An approach to representing spatial variability when evaluating model uncertainty in process-based watershed models

M. TISCARENO-LOPEZ & R. CLAVERAN
Secretaria de Agricultura, Ganaderia y Desarrollo Rural, Instituto Nacional de Investigaciones Forestales y Agropecuarias, Centro Nacional de Produccion Sostenible, Teniente Aleman 294, Morelia, Michoacan 58260, Mexico

M. A. WELTZ & M. R. KIDWELL
US Department of Agriculture, Agricultural Research Service, Southwest Watershed Research Center, 2000 E. Allen Road, Tucson, Arizona 85719, USA

Abstract Recent erosion prediction technology has overcome limitations of empirical models to estimate spatial variability of soil erosion in the watershed by applying process-based models. These new models, however, remain gross approximations of reality due to the complexity of natural systems. During model development, therefore, quantification of model uncertainty due to inaccuracies in representing natural spatial variability of the system is important. In this study, prediction uncertainties were obtained by Monte Carlo simulation with correlated deviates generation to produce distributions of uncertainties of model responses. A multiple normal distribution was assembled to generate correlated soil characteristics for model input to preserve the spatial variability of the watershed. Prediction intervals and the contribution of model and parameter error to total uncertainty were estimated for the WEPP model.

INTRODUCTION

The Water Erosion Prediction Project (WEPP) model of the US Department of Agriculture (USDA) is a major advance in erosion prediction technology to estimate the effects of land use practices on soil conservation for small agricultural watersheds. WEPP overcomes limitations of empirical models, such as USLE, to account for the effects of temporal and spatial variability of the system in estimating erosion and sedimentation in cropland and rangeland watersheds. Based on modern concepts of stochastic weather generation, hydrology, soil physics, plant science, hydraulics, and erosion mechanics (Lane & Nearing, 1989), WEPP will be used as a major planning tool for erosion estimation by the US Soil Conservation Service.

Unfortunately, WEPP remains an abstraction of reality and uncertainty arises in model components because it only approximates the represented processes. Under field applications, model complexity and natural variability of the system frequently induce uncertainty in predictions. Thus, the identification of model uncertainty is a major task during model testing. Uncertainties in model response depend on the range and spatial variability of parameters found in the field. This study presents an approach that (a) accounts for effects of spatial variability of a system when evaluating complex models, (b) quantifies the resulting uncertainty in predictions owing to parameter variability and structural errors, and (c) estimates prediction intervals for the simulated responses. These analyses were performed to evaluate the hillslope version (93.13) of the WEPP model using information from a small semiarid rangeland watershed (Mark Nearing, USDA, personal communication, 1993).

MODEL DESCRIPTION

In WEPP, characteristics of hillslope runoff from rainfall events provide the information needed to model erosion. Infiltration is computed using the Green & Ampt equation (Green & Ampt, 1911) as modified by Chu (1978) for unsteady rainfall events, in which the soil surface alternates between unponded and ponded. Infiltration parameters are updated daily to account for temporal variations of soil moisture. When rainfall intensity exceeds infiltration rate, the part of rainfall that does not infiltrate or become depression storage flows downslope. Overland flow is routed using either kinematic wave equations (Liggett & Woolhiser, 1967) or regression equations that approximate the kinematic wave.

Processes of detachment and deposition, shear stress, rill and inter-rill flow, and sediment transport capacity by flowing water, as described by Foster & Meyer (1972), serve as prototypes of the WEPP model. The steady-state continuity equation for sediment transport is:

$$\frac{dG}{dx} = D_i + D_r \tag{1}$$

with:

$$D_i = C_i\, K_i\, I_e^2\, S_f\, G_e \left[\frac{R_s}{w}\right] \tag{2}$$

and

$$D_r = C_r K_r (\tau - \tau_{cr})(1 - \frac{G}{T_c}) \tag{3}$$

where G is sediment load (kg m^{-1} s^{-1}), x is distance along the slope (m), D_i is delivery rate of sediment from inter-rill areas (kg m^{-2} s^{-1}), D_r is rill detachment rate (kg m^{-2} s^{-1}), C_i is an inter-rill canopy cover parameter (dimensionless), K_i is an inter-rill soil erodibility parameter (kg m^{-4} s^{-1}), I is rainfall intensity (m s^{-1}), S_f is an inter-rill slope adjustment factor (dimensionless), G_e is effective ground cover on inter-rill erosion (dimensionless), R_s is spacing of rills (m), w is rill width (m), C_r is a rill cover parameter (dimensionless), K_r is a rill soil erodibility parameter (s m^{-1}), τ is average cross-section shear stress (Pa), τ_{cr} is critical shear stress required for detachment (Pa), and T_c is transport capacity of flow (kg m^{-1} s^{-1}).

Equation (2) describes delivery to rills of soil detached by raindrop impact in inter-rill areas and equation (3) describes soil particle detachment in rills due to concentrated flow. Substituting equations (2) and (3) into (1), when the transport capacity is greater than the sediment load, the sediment continuity equation is:

$$\frac{\mathrm{d}G}{\mathrm{d}x} = C_i K_i I^2 S_f G_e \left[\frac{R_s}{w}\right] + C_r K_r (\tau - \tau_{cr})(1 - \frac{G}{T_c}) \tag{4}$$

The term $(1 - G/T_c)$ is a feedback term for rill detachment assuming that detachment in rills depends on the sediment load in the flow relative to the capacity of the flow to transport sediment. When the sediment load is greater than its transport capacity, net deposition occurs and the continuity equation is:

$$\frac{\mathrm{d}G}{\mathrm{d}x} = C_i K_i I^2 S_f G_e \left[\frac{R_s}{w}\right] + (\beta \frac{V_f}{q})(T_c - G) \tag{5}$$

where β is a deposition parameter equal to 0.5 (dimensionless), V_f is effective particle fall velocity (m s^{-1}), and q is discharge per unit width (m^2 s^{-1}). Four hydrologic variables drive the equations: peak runoff, derivation of effective runoff, intensity of effective rainfall, and effective rainfall excess duration.

A feature of the WEPP model is its capability of calculating and updating hydrologic and erosion parameters based on readily available soil and vegetation characteristics. Erosion parameters are approximated by regression equations developed from rainfall simulation experiments (Alberts *et al.*, 1989; Laflen *et al.*, 1991). With this capability, the model accounts for spatial variability and can be applied to ungaged watersheds in which null or scarce information about model parameters is available.

STUDY SITE

An analysis was performed for Kendall 2 (K2), a rangeland watershed of 1.86 ha in the Walnut Gulch Experimental Watershed near Tombstone, Arizona, operated by the Southwest Watershed Research Center, USDA Agricultural Research Service, in Tucson, Arizona. Vegetation is dominated by warm season short grasses with average canopy cover of 40%. Walnut Gulch is representative of 60 million ha of brush and grass rangeland in the semiarid southwestern United States and northern Mexico. It is a transitional zone between the Chihuahuan and Sonoran deserts (Renard *et al.*, 1993). Thunderstorms of extreme spatial variability, limited areal extent, and short duration occur from July through mid-September. These thunderstorms dominate the rainfall-runoff processes and cause nearly all of the soil erosion.

ASSESSMENT OF UNCERTAINTIES

Uncertainties in WEPP result from inaccurate estimates of erosion parameters by the regression equations and difficulty in representing natural processes. A Monte Carlo simulation with correlated random deviates generation was used to quantify uncertainty in predictions caused by errors in model parameters and structure. The errors were assessed by analyzing variability around observed and predicted variables based on prediction variance and model bias squared. Prediction variance is variance in model response due to parameter variability, and was obtained by varying model parameters in a Monte Carlo framework. Bias is the difference between expected behavior of the model when the parameters are uncertain and the mean of predictions. Total uncertainty was assessed as Mean Square Error (MSE):

$$\mathrm{MSE} = \mathrm{VARIANCE} + \mathrm{BIAS}^2 \tag{6}$$

$$\mathrm{BIAS}^2 = (y - \bar{y})^2 = (\text{observed} - \text{mean of predictions})^2 \tag{7}$$

