A first approach to global runoff simulation using satellite rainfall estimation

Yang Hong,1,2,3 Robert F. Adler,1 Faisal Hossain,4 Scott Curtis,5 and George J. Huffman1,6

Received 14 November 2006; revised 14 May 2007; accepted 23 May 2007; published 11 August 2007.

Motivated by the recent increasing availability of global remote sensing data for estimating precipitation and describing land surface characteristics, this note reports an approximate assessment of quasi-global runoff computed by incorporating satellite rainfall data and other remote sensing products in a relatively simple rainfall-runoff simulation approach: the Natural Resources Conservation Service (NRCS) runoff curve number (CN) method. Using an antecedent precipitation index (API) as a proxy of antecedent moisture conditions, this note estimates time-varying NRCS-CN values determined by the 5-day normalized API. Driven by a multiyear (1998–2006) Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis, quasi-global runoff was retrospectively simulated with the NRCS-CN method and compared to Global Runoff Data Centre data at global and catchment scales. Results demonstrated the potential for using this simple method when diagnosing runoff values from satellite rainfall for the globe and for medium to large river basins. This work was done with the simple NRCS-CN method as a first-cut approach to understanding the challenges that lie ahead in advancing the satellite-based inference of global runoff. We expect that the successes and limitations revealed in this study will lay the basis for applying more advanced methods to capture the dynamic variability of the global hydrologic process for global runoff monitoring in real time. The essential ingredient in this work is the use of global satellite-based rainfall estimation.


1. Introduction

Many hydrological models have been introduced in the hydrological literature to predict runoff [Singh, 1995], but few of these have become common planning or decision-making tools [Choi et al., 2002], either because the data requirements are substantial or because the modeling processes are too complicated for operational application. On the other hand, progress in regional or global rainfall-runoff simulation has been constrained by the difficulty of measuring spatiotemporal variability of the primary causative factor, i.e., rainfall fluxes, continuously over space and time. Building on progress in remote sensing technology, researchers have improved the accuracy, coverage, and resolution of rainfall estimates by combining imagery from infrared, passive microwave, and space-borne radar sensors.

[Adler et al., 2003]. Today remote sensing imagery acquired and processed in real time can provide near-real-time rainfall at hydrologically relevant spatiotemporal scales (tens of kilometers and subdaily [Hong et al., 2005; Huffman et al., 2007; Joyce et al., 2004; Sorooshian et al., 2000; Turk and Miller, 2005]). Over much of the globe, remote sensing precipitation estimates are the only available source of rainfall information, particularly in real time. Correspondingly, remote sensing has increasingly become a viable data source to augment the conventional hydrological rainfall-runoff simulation, especially for inaccessible regions or complex terrains, because remotely sensed imageries are able to monitor precipitation and identify land surface characteristics such as topography, stream network, land cover, vegetation, etc. Artan et al. [2007] demonstrated the improved performance of remotely sensed precipitation data in hydrologic modeling when the hydrologic model was recalibrated with satellite data rather than gauge rainfall over four subbasins of the Nile and Mekong rivers.

[4] Motivated by the recent increasing availability of global remote sensing data for estimating precipitation and describing land surface characteristics, this note attempts to obtain a ballpark assessment of global runoff by incorporating satellite rainfall data and other remote sensing products through a relatively simple rainfall-runoff simulation approach: the United States Natural Resources Conservation Service (NRCS) runoff curve number (CN) method [Natural Resources Conservation Service (NRCS), 1986; Burges et al., 1998]. Its simplicity is especially critical for
the vast ungauged regions and geopolitically transboundary basins of the world. Our effort is a first approach to understanding a challenging problem that lies ahead in advancing satellite-based global runoff monitoring. Thus the use of NRCS-CN should not be construed as a call for replacement of other more advanced methods for rainfall-runoff simulation. We expect that the successes and limitations revealed in this study will lay the basis for applying more advanced methods to capture the dynamic variability of the hydrologic process for global runoff monitoring in real time. The essential ingredient in this work is the use of global satellite-based rainfall estimation.

