J. For. 114(1):17–26
http://dx.doi.org/10.5849/jof.14-101

soils & hydrology

Can the Water Erosion Prediction Project Model Be Used to Estimate Best Management Practice Effectiveness from Forest Roads?

Kristopher R. Brown, Kevin J. McGuire, W. Cully Hession, and W. Michael Aust

The Water Erosion Prediction Project (WEPP) was used to predict event-based sediment yield and runoff for rainfall experiments on six stream-crossing approaches with different intensities of best management practice (BMP) implementation (i.e., different proportions of gravel on the road surface). WEPP was calibrated for three different BMP intensities at each site using a Markov chain Monte Carlo approach to explore parameter uncertainty and prediction performance. WEPP predictions of sediment yield showed clear differences among the different road surface treatments, but prediction intervals (or the range of possible simulation results) were wide, reflecting substantial parameter and prediction uncertainty. The posterior distribution analysis for rill erodibility, interrill erodibility, and critical shear indicated that we cannot recommend parameter ranges specific to different surface treatments. Results suggest that the utility of WEPP for estimating BMP effectiveness is limited to predicting relative differences in sediment yield among vastly different surface treatments (e.g., native surfaced versus completely graveled roads). Sediment predictions from models should always include information regarding the range of possible outcomes, given the many sources of uncertainty.

Keywords: forest roads, best management practice effectiveness, erosion modeling, uncertainty, watershed management

Forest roads at stream crossings are a major source of sediment delivery to streams (Lane and Sheridan 2002, Harris et al. 2008, Anderson and Lockaby 2011a). The management of channelized runoff from roads has been the focus of recent legislative debates in the United States regarding the protection of aquatic ecosystems in forests (Boston 2012). For the past 40 years, forestry best management practices (BMPs) have been used to manage runoff and sediment delivery from roads, but some environmental organizations have sought legislation to achieve the goal of water quality protection with National Pollution Discharge Elimination (NPDES) permits (Boston 2012). Currently, stormwater runoff (and sediment) from forest roads is managed as a nonpoint source pollutant in the United States. However, the potential shift to NPDES permits prompted the US Environmental Protection Agency to request that state forestry organizations evaluate the effectiveness of existing BMPs for reducing sediment delivery from major sources (i.e., roads and stream crossings) and provide guidance for enhanced BMPs (Jackson 2014, Loehle et al. 2014, MacDonald and Coe 2014).

Field studies provide valuable information about the effectiveness of different intensities of BMP implementations to reduce erosion and sediment delivery (Appelboom et al. 2002, Turton et al. 2009, Anderson

Received September 1, 2014; accepted January 6, 2015; published online February 12, 2015.

Affiliations: *Kristopher R. Brown (krisrb3@vt.edu), Virginia Tech, Forest Resources and Environmental Conservation, Blacksburg, VA. Kevin J. McGuire (kevin.mcguire@vt.edu), Virginia Tech. W. Cully Hession (chession@vt.edu), Virginia Tech. W. Michael Aust (waust@vt.edu), Virginia Tech.*

Acknowledgments: *Funding and support for this research were provided by the Department of Forest Resources and Environmental Conservation at Virginia Tech and the Virginia Water Resources Research Center. We thank Dr. Jasper Vrugt for providing access to DREAM_(ZS) software. Paolo Benettin assisted with computer programming in Matlab to run DREAM_(ZS) and the WEPP model. We also thank two anonymous reviewers whose comments improved the quality of this article.*

This article uses metric units; the applicable conversion factors are: centimeters (cm): 1 cm = 0.39 in.; meters (m): 1 m = 3.3 ft; millimeters (mm): 1 mm = 0.039 in.; kilograms (kg): 1 kg = 2.2 lb; grams (g) 1 g = 0.035 oz.

and Lockaby 2011b). "Intensity" is used throughout this article in the context of BMP implementations to describe the level of effort (e.g., time and money) involved in planning or practices to reduce soil erosion and sediment delivery. However, field experiments and monitoring of BMPs on roads is often an impractical option for managers because of cost, time, and site-specific road conditions (i.e., different climates, soils, slopes, road types, and traffic characteristics). This has led to increased interest in using models for evaluating BMP effectiveness. In particular, the Water Erosion Prediction Project (WEPP) model has been used extensively as a tool for this purpose (Sawyers et al. 2012, Wade et al. 2012, Elliot 2013) and is currently recommended by many forest management organizations.