$$\begin{aligned}\mathrm{VARIANCE} &= \Sigma(\hat{y} - \bar{y})^2 n^{-1} \\ &= \Sigma(\text{predicted} - \text{mean of predictions})^2 n^{-1}\end{aligned} \tag{8}$$

$$\mathrm{MSE} = \Sigma(y - \hat{y})^2 n^{-1} = \Sigma(\text{observed} - \text{predicted})^2 n^{-1} \tag{9}$$

where y = observed value, $\bar{y}$ = mean of prediction, $\hat{y}$ = predicted value, and n is the number of observations. Based on the definitions above, total model uncertainty is equal to uncertainty due to parameters plus uncertainty due to model structure.

Many responses were developed with Monte Carlo simulation. To preserve parameter correlation of spatial variability of soil characteristics at K2, random deviates were generated from a multiple normal distribution (MND) based on 25 soil samples from K2. Error sources and prediction intervals were identified by comparing model responses with observed values of runoff volume, peak runoff, and sediment yield. A climate input file was constructed using records from a meteorological station in the watershed. The approach is summarized as follows: (a) soil and vegetation were sampled to yield basic statistics; (b) the hillslope version of WEPP was modified to accept prior probability distribution functions of correlated variables; (c) a 14-year continuous simulation was performed using model inputs from corresponding probability distributions and the climate file (random deviates of uncorrelated variables were generated from univariate normal distributions); (d) model responses for 21 documented rainfall-runoff events occurring within the 14 years were saved at the end of simulation; (e) steps (c) and (d) were repeated for 1000 simulations; (f) sources of error were quantified for runoff volume, peak runoff, and sediment yield model response; and (g) using probability distributions of model responses, prediction intervals were obtained by rejecting the upper and lower 5% of simulations.

RESULTS

The first assessment of uncertainties was to identify behavior errors (variations and trends) caused by parameters and model structure. If errors tended to diminish as simulated years increased, it was considered necessary to identify a minimum time within a continuous simulation for which error is small. Figure 1 shows variations of total and parameter error for model response resulting from the 21 chronologically-ordered rainfall-runoff events. Model error is the difference between total error and

