (in mm) is then determined.

3.3.4 Baseflow extraction

The NRCS curve number method generates a uniform runoff depth (Q) for a watershed for an effectively uniform precipitation event. The resultant flowrate, determined by integrating Q over the watershed area and accounting for channel hydraulics, represents runoff but not necessarily streamflow, since streamflow includes both baseflow and runoff components. While the baseflow component can be isolated from a streamflow record through manual interpretation of the hydrographs, an automated method provided by the Baseflow Filter program, a software-based routine available online from USDA's Soil and Water Assessment Tool (SWAT) website (USDA – SWAT, 2006), was employed to extract an estimated baseflow. Using a methodology outlined by Arnold and Allen (1999), the Baseflow Filter program, through analysis of the recession limb of hydrographs, extracts an estimated baseflow from a streamflow record in 3 successive passes of the extraction algorithm. Depending on a user's conservative vs. aggressive interpretation, any one of the three passes can be selected as the final baseflow estimation. For this study, the baseflow for actual streamflow events of year 2004 was extracted as pass 1, the most conservative filter. On a day by day basis, these baseflow

12

estimations are then summed with the curve number runoff estimations, as described below, to produce a modeled streamflow for the watershed for 2004.

3.4 Remote sensing model development

With calendar year 2004 selected for the model, an 8-day mean datapoint interval was selected on the basis of MODIS 8-day surface temperature product. This interval appeared to be the best compromise of daily precipitation and streamflow data, and the 16-day MODIS vegetation indices and albedo products. An 8-day forward mean dataset was created for precipitation and gauged streamflow producing 46 individual data files for the year, each corresponding to the time frame of the MODIS 8-day data. The 16-day products were converted to an 8-day product by interpolation. The AMSR-E soil moisture product was the only dataset not based on an 8-day mean, but rather on individual days on an 8-day interval.

An examination of the daily precipitation and streamflow records for 2004 indicated a significant lag time between most rain event peaks and the subsequent response or streamflow peaks, typically 2 days. For most watersheds, this natural lag time can range from a few hours to days. In the creation of the 8-day mean precipitation periods, failing to include a data period offset for precipitation could preclude the precipitation events that caused the streamflow event. A separate correlation test was run on each of five offset periods for precipitation, from a 0-day to a 4-day offset. The highest correlation occurred when the 8-day mean period was offset or advanced 3 days – sufficient time to allow for the watershed's apparent lag time. Plots of both 8-day mean precipitation and streamflow records for offsets of 0 and 3 days are shown in Fig. 4. Note that the 3-day offset precipitation event means demonstrate a much improved peak-to-peak correspondence to streamflow means.

Given that the regression model was evaluating the effects of antecedent events and moisture states on streamflow response, it was necessary to evaluate the 4 primary MODIS parameters, NDVI, EVI, albedo, and surface temperature (Ts), at various time offsets or advances from the 8-day streamflow data period. Unlike the procedure uti-

13

lized to assess the best correlation of precipitation to streamflow, by offsetting the 8-day period means one day at a time, the offsetting of NDVI, EVI, albedo, and Ts had to occur in full 8-day increments. Three offsets were tested in the regression model for each parameter, a 2 period advance (16 day), a 1 period advance (8 day), and no advance (0 day). The AMSR-E soil moisture product was tested with a 0-day and with one 8-day period advance.

Prior to performing the forward stepwise regressions, a distribution analysis of the 8-day mean gauged streamflow record indicated a non-normal distribution, verified by the Shapiro-Wilks statistical test for normality. A log transformation of the streamflow data prior to the regression analysis satisfied the normal distribution assumptions for linear regression models.