WEPP is physically based and was designed to incorporate field observations and site-level information for predicting sediment yield and runoff. Previous studies indicate that WEPP can be a useful tool for estimating soil erosion from forest roads, where overland flow is the dominant hydrologic process (Laflen et al. 2004, Croke and Nethery 2006, Fu et al. 2010). However, model performance has not been evaluated for forests roads at stream crossings for a wide range of stream-crossing approach characteristics, BMP implementations, and rainfall conditions.

In addition, methodologies for evaluating soil erosion model performance have typically included model calibration and evaluation procedures without explicitly accounting for sources of model prediction uncertainty (Brazier et al. 2000, Beven 2008). Commonly, one or more objective functions (e.g., sum of squared errors and Nash-Sutcliffe efficiency) are used to identify the most acceptable model based on prediction performance for runoff and sediment yield. However, it is possible that multiple models, with unique combinations of parameter values, can generate equally acceptable model predictions. In the case of physically based models with many parameters, it can be very difficult or impossible to identify the most acceptable model due to complex interactions among the model parameters (Beven 2008). Further, it is an uncommon occurrence for the modeler to know the appropriate a priori values of all parameters used in model calibration because of the spatial and temporal heterogeneity of site physical characteristics, which translates to variability in measurements of runoff and sediment yield. Therefore, it appears prudent to embrace the concept of model equifinality, in which a number of different models (i.e., unique sets of parameter values) can produce satisfactory predictions and identify distinct ranges of model parameter values that are associated with acceptable model runs (Beven 2008, Ascough et al. 2013).

In light of the different types of uncertainty (Hession and Storm 2000) associated with soil erosion predictions (e.g., measurement error, model parameterization, and model structure), as well as the challenges associated with defining sediment criteria to maintain or improve aquatic habitat (Ice 2011), model utility need not be defined solely by prediction accuracy. However, useful models should facilitate the identification of problem road segments for water quality protection, or better, allow us to distinguish different treatments that represent increasing intensities of BMP implementation according to their respective soil erosion or sediment delivery rates.

The objectives of the study were to determine the overall prediction performance of WEPP and its ability to distinguish between different BMP intensities. This study focused on two research questions in the Piedmont physiographic region of southwestern Virginia, USA: How well does WEPP predict event-based runoff and sediment yield from forest roads at stream crossings? and Can distinct ranges of parameter values be identified in association with acceptable model runs and different road surface treatments?

WEPP was used to predict event-based sediment yield and runoff from rainfall simulation experiments on six stream-crossing approaches having different intensities of BMP implementation (i.e., different proportions of gravel on the road surface above the stream crossing). WEPP was calibrated for each of these stream-crossing approaches for three different BMP intensities using a Markov chain Monte Carlo (MCMC) approach to explore parameter identifiability, and prediction performance and uncertainty.

Materials and Methods

Study Area

Rainfall simulation experiments were performed on a reopened forest road at the Reynolds Homestead Forest Resources Research Center (RHFRRC), located in Critz, Virginia (Patrick County), USA (Figure 1) to measure event-based surface runoff and sediment yield associated with successive increases in gravel cover on stream-crossing approaches (Brown et al. 2014). Site topography is characterized by rolling hills, with side slopes generally ranging from 8 to 25% and a mean elevation of approximately 335 m above mean sea level (Natural Resources Conservation Service [NRCS] 2013). The mean annual rainfall is 1,250 mm, with a mean snow contribution of 270 mm to the total precipitation. The mean air temperature ranges from a low of −1.8° C in January to a high of 29.7° C in July (Sawyers et al. 2012). The predominant soil series is Fairview sandy clay loam (fine, kaolinitic, mesic typic Kanhapludults). The soil parent material is residuum from mica schist and mica gneiss. There is a severe erosion hazard rating for forest roads and trails at RH-

Management and Policy Implications

The complexity of data requirements for many physically based erosion models makes them best suited for academicians and state and federal agencies that have the resources to couple field monitoring with evaluations of model performance. Forestry practitioners can reduce sediment delivery from major sources (e.g., roads, trails, and associated stream crossings) through a careful emphasis on preharvest planning and the use of erosion models with readily attainable parameter values (e.g., the Universal Soil Loss Equation modified for forestland) to estimate the sediment-reduction efficacy of site-specific BMP implementations. Future research and extension programs should seek to improve road planning technology to reduce forest road length within a given watershed, minimize stream crossings, maintain gentle road gradients, and avoid locations where it is difficult to shed water from the road surface. In this way, water quality protection is not overly dependent on postconstruction BMP implementations to correct road deficiencies that resulted from poor planning. The field component of this study showed that completely graveled forest roads at stream crossings can reduce sediment delivery to streams. In light of potential policy shifts for the forest industry (i.e., NPDES permits for roads), continued research is needed to document the cost and effectiveness of site-specific BMP implementations to reduce sediment delivery.