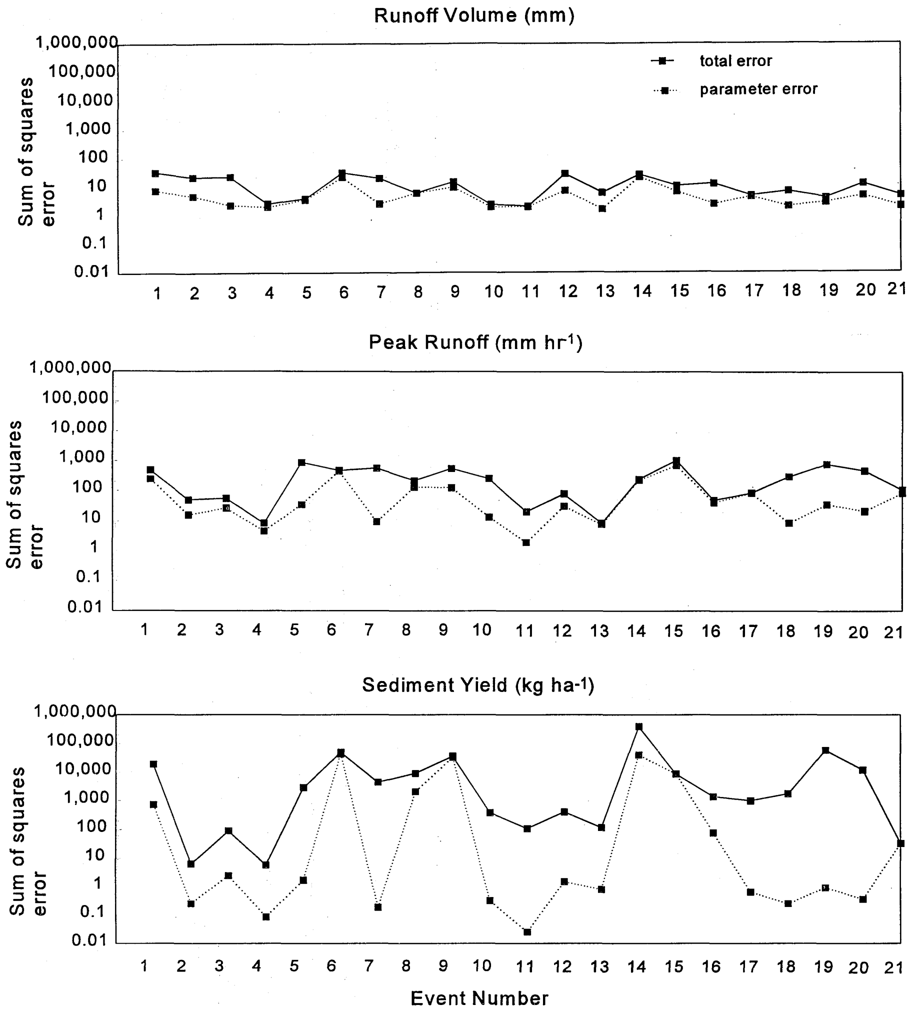

Fig. 1 Total and parameter error by event for runoff volume, peak runoff, and sediment yield.

parameter error. These graphs suggest that errors are uniform with time of simulation and that there is not a minimum number of years to diminish prediction error. Error propagation, however, was detected for some components of the WEPP model. Errors were largest in the model components with the highest levels of aggregation, with small error associated with runoff volume, intermediate error with peak runoff, and the largest error with sediment yield.

To identify problems with heteroscedasticity (i.e. larger events have larger associated prediction error), error for each rainfall event was plotted for all three response variables (Fig. 2). The difference between bars represents the contribution of model error to total model uncertainty. Error in runoff prediction is uniformly distributed for all rainfall events ranging from 9.6 to 36.6 mm. A tendency to larger errors was observed for peak runoff calculations. Model error in peak runoff calculations was larger for small rainfall events than for large events. Finally, largest error was observed for sediment yield calculations. All three error types tend to increase toward large rainfall events. Parameter error, however, was small for small rainfall events whereas model error was large. This revealed structural problems in erosion calculations.

Figure 3 shows prediction intervals of model response. The top graph shows each rainfall event. The runoff volume graph indicates that under continuous simulation, the WEPP model makes acceptable predictions because most observations were within the 90% prediction intervals. Empty circles represent model predictions when the model was run with calibrated parameters. Problems in estimating peak runoff are evident; only 13 of the 21 observed events were within the range of the prediction. Prediction intervals for sediment yield could not be delineated due to large prediction error.

