THE IMPORTANCE OF SEEPAGE ZONES IN PREDICTING SOIL MOISTURE CONTENT AND SURFACE RUNOFF USING GLEAMS AND RZWQM

A. Chinkuyu, T. Meixner, T. Gish, C. Daughtry

ABSTRACT. *Seepage zones have been shown to be of critical importance in controlling contaminant export from agricultural catchments. To date, no multi−purpose agricultural water quality model has incorporated seepage zones into its process−level representations. We chose to test two commonly used models of agricultural water quality, Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) and the Root Zone Water Quality Model (RZWQM), by seeing how well each predicted surface runoff and soil moisture content in two agricultural fields: one with and one without seepage zones. Daily simulated surface runoff and soil moisture content from both calibrated and default (or non−calibrated) GLEAMS and RZWQM were compared with three years of measured surface runoff and soil moisture content in the two fields. The results of the study show that GLEAMS and RZWQM, using default model parameters, were not capable of predicting surface runoff and soil moisture content in either field. Site−calibrated GLEAMS and RZWQM performed well in simulating surface runoff trends from the field with and without seepage zones, but they predicted soil moisture content poorly. Several statistical tests were used that showed that although both site−calibrated GLEAMS and RZWQM performed well, RZWQM performed better than GLEAMS and is better suited in assessing the effects of seepage zones on soil moisture content and surface runoff from agricultural fields.*

Keywords. *Field, Models, Runoff, Seepage zones, Sensitivity, Soil moisture.*

Although variable source area processes (e.g., seepage zones) have been well studied in forested and range land systems (Grayson and Blöschl, 2001; Walter et al., 2000), the impacts of these processes on agricultural processes has only recently been closely investigated and only in a limited number of settings (Gburek and Sharpley, 1998; Daughtry et al., 2001). Seepage zones occur when subsurface flow channels re−emerge on the surface and are common to agricultural lands bordering streams. Seepage zones can strongly influence surface runoff and chemical fluxes; however, their impacts on water quality have not been effectively modeled because of lack of data (Gburek et al., 2002; Gish et al., 2001).

Computational agricultural water quality models provide an opportunity to evaluate the response of soil and water resources to different farming practices, climatic conditions, soil, and topographic properties in an efficient and cost−effective way. However, the reliability of these models depends on how well each process is represented and on the accuracy of the model parameters used. Although most agricultural water quality models were developed to evaluate the effects of management practices on nonpoint−source pollutant loads, as opposed to the use of the models as absolute predictors of hydrology, the hydrologic component must be a reasonable representation of the processes occurring in a catchment. For this reason, agricultural water quality models are usually calibrated and evaluated for their hydrologic performance first before the chemical component of the model is addressed (Buchleiter et al., 1995; Hanson et al., 1999; Leonard et al., 1987; Truman et al., 1998). For example, if surface runoff, percolation, evaporation, and soil water storage can be quantified and simulated with reasonable accuracy, then it is assumed that the models can be used with confidence to evaluate nonpoint−source pollution from the field.

However, to determine if the model reasonably simulates the real conditions and to gauge the model's usefulness, an assessment of its performance for a variety of soil, crop, management practice, hydrologic, and climatic conditions is needed. Correlation and correlation−based measures (e.g., R^2) have been widely used to evaluate the "goodness−of−fit" of hydrologic and water quality models. However, these measures are oversensitive to extreme values (outliers) and are insensitive to additive and proportional differences between model predictions and observations (Legates and McCabe, 1999). In general, a single evaluation measure can indicate that a model is a good predictor, when in fact it is not. Because of these limitations, additional evaluation criteria, e.g., coefficient of determination (R^2), relative percent error (E_r), coefficient of efficiency (E), index of agreement (*d*), and absolute maximum error ($|E_{max}|$), have been proposed by different researchers to assess model performance (Buchleiter et al., 1995; Haan et al., 1993; Legates and McCabe, 1999).

Article was submitted for review in May 2003; approved for publication by the Soil & Water Division of ASAE in January 2004.

The authors are **Adion Chinkuyu, ASAE Member Engineer,** Post−Doctoral Research Associate, USDA−ARS Hydrology and Remote Sensing Lab, Beltsville, Maryland; **Tom Meixner, ASAE Member Engineer,** Assistant Professor, Department of Environmental Sciences, University of California, Riverside, California; and **Timothy Gish,** Soil Scientist, and **Craig Daughtry,** Research Agronomist, USDA−ARS Hydrology and Remote Sensing Lab, Beltsville, Maryland. **Corresponding author:** Adion Chinkuyu, USDA−ARS Hydrology and Remote Sensing Lab, 104 Building 007, Beltsville, MD 20705; phone: 301−504−5825; fax: 301−504−8931; e−mail: acinkuyu@ hydrolab. arsusda.gov.

  ISSN 0001−2351

In this study, two well−known and comprehensive models, Groundwater Loading Effects of Agricultural Management Systems (GLEAMS) v. 3.0.1, and Root Zone Water Quality Model (RZWQM) 98 v.1.0.2000.830, were calibrated and evaluated on two agricultural fields (field C with seepage zones, and field A without seepage zones) at the USDA research center in Beltsville, Maryland. The GLEAMS and RZWQM models were chosen for this study because: (1) they are common and widely used to evaluate agricultural management practices under different soil, climatic, and hydrologic conditions; (2) to our knowledge, there is no agricultural nonpoint−source pollution model that represents saturation and seepage zone processes in addition to the important biological and chemical controls on agricultural water quality (e.g., nutrient dynamics and pesticide processes); and (3) RZWQM and GLEAMS hydrologic concepts are passed on into other larger scale models, e.g., GLEAMS concepts are used in the Soil and Water Assessment Tool (SWAT) model.

The main objectives of this study were: (1) to assess the ability of GLEAMS and RZWQM models to simulate surface runoff and soil moisture content in agricultural fields under different hydrologic conditions (presence or absence of seepage zones), (2) to gain insight into how the models represent or fail to represent natural hydrologic processes (particularly seepage zone processes) and how they can be improved, and (3) to test the performance of GLEAMS and RZWQM using multi−evaluation techniques.

Background on the GLEAMS Model

The GLEAMS model was developed to simulate edge−of−field and bottom−of−root−zone loadings of water, sediment, pesticides, and plant nutrients from the complex climate−soil−management interactions (Knisel, 1993). As a field−scale water quality model, GLEAMS has been evaluated under different conditions and management practices with varied results (Chinkuyu and Kanwar, 2001; Bakhsh et al., 2000).

GLEAMS assigns soil physical and hydraulic properties into a maximum of 12 computational layers, taking input from one to five soil horizons based on the user's selection. Soil profile description and crop data are used to estimate effective rooting depth. Soils data are input by soil horizon, and the model assigns values for porosity, water retention characteristics, and organic matter into the appropriate computational layers. The GLEAMS model uses daily climatic data to calculate the water balance in the root zone. The hydrology component simulates runoff due to daily rainfall using a modified Natural Resources Conservation Service (NRCS) curve number method. Percolation in GLEAMS is estimated using a storage−routing technique whereby each layer in the soil is assumed to have a storage capacity, defined as the field capacity (Knisel, 1993). When infiltration from the above soil layer causes soil moisture to reach the field capacity, any excess infiltration is routed to the next lower layer in the profile. Percolation from the lowest layer is outflow from the system and is referred to as the "percolation component." Potential evaporation is estimated by the Penman−Monteith equation for arid and semi−arid regions or by the Priestly−Taylor equation for humid areas. Seepage zones and variable source area processes are not represented explicitly in the GLEAMS model. The ability of GLEAMS to represent variable source area processes is thus limited to generating compensating−hydrologic parameter values to represent seepage zone processes in a limited manner.