FRRC (NRCS 2013), which is due to the combination of moderate slopes and highly erodible soils. This underscores the importance of controlling road grade, water, and surface cover to reduce erosion and sediment delivery.

Field Methods

Before road reopening, a Sokkia model SET-520 total station was used to measure the length of the stream-crossing approach study plots, as well as the approach slope and mean width of the running surface. Length was defined as the distance between the nearest water control structure (i.e., water bar and turnout) and the stream (Figure 2). In late July 2011, six stream-crossing approaches were reopened by bulldozer blading, creating initial conditions of approximately 100% bare soil on the approach running surfaces. In October 2011, Kadak used double-ring infiltrometers to estimate the infiltration capacities of the reopened stream-crossing approaches (M. Kadak, undergraduate student researcher, unpubl. data, Feb. 11, 2012). The infiltration capacities ranged from 0.6 to 7.2 mm hour^{-1}. Bulk density samples (n = 4–7 per site) were obtained from the running surface via the soil extraction method (Soil Science Society of America 1986).

Rainfall simulation experiments were conducted for a succession of gravel surfacing treatments that represented increasing intensities of BMP implementations on the stream-crossing approaches (Brown et al. 2014) (Figure 3). All rainfall experiments were conducted between February and August 2012. The unsurfaced approaches were trafficked with a bulldozer immediately before the first series of rainfall experiments to mimic newly disturbed conditions associated with road reopening. After this treatment ("no gravel"), the stream-crossing approaches had 10–19% surface cover, which consisted of residual leaf litter and other debris. After the no gravel treatment rainfall experiments, a dump truck was used to tailgate spread a mixture of size 3, 5, and 7 (ranging from 5.1- to 1.9-cm diameter) granite gravel beginning at the lower plot boundary (Figure 2) and continuing uphill for a distance of 9.8 m ("low gravel" treatment). The mean gravel depth was approximately 0.08 m, and the width of gravel application extended across the width of the road between the outer edges of the running surfaces, which averaged 2.8 m. Gravel was not washed before application to the stream-

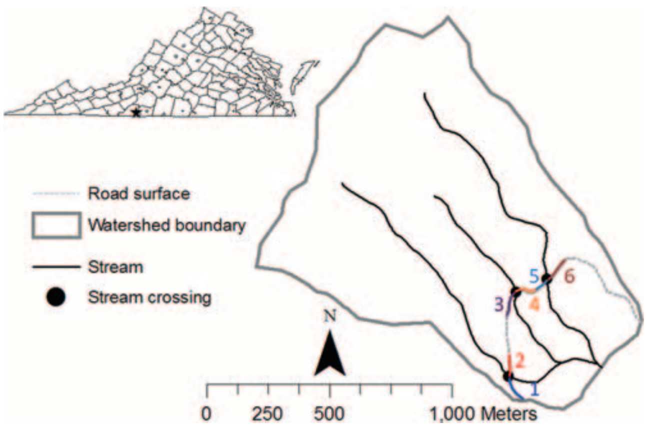

Figure 1. Location map adapted from Brown et al. (2014) of the RHFRRC in Critz, Virginia (Patrick County), USA, and a schematic diagram showing the road location within the second-order watershed that contains three unimproved ford stream crossings. Stream-crossing approaches are labeled 1–6.