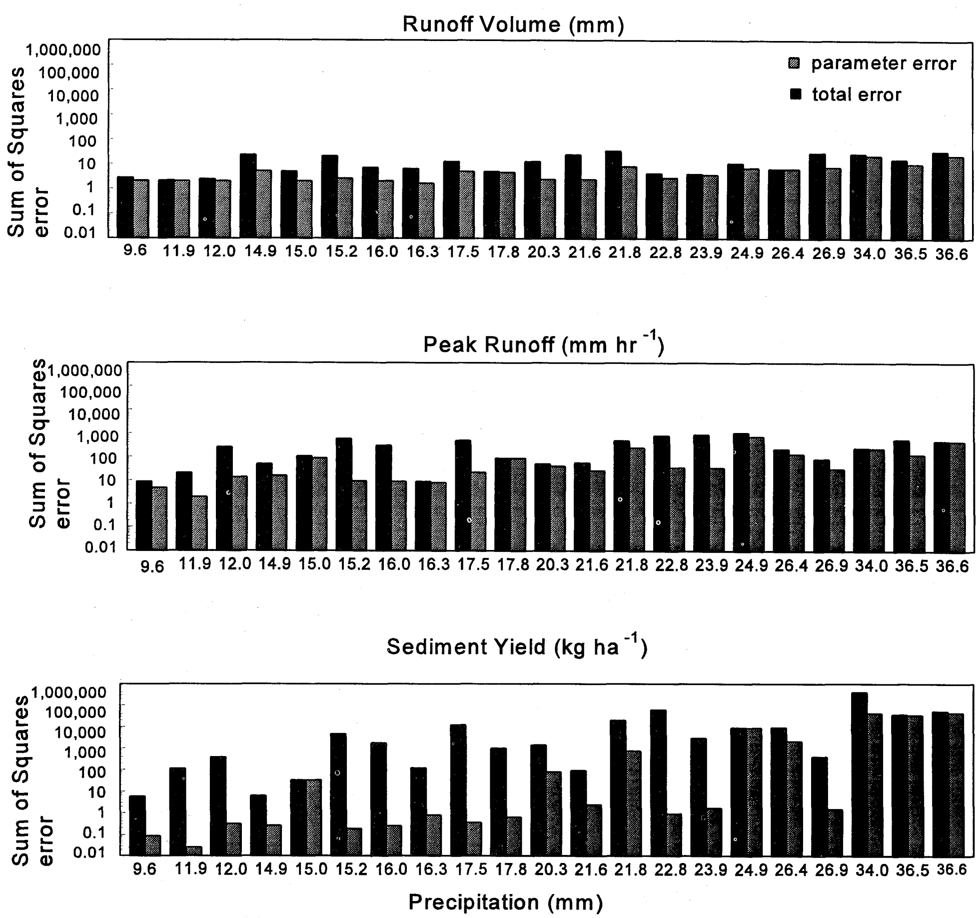

Fig. 2 Error by magnitude of precipitation event for runoff volume, peak runoff, and sediment yield model response.