Although Ponce and Hawkins [1996] indicated that the NRCS-CN method is widely used in the United States and other countries, they also criticized it as a simplistic methodology to simulate the sophisticated hydrological system. As an example this method is imprecise for the monsoon-type climate in Ethiopia [Mohammed et al., 2004]. Taylor et al. [2006] also show that the annual runoff in the Volta River basin is a linear function of cumulative rainfall during the wet season when more than approximately 700 mm of rain has fallen. In a literature review, Choi et al. [2002] concluded that NRCS-CN has useful skill because it responds to major runoff-generating properties including soil type, land use/treatment, and soil moisture conditions. They point out that it has been successfully applied to situations that include simple runoff calculation [Heaney et al., 2001], assessment of long-term hydrological impact on land use change [Harbor, 1994] for tens of years, streamflow estimation for watersheds with no streamflow records [Bhaduri et al., 2000], and comprehensive hydrologic/water quality simulation [Srinivasan and Arnold, 1994; Engel, 1997; Burges et al., 1998; Rietz and Hawkins, 2000]. Recently, Curtis et al. [2007] used satellite remote sensing rainfall and gauged runoff data to estimate the CN for basins in eastern North Carolina. A. Harris and F. Hossain (Investigating the optimal configuration of conceptual hydrologic models for satellite rainfall-based flood prediction in the upper Cumberland River, submitted to Journal of Hydrometeorology, 2007) found simpler approaches such as the NRCS-CN method to be more robust than more complicated schemes for the levels of uncertainty that exist in current satellite rainfall data products. On the other hand, we note the risks of implementing this or any other method without fully understanding its associated “uncertainty.” As such, we adopt the NRCS-CN method to estimate a first-cut global runoff by taking advantage of the first 9 years of rainfall estimates from the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) [Huffman et al., 2007].

In this note we first develop spatially distributed and time-variant CN maps for the global land surface. Driven by multiyear remote sensing rainfall, the NRCS-CN method is then used to compute the surface runoff for each grid independently and to subsequently route the surface runoff to the watershed outlet through downstream cells [U.S. Army Corps of Engineers, 2000]. Finally, simulated quasi-global runoff is evaluated with Global Runoff Data Center (GRDC) observed runoff (B. M. Fekete et al., Global Composite Runoff Data Set (v1.0), Complex System Research Center, University of New Hampshire, Durham, 2000, available at http://www.grdc.unh.edu, hereinafter referred to as Fekete et al., Global Composite Runoff Data Set (v1.0), 2000) and water balance model–simulated runoff [Thorntwaite and Mather, 1955; Steenhuis and Van der Molen, 1986; Vorosmarty et al., 1998].

2. Mapping NRCS-CN

2.1. Data

The data sets (i.e., precipitation, soil information, and land cover) required by the NRCS-CN runoff generation scheme are all available globally with a well-established record in Earth system analysis (Fekete et al., Global Composite Runoff Data Set (v1.0), 2000). Information on soil properties is obtained from the Food and Agriculture Organization [2003] (see http://www.fao.org/AG/agl/agll/Natural Resources Conservation Service (NRCS) is runoff generated by rainfall accumulation (mm/d); IA is initial abstraction; Q is runoff generated by P; PR is potential retention; CN is the runoff curve number, with a higher CN associated with higher runoff potential; and IA is approximated by 0.2PR.

CN values are approximated from the area’s hydrologic soil group (HSG), land use/cover, and hydrologic condition, the two former factors being of greatest importance in determining its value [NRCS, 1986]. First, following the Natural Resources Conservation Service (NRCS) [1997], a global HSG map is derived from the digital soil classification which includes 13 textural classes, an important indicator for infiltration rate (Table 1). Given “fair” moisture condition (defined below), the MODIS land cover classification and the HSG map are used to estimate the CN

\[ Q = \frac{(P - IA)^2}{(P - IA + PR)} \]  
\[ PR = \frac{25,400}{CN} - 254, \]
by indexing in the standard lookup tables of the NRCS [1986, 1997]. Figure 1 shows the estimated climatological global CN map for fair moisture conditions, with higher value associated with larger runoff potential. Thus, for a watershed on a coarse grid a composite CN can be calculated as

$$CN_{com} = \frac{\sum A_iCN_i}{\sum A_i},$$  \hspace{1cm} (3)

in which $CN_{com}$ is the composite CN used for runoff volume computations; $i$ is the index of subgrids or watershed subdivisions; and $A_i$ is the drainage area of area $i$. The composite CN values for several watersheds are listed in Table 2.

