JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
Vol. 45, No. 4 AMERICAN WATER RESOURCES ASSOCIATION August 2009

GIS-BASED PREDICTIVE MODELS OF HILLSLOPE RUNOFF GENERATION PROCESSES[1]

Mansour D. Leh and Indrajeet Chaubey

ABSTRACT: Successful nonpoint source pollution control using best management practice placement is a complex process that requires in-depth knowledge of the locations of runoff source areas in a watershed. Currently, very few simulation tools are capable of identifying critical runoff source areas on hillslopes and those available are not directly applicable under all runoff conditions. In this paper, a comparison of two geographic information system (GIS)-based approaches: a topographic index model and a likelihood indicator model is presented, in predicting likely locations of saturation excess and infiltration excess runoff source areas in a hillslope of the Savoy Experimental Watershed located in northwest Arkansas. Based on intensive data collected from a two-year field study, the spatial distributions of hydrologic variables were processed using GIS software to develop the models. The likelihood indicator model was used to produce probability surfaces that indicated the likelihood of location of both saturation and infiltration excess runoff mechanisms on the hillslope. Overall accuracies of the likelihood indicator model predictions varied between 81 and 87% for the infiltration excess and saturation excess runoff locations respectively. On the basis of accuracy of prediction, the likelihood indicator models were found to be superior (accuracy 81-87%) to the predications made by the topographic index model (accuracy 69.5%). By combining statistics with GIS, runoff source areas on a hillslope can be identified by incorporating easily determined hydrologic measurements (such as bulk density, porosity, slope, depth to bed rock, depth to water table) and could serve as a watershed management tool for identifying critical runoff source areas in locations where the topographic index or other similar methods do not provide reliable results.

(KEY TERMS: binary logistic regression GIS; hillslope; runoff indicator modeling; saturation excess runoff; infiltration excess runoff; topographic index.)

Leh, Mansour D. and Indrajeet Chaubey, 2009. GIS-Based Predictive Models of Hillslope Runoff Generation Processes. *Journal of the American Water Resources Association* (JAWRA) 45(4):844-856. DOI: 10.1111/j.1752-1688.2009.00328.x

INTRODUCTION

Watershed simulation models are widely used in evaluating the hydrological and water quality responses from various land use and land management practices. Watershed and water resource managers use simulation models to analyze the effects that different watershed management practices may have on the hydrological regime of a catchment. As runoff is the dominant driving mechanism for nonpoint source pollutants in agricultural and mixed land use watersheds, partitioning of rainfall into runoff has received the bulk of attention in the scientific community involved with hydrologic model development and application, while little attention has been

[1]Paper No. JAWRA-08-0072-P of the *Journal of the American Water Resources Association* (JAWRA). Received April 18, 2008; accepted December 24, 2008.

[2]Respectively, Graduate Student, Department of Biological and Agricultural Engineering, University of Arkansas, Fayetteville, Arkansas; and Associate Professor, Departments of Agricultural and Biological Engineering, and Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907 (E-Mail/Chaubey: ichaubey@purdue.edu).

paid to the runoff generation mechanism and locations of runoff producing areas in a watershed. Johnson *et al.* (2003) categorized watershed models into four types based on the runoff generating mechanism: (1) infiltration excess based models, (2) saturation excess based models, (3) empirically developed models that do not differentiate between runoff mechanisms, and (4) models that may be a combination of any of the above. The infiltration excess runoff generating mechanism is based on the assumption that runoff will be produced when rainfall intensity rates exceed the infiltration capacity of the soil (Horton, 1933). Saturation excess based models were developed to describe runoff mechanisms in areas where infiltration excess did not occur and the soil gradually saturated to the surface to generate runoff. It is generally recognized that for many rainfall storm events, the entire watershed does not contribute to runoff generation (Engman, 1974) and this has led to the partial area concept (Betson, 1964) and the variable source area model of Hewlett and Hibbert (1967).

Decisions related to nonpoint source pollution control to improve water quality depends on the accuracy of runoff prediction as well as details about the mechanisms and locations of runoff source areas. The accurate simulation of the locations of a watershed that contribute dynamically to runoff will enable watershed managers to implement best management practices in those areas for effective nonpoint source pollution control. Current water quality research has focused on phosphorous (P) transport from agricultural areas because of the role that P plays in anthropogenic eutrophication of freshwater systems. In the Ozark Highlands of the United States (U.S.), annual land application of poultry litter to pasture land has been a standard practice for over three decades now (Steele and Adamski, 1987; Steele and McCalister, 1991; USDA-ARS, 1998). This waste management practice has led to environmental concerns of the transport of P to nearby water bodies. Several studies have focused on runoff water quality from areas that have received poultry litter (Edwards and Daniel, 1993, 1994; Edwards *et al.*, 1996a; Sauer *et al.*, 1999, 2000), and hydrologic models have been used to simulate runoff volumes and pollutant loads coming from these areas (Chaubey *et al.*, 1995; Edwards *et al.*, 1996b; White and Chaubey, 2005). However, many of these hydrologic models still consider watersheds as a single homogenous unit with very little recognition of the spatial variability (Wood *et al.*, 1990). The studies that have focused on locating and identifying runoff contributing areas have mostly been based on the variable source area hydrology, where saturation excess runoff plays a major role in the runoff generating process (Zollweg *et al.*, 1995; Walter *et al.*, 2000; Lyon *et al.*, 2004, 2006a; Mehta *et al.*, 2004). These models may be applied in management decision making where saturation excess runoff is the dominant runoff process by focusing only on the watershed areas, which produce saturation excess runoff. However, it is shown that in the Ozark hillslopes, both saturation excess and infiltration excess runoff can occur and watershed management decisions must consider both runoff mechanisms (Leh *et al.*, 2008). There is a need to develop and evaluate the infiltration excess runoff model or a combination of the infiltration excess and saturation excess runoff generating mechanisms to successfully locate runoff source areas.

The overall goal of this study is to develop a procedure for quantifying critical runoff source areas in a watershed. The specific objective is to apply and validate Geographic Information System (GIS)-based models to simulate the likely locations of the two dominant runoff generating mechanisms (infiltration excess runoff and saturation excess runoff) in a pasture hillslope typical of the Ozark Highland agricultural watersheds. Successful identification of runoff mechanisms and source areas could improve watershed management techniques by providing an easy to use methodology to identify runoff "hot spots" for nonpoint source pollution control measures.

MATERIALS AND METHODS

Study Area

This study was conducted within Basin 1 of the 1,250 ha Savoy Experimental Watershed (SEW) located in northwest Arkansas. The SEW is about 20 km west of the University of Arkansas main campus at Fayetteville and serves as a collaborative research site designed for long-term hydrologic monitoring (Brahana *et al.*, 2005). The SEW is comprised of six catchment basins including Basin 1, a 147 ha subbasin, located in the southern part of the watershed, which drains into the Illinois River (Figure 1). The Illinois River, which serves as the western boundary of the SEW is a transboundary river that originates approximately 24 km southwest of the city of Fayetteville, Arkansas, and flows into Oklahoma. The mean annual (30 years) precipitation measured in the city of Fayetteville (approximately 12 km east of the study location) is 117 cm (NOAA, 2002). A low average monthly precipitation of 5 cm occurs in January and an average monthly high of 13 cm occurs in June. Winters are relatively short, with brief periods of snow cover and average January temperature of 1.3°C, whereas summers are long,