4 Results

4.1 Curve number model

From the assemblage of LULC and soil hydrologic groups encompassed within Sandies Creek watershed, the spatially weighted composited curve number was calculated as 71.8. This equates to an adjusted CN of 51.7 for "dry" AMC(I) conditions and a CN of 85.4 for "wet" AMC(III) conditions using Eqs. (2) and (3). Direct comparisons of curve number 8-day aggregated modeled streamflows for a single composite curve number, the three AMC adjusted curve numbers, and a smoothed AMC continuum of curve numbers are shown in Fig. 5. Since individual CN-derived runoff depths are synchronous with the precipitation record, the record of all runoff events were offset 3 calendar days from the precipitation events (prior to being aggregated as 8-day means) to account for the watershed's apparent lag time and to facilitate comparison to the RS model.

An examination of Fig. 5a strongly suggests that a single composite but lumped CN parameter is insufficient to reasonably estimate streamflow (runoff plus estimated

14

baseflow). The cumulative gaged flow for 2004 was $112.2 \times 10^6 \, \text{m}^3$. The lumped CN modeled flow was $152.6 \times 10^6 \, \text{m}^3$, representing an overestimate of 36.1%. While this overestimate could be considered reasonable on an annual basis, individual 8-day comparisons are poor both in terms of magnitude and peak to peak correspondence, strongly suggesting the importance of modeling antecedent moisture conditions.

The AMC modification of the lumped curve number approach, by including 3 moisture states and their respective adjustments to CN, offers a considerable improvement in both individual 8-day magnitudes and peak to peak correlation (Fig. 5b). The cumulative modeled flow was $111.4 \times 10^6 \, \text{m}^3$; a 6.9% underestimate of cumulative gauged flow. However, there are noticeable discrepencies in individual 8-day events, such as the near 100% overestimate in peak flow for the period 21 and 42 events (9 June and 16 November data periods, respectively) and peak to peak mismatches at events 24 and 36.

The smoothed curve number model (Fig. 5c), the more hydrologically realistic one considering temporally varying AMC, underestimates cumulative annual flow by 28.7%, with a total flow of $80.0 \times 10^6 \, \text{m}^3$. Overall, it appears to be the best fitting model in terms of 8-day event magnitudes and peak to peak correspondence.

4.2 Remote sensing model

A multiple linear regression model was developed to examine the potential of remote sensing to estimate an 8-day mean streamflow response of Sandies Creek watershed to 8-day mean precipitation events for 2004. The conventional CN derived streamflows were used for comparison purposes.

Prior to regression, all RS parameters were checked for colinearity. Table 3 shows all pairwise parameter correlations, along with significance probabilities and a correlation chart. From this table it is apparent that there is a weak correlation of the two vegetation indices and albedo to temperature, but a much stronger correlation of albedo to both NDVI and EVI. Any r^2-value above 70% in a pairwise correlation is cause for concern.

15

Rather than exclude the albedo parameter a-priori, it was decided to leave it in the model, check for its level of significance, and then exclude it if necessary.

With 15 regression parameters now included (6 original parameters plus 9 variants based on time offsets), the "best" parameter subset for model inclusion was obtained using forward stepwise multiple regression. This regression method involves evaluating the significance of a parameter in reducing the sequential sum of squares of the model as that parameter enters the model. The significance probability criteria for entering the model was set at $p = 0.250$. A total of 6 parameters successfully entered the model giving the following estimates of model success (Table 4).