Background on the RZWQM Model

The USDA−ARS Root Zone Water Quality Model (RZWQM) is a physically based simulation model designed to predict hydrologic and chemical responses, including potential groundwater contamination of agricultural management systems (Ahuja et al., 2000; RZWQM Team, 1995). RZWQM is sufficiently comprehensive to predict the relative response of plants and interactions among system processes to changes in water balance, temperature, nutrient cycling, plant growth, and soil chemistry as management practices change. Such practices include the application of manure, crop residue, and tillage that are introduced into the system. Details of the model components are given in the model documentation (Ahuja et al., 2000) and are not repeated here. The RZWQM model used in this study is the RZWQM98 v. 1.0.2000.830, which is a Windows version of the original RZWQM v.3.25.

Water flow and chemical transport processes in RZWQM are divided into two phases: (1) infiltration into the soil matrix during rainfall or irrigation, and (2) redistribution of water and chemicals following infiltration. A modified form of the Green−Ampt model is used to calculate infiltration, which incorporates initial soil moisture content for each rainfall event (Ahuja et al., 1993). Rainfall in excess of infiltration becomes surface runoff. Redistribution of water and chemicals between rainfall or irrigation events is modeled by Richard's equation, and the subsurface drainage rate is calculated from Hooghoudt's steady−state equation. RZWQM simulates potential ET using a modified Penman−Monteith model, and actual ET is constrained by stomatal resistance and water availability, as estimated from Richards' equation. As in GLEAMS, seepage zone processes are not explicitly represented in RZWQM model and must be represented by adjusting the effective hydrologic parameters of the model.

Methods

Experimental Site

Four years of data (1999 to 2002) for calibration and evaluation of GLEAMS and RZWQM were obtained from two neighboring fields (fig. 1) at the USDA research center in Beltsville, Maryland. The 7−year average annual precipitation at the research site is about 87 cm. Annual total precipitation values in 1999, 2000, 2001, and 2002 were approximately 95, 91, 87, and 89 cm, respectively. In 1999, no significant amount of rainfall fell until September, when major storms including Hurricane Floyd generated significant surface runoff. During the other years, precipitation was uniformly distributed during the growing season (between April to mid November), which resulted in some surface runoff throughout the season.

The two fields have similar soils, climate, and agricultural management practices. However, field C (4.0 ha) has large natural seepage zones, and field A (3.6 ha) has no seepage zones. Seepage zones occur when subsurface flow channels re−emerge on the surface and are common to agricultural lands bordering streams. Each field drains into a riparian wetland forest that contains a first−order stream. Surface runoff water (including runoff from seepage zones) from each field was measured with a 45.7 cm H−flume equipped with a flowmeter

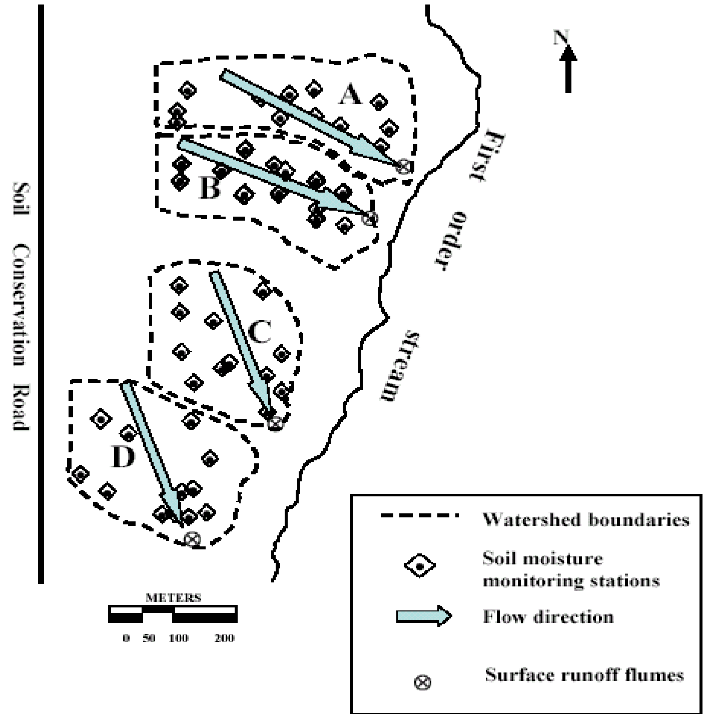

Figure 1. Field layout at the research site.

Table 1. Selected physical soil properties measured at the study site and used as inputs in the models.

Soil Depth (cm)	Clay (%)	Silt (%)	Sand (%)	Organic Matter (%)
0 − 15	5	15	80	3.5
15 − 30	11	18	71	3.2
30 − 75	7	10	83	3.0
75 − 90	6	16	78	2.0
90 − 120	10	25	65	2.0

Table 2. Dates of field activities at the study site.

	Year and Date			
Activity	1999	2000	2001	2002
Chisel plowing	8 May	17 May	12 May	20 May
Applying manure	15 May	2 June	25 May	10 June
Incorporating manure	15 May	2 June	25 May	10 June
Planting corn	28 May	9 June	29 May	12 June
Applying herbicides	1 June	13 June	10 June	24 June
Applying fertilizer	26 June	13 July	28 June	15 July
Harvesting corn	7 Nov.	5 Nov.	1 Nov.	10 Nov.

and a water sampler. Amount of surface runoff was measured automatically and continuously recorded whenever there was a runoff event. Field slopes ranged from 0% to 5%, becoming greater as they approached the riparian wetland. Soils at the site were formed from sandy fluvial deposits with a predominantly sandy loam texture for the first 60 cm followed by loam from 60 to 120 cm. Selected physical soil properties measured at the research site are presented in table 1 (Daughtry et al., 2001).

Capacitance probes (EnviroScan, Sentek Pty Ltd., Adelaide, South Australia) were used to measure soil moisture content in both fields. Twelve soil moisture probes were used to distribute 64 soil moisture sensors within each field (fig. 1). In each field, 12 sensors were located at 10 and 30 cm, with eight additional sensors at 50, 80, 120, 150, and 180 cm deep. Probes were installed throughout the field in areas with low, medium, and high infiltration capacities based on electromagnetic (EM−38) measurement values (Daughtry et al., 2001). Although each soil moisture sensor was activated at 10 min intervals, daily volumetric soil moisture contents for each field were calculated by averaging across probes and then averaging over the top 120 cm. An effective rooting depth (soil profile) of 120 cm was chosen because it gave reasonable amount of surface runoff, evapotranspiration, soil moisture storage, and deep percolation by both GLEAMS and RZWQM.

Both fields were tilled in early spring using chisel plow. Digested dairy cow liquid manure was applied to both fields from 1999 to 2002. Immediately after application of manure, the soil was disked to incorporate the manure and minimize N loss through volatilization. During the four growing seasons, corn (*Zea mays* L.) was planted in both fields after the application of manure. Additional urea ammonium nitrate (UAN) fertilizer was side−dressed in both fields according to the pre−sidedress nitrate test (PSNT) (Meisinger et al., 1992). Dates of management activities for both fields are given in table 2.