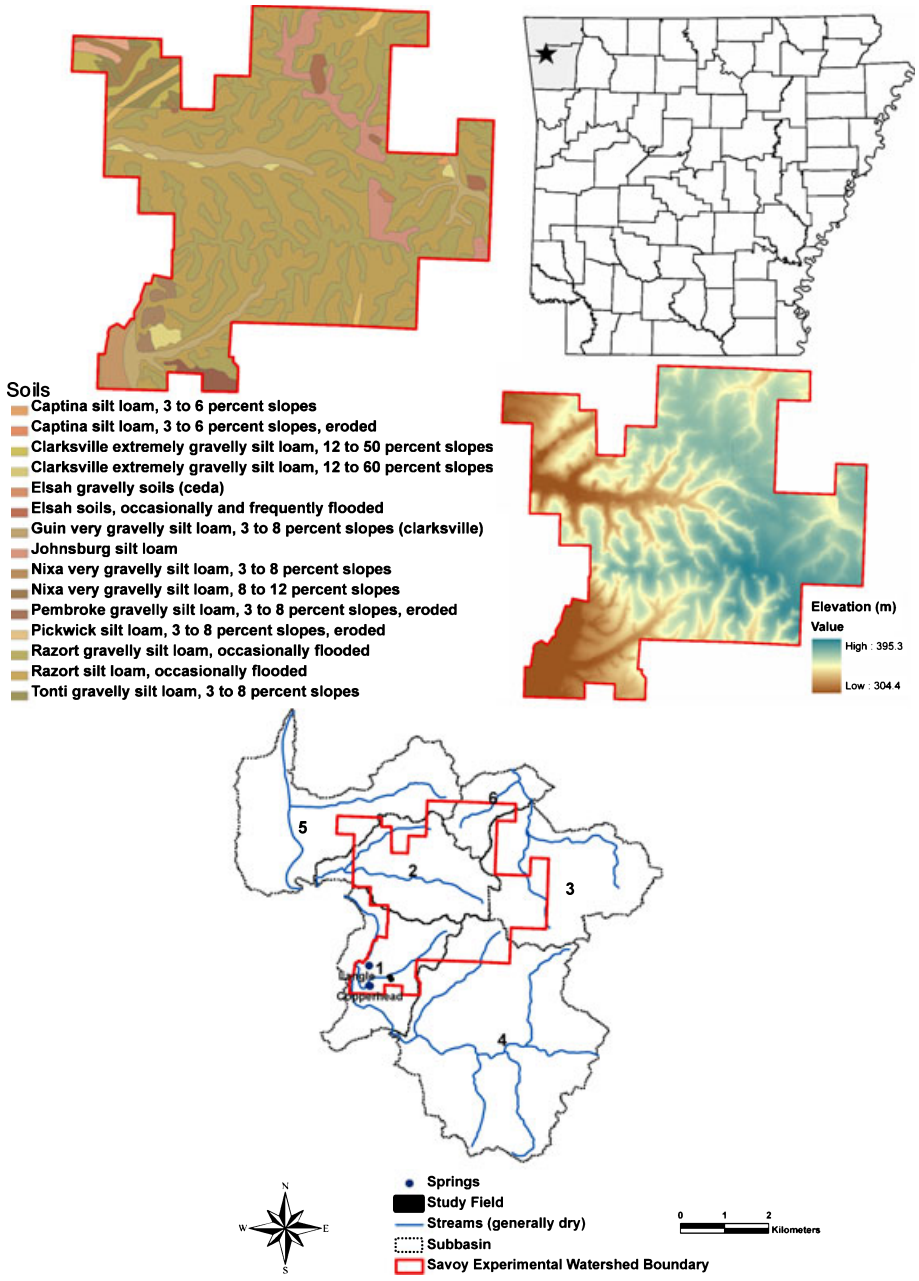

FIGURE 1. Maps of the Major Soils, Elevation, and Location of the Study Site at the Savoy Experimental Watershed (SEW). Subbasins are numbered 1 through 6.

warm, and humid with average July temperature of 26.1°C (NOAA, 2002).

Land use in the SEW represents typical pasture dominated agricultural activities in the southeast U.S. Land use within Basin 1 is 57% hardwood forest and 43% pasture (Sauer *et al.*, 2005). The two major soils within Basin 1 are the Clarksville cherty silt loam and Nixa cherty silt loam (Figure 1) that account for 49 and 30%, of the area respectively (Sauer and Logsdon, 2002). The Captina silt loam and Nixa cherty silt loam are the major soils within the study field. The Captina silt loam soil is characterized by strong brown subsoil that is 25-50 cm thick typically found on stream terraces and the Nixa chert silt loam typically have slow permeable fragipans that occur at the 36-60 cm depth (Harper *et al.*, 1969).

Methods

Three 23×23 m adjacent bermed plots (labeled Plots 1, 2, and 3) were instrumented with a grid of

paired subsurface saturation sensors (SSS) and surface runoff sensors (SRS) (hence forth called saturation and runoff sensor respectively) and installed at 33 points on the plots (12 on Plot 1, 11 on Plot 2, and 10 on Plot 3, Figure 2). The sensors used in this study were similar to that of Srinivasan *et al.* (2000). The saturation sensors are printed circuit boards with sensing pins that indicate the level of soil saturation at preset depths (1 cm, 5 cm, 10 cm, 20 cm, 31 cm, and 46 cm). The runoff sensors are miniature v-notch weirs made of 2 mm thick galvanized sheet metal with a sensor pin and ground pin set 2 cm apart and 2.5 cm away from the v-notch, at the same level as the bottom of the v-notch. The runoff sensor operates on a "yes-no" basis to indicate the presence or absence of surface runoff. A detailed description of both sensors can be obtained from Srinivasan *et al.* (2000). A 0.305 m H-flume at the downslope end of each plot measured the total volume of runoff and hydrograph from each plot. Each H-flume had a Keller pressure transducer (Model 173; Keller America Inc., Newport News, Virginia) that was used to record water depth; the predetermined stage-discharge relation was used to estimate runoff rate. Rainfall occurring on each plot was determined by installing a tipping bucket rain gage (HOBO RG2 model, Onset Corp., Bourne, Massachusetts) on each plot. The rainfall data collected on each plot was supplemented by data collected from a weather station located approximately 0.4 km north of the study site (Figure 2). The weather station was instrumented with a Campbell CR-10X (Campbell Scientific Inc., Logan, Utah) data logger that recorded rainfall, air temperature, soil temperature, relative humidity, global radiation, wind speed, and wind direction at 5-min intervals. Finally, a shallow ground-water well with a Keller pressure transducer was installed (0.6-0.7 m deep) near each saturation sensor upslope of each H-flume to monitor depth to ground water.

Figure 2 shows the locations of all instruments used for data collection in this study. Data collected by each sensor was logged at 5-min intervals with a Campbell CR-23X data logger on Plots 1 and 2, and a Campbell CR-10X data logger on Plot 3. The data presented in this study covers the period between 2004 (April) and 2005 (December).

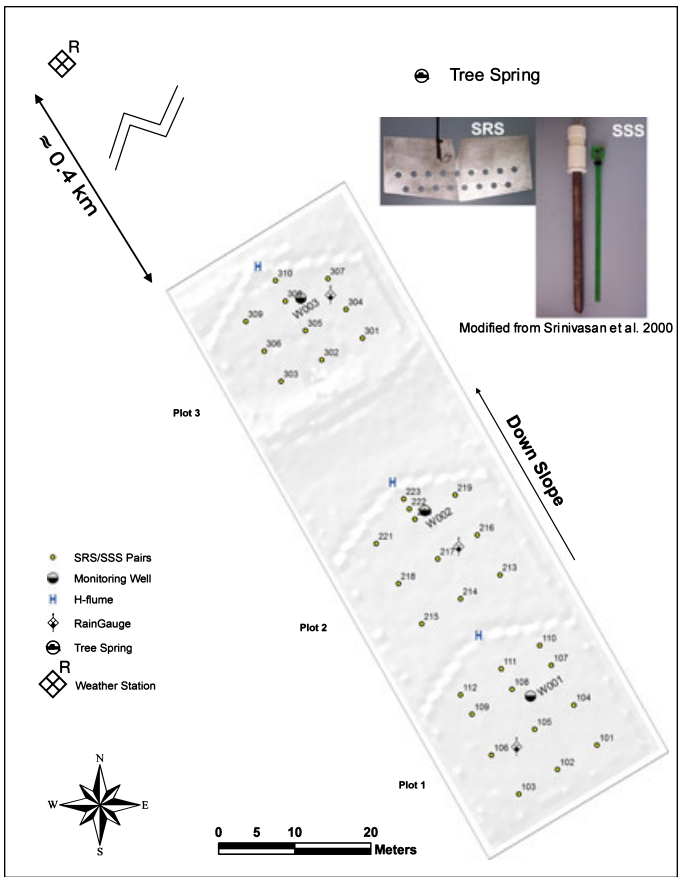

FIGURE 2. Site Layout and Orientation Showing Instrumentation Infrastructure Including Weather Stations, Spring, H-Flumes (H), Surface Runoff Sensors (SRS) and Subsurface Sensors (SSS), and Monitoring Wells (W) on Plots 1 Through 3.