An examination of the step history revealed that the first parameter to enter the model was precipitation (3-day advance or offset) with 46.6% of the streamflow variability explained. Surface temperature, with a 1 period or 8-day offset, accounted for an additional 14% of the streamflow variability. EVI with a 0 period offset accounted for another 20% of streamflow variability. Altogether, these first three parameters accounted for 80.2% of the variability in log transformed streamflow (Y), while the subsequent 3 parameters only improved the model by an additional 6%. Despite the "best" Mallow's Cp criteria – a factor indicating model adequacy – occurring at step 4 or 5 (point at which Cp is closest to p, the number of parameters), the parameters associated with steps 4 and 5 are repeats of the parameters of steps 2 and 3, with different offsets. The removal of these 2 parameters from the model do not significantly degrade the results ($r^2_{adj} = 0.782$), but once removed, AMSR-E derived soil moisture is no longer significant at $P = 0.796$. With the soil moisture parameter removed, the final multiple regression model includes precipitation (3 day offset), temperature (8 day offset), and EVI (no offset), with the final $r^2_{adj} = 0.787$. All parameter estimates were significant at $P < 0.0001$ (Fig. 6), as indicated by the individual leverage plots for each parameter. A leverage plot is a plot of parameter values against the predicted residuals. The 95% confidence limit curves, as displayed in the plots, indicate parameter significance if the curve region does not contain the line denoting the zero value residual (horizontal dashed line).

16

The final prediction equation for this regression model follows.

$$\hat{Y} = e^{0.599+0.326P+9.112EVI-0.201T} \tag{6}$$

Where $\hat{Y}$ is predicted streamflow ($m^3 s^{-1}$), EVI is enhanced vegetation index (unitless) with no advance, P is precipitation (mm) with a 3-day advance, and T is surface temperature ($^{\circ}$C) with an 8-day advance. A plot of gauged versus modeled streamflow for both curve number and remote sensing methods can be seen in Fig. 7. The remote sensing multiple regression approach, in an 8-day mean model, estimated total annual cumulative streamflow at $93.3 \times 10^6 \, m^3$, an underestimate of 16.8% compared to a 28.7% underestimate for the optimal AMC adjusted curve number model.

5 Discussion

The remote sensing model more closely corresponded to gauged peak flow, in both magnitude and alignment, than did any of the three curve number models evaluated. The significance of the two primary remote sensing parameters as indicators of landscape moisture status, EVI and surface temperature, confirm the results of other studies that both vegetation and ambient surface temperature retain a "memory" of antecedent moisture. The non-significance of the AMSR-E derived soil moisture in the final model can perhaps be attributed to a combination of factors, such as the masking effect of the vegetation canopy, and the fact that this dataset represented single day moisture estimations every 8 days and not 8-day continuous means.

The surprising and original result from this study indicates that a watershed's streamflow can be reasonably estimated on an 8-day basis from readily available and temporally continuous remote sensing parameters and a precipitation record. Much work remains in this study in terms of calibration and validation of the estimation equation for other time periods in the data record for this particular watershed as well as the derivation of estimation equations for other regionally proximate watersheds. A challenging issue in ultimately developing a reasonable streamflow prediction equation is

17

to derive a spatial scaling coefficient for incorporation in the estimation equation such that streamflow can be estimated for a range of watershed dimension. Another goal of future research in this streamflow model is to assess and evaluate spatially distributed precipitation estimates such as those offered by weather radar (NEXRAD) derived or satellite derived precipitation products. If these products can be utilized in lieu of gauged data, this modeling approach could have far-reaching implications for water balance models and flood vulnerability forecasting in regions of the world with sparse hydrologic data and knowledge.

6 Conclusions

A multiple regression statistical approach to streamflow estimation utilizing remotely sensed biophysical parameters sensitive to a watershed's soil moisture condition was presented in this research study. A method was developed to parameterize a statistical model for streamflow estimation in a rural Texas watershed using surface temperature and the Enhanced Vegetation Index imagery products from the MODIS/Terra satellite, combined with a gauged precipitation record for calendar year 2004. For comparison, streamflow or runoff for this same data period was determined with the NRCS curve number method based on soil hydrologic group type and land cover type information. Estimated 8-day mean streamflows from the remote sensing model represented an improvement over curve number modeled streamflows, developed with the standard 5-day antecedent moisture condition model. From this study, it is concluded that multitemporal satellite imagery holds great promise for analysis of the role of landscape moisture status in the subsequent modeling of a watershed's streamflow response to precipitation events. It is anticipated that future research will demonstrate that streamflow estimation equations can be determined for watersheds of varying spatial dimension under similar environmental and climatological conditions.