GLEAMS AND RZWQM DATA INPUT

Climatic Data

The GLEAMS model requires mean daily air temperature, daily precipitation, mean monthly maximum and minimum temperatures, solar radiation, wind speed, and dewpoint temperature data. The RZWQM model also requires input values of daily minimum and maximum air temperature, wind speed, shortwave radiation, and relative humidity. Breakpoint rainfall (amount of rainfall where rainfall intensity changes) is also required as input to RZWQM. Climatic data measured at the experimental site were used as input to both models.

Soil Physical Parameters

In both GLEAMS and RZWQM, an effective rooting depth of 120 cm was used and divided into five horizons based on soil texture. Data on clay, silt, sand, and organic matter contents were measured at the site and are presented in table 2. Based on the soil properties at the site, the soils in the two fields were classified as belonging to the hydrological soil Group B according to GLEAMS user's manual. Porosity, field capacity, wilting point, and hydraulic conductivity were obtained from the GLEAMS database (default values). Soil bulk density, porosity, saturated hydraulic conductivity, and soil water content were taken from the RZWQM database (default values) and used as inputs to RZWQM.

Management Practices

All management (tillage, planting, harvesting) information was collected each year at the site and used as input to both models (table 2). Other crop characteristics data such as leaf area index, crop height, dry matter ratio, residue C: N ratio, and N: P ratio were taken from the model's respective databases.

MODEL CALIBRATION

All process−level, comprehensive models (such as GLEAMS and RZWQM) require a detailed set of parameters. Some of these parameters cannot be easily determined or measured. Both GLEAMS and RZWQM simulations were first performed using best available estimates for the required parameters. These default parameters were obtained from the GLEAMS databases (Knisel et al., 1993) and RZWQM databases (Ahuja et al., 2000) based on local site information from the Prince Georges County soil survey and field and laboratory measurements made at this research site. Simulations were first run (from 1 January 1999 to 31 December 2002) using default model input parameters to compare model simulations based on default parameters. Sensitive parameters of GLEAMS and RZWQM models were site−calibrated using field runoff and soil moisture data for 1999. Data collected in 2000, 2001, and 2002 were used for model evaluation (verification).

GLEAMS Model Calibration

GLEAMS calibration focused mainly on parameters controlling surface runoff and soil moisture content. Calibration of model parameters included the NRCS curve number for soil moisture condition II (CN2), field capacity, permanent wilting point, and effective rooting depth (RD). Curve numbers of 78 and 82 for hydrologic soil Group B were selected from the user's manual and calibrated for field A (without seepage zones) and field C (with seepage zones), respectively. Although field C had large natural seepage zones in the middle and close to the outlet, the whole field was assumed to be homogeneous, thus using one curve number as in field A. This decision was made since subdividing the field would increase model complexity and limit our ability to make a fair comparison between field A and field C. Our intent was to compare these two models based on how well they did or did not simulate conditions for these two fields.

The RD, field capacity, permanent wilting point, and CN2 calibrations were manually calibrated until the simulated surface runoff and soil moisture content closely matched the measurements of these two variables for 1999. The model was considered calibrated when the relative percent error was less than ±25% for the two objectives. The initial and final calibrated parameter values are given in table 3.

RZWQM Model Calibration

Saturated hydraulic conductivity, field capacity, and rooting depth were identified as sensitive parameters controlling surface runoff, evapotranspiration, and soil moisture content. If predicted soil moisture content and surface runoff did not match the observed values (within ±25% range), then saturated hydraulic conductivity, field capacity, and rooting depth were adjusted repeatedly until suitable responses were obtained. Initial and final calibrated parameter values for RZWQM are given in table 3.

MODEL EVALUATION AND DATA ANALYSIS

Default and site−calibrated GLEAMS and RZWQM models were tested using measured data from field A (without seepage zones) and field C (with seepage zones) over a three−year period (2000 to 2002). Comparisons were made only for periods with measured data, although simulations were continuous from 1 January 1999 through 31 December 2002. Several techniques, based on objective and subjective approaches, were used to test the performance of the models (Bakhsh et al., 2000; Chinkuyu and Kanwar, 2001). Subjective criteria included graphical display of simulated and measured surface runoff amount and soil moisture content. The subjective criteria were used to locate anomalies in model predic−

Table 3. Sensitive parameters used in calibrating GLEAMS and RZWQM models.

Model and parameter[a]	Field A Def.[b]	Field A Cal.[c]	Field C Def.	Field C Cal.
GLEAMS				
NRCS curve number	78	74	82	80
Eff. root depth (cm)	100	120	100	120
Field capacity at depth (%):				
15 cm	19	23	22	29
30 cm	22	23	19	28
45 cm	30	26	22	29
90 cm	22	23	16	28
120 cm	19	23	22	29
Wilting point at depth (%):				
15 cm	5	9	8	7
30 cm	8	11	5	6
45 cm	18	17	8	8
90 cm	8	11	3	6
120 cm	5	9	8	8
RZWQM				
Rooting depth (cm)	100	120	100	120
Sat. hydraulic conductivity (cm/h) at depth:				
15 cm	6.11	5.22	2.59	2.11
30 cm	2.59	1.70	6.11	2.77
45 cm	0.43	0.35	2.59	2.11
90 cm	2.59	1.70	21.00	12.00
120 cm	6.11	5.22	2.59	2.11
Field capacity at depth (%):				
15 cm	10.64	10.74	19.17	20.19
30 cm	19.17	19.27	10.64	11.65
45 cm	24.58	24.68	19.17	20.19
90 cm	19.17	19.27	6.33	7.45
120 cm	10.64	10.74	19.17	20.19

[a] No other parameters were calibrated in the model.
[b] Def = default, i.e, initial parameter values obtained from models' databases before calibration.
[c] Cal. = calibrated, i.e., final parameter value after calibration. Calibrated values were used to simulate 2000, 2001, and 2002.

tions and to provide an insight into temporal response of the model for the entire simulation period. Several statistical criteria were used that account for differences over the whole simulation, ignoring differences between simulations and observations over time.

Objective criteria included: coefficient of determination (R^2), regression analysis slope (m) and intercept (b), relative percent error (E_r), coefficient of efficiency (E), index of agreement (d), absolute maximum error ($|E_{max}|$), and probability for the mean difference ($P \leq |t|$) at significance level of 0.05 (table 4). For each evaluation technique, a benchmark that the model should outperform for each of these statistics was established based on other studies (rightmost column in table 4). A model was considered to have performed well when: (1) relative percent error was between −25% and +25%, (2) coefficient of efficiency was greater than 0, (3) coefficient of determination was greater than 0.5, (4) the index of agreement was greater than 0, (5) probability for the mean difference (using Student's t−test) was equal to or greater than 0.05, (6) absolute maximum error was between −25% and +25% of the total or average observed values, and (7) slope of the regression line was between +0.5 and +1.4, and intercept of the regression line was between −5% and +5% around the origin for the predicted data (see table 4 for specific comments). These benchmark values were chosen based on other studies that gave similar "acceptable" values showing good model performance (Leavesley et al., 1983; Wilcox et al., 1990).

For each variable (surface runoff or soil moisture), the model would get a score of 1 for each of the evaluation techniques meeting the conditions specified in table 4. The scores were then added for each model to get the total score. There was a possible maximum total score of 8 and a possible minimum score of zero for each combination of model, field, and variable of interest (runoff and soil moisture). The model with the highest total score for a given variable of interest was considered to be superior to the other. By using graphical comparison, several statistical tests, and the combined scoring from the statistical tests, we hoped to get a more robust picture of model performance than if we had just used graphical comparison and a single statistical test, as is done for most water quality modeling studies.