### 3. Time-Variant NRCS-CN and Runoff Simulation

#### 3.1. Time-Variant NRCS-CN

Note that the CN values displayed in Figure 1 are for the “fair” hydrologic condition from standard lookup tables, which are used primarily for design applications. However, for the same rainfall amount, there will be more runoff under wet conditions than under dry. In practice, lower and upper enveloping curves can be computed to determine the range of the CN according to the antecedent moisture conditions (AMC):

$$CN_{I}^{I} = \frac{CN_{II}^{II}}{2.281 - 0.01281CN_{II}^{II}},$$  \hspace{1cm} (4)

$$CN_{III}^{III} = \frac{CN_{II}^{II}}{0.427 + 0.00573CN_{II}^{II}},$$  \hspace{1cm} (5)

where superscripts indicate the AMC, I being dry, II being normal (average), and III being wet [Hawkins, 1993]. The change of AMC is closely related to antecedent precipitation [NRCS, 1997]. We apply the concept of an antecedent precipitation index (API) to provide guidance on how to estimate the variation of CN values under dry or wet antecedent precipitation conditions. Kohler and Linsley [1951] define API as

$$API = \sum_{j=1}^{n} P_j k^{-j},$$  \hspace{1cm} (6)

### Table 1. Hydrological Soil Group Derived From Soil Property

<table>
<thead>
<tr>
<th>HSG</th>
<th>USDA SOIL TEXTURE CLASS</th>
<th>SOIL CONTENTS</th>
<th>EARTH’S SURFACE, %</th>
<th>PROPERTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1, 2, 3</td>
<td>sand, loamy sand or sandy loam types of soils</td>
<td>4.69</td>
<td>low runoff potential and high infiltration rates even when thoroughly wetted and consists chiefly of deep, well to excessively drained sands or gravels</td>
</tr>
<tr>
<td>B</td>
<td>4, 5, 6</td>
<td>silt loam, loam, or silt</td>
<td>8.41</td>
<td>moderate infiltration rate and consists chiefly of soils with moderately fine to moderately coarse textures</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>sandy clay loam</td>
<td>3.98</td>
<td>low infiltration rates when thoroughly wetted and consists chiefly of soils with moderately fine to fine structure</td>
</tr>
<tr>
<td>D</td>
<td>8, 9, 10, 11, 12</td>
<td>clay loam, silty clay loam, sandy clay, silty clay or clay</td>
<td>5.78</td>
<td>highest runoff potential, very low infiltration rates when thoroughly wetted and consists chiefly of clay soils</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>water bodies</td>
<td>65.55</td>
<td>NA</td>
</tr>
<tr>
<td>-1</td>
<td>13</td>
<td>permanent ice/snow</td>
<td>11.59</td>
<td>NA</td>
</tr>
</tbody>
</table>

aAbbreviations are HSG, hydrological soil group; USDA, U.S. Department of Agriculture; and NA, not applicable.

Figure 1. Global NRCS runoff curve number map derived from U.S. Department of Agriculture hydrological soil groups and land cover classification for fair hydrological conditions.
where $T$ is the number of antecedent days, $k$ is the decay constant, and $P$ is the precipitation during day $t$. The model is also known as “retained rainfall” [Singh, 1989]. Decay constant $k$ is the antilog of the slope on a semilog plot of soil moisture and time [Heggen, 2001]. API practice suggests that $k$ is generally between 0.80 and 0.98 [Viessman and Lewis, 1996]. Here we use decay constant $k$ as 0.85 for demonstration purposes. API generally includes moisture conditions for the previous 5 days (or pentad) [NRCS, 1997]. In order to obtain a time-variant CN the site-specified API is first normalized as

$$\text{NAPI} = \frac{\sum_{t=1}^{T} P_t k^{-t}}{P \sum_{t=1}^{T} k^{-t}}, \quad (7)$$

where $T = 5$ for pentads, the numerator is API, and the denominator is a normalizing operator with two components: average daily precipitation $P$ and the $\sum k^{-t}$ series. The “dry” condition is defined as normalized antecedent precipitation index (NAPI) $< 0.33$, the “wet” condition is defined as NAPI $> 3$, and the intermediate range $0.33 \sim 3$ is the “fair” hydrological condition. By definition, the surface moisture conditions are delineated as dry (or wet) if any pentad API is less than one third of (or larger than 3 times) the climatologically averaged pentad API, and fair conditions are designated for all others. Summarizing, the CN can be converted to dry, fair, or wet condition using equations (4)–(7) according to the moisture conditions approximated by the pentad NAPI.