18

References

Arnold, J. G. and Allen, P. M.: Automated methods for estimating baseflow and groundwater recharge from streamflow records, J. Am. Water Resour. Assoc., 35(2), 411–424, 1999.

Adegoke, J. O. and Carleton, A. M.: Relations between soil moisture and satellite vegetation indices in the U.S. Corn Belt, J. Hydrometeorol., 3(4), 395–405, 2002.

Cashion, J., Lakshmi, V., Bosch, D., and Jackson, T. J.: Microwave remote sensing of soil moisture: evaluation of the TRMM Microwave Imager (TMI) satellite for the Little River Watershed, Tifton, Georgia, J. Hydrol., 307, 242–253, 2005.

Centolani, M. and Savi, F.: Regional estimation of CN on Tiber Basin, Proceedings of Applied Simulation and Modeling, 410, 2003.

Chow, V. T., Maidment, D. R., and Mays, L. W.: Applied hydrology, McGraw-Hill, Inc., New York, New York, 1988.

Durre, I. and Wallace, J. M.: Dependence of extreme daily maximum temperatures in antecedent soil moisture in the contiguous United States during summer, J. Climate, 1, 2641–2651, 2000.

Fennessey, L. A. J., Miller, A. C., and Hamlett, J. M.: Accuracy and prediction of NCRS models for small watersheds, J. Am. Water Resour. Assoc., 37(4), 899–912, 2001.

Gao, X., Huete, A. R., Ni, W., and Miura, T.: Optical-biophysical relationships of vegetation spectra without background contamination, Remote Sens. Environ. 74, 609–620, 2000.

Grove, M., Harbor, J. M., Engel, M.: Composite vs. distributed curve numbers: effects on estimates of storm runoff depths, J. American Water Resources Association, 34(5), 1015–1023, 1998.

Heggen, R. J.: Normalized antecedent precipitation index, J. Hydrol. Eng., 6(5), 377–381, 2001.

19

Jacobs, J. M., Myers, D. A., and Whitfield, B. M.: Improved rainfall/runoff estimates using remotely sensed soil moisture, J. Am. Water Resour. Assoc., 39(2), 313–324, 2003.

Larkin, T. J. and Bowmar, G. W.: Climatic atlas of Texas, Texas Department of Water Resources, LP-192, 151 p, 1983.

McCuen, R. H.: Approach to confidence interval estimation for curve numbers, J. Hydrol. Eng., 7(1), 43–48, 2002.

Melesse, A. M. and Graham, W. D.: Storm runoff prediction based on a spatially distributed travel time method utilizing remote sensing and GIS, J. Am. Water Resour. Assoc., 40(4), 863–879, 2004.

Miserocchi, F. and Savi, F.: Antecedent soil moisture conditions in the estimation of the SCS-CN values, Geoph. Res. Abstracts, 7(06241), 2005.

Mishra, S. K., Jain, M. K., and Singh, V. P.: Evaluation of the SCS-CN-based model incorporating antecedent moisture, Water Resour. Manag., 18, 567–589, 2004.

Miura, T., Huete, A. R., Yoshioka, H., and Holben, B. N.: An error and sensitivity analysis of atmospheric resistant vegetation indices derived from dark target-based atmospheric correction, Remote Sens. Environ., 78, 284–298, 2001.

Moglen, G. E.: Effect of orientation of spatially distributed curve numbers in runoff calculations, J. Am. Water Resour. Assoc., 36(6), 1391–1400, 2000.

Mustafa, Y. M., Amin, M. S. M., Lee, T. S., and Shariff, A. R. M.: Evaluation of land development impact on a tropical watershed hydrology using remote sensing and GIS, J. Spatial Hydrol., 5(2), 16–30, 2005.