RESULTS AND DISCUSSION

SIMULATIONS WITH DEFAULT MODEL PARAMETERS

The results of GLEAMS and RZWQM with default input parameters from 1999 to 2002 are presented in tables 5 through 7 and figures 2 through 7. Note that data and predictions for the year 1999 were used to adjust model parameters so that they would match the observations. In field A, both GLEAMS and RZWQM using default parameters overpredicted surface runoff in each year with relative percent errors of over 60% (table 5). When all data from the evaluation period (2000 to 2002) were pooled together, the results show that GLEAMS using default parameters overpredicted surface runoff with a relative percent error of 109% and a probability for a mean difference of less than 0.001 (table 6). RZWQM overpredicted surface runoff from field A with a relative percent error of 377% (table 6). The results also show that both GLEAMS and RZWQM simulations using default parameters were not capable of predicting soil moisture content in field A (tables 5 and 6). The poor performance of the default models is supported by the low total score of the evaluation criteria, indicating that few of the performance criteria for these models were surpassed (table 6).

In field C, GLEAMS and RZWQM with default input parameters underpredicted surface runoff by −53% and −29%, respectively (table 7). RZWQM performed relatively better than GLEAMS in simulating surface runoff from field C (table 7 and figs. 5 and 6). GLEAMS and RZWQM underpredicted soil moisture content in field C by −31% and −37%, respectively. The results in table 7 show that both models had low total scores for the evaluation techniques for both soil moisture and surface runoff. The overall results show that both GLEAMS and RZWQM with default input parameters were not capable of predicting surface runoff and soil moisture content in either field.

Table 4. Techniques used to evaluate performance of GLEAMS and RZWQM models.

Evaluation Technique	Comments	Scoring Criteria[a]
1. Relative percent error (E_r)	E_r = 0 indicates perfect model. E_r = − indicates model is underpredicting. E_r = + indicates model is overpredicting.	$-25\% \leq E_r \geq +25\%$
2. Coefficient of efficiency (E)	E = − indicates model does not simulate data as well as average of observations. E = + indicates that model performs better at estimating observations than the average.	E > 0 (i.e., positive)
3. Coefficient of determination (R^2)	R^2 = 1 indicates perfect. R^2 = 0 indicates low agreement.	$R^2 \geq 0.5$
4. Index of agreement (d)	d = 1.0 indicates good agreement. d = − indicates model performs poorly.	d > 0 (i.e., positive)
5. Mean difference (M_d), tested with Student's t−test at level of significance of P = 0.05.	P > 0.05 indicates M_d = 0 (i.e., no significant difference between model and observed values). P < 0.05 indicates $M_d \neq 0$ (i.e., model not predicting well).	$P \geq 0.05$
6. Absolute maximum error ($\lvert E_{max} \rvert$)	Large value indicates poor agreement between model and observed data.	$-25\% = \lvert E_{max} \rvert \geq +25\%$ of total or average observed values.
7. Regression analysis: slope (m) and intercept (b)	m = 1 and b = 0 indicate model is predicting well.	$0.5 = m \geq +1.4$ $-5\% \leq b \geq +5\%$ around the origin for the predicted runoff or soil moisture values.

[a] For each scoring criteria/condition met, the model gets a score of 1. The scores for each criterion are then summed and reported as a total score. A model with the highest total score (possible maximum of 8 and minimum of 0) was considered superior to the other.

Table 5. Annual measured and predicted surface runoff and soil moisture contents in the two fields.

Year	Measured	GLEAMS Def.[a]	GLEAMS Def. E_r[b]	GLEAMS Cal.[c]	GLEAMS Cal. E_r	RZWQM Def.	RZWQM Def. E_r	RZWQM Cal.	RZWQM Cal. E_r
Total surface runoff (cm)									
Field A									
1999[d]	1.91	3.11	63	2.14	12	3.40	78	2.22	16
2000	0.20	0.78	−170	0.25	19	2.66	1167	0.25	19
2001	2.06	3.54	72	1.73	−16	6.44	213	1.95	−5
2002	0.20	0.79	295	0.19	−5	2.62	1210	0.24	20
Field C									
1999	6.23	3.92	−37	6.72	8	4.21	−32	6.24	0.2
2000	6.59	2.93	−56	6.04	−8	5.01	−24	6.53	−0.9
2001	8.87	4.53	−49	8.42	−5	7.42	−16	10.49	18
2002	4.69	2.10	−55	4.42	−5	3.20	−32	5.52	18
Mean soil moisture content (% vol.)									
Field A									
1999	24.98	17.85	−32	20.33	−19	17.27	−31	27.75	11
2000	27.41	18.66	−34	19.69	−31	17.72	−38	29.20	3
2001	25.68	17.75	−31	18.48	−28	13.44	−48	24.75	−4
2002	24.87	19.24	−23	20.30	−18	14.15	−43	24.61	−1
Field C									
1999	24.51	17.41	−29	24.33	−1	18.11	−26	21.94	−10
2000	24.89	19.34	−23	25.58	1	19.19	−24	23.59	−5
2001	28.21	16.40	−42	22.97	−19	14.27	−49	20.26	−28
2002	27.46	18.74	−32	24.80	−10	16.38	−40	20.89	−24

[a] Def. = model using default parameters.
[b] E_r = relative percent error (%).
[c] Cal. = calibrated model.
[d] Note that 1999 measured and predicted data were used for model calibration.

Table 6. Statistical comparison of measured and predicted surface runoff and soil moisture in field A from 2000 to 2002.

Parameter	Surface Runoff GLEAMS Def.[a]	Surface Runoff GLEAMS Cal.[b]	Surface Runoff RZWQM Def.	Surface Runoff RZWQM Cal.	Soil Moisture Content GLEAMS Def.	Soil Moisture Content GLEAMS Cal.	Soil Moisture Content RZWQM Def.	Soil Moisture Content RZWQM Cal.
Field observed data[c]	2.46	2.46	2.46	2.46	25.92	25.92	25.92	25.92
Simulated value	5.12	2.15	11.71	2.43	18.59	19.54	15.04	26.09
Prob. for mean difference (P > |*t*|)	<0.0001	0.63	<0.0001	0.92	<0.0001	<0.0001	<0.0001	0.15
Number of observations	73	73	73	73	777	777	777	777
Coefficient of determination (R^2)	0.66	0.76	0.64	0.96	0.03	0.05	0.54	0.58
Index of agreement (*d*)	0.88	0.88	0.72	0.99	0.30	0.32	0.36	0.82
Coefficient of efficiency (E)	0.50	0.69	−1.62	0.96	−7.08	−5.80	−12.17	−0.15
Relative percent error (E_r)	109	−13	377	−1	−28	−24	−42	1
Maximum error (E_{max})	0.45	0.48	0.75	0.17	16.61	15.52	16.46	11.61
Regression analysis: slope (*m*)	0.88	0.53	1.59	0.92	0.22	0.31	0.88	1.23
intercept (*b*)	0.05	0.02	0.11	0.003	12.86	11.39	−7.84	−5.82
Total score[d]	6	8	3	8	1	2	3	5

[a] Def. = default model with input parameters based on databases, soil surveys, and site−specific information.
[b] Cal. = calibrated model.
[c] Data observed or measured in the field.
[d] Total number of evaluation techniques that meet the scoring criteria in table 4.

Because of the poor performances of both GLEAMS and RZWQM with default parameters, these models were calibrated by adjusting several soil parameters and hydrologic characteristics and best fitting the simulations to the 1999 data sets (table 3). The results of site−calibrated GLEAMS and RZWQM are presented in the following sections.