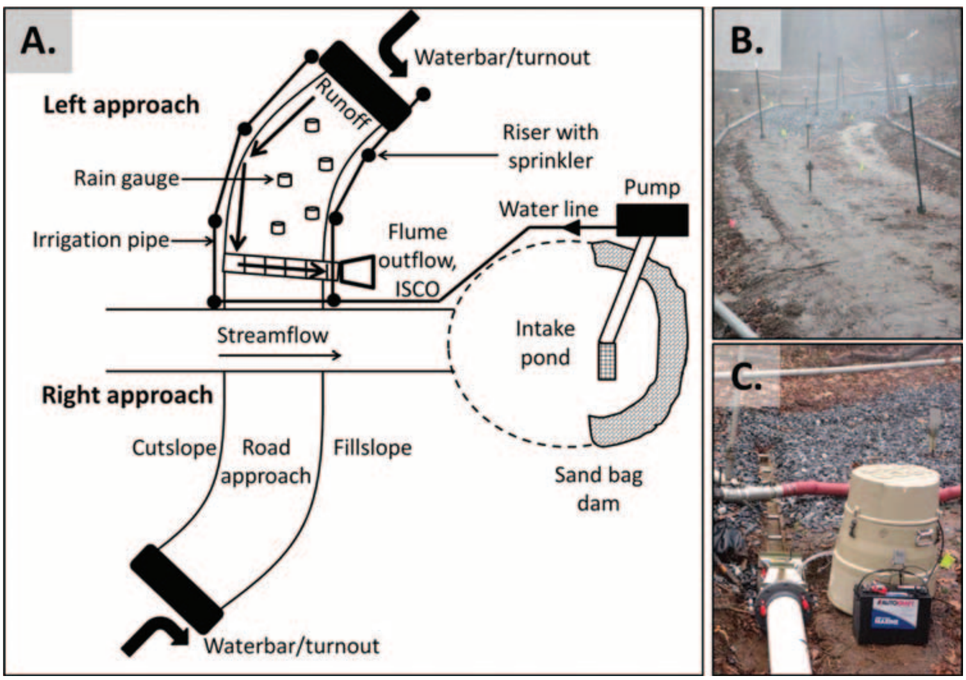

Figure 2. (A) Plan view of two idealized stream-crossing approaches with rainfall simulator equipment and monitoring instrumentation (adapted from Brown et al. 2014). Open-top box culverts collected surface runoff at the bottom of the plot, immediately upslope of the stream. Photographs depicting a rainfall experiment on Mar. 12, 2012 (B) and the equipment used to measure surface runoff quantity and quality (C).

Figure 3. Measurements of event-based rainfall, surface runoff, and sediment yield were used to evaluate the performance of the Water Erosion Prediction Project (WEPP) model to estimate the sediment-reduction efficacy of increasing gravel cover at road-stream crossings. (Photo credit: Kristopher Brown.)

crossing approaches as is typical during forest road construction and graveling.

The near-stream 9.8-m gravel section was chosen for the low gravel treatment because this length approximated half the distance of the shortest approach used in this study. The low gravel treatment resulted in different proportions of cover on the running surface area of the study plots primarily because each approach length was different, ranging from 19.2 to 41.3 m. For example, the low gravel treatment resulted in 60% surface cover for the shortest approach and 40% surface cover for the longest approach used in this study. After the low gravel rainfall experiments, no additional gravel was applied to the initial 9.8-m-long segment, but gravel was applied to the adjacent (uphill) 9.8-m section of the road approach ("high gravel"). This treatment effectively doubled the length of the first gravel application and resulted in an overall range of 50–99% surface cover on the approach running surfaces. The succession of treatments at each site (no gravel, low gravel, and high gravel) facilitated the evaluation of a wide range of surface cover on the stream-crossing approaches (10–99% cover) in reducing sediment delivery to streams during simulated rainfall events (Figure 4). Three to four rainfall experiments were completed in succession within a given treatment at each site ($n = 6$), which resulted in a total of 58 rainfall experiments.

Overall, sediment yield was reduced at each stream-crossing approach as a result of the combined effect of decreased soil erodibility from successive rainfall events within a given treatment and increased gravel surface cover between treatments (Brown et al. 2014). Brown et al. (2014) found that sediment yield per unit rainfall (g m^{-2} mm^{-1}) was commonly greatest during the first rainfall experiment within the no gravel treatment, whereas subsequent no gravel treatment experiments (e.g., no gravel experiments 2, 3, and 4) were similar. The authors concluded that the supply of loose sediment was approaching depletion after the first no gravel treatment rainfall experiment and that the treatment effect of gravel application was evidenced by further declines in sediment yield with increasing gravel cover. Renewed sediment sources associated with the application of the gravel treatments included truck traffic on the stream-crossing approaches and dust associated with the unwashed gravel. Consequently, sediment yields were also greatest for the first rainfall experiments within the low gravel and high gravel treatments, whereas subsequent experiments within each treatment were similar.

Applied rain event characteristics (amount, duration, and intensity), total runoff, and total sediment yield were quantified for each rainfall experiment. These data were used to evaluate WEPP model predictions of event-based runoff and sediment yield.