Topographic and Geophysical Surveys

A survey grade (mm level accuracy) Global Positions System (GPS) unit (Leica 500 GPS system, Leica Geosystems AG, St. Gallen, Switzerland) was used to conduct a topography survey of the field area. The GPS survey was performed in 2004 by establishing approximately 1,500 grid points across the plots. The data obtained was used to derive a 1 m resolution digital elevation model (DEM) of the field.

Two surface geophysical techniques [Ground Penetrating Radar (GPR) and electrical resistance] were employed in characterizing the subsurface attributes beneath the plots that could potentially affect the runoff response of each plot (Ernenwein and Kvamme, 2004). The GPR and resistance methods are well known techniques used to determine bedrock depth and integrity (Ernenwein and Kvamme, 2004). The electrical resistance was determined by measuring the potential voltage after a weak electrical current was introduced into the ground. The amount of electric current that flows through the soil is directly related to soil moisture, clay content, and solutes present in the soil and rock. GPR data collection was accomplished by dragging a 400 MHz antenna connected to a Subsurface Imaging Radar system over the surface of equally spaced profiles and recording the reflections obtained.

Soil depth, porosity, and bulk density were measured on a grid of 65 points located across the plots.

Undisturbed soil core samples were taken from the 0-5 cm depth of the points to determine soil bulk density and total porosity. A 1 cm diameter graduated metal rod was used to determine the depth to subsurface material by insertion into the soil until refusal at the 65 measuring points. Three measurements were made at each location and the mean value was recorded for each point. Measurements for the 65 points were imported into ArcGIS software (ESRI Inc., Redlands, California) and interpolated using the inverse weighted distance algorithm to derive the spatial distribution of soil properties across the plots. Soil water saturation levels were determined using sensor readings at each location. As rainfall events of small magnitudes were less likely to induce soil saturated conditions, a 1 cm threshold was used to select rainfall events that fell within this criterion for the entire 2004-2005 study period. Based on these rainfall events, preevent depth to saturated soil levels were determined for each sensor location and the points were interpolated to obtain a surface of preevent saturated soil water level.

Modeling Runoff Mechanisms

Two different modeling approaches (topographic index model and a likelihood indicator model) were compared to identify runoff mechanism and spatial locations of runoff events that occurred between 2004 and 2005.

The Topographic Index Model. The spatial distribution of the TOPMODEL topographic index (TI) of Beven and Kirkby (1979) was computed for each plot. In its simplest form, TI is calculated as:

$$\mathrm{TI} = \ln\left(\frac{a}{\tan B}\right), \tag{1}$$

where a is the upslope contributing area per unit contour length, and tan B is the slope. Upslope contributing area was determined using the D-infinity algorithm (Tarboton, 1997). The D-infinity algorithm computes multiple flow directions by a recursive procedure that divides slope angles in all directions. The upslope contributing area is then found by accumulating the area of the contributing cell and the fraction drained into it by its adjacent grid cells. The TI values were normalized with the highest value to obtain a probability surface of TI values that ranged between 0 and 1. High values of TI indicate areas that are most likely to induce saturation excess overland flow. To identify areas of high likelihood of saturation excess runoff generation, a threshold of 0.6 was used to highlight the most likely saturated regions that produced saturation excess runoff. The 0.6 threshold is a soil dependent parameter that was empirically derived based on field observations after major storm events (Leh *et al.*, 2008).

The Likelihood Indicator Model. The dominant runoff mechanisms observed during each runoff event for the period of study (2004-2005) were quantified as

$$\mathrm{RRR_i} = \frac{\sum_{i=1}^{n} R_\mathrm{p}}{n}, \tag{2}$$

where RRR was the runoff response ratio, R_p was the number of 5-min occurrences of a particular runoff process p (i.e., p = saturation excess, infiltration excess or no runoff) at a sensor location and n was the total number of 5-min readings taken (nonrunoff events inclusive). The runoff response ratio for the period of the study (2004-2005) was used to delineate the locations of the two controlling runoff mechanisms – saturation excess runoff and infiltration excess runoff. The delineated runoff mechanism maps were converted to probability of occurrences by normalizing with the maximum runoff response ratio. A threshold probability of occurrence was determined for each mechanism to create a training map of binary indicator values. The training maps represented areas of high probability of occurrence of each runoff mechanism and were used by the algorithm to identify locations of similar characteristics. Locations greater than or equal to the threshold were assigned a value of 1 and locations below the threshold were assigned a value of 0.

Binary logistic regression procedures were performed on the observed saturation excess and infiltration excess maps (Figure 3) by linkage to JMP statistical software (SAS Institute Inc., Cary, North Carolina). This methodology is similar to that employed by Lyon *et al.* (2006b) to determine factors that influenced saturated regions in the Town Brook watershed of New York. Binary logistic regression employs an iterative maximum likelihood procedure, which determines values of unknown parameters and maximizes the probability of occurrence of the observed data (location of infiltration excess and saturation excess mechanism areas) by a logistic regression equation (Hosmer and Lemeshow, 2000). Assuming I = 1 is the event that a particular runoff mechanism occurs at location i, and I = 0 is the event that the runoff mechanism does not occur at that location, then the logistic regression could be expressed as:

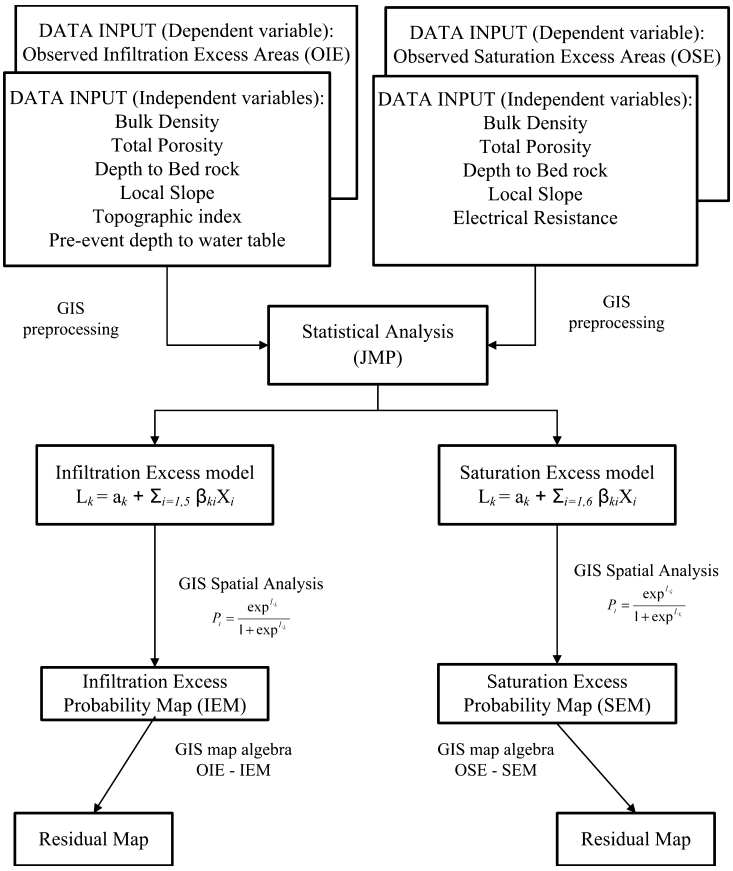

FIGURE 3. Flow Chart of the Modeling Approach Used in the Likelihood Indicator Models to Determine Runoff Generation Processes.

$$\text{Logit}(p) = \ln\left(\frac{p}{1-p}\right) = a + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_i X_i, \tag{3}$$

where p is the probability that $I = 1$, a is the intercept and β_1, $\beta_2 \ldots \beta_i$ are coefficients of the independent variables X_1, X_2, $X_3 \ldots X_i$. Using the logistic transformation;

$$P_i = \frac{\exp^{\text{Logit}(p)}}{1 + \exp^{\text{Logit}(p)}} \tag{4}$$

GIS map algebra operations were employed to map the probability surface of the occurrence of infiltration excess or saturation excess mechanism locations once the respective coefficients of the independent variables were obtained.