Simulations with Calibrated Models

Surface Runoff and Soil Moisture Content in Field without Seepage Zones

The data in figures 2 and 3 show that site−calibrated GLEAMS and RZWQM performed relatively well in predicting surface runoff from the field without seepage zones (field A). The probability for the mean difference of 0.63 (at significance level of P = 0.05), index of agreement of 0.88, coefficient determination of 0.69, and relative percent error of −13% for the three−year evaluation period (2000 to 2002) show that GLEAMS adequately predicted surface runoff from field A and did much better than using the model with default parameters (table 6). The results also show that over the three−year evaluation period, RZWQM predicted surface runoff from field A with the probability for the mean difference of 0.92, index of agreement of 0.99, coefficient of determination of 0.96, coefficient of efficiency of 0.96, and

Table 7. Statistical comparison of measured and simulated surface runoff and soil moisture in field C from 2000 to 2002.

	Surface Runoff				Soil Moisture Content			
	GLEAMS		RZWQM		GLEAMS		RZWQM	
Parameter	Def.[a]	Cal.[b]	Def.	Cal.	Def.	Cal.	Def.	Cal.
Field observed data[c]	20.15	20.15	20.15	20.15	26.70	26.70	26.70	26.70
Simulated value	9.56	18.90	15.63	22.54	18.36	24.62	16.88	21.73
Prob. for mean difference (P>|t|)	<0.0001	0.67	0.01	0.12	<0.0001	<0.0001	<0.0001	<0.0001
Number of observations	267	267	267	267	984	984	984	984
Coefficient of determination (R^2)	0.39	0.50	0.55	0.72	0.10	0.12	0.17	0.22
Index of agreement (d)	0.75	0.79	0.85	0.92	0.38	0.51	0.38	0.52
Coefficient of efficiency (E)	0.30	0.06	0.53	0.65	−5.97	−1.03	−7.33	−2.39
Relative percent error (E_r)	−53	−6	−29	12	−31	−8	−37	−19
Maximum error (E_{max})	1.00	0.97	0.93	0.55	18.53	10.56	20.03	15.95
Regression analysis: slope (m)	0.48	0.84	0.64	0.93	0.43	0.35	0.46	0.64
intercept (b)	−0.001	0.007	0.01	0.014	6.91	15.20	4.56	4.62
Total score[d]	5	8	6	8	1	2	1	3

[a] Def. = default model with input parameters based on databases, soil surveys, and site−specific information.
[b] Cal. = calibrated model.
[c] Data observed or measured in the field.
[d] Total number of evaluation techniques that meet the scoring criteria in table 4.

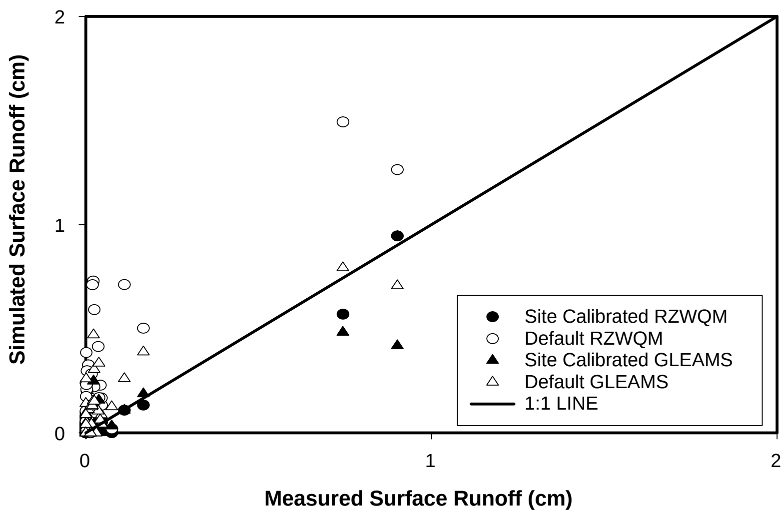

Figure 2. Comparison of daily measured and simulated surface runoff from field A from 2000 to 2002.

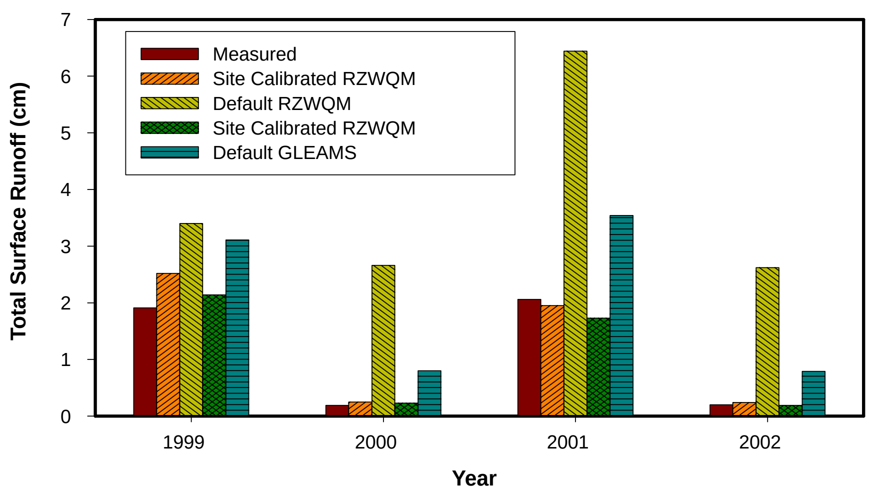

Figure 3. Total growing−season measured and simulated surface runoff from field A.

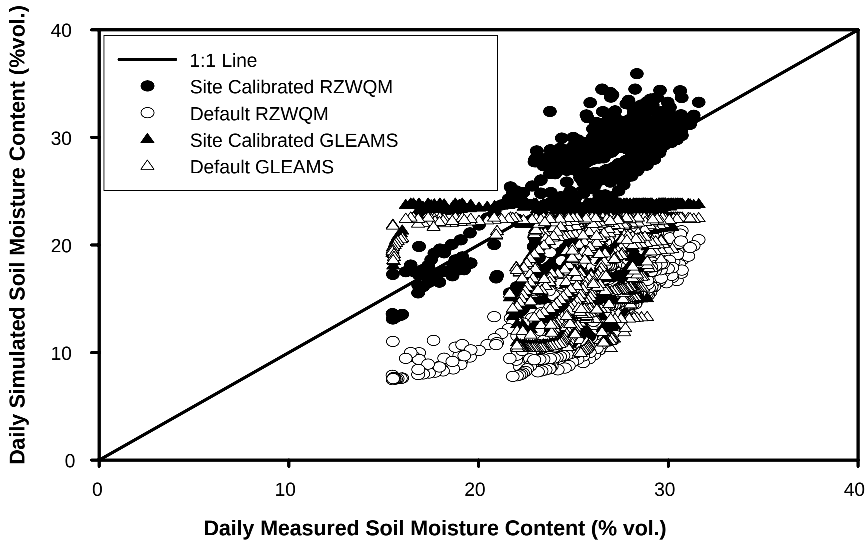

Figure 4. Comparison of average daily measured and simulated soil moisture contents in field A from 2000 to 2002.