Nachebe, M. H.: Equivalence between topmodel and the NCRS curve number method in predicting variable runoff source areas, J. Am. Water Resour. Assoc., 42(1), 225–235, 2006.

Ponce, V. M. and Hawkins, R. H.: Runoff curve number: has it reached maturity?, J. Hydrol. Eng., 1(1),11–19, 1996.

Schmugge, T. J., Kustas, W. P., Ritchie, J. C., and Jackson, T. J.: Remote sensing in hydrology, Adv. Water Resour., 25, 1367–1385, 2002.

Shirmohammadi, A., Yoon , K. S., Rawls, W. J., and Smith, O. H.: Evaluation of curve number procedures to predict runoff in GLEAMS, J. Am. Water Resour. Assoc., 33(5), 1069–1076, 1997.

Smith, J. A., Baeck, M. L. Morrison, J. E., and Sturdevant-Rees, P.: Catastrophic rainfall and flooding in Texas. J. Hydrometeorology, Am. Meteorol. Soc., 1, 5–25, 2000.

Texas Commission on Environmental Quality: Improving water quality in southcentral Texas:

20

four TMDLs for bacteria and dissolved oxygen in Elm and Sandies Creeks, http://www.tceq.state.tx.us/assets/public/implementation/water/tmdl/31-elmsandies.pdf, 2006.

U.S. Department of Agriculture, Grassland, Soil, and Water Research Laboratory: Soil and Water Assessment Tool (SWAT), Temple, Texas, http://www.brc.tamus.edu/swat/index.html, 2006.

U.S. Department of Agriculture, Soil Conservation Service: National Engineering Handbook, Section 4 – Hydrology, Washington D.C., 1985.

U.S. Department of Agriculture, Soil Conservation Service: Urban hydrology for small watersheds, Technical Release 55, Washington D.C., 1986.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.-L.: Quality assessment and validation of the MODIS global land surface temperature, Int. J. Remote Sens., 25, 261–274, 2004.

Xie, H., Zhou, X., Vivoni, E. R., Hendrickx, J., and Small, E. E.: GIS-based NEXRAD stage III precipitation database: automated approaches for data processing and visualization, Computers Geosci., 31, 65–76, 2005.

Zhou, X., Xie, H., and Hendrickx, J.: Statistical evaluation of MODIS snow cover products with constraints from streamflow and SNOTEL measurement, Remote Sens. Environ., 94(2), 214–231, 2005.

21

Table 1. Pearson's r^2 correlations for Gonzales 1N and 10SW weather station monthly rainfall by year.

Year	1997	1998	1999	2000	2001	2002	2003	2004	2005
r^2	0.71	0.94	0.83	0.90	0.60	0.86	0.87	0.98	0.80

Table 2. Antecedent moisture condition assessment, modified for metric units, from National Engineering Handbook 4 (USDA – SCS, 1985).

		Total rain previous 5 days (mm)	
Condition	Soil wetness	Dormant season	Growing season
I	Dry but above wilting point	<12	<35
II	Average	12–28	35–53
III	Near saturation	>28	>53

Table 3. All pairwise parameter correlations with significance probability.

Variable	by Variable	Correlation	Significance Probability
TEMP	PREC	–0.1970	0.1945
NDVI	PREC	0.1285	0.4000
NDVI	TEMP	0.6733	0.0000
EVI	PREC	0.1235	0.4189
EVI	TEMP	0.6174	0.0000
EVI	NDVI	0.9468	0.0000
ALB	PREC	–0.1322	0.3868
ALB	TEMP	–0.7259	0.0000
ALB	NDVI	–0.9666	0.0000
ALB	EVI	–0.9066	0.0000
SM	PREC	–0.0588	0.7044
SM	TEMP	–0.3042	0.0447
SM	NDVI	–0.1523	0.3237
SM	EVI	–0.1466	0.3424
SM	ALB	0.1253	0.4177

TEMP = temperature (^{o}C), NDVI = Normalized Difference Vegetation Index, EVI = Enhanced Vegetation Index, ALB = albedo, SM = soil moisture

Table 4.1. Forward stepwise regression step history and initial model results.