[10] Using the multiyear (1998–2006) satellite-based precipitation data set from NASA TRMM, the 9-year climatological pentad API is shown in Figure 2a. Thus, given any date, the pentad NAPI can be determined, and thus the CN can be updated with equations (4)–(7). For example, on 25 August 2005 the pentad rainfall accumulation, pentad NAPI, resulting hydrological conditions (dry, fair, or wet), and the updated CN on the same date are shown in Figures 2b, 2c, 2d, and 2e, respectively.

### 3.2. Runoff Simulation

[11] Using the concept of NAPI and the NRCS-CN method [NRCS, 1997], the TRMM-simulated runoff (TRMM-CN) can be calculated and compared with three sets of GRDC annual climatological runoff fields: observed (OBS), water balance model (WBM)-simulated, and composite (CMP) from the OBS and WBM (Fekete et al., Global Composite Runoff Data Set (v1.0), 2000). The WBM used the water balance model of Thornthwaite and Mather [1955] with a modified potential evaporation scheme from Vorosmarty et al. [1998], driven by input monthly air temperature and precipitation from Legates and Willmott [1990a, 1990b].

Note that the three GRDC runoff climatologies span a period of incomplete data records (1950–1979), while the TRMM-CN runoff is simulated for 9 years (1998–2006) of satellite rainfall with complete spatiotemporal coverage. One assumption here is that the change of rainfall between the two time periods is small enough so that the resulting runoff climatology is spatially consistent. Table 3 shows that the TRMM-CN runoff corresponds more closely with the WBM, having a relatively high correlation and low error. An intercomparison with the GRDC runoff observation demonstrates that the WBM has a moderate advantage over the TRMM-CN runoff. The correlation and root-mean-square difference (RMSD) between the GRDC OBS and WBM are 0.81 and 159.7 mm/yr (or 0.44 mm/d), respectively, which is slightly better than for the TRMM-CN case (Table 3).

[12] Figure 3a shows the annual mean runoff (mm/yr) driven by TRMM daily precipitation for the same 9-year period in comparison with the GRDC-observed runoff climatology (Figure 3b). Note that the gray areas indicate no data or water surface in Figures 3a and 3b. By averaging areas covered by both TRMM-CN and GRDC runoff data, Figure 3c shows the TRMM-CN runoff zonal mean profile against the OBS, WBM, and CMP. In general, the TRMM-CN zonal mean runoff follows more closely the three GRDC runoff profiles in the Northern Hemisphere than in the Southern Hemisphere. We believe that this difference is the result of having many more samples in the Northern Hemisphere as well as more accurate GRDC data. Considering the TRMM-CN runoff difference as a function of basin area shows the TRMM-CN performance deviates more for basins smaller than 10,000 km$^2$, with significantly better agreement for larger basins (Figure 4).

### 4. Summary and Discussion

[13] Given the increasing availability of global geospatial data describing land surface characteristics, this note estimated a global CN map primarily based on soil property and land use/cover information under the “fair” moisture condition. Then, using API as a proxy of AMC, this note further estimated time-variant CN values bounded by dry and wet AMC approximated by pentad-normalized API. Finally, driven by satellite-based TMPA precipitation estimates, quasi-global runoff was simulated with the NRCS-CN method and was compared with GRDC runoff measurements for climatology and at the basin scale.