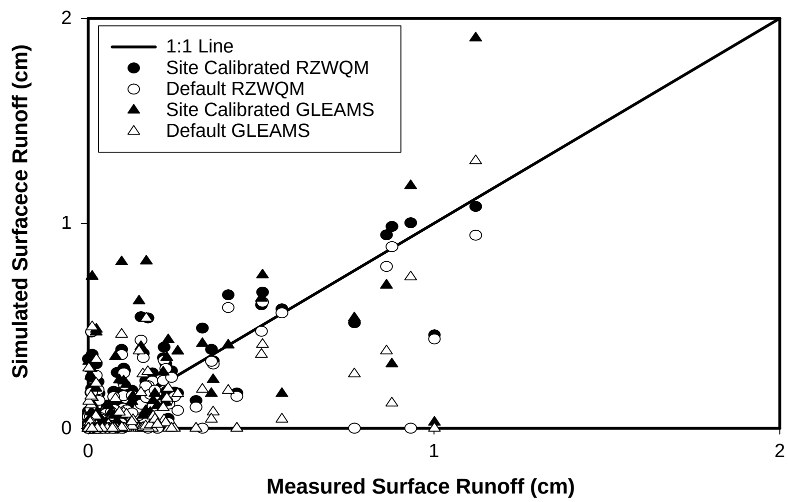

Figure 5. Comparison of daily measured and simulated surface runoff from field C from 2000 to 2002.

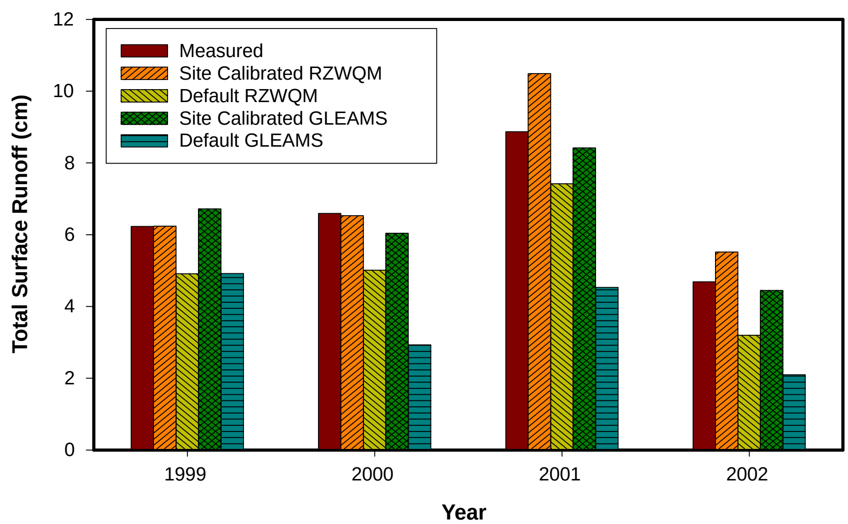

Figure 6. Total growing−season measured and simulated surface runoff from field C.

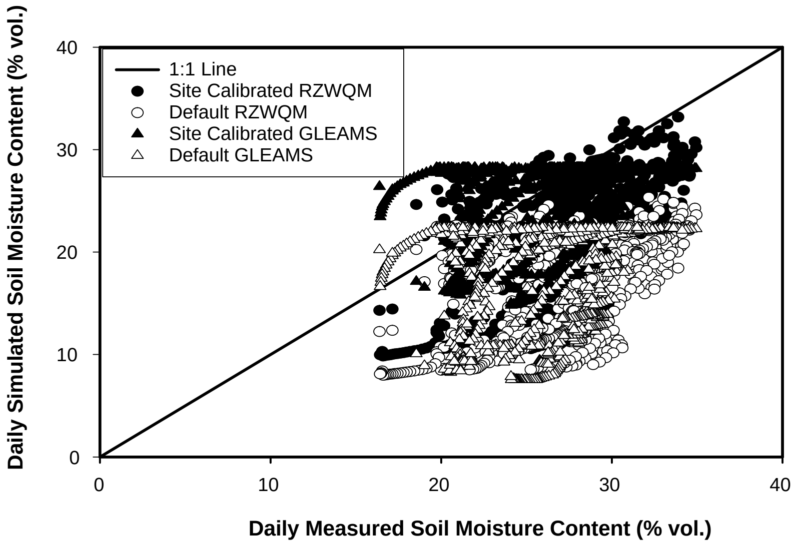

Figure 7. Comparison of average daily measured and simulated soil moisture content in field C from 2000 to 2002.

relative percent error of −1% (table 6). These results show that site−calibrated GLEAMS and RZWQM predicted surface runoff trends from field A fairly well and much better than the models with default input parameters. For example, the results in table 6 show that site−calibrated GLEAMS had higher total score (8) than GLEAMS using default input parameters (6). Similar results were also observed for RZWQM (total score of 8 for calibrated model, as opposed to 3 for default models).

The results in tables 5 and 6 and figure 4 show that GLEAMS consistently underpredicted soil moisture content in field A. The coefficient of determination, coefficient of efficiency, and regression analysis also show that GLEAMS underpredicted soil moisture content in field A (table 6). Results presented in figure 4 show that RZWQM performed relatively better than the default model in predicting soil moisture content in field A. The relative percent error (1%), probability for the man difference (0.15), coefficient of determination (0.58), index of agreement (0.82), and slope show that RZWQM performed well in predicting soil moisture content from field A (table 6). Although the index of agreement (0.32) and relative percent error (−24%) show improved performance of GLEAMS, the other statistical analysis criteria indicate that the calibrated GLEAMS, while improved over the model using default input parameters, was a poor predictor of soil moisture (e.g., a total score of 2 was the best performance for calibrated GLEAMS, table 6).

Gentle slopes (typical of agricultural areas), local depression storage, and lack of seepage zones in field A minimize the role of surface runoff in distributing and transporting the total flux of water and chemicals to neighboring systems. Although most of the water was percolating into the soil, GLEAMS could not accurately predict the amount of soil moisture in this field, even when site−calibrated with one year of intense soil moisture data. This result could be due to many factors, one being the comparison of information from different scales of observation. The model calculates integrated soil moisture content for an entire field based on several soil and landscape parameters, while the measured data were the averages of point measurements, each having a sampling radius of only 10 cm. Additionally, the lack of spatial organization in field A causes problems in modeling because most lumped models were fundamentally developed to handle and interpret fields that are homogeneous and have organized patterns of surface and subsurface hydrology (Grayson. and Blöschl, 2001). The poor performance of GLEAMS may also be due to inadequate spatial representation of measured soil properties and the fact that the structure of the model may not properly integrate and represent the spatial variability of hydrological processes at scales smaller than the one the model is representing. Similarly, lack of accurate measurement or determination of theoretical parameters, which is required by all process−level, comprehensive models like GLEAMS and RZWQM, can also result in poor performance of models (RZWQM Team, 1995).

Surface Runoff and Soil Moisture Content in Field with Seepage Zones

The data presented in figures 5 and 6 show that the amount of daily and annual surface runoff predicted by GLEAMS consistently followed the pattern of measured surface runoff from the field with seepage zones (field C). GLEAMS predicted surface runoff from field C with an index of agreement of 0.79, probability for the mean difference of 0.67, coefficient of determination of 0.50, relative percent error of −6%, slope of 0.84, intercept of 0.007, and coefficient of efficiency of −0.06 (table 7). Predicted surface runoff by RZWQM was also similar to the measured surface runoff from field C (figs. 5 and 6). The results in table 7 show that RZWQM−predicted and measured surface runoff amounts from field C were in fair agreement, showing an index of agreement of 0.92. The following statistics for the calibrated model indicate that RZWQM adequately predicted surface runoff for field C: coefficient of efficiency (0.65), relative percent error (12%), probability for the mean difference (0.12), coefficient of determination (0.72), absolute maximum error (0.55), slope (0.93), and intercept (0.014) (table 7). The results in table 7 also show that RZWQM generated more surface runoff than GLEAMS and thus better represented the processes controlling surface runoff in the field. The greater field slopes, concentration of runoff in specific drainage areas, and the presence of seepage zones in field C contributed to high measured surface runoff at the flume (Daughtry et al., 2001).