Step History

Step	Parameter	Action	P	SS_{seq}	r^2	Cp	p
1	Precipitation (3 day offset)	Entered	0.000	57.18	0.466	90.8	2
2	Temp (1 period offset)	Entered	0.001	16.97	0.605	59.2	3
3	EVI (0 period offset)	Entered	0.000	24.20	0.802	13.1	4
4	Temp (0 period offset)	Entered	0.008	4.11	0.836	6.9	5
5	EVI (2 period offset)	Entered	0.055	1.93	0.852	5.1	6
6	Soil moisture (1 period offset)	Entered	0.117	1.21	0.862	4.7	7

P – significance probability, SS_{seq} – sequential sum of squares, r^2 – Coefficient of Multiple Determination, Cp – model "adequacy" factor, p = # of parameters in model

Table 4.2. Continued.

Model Results

SSE	DF	MSE	r^2	r^2_{adj}	Mallows Cp
16.974	36	0.472	0.862	0.839	4.694

SSE – sum square error, DF – degrees of freedom, MSE – mean square error, r^2 – Coefficient of Determination, r^2_{adj} – Coefficient of Determination adjusted for multiple parameters, Mallows Cp – model "adequacy" factor
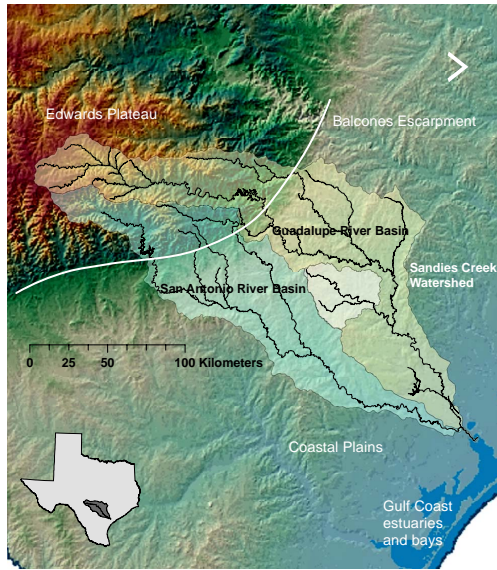
Fig. 1. Location of Guadalupe and San Antonio River basins of southcentral Texas and Sandies Creek watershed, a 1420 km^2 rural catchment of predominately oak and hickory woodlands, shrubland, and agricultural grasslands for grazing and hay production.
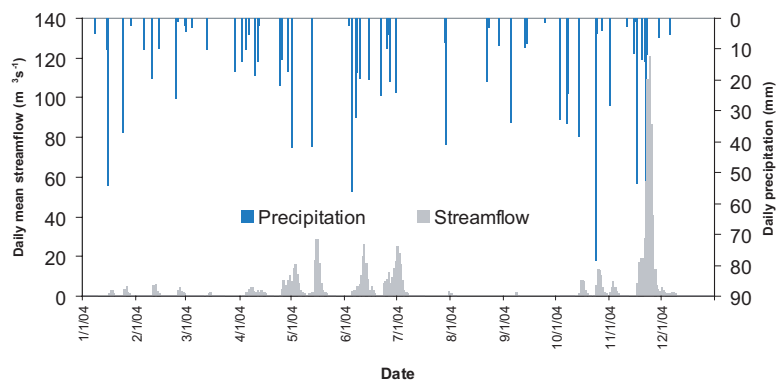
27