Statistical Analysis and Accuracy Assessment

The degree of fit of the model predictions were evaluated using residual plots and other goodness of fit statistics such as User's, Producer's, and Overall accuracy, as described by Story and Congalton (1986). The accuracies were represented in an error matrix (sometimes called confusion matrix or contingency table), which provides an effective way to evaluate errors and to determine the performance of individual categories (Congalton *et al.*, 1983).

The overall accuracy was computed as the ratio of the total number of correct classifications (leading diagonal values in the error matrix) to the total number of observations (Congalton and Green, 1999). The producer's accuracy (also referred to as error of omission) is the probability that a ground sample is correctly classified (Story and Congalton, 1986) and was computed as the ratio of the number of areas correctly classified for a particular runoff generating process to the total number of areas in that runoff generating process. The user's accuracy (also referred to as error of commission) indicates how well a map actually represents the observed data (Story and Congalton, 1986) and was computed as the number of areas correctly classified as a particular runoff generating process divided by the total number of areas that were classified in that runoff generating process. A residual surface

for each prediction was computed by subtracting the predicted surface from the observed surface using GIS map algebra. Residuals ranged between −1 and 1, where greatest error of fit occurred at the maximum residual of 1 or a minimum residual of −1, and a perfect fit would indicate a residual of 0 for each pixel location. The Likelihood Ratio statistic was used to compute a pseudo-R^2 (U) (also called the uncertainty coefficient), which is similar to R^2 expressed as

$$pseudo - R^2 = 1 - \frac{L(B)}{L(0)}, \quad (5)$$

where $L(0)$ is the value of the log-likelihood function if all coefficients except the intercept are 0 and L(B) is the value of the log-likelihood function for the full model (Clark and Hosking, 1986). Psuedo-R^2 (U) ranges from zero for no improvement to 1 for a perfect fit.

RESULTS AND DISCUSSION

Runoff Generating Mechanisms

Rainfall depth recorded in 2004 was 106 cm, representing 106 rainfall events; in 2005, 95 cm rainfall depth was recorded, which represented 111 rainfall events. These rainfall events were defined as any 24 h period with measurable rainfall. The dominant runoff mechanisms that were measured in the field during the monitoring period are shown in Figure 4. Both infiltration excess and saturation excess runoff were observed. Infiltration excess seemed to be the dominant runoff process that occurred in about 58% of the total area and was mainly located in the southern portions of the field. Saturation excess runoff occurred in about 26% of the total field area and appeared to be dominant on Plot 3 and along the southwestern boundary of Plot 1. These findings validate previous studies conducted in Basin 1 (Sauer *et al.*, 1998, 2000, 2005; Sauer and Logsdon, 2002) where infiltration excess is thought to occur frequently in upland areas and saturation excess in the valley bottom. Other field-scale studies conducted on hillslopes have also reported the occurrence of saturation excess runoff in the lower portions of hillslopes and near streams (e.g., Hewlett, 1961; Dunne and Black, 1970a,b; Srinivasan *et al.*, 2002; Meyles *et al.*, 2003; Rezzoug *et al.*, 2005; Badoux *et al.*, 2006).

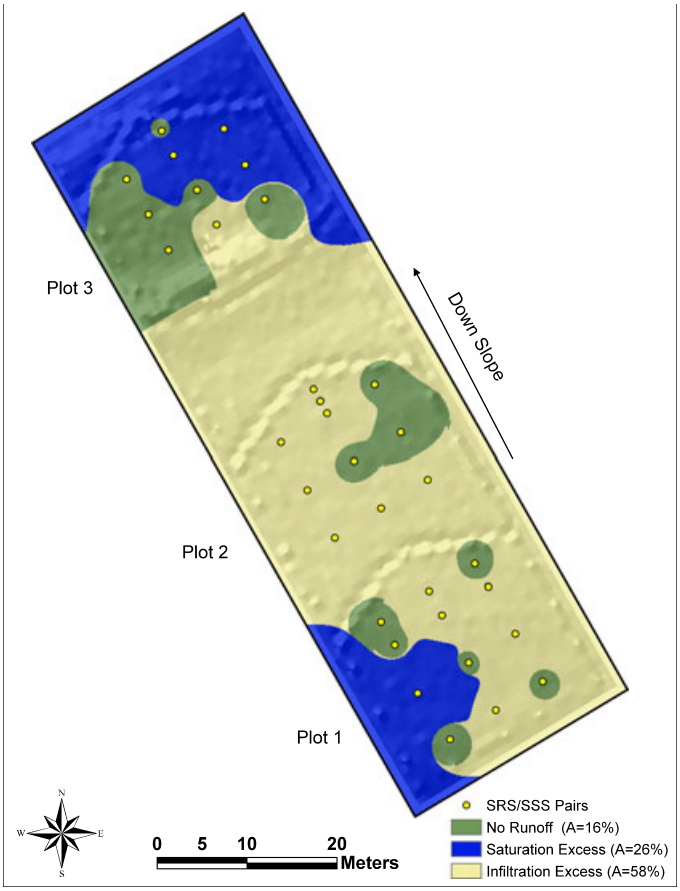

FIGURE 4. Map of Dominant Runoff Process That Occur at Study Site, Savoy, Arkansas. *A* is the total percentage area of a runoff process over the area of the three plots.

The Topographic Index Model

Figure 5 shows the locations of the plots most likely to contribute to saturation excess runoff as indicated by the TI model. The locations of saturation excess runoff appeared to be spatially distributed across the plots and did not correlate well with the observed data (Figure 4). The percent total area most likely to generate saturation excess runoff with a threshold of 0.60 is 19%. Table 1 shows the model accuracies when the predicted locations of saturation excess runoff are tabulated against the observed saturation excess locations. The producer's accuracy is the probability that a pixel location of a saturation excess runoff generating mechanism is correctly represented spatially; and the user's accuracy indicates the probability that a pixel location correctly corresponds to the saturation excess runoff source area as it exists in the field. Overall accuracy was 69.5% with an average producer's accuracy of 48.8%. The model predicted nonsaturation excess areas relatively well (77-86%), whereas the prediction of saturation excess

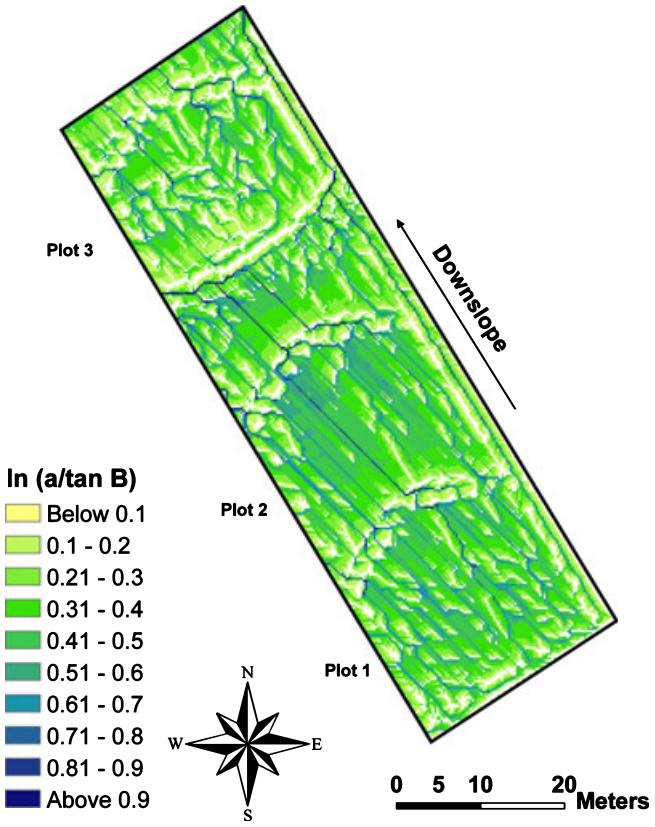

FIGURE 5. Normalized Distribution of Topographic Index Across Plots Showing Areas Most Likely to Generate Saturation Excess Runoff.