The results presented in figure 7 show that the GLEAMS−predicted soil moisture content slightly varied from the measured soil moisture content in field C. Overall statistical results in table 7 show that GLEAMS performance improved relative to GLEAMS using default parameters with relative percent error of −8%. In addition, the index of agreement (0.51) indicates that GLEAMS predicted soil moisture content in field C with increased accuracy. However, the other statistical tests, e.g., coefficient of efficiency (−1.03), absolute maximum error (10.56), and intercept (15.20), show that the model was not perfect (table 7). The results show that RZWQM predicted soil moisture content in field C with a relative percent error of −19%, index of agreement of 0.52, and slope of 0.64 (table 7). However, the coefficient of efficiency of −2.39, absolute maximum error of 15.95, coefficient of determination of 0.22, and intercept of 4.62 show that RZWQM underpredicted soil moisture content in field C (table 7). These results indicate that both site−calibrated GLEAMS and RZWQM were capable of adequately predicting surface runoff, but that soil moisture prediction is still in its infancy. The results in tables 6 and 7 show that in either field, both models had very low total scores for soil moisture content. However, RZWQM had higher total scores for surface runoff and soil moisture than GLEAMS. These data suggest that RZWQM was a superior model to GLEAMS. RZWQM may have performed better than GLEAMS due the representation of the infiltration process using the Green−Ampt infiltration equation, which incorporates initial soil moisture content for each rainfall event, thus being able to separate infiltration and runoff components slightly better than the NRCS curve number method in GLEAMS.

Implications of Model Results

This study (tables 5 through 7 and figs. 2 through 7) shows that both GLEAMS and RZWQM performed well in predicting surface runoff from the two fields. However, GLEAMS performed relatively well in field C and poorly in field A in predicting soil moisture content. RZWQM performed well in both fields, although not perfectly. Overall results show that both GLEAMS and RZWQM underpredicted soil moisture for much of the study period in both fields, as indicated by the low total scores. Although not determined, this latter deficiency may be due to our inability to scale up point measurements of soil moisture to a field scale. The poor simulation might also be due to the fact that field A did not have the same degree of hydrologic connectivity as field C (interaction of seepage zone and surface runoff water). With its seepage zones and the spatial organization of the flow regime in field C, the field as a whole is relatively well connected (Gish et al., 2001). Field A does not have the same sort of spatial organization, and this difference means it has a lower degree of hydrologic connectivity than field C.

Both models performed better on field C, indicating that both models perform better in situations with a high degree of hydrologic connectivity. Lumped parameter models by nature are designed to assume that the area modeled (in this case, a field) is completely hydrologically connected and homogenous. Therefore, it can be assumed that most lumped agricultural nonpoint−source models will perform best when simulating fields with a high degree of hydrologic connectivity, as seen in this study on field C. Our results indicate that the current structures of RZWQM and GLEAMS can represent seepage zones adequately by adjusting model parameters; thus, seepage zones appear to be less of a simulation problem for these models than might have been expected. This result should not be read to mean that seepage zone processes could not be better represented by an alternative model structure to that of GLEAMS and RZWQM. Most likely, an explicit representation of seepage zones (adding a parameter to represent seepage zone processes) would improve the performance of these two models on field C. The exact formulation of such a structure awaits further research and could be tested by seeing if such a reformulated model did a better job at simulating runoff and soil moisture in these two fields than the model formulations of GLEAMS and RZWQM.

Additionally, our results indicate that although two fields may be next to each other (like these two fields), it may not be appropriate to model them as one homogenous area because of differences in surface and subsurface hydrological properties, as our results for fields A and C have demonstrated. These results mean that with large watershed−scale models like SWAT, modelers need to delineate the field properly by using both surface and subsurface hydrologic properties. The key to understanding and developing solutions to water quality modeling problems is quantifying and incorporating surface runoff dynamics: rates, volumes, and most importantly, source areas of surface runoff within the field into the models. These model improvements will require incorporation of better field data on seepage zones and the geomorphic controls on seepage zones and methodologies to incorporate these dynamics into agricultural water quality models.

Multiple Criteria Evaluation Techniques

Correlation and correlation−based measures (e.g., R^2) (such as figs. 2, 4, 5, and 7) have been widely used to evaluate the performance of hydrologic and water quality models. However, these measures are sensitive to extreme values (outliers) and are insensitive to additive and proportional differences between model predictions and observations (Legates and McCabe, 1999). A single evaluation measure can indicate that a model is a good predictor, when in fact it is not. Because of these limitations, additional evaluation measures, as reported in this study, should be used to vigorously evaluate the performance and usefulness of water quality models. The results of this study show that by using graphical comparison, several statistical tests, and the combined scoring from the statistical tests, we received a more robust picture of model performance than if we had just used graphical comparison and a single statistical test, as is done for most water quality modeling studies.

Importance of Seepage Zones in Modeling

Field C generated more surface runoff than field A even though the climatic and soil parameters were similar. The higher surface runoff from field C was due to concentration of runoff in specific drainage areas and the presence of seepage zones. In field C, the seepage zones were close to the runoff flume, which often resulted in continued surface runoff for several hours or days after significant rainfall events. On the other hand, field A had very short durations of runoff events, typically ending shortly after precipitation stopped because it had no seepage zones. Thus, field A generated limited surface runoff, with infiltration and groundwater recharge being the dominant hydrologic processes.

Most lumped computer models were developed to handle homogeneous fields, and they do not integrate all aspects of hydrologic controls from the runoff flow perspective and much less from that of their interactions with water quality

(Grayson and Blöschl, 2001). Although seepage zone processes are not represented in GLEAMS and RZWQM, these models performed better in predicting surface runoff from the field with seepage zones than without seepage zones. Modeling of variable source area hydrology has been hindered by lack of field data documenting the controlling mechanisms and characterizing their spatial and temporal variability (Gburek and Sharpley, 1998; Gish et al., 2001). The research project at the Beltsville Research Center is a good starting point for identifying and documenting these processes for eventual incorporation and use in nonpoint−source pollution models.

Two approaches could be taken to solve the shortcomings of RZWQM and GLEAMS in simulating soil moisture conditions in these two fields. First, each model could become a fully distributed representation of a farm field. This approach seems extreme since the detailed information needed to parameterize these models is already hard to produce for most locations. Producing it on a distributed basis for the purpose of making predictions of water quality runoff seems like an unrealistic task for most applications. Second, each model could be modified to represent seepage zone processes by representing the sub−field scale variability of seepage zone processes in some manner (e.g., the method of Jolley and Wheater, 1997). By changing the structure of the models in this way, the models might allow some water that percolates downward to later mix with surface runoff in the modeled field. This approach might use information on saturated surface area to parameterize this portion of the model and thus improve model performance without greatly increasing model complexity, since more complex models do not necessarily improve model performance (Loague and Freeze, 1985). This second approach maintains the relatively simple model structures that have been preferred when simulating agricultural water quality, since we do not typically have the robust detailed data sets needed to calibrate a detailed model.