[14] Although we were able to demonstrate the potential for using the NRCS-CN runoff model when predicting

### Table 2. Composite CN for Several Watersheds for “Fair” Hydrological Conditions*  

<table>
<thead>
<tr>
<th>Watershed</th>
<th>Amazon</th>
<th>Mississippi</th>
<th>Yangtze</th>
<th>Colorado</th>
<th>Mekong</th>
<th>Uruguay</th>
<th>Sacramento</th>
<th>Albany</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite CN</td>
<td>75.484</td>
<td>73.165</td>
<td>81.787</td>
<td>78.621</td>
<td>62.355</td>
<td>83.7</td>
<td>77.425</td>
<td>54.702</td>
</tr>
<tr>
<td>Basin length, km</td>
<td>4,327</td>
<td>4,184</td>
<td>4,734</td>
<td>1,807</td>
<td>3,977</td>
<td>1,424</td>
<td>926</td>
<td>951</td>
</tr>
<tr>
<td>Area, km$^2$</td>
<td>5,853,804</td>
<td>3,202,958</td>
<td>1,794,242</td>
<td>807,573</td>
<td>773,737</td>
<td>355,505</td>
<td>192,563</td>
<td>132,799</td>
</tr>
</tbody>
</table>

*Global surface-averaged CN is 72.803.
Figure 2. (a) Climatological pentad antecedent precipitation index (API) averaged over 9 years (1998–2006). (b) Pentad antecedent rainfall accumulation (mm) ending on 25 August 2005. (c) Pentad-normalized API (NAPI) on 25 August 2005. (d) Hydrological condition, with $-1$, $0$, $1$, and $2$ corresponding to no data, dry, fair, and wet conditions, respectively, determined by NAPI as of 25 August 2005. (e) Updated CN on 25 August 2005.
approximate runoff values from satellite rainfall for the globe and medium to large river basins, there remain several unanswered questions. First, among many methods to estimate CN values, Hawkins [1993] recognized that remote sensing data may not be adequate to define the “true” value of a CN. Thus field surveys of basin characteristics should be conducted where feasible in order to obtain “true” soil and land cover data. Second, while this study recognized the uncertainty of the estimates of actual CN values and assumed that they likely fall within the enveloping wet (upper) and dry (lower) conditions approximated by the 5-day-normalized API, it may be possible to adjust the CN more precisely to account for local or regional information. Finally, one major unaddressed hydrological concern for rainfall-runoff applications of remotely sensed precipitation is the thorough evaluation of satellite-based rainfall estimation error and its nonlinear influence on rainfall-runoff modeling uncertainty in varying landscapes and climate

<table>
<thead>
<tr>
<th>Statistics</th>
<th>OBS</th>
<th>WBM</th>
<th>CMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>0.75</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>Bias ratio</td>
<td>1.28</td>
<td>1.12</td>
<td>1.12</td>
</tr>
<tr>
<td>RMSD, mm/d</td>
<td>0.56</td>
<td>0.48</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Abbreviations are OBS, observed; WBM, water balance model; CMP, composite; and RMSD, root-mean-square difference.

Figure 3. (a) Annual mean runoff (mm/yr) simulated using NRCS-CN methods from TRMM estimates for the period 1998–2006. (b) GRDC-observed runoff (mm/yr). (c) Runoff zonal mean profiles comparing TRMM precipitation (green) and simulated runoff (red) to GRDC runoff (blue) from the (left) observed, (middle) WBGS, and (right) composite data sets. Note the gray areas in Figures 3a and 3b indicate no data or water surface.
regimes [Hong et al., 2006; Hossain and Anagnostou, 2006; Villarini and Krajewski, 2007]. Thus, while we conclude that this simple approach seems to provide a reliable tool when using coarse-resolution satellite precipitation data, we also urge similar studies using more sophisticated hydrological models, particularly seeking to serve the vast ungauged regions and geopolitically transboundary basins of the world [Hossain et al., 2007].

Acknowledgments. This research is carried out with support from NASA’s Precipitation Measurement Mission (PMM) and Applied Sciences program under Ramesh Kakar and Stephen Ambrose of NASA Headquarters. The authors thank Tammo Steenhuis and the two anonymous reviewers for constructive comments. The Editor-in-Chief Marc Parlange is also thanked for his encouragement.

References


Hossain, F., N. Katiyar, Y. Hong, and A. Wolf (2007), The emerging role of satellite rainfall data in improving the hydro-political situation of flood monitoring in the under-developed regions of the world, Nat. Hazards, in press.


Figure 4. TRMM-CN (a) runoff difference distribution and (b) root-mean-square difference (RMSD) as a function of basin area.