TABLE 1. Classification Accuracies of the Saturation Excess (SE) Locations as Indicated by the Topographic Index Model.

	Observed			
Predicted	**Non-SE Areas (pixels)**	**SE Areas (pixels)**	**Total (pixels)**	**User's Accuracy (%)[1]**
Non-SE areas (pixels)	***23,739***	6,977	30,716	77.3
SE areas (pixels)	3,829	***912***	4,741	19.2
Total (pixels)	27,568	7,889	**35,457**	
Producer's accuracy (%)[2]	86.1	11.6		

Notes: Overall accuracy = (23,739 + 912)/35,457 = 69.5%; Average producer's accuracy = 48.8%.

Overall accuracy was obtained by dividing total correct predictions (diagonal, bold) by the total number of pixel predictions (bold). User's accuracies were obtained by dividing the number of correct predictions in a row by the total number of pixel predictions in that row. Producer's accuracies were obtained by dividing the number of correct predictions in a column by the total number of pixel predictions in that column.

[1]Non-SE areas = 23,739/30,716 = 77.3%; SE areas = 912/4,741 = 19.2%

[2]Non-SE areas = 23,739/27,568 = 86.1%; SE areas = 912/7,889 = 11.6%

locations were rather poor (11-19%) (Table 1). While several studies have applied the TI as a surrogate for critical source area identification (e.g., Juracek, 1999, 2000; Endreny and Wood, 2003), the results obtained from this study show that only 19% (Table 1) of the saturation excess area was correctly identified, hence this method may not be suitable for the investigated site. Lyon *et al.* (2006b) used binary logistic regression models to determine that the TI was inadequate in describing the spatial distribution of saturated areas in the Townbrook watershed. Page *et al.* (2005) reported a similar conclusion in an attempt to find a relation between TI and the distribution of soil P for two contrasting soil types, and concluded that TI alone could not be used as an indicator of critical source areas in two catchments in the United Kingdom.

The Likelihood Indicator Model

Figure 6 shows the probability map and a residual surface of the location of saturation excess mechanism. Visual comparison with the observed data indicates a relatively good prediction of the location of saturation excess areas. Table 2 shows the explanatory variables that were used in the model along with a Chi-square value, which could be interpreted as the p-value for each variable, and a pseudo-R^2 (U) which is analogous to the R^2 in linear regression. Positive parameter estimates indicate that high values of the corresponding variables are related to the saturation excess mechanism location and negative coefficients link low variables to the location of saturation excess areas. Thus saturation excess runoff source areas are predicted to be located in areas of low bulk density, total soil porosity, depth to bedrock, and low preevent depth to saturated soil. This prediction is consistent with the saturation excess theory. Typically, water would easily infiltrate the soil in areas with low soil bulk density while areas of a low total porosity indicate that the pore space available for storage of water is limited. The areas of low depth to bedrock would have a limited volume of soil available to store water and a low preevent depth to saturated soil level indicates near soil surface water table, hence a high propensity of generating saturation excess runoff. The local slope and the TI were not found to be significant predictors of the location of saturation excess runoff. The pseudo-R^2 for the overall model was 0.48. For logistic regression, pseudo-R^2 (U) greater than 0.2 may be considered a relatively good fit (Clark and Hosking, 1986). GIS cross tabulation with the observed locations revealed an overall accuracy of 86.7% in model predictions (Table 3).

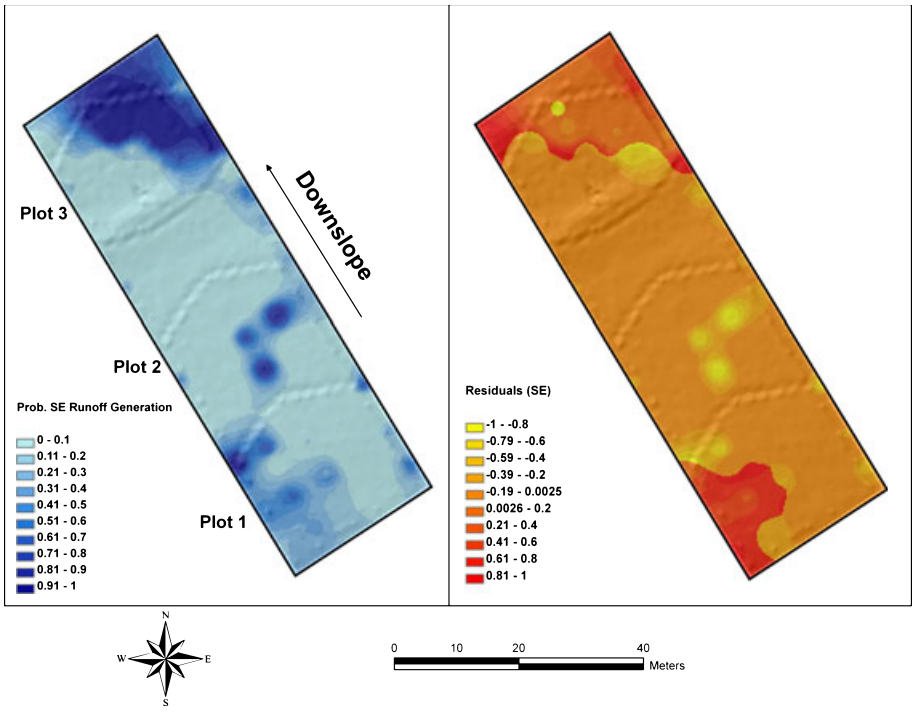

FIGURE 6. Probability of Occurrence of Saturation Excess (SE) Runoff Mechanism and Residual Surface Performed Using Binary Logistic Regression.

TABLE 2. Parameter Estimates and Fit Statistics for the Saturation Excess Runoff Mechanism Binary Logistic Regression Model: $L_k = \alpha_k + \Sigma_{i=1,6}\ \beta_{ki}X_i$ Applied to Each Pixel Location k.

Parameter Estimates	Estimate	Units	Prob > Chi-Sq
Intercept (a)	113.03	-	<0.0001
Bulk density (β_1)	−35.84	g/cm^3	<0.0001
Total porosity (β_2)	−79.85	g/cm^3	<0.0001
Depth to bedrock (β_3)	−0.095	cm	<0.0001
Local slope (β_4)	0.007	%	0.0531[1]
Topographic index (β_5)	0.020	-	0.0594[1]
Preevent depth to saturated soil (β_6)	−0.623	cm	<0.0001

Notes: Pseudo-R^2 (U) = 0.48.
[1]Parameters not significant at α = 0.05.

The probability map developed for the location of infiltration excess runoff areas and its associated residual plot are shown in Figure 7. Table 4 indicates that infiltration excess runoff areas are located in areas of low soil bulk density, total porosity, and soil electrical resistance while areas of high local slopes and depth to bedrock influence the location of infiltration excess areas.