Summary and Conclusions

GLEAMS v. 3.0.1 and RZWQM98 v. 1.0.2000.830 models were used to predict the amount of surface runoff and soil moisture content in two agricultural fields: with and without seepage zones. The results of this study have particular importance in using the models to assess the impacts of various management practices on agricultural fields that have seepage zones. First, daily simulated surface runoff and soil moisture content from both default and site−calibrated GLEAMS and RZWQM were compared with measured surface runoff and soil moisture content from the two fields from 2000 to 2002. The results show that GLEAMS and RZWQM using default input parameters were not capable of predicting surface runoff and soil moisture content in fields with and without seepage zones. The poor performance of both models was probably due to poor representation of measured soil properties, or the structure of the two models may not properly represent hydrological processes occurring in the fields. Additional field data on the spatial distribution of physical soil and hydrologic properties might improve the models using default input parameters, but there is little question that field−based parameterization would need to be incorporated into the application of these water quality models to ensure optimal model performance.

Second, the results show that site−calibration of GLEAMS and RZWQM improved the performance of both models. Site−calibrated GLEAMS and RZWQM performed fairly well in simulating surface runoff from both fields. Site−calibrated GLEAMS and RZWQM did a better job of predicting soil moisture when seepage zones were present; however, the coefficients of determination were still quite low (<0.4).

Third, based on the different model evaluation techniques used in this study, the performance of both GLEAMS and RZWQM was fair, but RZWQM performed better than GLEAMS in both fields, possibly because it uses the Green−Ampt infiltration model as opposed to the NRCS curve number method used in GLEAMS. Therefore, RZWQM is better suited to assess the effects of seepage zones on soil moisture and surface runoff from agricultural fields than GLEAMS. This is not to say that RZWQM is a perfect model for representing seepage zone dynamics, but merely the better of the two models studied here.

Acknowledgements

The University of California and the National Science Foundation (EAR−0094312) and in part, the CSREES−NRI Grant No. 2001−01091 “Quantification and Evaluation of Subsurface Water Dynamics for Determining Water and Chemical Fluxes on Adjacent Fields” supported this research.

Reference

Ahuja, L. R., D. G. DeCoursey, B. B. Barnes, and K. W. Rojas. 1993. Characteristics of macropore transport studied with the USDA−ARS root zone water quality model. *Trans. ASAE* 36(2): 369−380.

Ahuja, L. R., K. W. Rojas, J. D. Hanson, M. J. Shaffer, and L. Ma, eds. 2000. *Root Zone Water Quality Model: Modeling Management Effects on Water Quality and Crop Production*. Highlands Ranch, Colo.: Water Resources Publications.

Bakhsh, A., R. S. Kanwar, D. B. Jaynes, T. S. Colvin, and L. R. Ahuja. 2000. Prediction of NO_3−N losses with subsurface drainage water from manured and UAN−fertilized plots using GLEAMS. *Trans. ASAE* 43(1): 69−77.

Buchleiter, G. W., H. J. Farahani, and L. R. Ahuja. 1995. Model evaluation of groundwater contamination under center−pivot irrigated corn in eastern Colorado. In *Int. Symp. on Water Quality Modeling*, 41−50. St. Joseph, Mich.: ASAE.

Chinkuyu, A. J., and R. S. Kanwar. 2001. Predicting soil nitrate−nitrogen losses from incorporated poultry manure using GLEAMS model. *Trans. ASAE*44(6): 1643−1650.

Daughtry, C. S. T., T. J. Gish, W. P. Dulaney, C. L. Walthall, K. J. S. Kung, G. W. McCarty, J. T. Angier, and P. Buss. 2001. Surface and subsurface nitrate flow pathways on a watershed scale. In *Optimizing Nitrogen Management in Food and Energy Production and Environmental Protection: Proc. 2nd International Nitrogen Conference on Science and Policy. The Scientific World*1(S2): 155−162.

Gburek, W. J., and A. N. Sharpley. 1998. Hydrologic controls on phosphorus loss from upland agricultural watersheds. *J. Environ. Quality* 27(2): 267−277.

Gburek, W. J., C. C. Drungil, M. S. Srinivasan, B. A. Needelman, and D. E. Woodward. 2002. Variable−source−area controls on phosphorus transport: Bridging the gap between research and design. *J. Soil and Water Conservation* 57(6): 534−544.

Gish, T. J., W. P. Dulaney, C. S. T. Daughtry, and K.−J. S. Kung. 2001. Influence of preferential flow on surface runoff fluxes. In *Proc. 2nd International Preferential Flow Symposium*, 205−209. St. Joseph, Mich.: ASAE.

Grayson, R., and G. Blöschl, eds. 2001. Spatial modelling of catchment dynamics. In *Spatial Patterns in Catchment Hydrology: Observations and Modeling*, 51−81. Cambridge, U.K.: Cambridge University Press.

Haan, C. T., B. Allred, D. E. Storm, G. Sabbagh, and S. Prabhu. 1993. Evaluation of hydrologic/water quality models: A statistical procedure. ASAE Paper No. 932505. St. Joseph, Mich.: ASAE.

Hanson, J. D., K. W. Rojas, and M. J. Shaffer. 1999. Calibrating the root zone water quality model. *Agronomy J.* 91(2): 171−177.

Knisel, W. G., ed. 1993. *GLEAMS: Groundwater Loading Effects of Agricultural Management Systems*. Ver. 2.10. UGA−CPES−BAED Publ. No. 5. Tifton, Ga.: University of Georgia, Coastal Plain Experiment Station.

Jolley, T. J., and H. S. Wheater. 1997. An investigation into the effect of spatial scale on the performance of a one−dimensional water balance model. *Hydrol. Processes* 11(15): 1927−1944.

Leavesley, G. H., R. W. Lichty, B. M. Troutman, and L. G. Saindon. 1983. *Precipitation−Runoff Modeling System: User's Manual*. Water Resources Invest. Report No. 83−4238. Denver, Colo.: USGS.

Legates, D. R., and G. J. McCabe, Jr. 1999. Evaluating the use of "goodness−of−fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* 35(1): 233−241.

Leonard, R. A., W. G. Knisel, and D. A. Still. 1987. GLEAMS: Groundwater loading effects of agricultural management systems. *Trans. ASAE* 30(5): 1403−1418.

Loague, K. M., and R. A. Freeze. 1985. A comparison of rainfall−runoff modeling techniques on small upland catchments. *Water Resources Research* 21(2): 229−248.

Meisinger, J. J., V. A. Bandel, J. S. Angle, B. E. O'Keefe, and C. M. Reynolds. 1992. Pre−sidedress soil nitrate test evaluation in Maryland. *SSSA J.* 56: 1527−1532.

RZWQM Team. 1995. *RZWQM User's Manual*. GPSR Technical Report No. 5. Ft. Collins, Colo.: USDA−ARS Great Plains Systems Research.

Truman, C. C., R. A. Leonard, and F. M. Davis. 1998. GLEAMS−TC: A two−compartment model for simulating temperature and soil water content effects on pesticide losses. *Soil Science* 163(5): 362−373.

Walter, M. T., M. F. Walter, E. S. Brooks, T. S. Steenhuis, J. Boll, and K. Weiler. 2000. Hydrologically sensitive areas: Variable source area hydrology implications for water quality risk assessment. *J. Soil and Water Conservation* 55(3): 277−284.

Wilcox, B. P., W. J. Rawls, D. J. Brakensiek, and J. R. Wight. 1990. Predicting runoff from rangeland catchments: A comparison of two models. *Water Resources Research* 26: 2401−2410.