Fig. 2. Daily mean streamflow (m^{3}s^{-1}) at the outlet of Sandies Creek watershed (USGS gauge station 8175000) and precipitation (mm) for 2004 at Gonzales 10 SW weather station on the northern edge of the watershed.
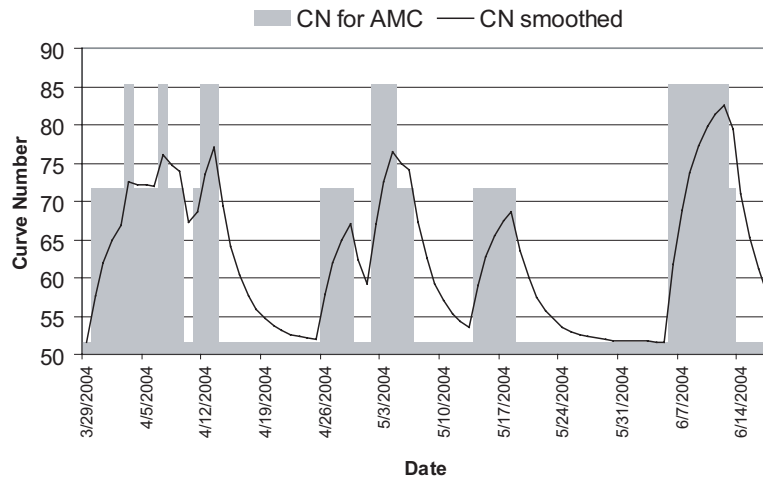
Fig. 3. Example of curve number adjustments for AMC I (51.6), AMC II (71.8), and AMC III (85.4) states (denoted as stairsteps) and exponentially smoothed adjustments.
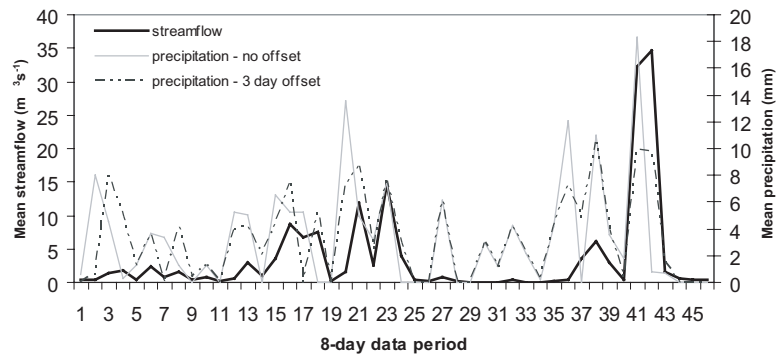
29

Fig. 4. Comparison of the 8-day mean daily precipitation (mm) for zero and 3-day time offsets, and 8-day mean streamflow (m^3s^{-1}) for 2004 at Sandies Creek watershed. Note the improved peak to peak fit of the 3-day offset precipitation means to streamflow event means.
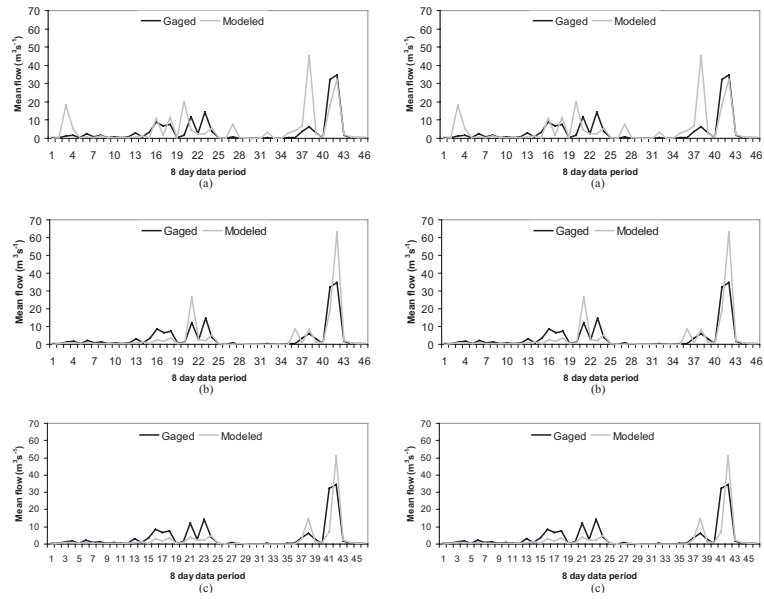
30