The location of infiltration excess areas in low soil bulk density areas does not seem to follow the infiltration excess theory. One would rather expect that areas of high bulk density would result in low water storage capacity which is directly related to areas of lower infiltration rates (Holtan, 1961) and hence infiltration excess areas. One important infiltration excess variable that was not included in the model is soil infiltration rates. Due to the limited data (three measurements per plot) on the spatial distribution of infiltration rates across the plots, this variable could not be included in the model. In order to obtain the spatial distribution of infiltration rates, one would need approximately 65 point measurements of infiltration capacity to obtain data of the same resolution as the other measured variables from this site. Given the complexity of the karst geologic features beneath the plots (Leh *et al.*, 2008) and the heterogeneous nature of the soil properties at SEW (Sauer *et al.*, 2005), a rather large dataset would be required to obtain reliable estimates of the spatial variability of infiltration capacity across the field. Perhaps its inclusion in the model would have identified the infiltration excess areas in locations with higher bulk density and increased the accuracy of the predictions. However, a low total porosity area has very little pore space available for storage of water. Also, soil electrical resistance is inversely related to the amount of moisture present in pore spaces. Areas of low electrical resistance will therefore indicate soil pore spaces that are already filled with water and hence any excess water would runoff. The local slope and the TI were not found to be significant predictors of the location of infiltration excess runoff.

TABLE 3. Classification Accuracies of the Saturation Excess (SE) Model Applied to the Observed Data.

	Observed			
Predicted	**Non-SE Areas (pixels)**	**SE Areas (pixels)**	**Total (pixels)**	**User's Accuracy (%)[1]**
Non-SE areas (pixels)	***25,302***	2,266	27,568	91.8
SE areas (pixels)	2,444	***5,445***	7,889	69.0
Total (pixels)	27,746	7,711	**35,457**	
Producer's accuracy (%)[2]	91.2	70.6		

Notes: Overall accuracy = (25,302 + 5,445)/35,457 = 86.7%; Average producer's accuracy = 80.9%.
Overall accuracy was obtained by dividing total correct predictions (diagonal, bold) by the total number of pixel predictions (bold). User's accuracies were obtained by dividing the number of correct predictions in a row by the total number of pixel predictions in that row. Producer's accuracies were obtained by dividing the number of correct predictions in a column by the total number of pixel predictions in that column.
[1]Non-SE areas = 25,302/27,568 = 91.8%; SE areas = 5,445/7,889 = 69.0%.
[2]Non-SE areas = 25,302/27,746 = 91.2%; SE areas = 5,445/7,711 = 70.6%.

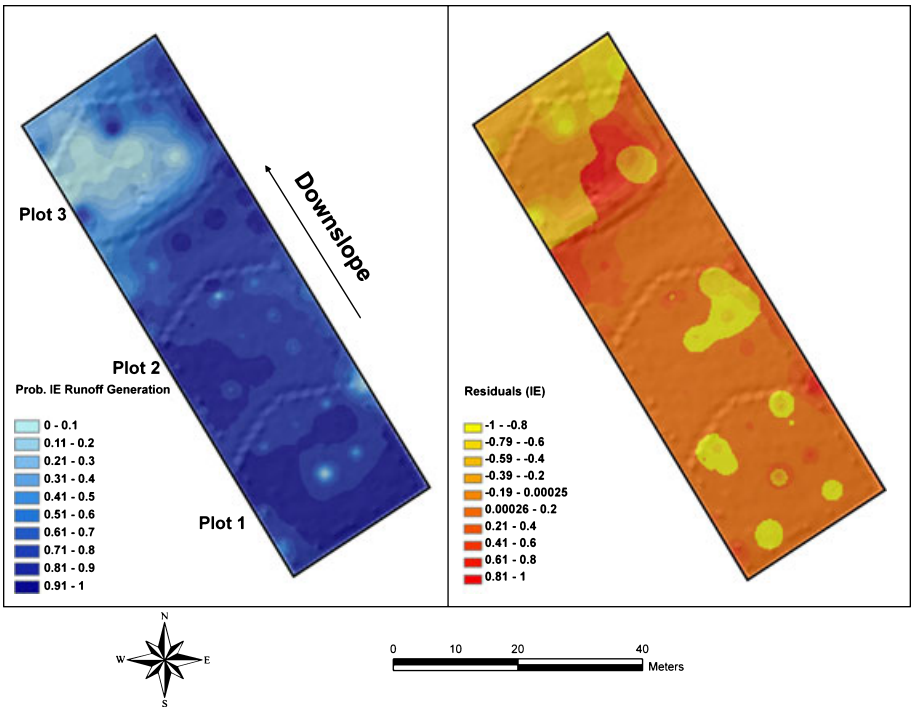

FIGURE 7. Probability of Occurrence of Infiltration Excess (IE) Runoff Mechanism and Residual Surface Performed Using Binary Logistic Regression.

A visual examination of the results indicated that the infiltration excess model did not perform as well as the saturation excess model. Areas located in the "no runoff" zones were classified as infiltration excess areas. Nonetheless the model performed satisfactorily as indicated by an overall accuracy of 81% (Table 5). While individual classes varied between 71 and 91% in the case of saturation excess runoff location (Table 3) and between 74 and 83% in the case of infiltration excess mechanism (Table 5), it is obvious that a greater percentage of the observed data was correctly represented spatially.

Extension to Other Watersheds

The modeling procedures outlined in this manuscript could be easily extended to other agricultural watersheds by following the three main stages of the likelihood modeling procedure suggested by Dalla Bona (1994): (1) primary stage, which includes hypothesis building, data collection strategies and collection; (2) secondary stage, which is the deductive phase and involves association between environmental variables, literature review and integration into model, development or application of initial model,

TABLE 4. Parameter Estimates and Fit Statistics for the Infiltration Excess (IE) Runoff Mechanism Binary Logistic Regression Model: $L_k = a_k + \Sigma_{i=1,5} \beta_{ki} X_i$ Applied to Each Pixel Location k.

Parameter Estimates	Estimate	Units	Prob > Chi-Sq
Intercept (α)	149.9	-	<0.0001
Bulk density (β_1)	−63.6	g/cm^3	<0.0001
Total porosity (β_2)	−141.6	g/cm^3	<0.0001
Electrical resistance (β_3)	−0.004	Ω	<0.0001
Depth to bedrock (β_4)	0.055	cm	<0.0001
Local slope (β_5)	0.003	%	0.2158[1]

Notes: Pseudo-R^2 (U) = 0.25.
[1]Parameter not significant at α = 0.05.

testing of model on previously surveyed areas; and (3) tertiary stage, which is a continuous application and refinement of the model and incorporation of new data into the process and into the existing model. Decisions regarding the scale at which modeling will take place, the spatial boundaries within which the model is applicable and the temporal scope of the model are usually made during the primary stage. The collected baseline data should be at relatively similar scale. Also, the environmental characteristics that are determined to be the independent variables should be representative of the whole study area to be modeled (Dalla Bona, 1994).

The application of this predictive modeling procedure to other watersheds will vary according to the environmental data available, however readily available spatial data such as elevation data (DEM) and soil data (e.g., SSURGO) could be used to supplement field measurements if data such as those used in this study are not available. Soil properties (bulk density, total porosity, hydraulic conductivity, and soil depth) could be estimated from the soil database. Juracek (2000) used soil permeability data derived from the U.S. Department of Agriculture's 1:24,000-scale soils database (SSURGO) to identify infiltration excess runoff locations in Kansas. In well-instrumented watersheds, such as the SEW with relatively well-documented data such as the one used in this study, more accurate predictive models with greater spatial resolution could be possible. However, this modeling procedure may actually prove more useful in ungaged watersheds where the occurrences of runoff processes are unknown due to inadequate data collection.