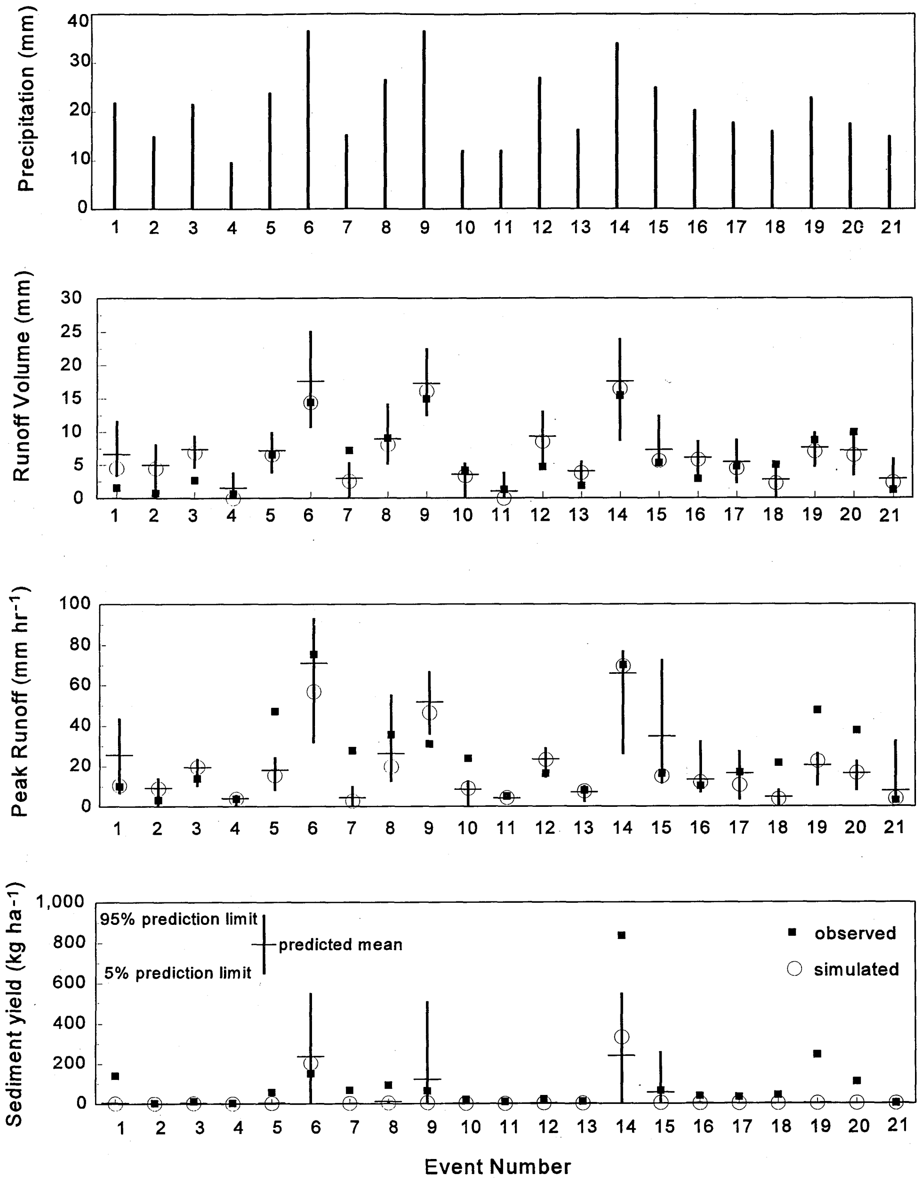

Fig. 3 Prediction intervals for runoff volume, peak runoff, and sediment yield.

CONCLUSIONS

Based on these analyses it is concluded that Monte Carlo simulation with correlated parameter deviates generation is a good approach for identifying uncertainty in process-based models during model development. For the WEPP model (93.13), three major conclusions are drawn:

(a) Error is uniform during continuous simulations, never changing with time of simulation and suggesting that there is not a minimum period required to diminish prediction error.
(b) Problems of lack of homoscedasticity are observed with the largest errors and rainfall events, which is evident for components of higher levels of aggregation such as peak runoff and sediment yield. The large contribution of model error to total error in peak runoff and sediment yield predictions reveal structural problems with the model.
(c) Prediction intervals of runoff volume show that the WEPP model provides acceptable response in estimating infiltration variables. Because of large error in estimating sediment yield, most sediment yield observations are outside the prediction intervals. This difficulty is attributed to under-prediction of inter-rill erosion.

REFERENCES

Alberts, E. E., Laflen, J. M., Rawls, W. J., Simanton, J. R. & Nearing, M. A. (1989) Soil component. In: *USDA Water Erosion Prediction Project: Hillslope Profile Model Documentation* (ed. by L. J. Lane & M. A. Nearing), 6.1-6.15. USDA-ARS National Soil Erosion Research Laboratory Report no. 2, West Lafayette, Indiana.

Chu, S. T. (1978) Infiltration during steady and unsteady rain. *Wat. Resour. Res.* **14**, 461-466.

Green, W. H. & Ampt, G. A. (1911) Studies in soil physics. I. The flow of air and water through soils. *J. Agric. Sci.* **4**, 1-24.

Foster, G. R. & Meyer, L. D. (1972) Transport of soil particles by shallow flow. *Trans. Am. Soc. Agric. Engrs* **15**, 99-102.

Laflen, J. M., Elliot, W. J., Simanton, J. R., Holzhey, C. S. & Kohl, K. D. (1991) WEPP soil erodibility experiments for rangeland and cropland soils. *J. Soil Wat. Conserv.* **46**, 39-44.

Lane L. J. & Nearing, M. A. (editors) (1989) *USDA Water Erosion Prediction Project: Hillslope Profile Model Documentation*. USDA-ARS National Soil Erosion Research Laboratory Report no. 2, West Lafayette, Indiana.

Liggett, J. A. & Woolhiser, D. A. (1967) The use of the shallow water equations in runoff computation. In: *Proc. Third Annual American Water Resources Conf.* (San Francisco), 117-126. Am. Wat. Resourc. Ass.

Renard, K. G., Lane, L. J., Simanton, J. R., Emmerich, W. E., Stone, J. J., Weltz, M. A., Goodrich, D. C. & Yakowitz, D. S. (1993) Agricultural impacts in an arid environment: Walnut Gulch Studies. In: *Proc. Second US/CIS Joint Conf. on Environmental Hydrology and Hydrogeology* (Arlington, Virginia, May 1993).