Fig. 5. (a) Sandies Creek gauged and modeled streamflow for a single composite CN of 71.8, **(b)** Sandies Creek gauged and modeled streamflow for AMC states I (CN = 51.7), II (CN = 71.8) and III (CN = 85.4), **(c)** Sandies Creek gauged and modeled streamflow for a smoothed AMC continuum.
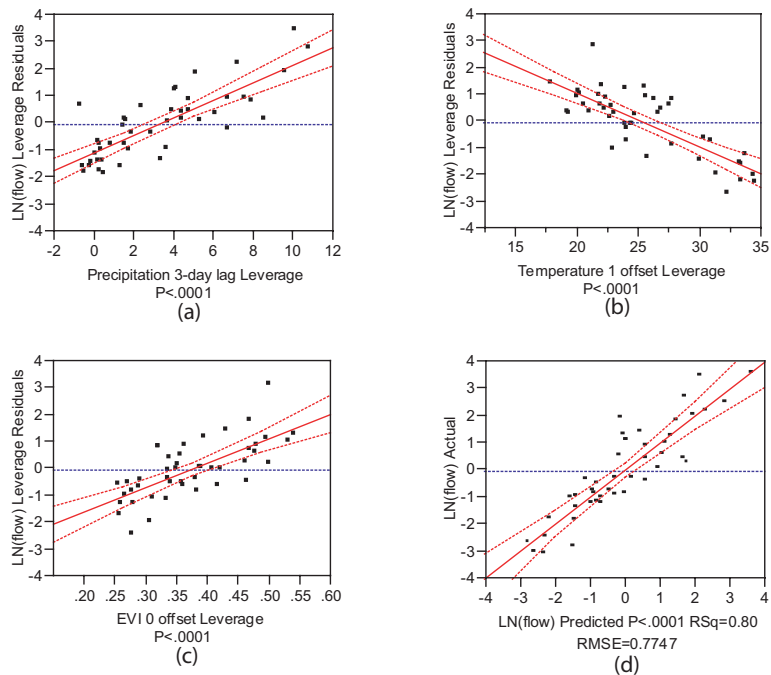
31

Fig. 6. Leverage plots of precipitation **(a)**, temperature **(b)**, EVI **(c)** and the overall model **(d)**. If a parameter's significance curves envelop or contain the dashed line at 0, then that parameter is not considered significant in the model.
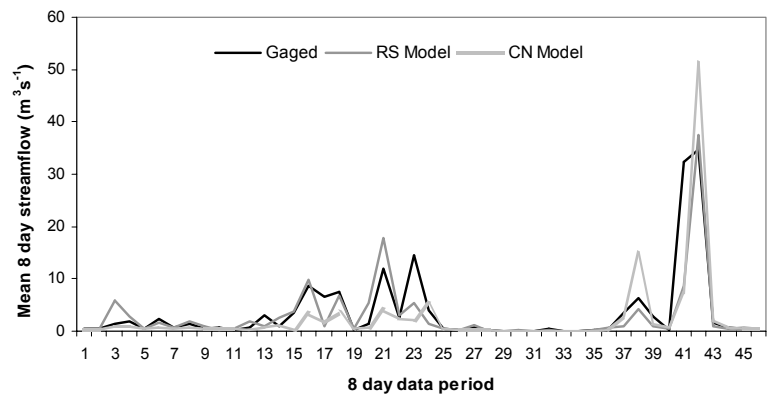
32

Fig. 7. Plot of actual versus estimated 8-day streamflow means for the remote sensing and curve number models for Sandies Creek watershed, calendar year 2004.

33