Generating information about runoff processes and spatial locations can greatly increase the success of a selected BMP plan. Nonpoint sources from agricultural watersheds can be minimized by implementing BMPs that: (1) reduce the source of pollutants, (2) minimize off-site losses by reducing pollutant transport from source areas, or (3) treat the water body affected by specific pollutants of interest. The minimization of off-site losses of pollutants directly depends on the identification of runoff processes and the location of runoff generating areas since runoff is the primary pathway of NPS pollution from agricultural fields. Lyon *et al.* (2006a) demonstrated the importance of applying the appropriate BMP plan to the proper runoff generating areas. A saturation excess based BMP implementation in areas indicated by the TI model in this study would only be effective in 19% of the area whereas a saturation excess based BMP implemented in areas indicated by the likelihood indicator model would be effective in 69% of the area. The infiltration excess likelihood model indicates that an infiltration excess based BMP would be effective in 93% of the area.

By explicitly outlining the variables associated with specific sites and processes, the likelihood indicator model may actually identify specific hydrological pathways. The principal advantage of the likelihood indicator model is its ability to identify processes beyond those indicated in the training data.

TABLE 5. Classification Accuracies of the Infiltration Excess (IE) Model Applied to the Observed Data.

	Observed			
Predicted	**Non-IE Areas (pixels)**	**IE Areas (pixels)**	**Total**	**User's Accuracy (%)[1]**
Non-IE areas (pixels)	***4,715***	5,058	9,773	48.2
IE areas (pixels)	1,687	***23,997***	25,684	93.4
Total (pixels)	6,402	29,055	**35,457**	
Producer's accuracy (%)[2]	73.6	82.6		

Notes: Overall accuracy = (4,715 + 23,997)/35,457 = 81.0%; Average producer's accuracy = 78.1%.
Overall accuracy was obtained by dividing total correct predictions (diagonal, bold) by the total number of pixel predictions (bold). User's accuracies were obtained by dividing the number of correct predictions in a row by the total number of pixel predictions in that row. Producer's accuracies were obtained by dividing the number of correct predictions in a column by the total number of pixel predictions in that column.
[1]Non-IE areas = 4,715/9,773 = 48.2%; IE areas = 23,997/25,684 = 93.4%.
[2]Non-IE areas = 4,715/6,402 = 73.6%; IE areas = 23,997/29,055 = 82.6%.

Specifically, runoff generating areas are identified that possess characteristics similar to the known site features but that are not apparent in the independently measured data.

SUMMARY AND CONCLUSIONS

Data from a two-year field study of runoff generating mechanisms in a pasture hillslope underlain by karst terrain were used to quantify the spatial distribution of occurrence of two runoff generating mechanisms – infiltration excess and saturation excess. Two approaches (TI and binary logistic regression procedures) were used to develop models that predicted the locations of the runoff mechanisms based on the significant hydrologic features that influenced each runoff process. By combining spatial analysis with statistical analysis, the accuracy of predicting the location of the major runoff processes was improved. Results suggest that the binary logistic regression model for the saturation excess mechanism provided more realistic results than the common TI model. Even though the models developed for both runoff mechanisms produced fairly accurate results, the saturation excess mechanism model was a better fit than the infiltration excess model. This method combines complex hydrologic information to produce better estimates of runoff contributing areas and may be applied in areas where the TI does not provide reasonable estimates of critical source areas.

ACKNOWLEDGMENTS

This research project was made possible by funding from the USDA-CSREES (NRI Program) Grant No. 2003-35102-13599, and USGS through Arkansas Water Resources Center (AWRC). The authors would also like to thank P. Srivastava, J. Murdoch, K. Migliaccio, V. Garg, B. Haggard, and J.V. Brahana for assistance in completing this project. Comments provided by three anonymous reviewers greatly improved an earlier version of this manuscript.

LITERATURE CITED

Badoux, A., J. Witzig, P.F. Germann, H. Kienholz, P. Lüscher, R. Weingartner, and C. Hegg, 2006. Investigations on the Runoff Generation at the Profile and Plot Scales, Swiss Emmental. Hydrological Processes 2:377-394, doi: 10.1002/hyp.6056.

Betson, R.P., 1964. What is Watershed Runoff? Journal of Geophysical Research 69:1541-1552.

Beven, K.J. and M.J. Kirkby, 1979. A Physically Based Variable Contributing Area Model of Basin Hydrology. Hydrologic Sciences Bulletin 24:43-69.

Brahana, J.V., T.E. Ting, M. Al-Qinna, J.F. Murdoch, R.K. Davis, J. Laincz, J.J. Killingbeck, E. Szilvagyi, M. Doheny-Skubic, I. Chaubey, P.D. Hays, and G. Thoma, 2005. Quantification of Hydrologic Budget Parameters for the Vadose Zone and Epikarst in Mantled Karst. *In:* Proceedings of the U.S. Geological Survey Karst Interest Group, Rapid City, South Dakota, September 12-15, 2005, E.L. Kuniansky (Editor). U.S Geological Survey Scientific Investigations Report 2005-5160, pp. 144-152.

Chaubey, I., D.R. Edwards, T.C. Daniel, P.A. Moore Jr., and D.J. Nichols, 1995. Effectiveness of Vegetative Filter Strips in Controlling Losses of Surface-Applied Poultry Litter Constituents. Transactions of the American Society of Agricultural Engineers 38:1687-1692.

Clark, W.A. and P.L. Hosking, 1986. Statistical Methods for Geographers. John Wiley & Sons, New York.

Congalton, R.G. and K. Green, 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publishers, Boca Raton, Florida.

Congalton, R.G., R.G. Oderwald, and R.A. Mead, 1983. Assessing Landsat Classification Accuracy Using Discrete Multivariate Statistical Techniques. Photogrammetric Engineering and Remote Sensing 49:1671-1678.

Dalla Bona, L. 1994. Cultural Heritage Resource Predictive Modeling Project: Methodological Considerations. A Report Prepared for the Ontario Ministry of Natural Resources. Lakehead University: Center for Archaeological Resource Prediction, Thunder Bay, Ontario, 3:1-45.

Dunne, T. and R.D. Black, 1970a. An Experimental Investigation of Runoff Production in Permeable Soils. Water Resources Research 6:78-490.

Dunne, T. and R.D. Black, 1970b. Partial Area Contributions to Storm Runoff in a Small New England Watershed. Water Resources Research 6:1296-1311.

Edwards, D.R. and T.C. Daniel, 1993. Runoff Quality Impacts of Swine Manure Applied to Fescue Plots. Transactions of the American Society of Agricultural Engineers 36:81-86.

Edwards, D.R. and T.C. Daniel, 1994. Quality of Runoff From Fescuegrass Plots Treated With Poultry Litter and Inorganic Fertilizer. Journal of Environmental Quality 23:579-584.

Edwards, D.R., C.T. Haan, J.F. Murdoch, A.N. Sharpley, T.C. Daniel, and P.A. Moore Jr., 1996b. Application of Simplified Phosphorus Transport Models to Pasture Fields in Northwest Arkansas. Transactions of the American Society of Agricultural Engineers 39:489-496.

Edwards, D.R., P.A. Moore Jr., T.C. Daniel, and P. Srivastava, 1996a. Poultry Litter-Treated Length Effects on Quality of Runoff From Fescue Plots. Transactions of the American Society of Agricultural Engineers 39:105-110.

Endreny, T.A. and E.F. Wood, 2003. Watershed Weighting of Export Coefficients to map Critical Phosphorous Loading Areas. Journal of the American Water Resources Association 39:165-181.

Engman, E.T., 1974. Partial Area Hydrology and its Implications to Water Resources. Water Resources Bulletin 10:512-521.

Ernenwein, E.G. and K.L. Kvamme, 2004. Geophysical Investigations for Subsurface Fracture Detection in the Savoy Experimental Watershed, Arkansas. Archeo-Imaging Lab unnumbered report, University of Arkansas, Fayetteville, Arkansas, pp. 1-48.

Harper, M.D., W.W. Phillips, and G.J. Haley, 1969. Soil Survey of Washington County, Arkansas. U.S. Department of Agriculture, Soil Conservation Service, U.S. Government Printing Office, Washington, D.C.

Hewlett, J.D. 1961. Soil Moisture as a Source of Base Flow Seep Mountain Watersheds. Southeast Forest Experiment Station Paper 132. U.S. Forest Service, U.S. Government Printing Office, Washington, D.C.

Hewlett, J.D. and A.R. Hibbert, 1967. Factors Affecting the Response of Small Watersheds to Precipitation in Humid Regions. In: Forest Hydrology. W.E. Sopper, and H.W. Lull (Editors). Pergamon Press, Oxford, England. pp. 275-290.

Holtan, H.N., 1961. A Concept of Infiltration Excess Estimates in Watershed Engineering. USDA-ARS, Washington, D.C., pp. 41-51.

Horton, R.E., 1933. The Role of Infiltration in the Hydrologic Cycle. Transaction of the American Geophysical Union 14:446-460.

Hosmer, D.W. and S. Lemeshow, 2000. Applied Logistic Regression (Second Edition). John Wiley, New York.

Johnson, M.S., W.F. Coon, V.K. Mehta, T.S. Steenhuis, E.S. Brooks, and J. Boll, 2003. Application of two Hydrologic Models With Different Runoff Mechanisms to a Hillslope Dominated Watershed in the Northeastern U.S.: A Comparison of HSPF and SMR. Journal of Hydrology 284:57-76, doi: 10.1016/j.jhydrol.2003.07.005.

Juracek, K.E. 1999. Estimation of Potential Runoff-Contributing Areas in the Kansas-Lower Republican River Basin, Kansas. USGS Water Resources Investigation Report 99-4089.

Juracek, K.E. 2000. Estimation and Comparison of Potential Runoff-Contributing Areas in Kansas Using Topographic, Soil, and Land-use Information. USGS Water Resources Investigation Report 00-4177.

Leh, M.D., I. Chaubey, J. Murdoch, J.V. Brahana, and B.E. Haggard, 2008. Delineating Runoff Processes and Critical Runoff Source Areas in a Pasture Hillslope of the Ozark Highlands. Hydrological Processes 22:4190-4204.

Lyon, S.W., A.J. Lembo Jr., M.T. Walter, and T.S. Steenhuis, 2006b. Defining Probability of Saturation With Indicator Kriging on Hard and Soft Data. Advances in Water Resources 28:181-193.

Lyon, S.W., M.R. McHale, M.T. Walter, and T.S. Steenhuis, 2006a. The Impact of Runoff-Generation Mechanisms on the Location of Critical Source Areas. Journal of the American Water Resources Association 42:793-804.

Lyon, S.W., M.T. Walter, P. Gérard-Marchant, and T.S. Steenhuis, 2004. Using a Topographic Index to Distribute Variable Source Area Runoff Predicted With the SCS-Curve Number Equation. Hydrological Processes 18:2757-2771.

Mehta, V.K., M.T. Walter, E.S. Brooks, T.S. Steenhuis, M.F. Walter, M. Johnson, J. Boll, and D. Thongs, 2004. Application of SMR to Modeling Watersheds in the Catskill Mountains. Environmental Modeling and Assessment 9:77-89.

Meyles, E.W., A.G. Williams, J.F. Dowd, and J.L. Ternan, 2003. Soil Moisture Patterns and Runoff Generation in a Small Dartmoor Catchment, Southwest England. Hydrological Processes 17:251-264.

NOAA, 2002. Climatography of the United States No. 81. Monthly Station Normals of Temperature, Precipitation, and Heating and Cooling Degree Days 1971-2000: Arkansas, Asheville, North Carolina.

Page, T., P. Haygarth, K.J. Beven, A. Joynes, T. Butler, C. Keeler, J. Freer, P. Owens, and G.A. Wood, 2005. Spatial Variability of Soil Phosphorus in Relation to the Topographic Index and Critical Source Areas: Sampling for Assessing Risk to Water Quality. Journal of Environmental Quality 34:2263-2277, doi: 10.2134/jeq2004.0398.

Rezzoug, A., A. Schumann, P. Chifflard, and H. Zepp, 2005. Field Measurement of Soil Moisture Dynamics and Numerical Simulation Using the Kinematic Wave Approximation. Advances in Water Resources 28:917-926.

Sauer, T.J., T.C. Daniel, P.A. Moore Jr., K.P. Coffeey, D.J. Nichols, and C.P. West, 1999. Poultry Litter and Grazing Animal Waste Effects on Runoff Water Quality. Journal of Environmental Quality 28:860-865.

Sauer, T.J., T.C. Daniel, D.J. Nichols, C.P. West, P.A. Moore Jr., and G.L. Wheeler, 2000. Runoff Water Quality From Poultry Litter-Treated Pasture and Forest Sites. Journal of Environmental Quality 29:515-521.

Sauer, T.J. and S.D. Logsdon, 2002. Hydraulic and Physical Properties of Stony Soils in a Small Watershed. Soil Science Society of America Journal 66:1947-1956.

Sauer, T.J., S.D. Logsdon, J.V. Brahana, and J.F. Murdoch, 2005. Variation in Infiltration With Landscape Position: Implications for Forest Productivity and Surface Water Quality. Forest Ecology and Management 220:118-127.

Sauer, T.J., P.A. Moore Jr, K.P. Coffeey, and E.M. Rutledge, 1998. Characterizing the Surface Properties of Soils at Varying Landscape Positions in the Ozark Highlands. Soil Science 163:907-915.

Srinivasan, M.S., W.J. Gburek, and J.M. Hamlett, 2002. Dynamics of Storm Flow Generation-A Hillslope Scale Field Study in East-Central Pennsylvania, USA. Hydrological Processes 16:649-665.

Srinivasan, M.S., M.A. Wittman, J.M. Hamlett, and W.J. Gburek, 2000. Surface and Subsurface Sensors to Record Variable Runoff Generation Areas. Transactions of the American Society of Agricultural Engineers 43:651-660.

Steele, K.F. and J.C. Adamski, 1987. Land Use Effect on Ground Water Quality in Carbonate Rock Terrain. Publication Number 129. Arkansas Water Resource Center, Fayetteville, Arkansas.

Steele, K.F. and W.K. McCalister, 1991. Potential Nitrate Pollution of Groundwater in Limestone Terrain by Poultry Litter, Ozark Region, U.S.A. *In:* Nitrate Contamination Exposure, Consequence, and Control. I. Bogardi and R.D. Kuzelka (Editors). NATO ASI Series. Springer-Verlag, Heidelberg, Berlin, pp. 30: 209-218.

Story, M. and R. Congalton, 1986. Accuracy Assessment: A User's Perspective. Photogrammetric Engineering and Remote Sensing 52:397-399.

Tarboton, D.G., 1997. A new Method for the Determination of Flow Directions and Upslope Areas in Grid Digital Elevation Models. Water Resources Research 33:309-319.

USDA-ARS, 1998. Managing Poultry Manure Nutrients. Agricultural Research Magazine 46:12-13.

Walter, M.T., M.F. Walter, E.S. Brooks, T.S. Steenhuis, J. Boll, and K.R. Weiler, 2000. Hydrologically Sensitive Areas: Variable Source Area Hydrology Implications for Water Quality Risk Assessment. Journal of Soil and Water Conservation 3:277-284.

White, K.L. and I. Chaubey, 2005. Multi-Site and Multi-Variable Calibration of the SWAT Model. Journal of American Water Resources Association 41:1077-1089.

Wood, E.F., M. Sivapalan, and K. Beven, 1990. Similarity and Scale in Catchment Storm Response. Reviews of Geophysics 28:1-18.

Zollweg, J.A., W.J. Gburek, H.B. Pionke, and A.N. Sharpley, 1995. GIS-Based Delineation of Source Areas of Phosphorous Within Agricultural Watersheds of the Northeastern U.S. *In:* S.P.K. Simonovic, and Z. Takeuchi (eds.) Proceedings of a Boulder Symposium on Modeling and Management of Sustainable Basin-Scale Water Resource Systems. IAHS Press, Wallingford, United Kingdom, pp. 31-39.