WEPP Model Setup

The WEPP model (version 2012.8) was used to build unique hillslope profiles for each rainfall experiment ($N = 58$). Each hillslope profile contained site-specific details related to slope, soil type, vegetation management, and applied rain event characteristics (amount, intensity, and duration). Data regarding stream-crossing approach length, slope, running surface width, road vertical shape (concave, convex, linear, or S-shaped), and aspect were used to create six slope files corresponding to the six stream approaches used in this study. Significant changes in road grade (e.g., a stream approach with a 12% slope at the top of the approach, which transitions to a 4% slope near the stream) were included in the slope profiles as breakpoints by using the slope profile editor.

Unique breakpoint climate files were created for each rainfall experiment so that WEPP could be run in single-storm mode. Breakpoint climate files contained cumulative rainfall amounts in 1-minute intervals (i.e., breakpoints) for the duration of each rainfall experiment. We selected the "skid-clay loam" soil file because the parameter values were representative of a low-volume forest road with a clay loam soil texture. We selected the "insloped road-unrutted bare" vegetation management file because it was representative of road surface conditions after road reopening by bulldozer blading. These files for soil and vegetation management were used at each of the study sites.

The initial plant file, which is part of the initial conditions database in the vegetation management file, was "insloped road-bare." This file was used without alteration for each of the stream-crossing approaches. However, it was necessary to create unique vegetation management files for each rainfall experiment to reflect changes in antecedent rainfall, as well as surface cover. Specifically, cumulative rainfall amounts since bulldozer trafficking were calculated for each rainfall experiment, and these values were used for the parameter, "cumulative rainfall since last tillage." Field estimates of surface cover on the running surface component of the stream-crossing approaches were made before each rainfall experiment. These surface-cover estimates were used as the parameter values for initial rill and interrill cover. Rill width type was set to "permanent" because

Figure 4. Photographs depicting rainfall experiments for a succession of gravel surfacing treatments at site 5. (A) No gravel. (B) Low gravel. The yellow lines approximate the upper boundary of the first gravel application. The lower boundary of gravel application was the open-top box culvert (Figure 2C), immediately uphill of the stream. (C) High gravel. The yellow lines indicate the additional coverage afforded by the second gravel application. Surface cover for the rainfall experiments at site 5 was 14, 47, and 63% for the no gravel, low gravel, and high gravel treatments, respectively.

Table 1. Description of model parameters and the ranges of values used in the generation of unique sets of parameters by way of MCMC sampling of the model parameter ranges.

Parameter	Description	Units	Minimum	Maximum	WEPP file
RRINIT	Initial ridge roughness	m	0	0.08	Management
WIDTH	Initial rill width	m	0	0.2	Management
K_i	Interrill erodibility	kg s m^{-4}	2×10^6	11×10^6	Soil
K_r	Rill erodibility	s m^{-1}	0.0001	0.01	Soil
SHCRIT	Critical shear	N m^{-2}	0.4	2.6	Soil
AVKE	Effective hydraulic conductivity	mm $hour^{-1}$	0.1	10	Soil

tillage was not recurring. The mean of the soil bulk density measurements for each stream-crossing approach was used as a constant parameter value in the vegetation management files.

Forest cover adjacent to the stream-crossing approaches consisted of mature hardwood forests. Therefore, the parameter, "days since last harvest", was set to 3,650 (10 years), which reduced the effect of this parameter on soil erosion predictions. The parameter, "days since last tillage" was calculated as the cumulative number of days since the stream-crossing approaches were trafficked with a bulldozer (i.e., before the first series of no gravel treatment rainfall experiments). The remaining model parameters associated with the initial conditions section for insloped road-unrutted bare were not changed, with the exception of initial ridge roughness and initial rill width.

We defined ranges of values for model calibration parameters that are integral to WEPP predictions of runoff and sediment yield from forest roads. Specifically, we defined ranges of values for effective hydraulic conductivity, initial ridge roughness, initial rill width, rill erodibility, interrill erodibility, and critical shear (Table 1). Field observations of antecedent soil water content expressed relative to saturation (i.e., determined from field conditions at each stream-crossing approach) were used for the parameter, initial saturation, in the soil input files. The ranges for model parameters were chosen to reflect site conditions of the stream-crossing approaches used in this study and were based on our own field observations, when possible. In other cases, ranges for model parameter values were based on field experiments by Foltz et al. (2008) or WEPP's technical documentation (National Soil Erosion Research Laboratory [NSERL] 1995).

Uncertainty Analyses

A MCMC algorithm, DREAM_(ZS) (ter Braak and Vrugt 2008, Vrugt et al. 2008, Laloy and Vrugt 2012), was used to efficiently select parameter ranges for WEPP that minimize the discrepancy between model predictions and observations based on a simple least-squares objective function. In this case, WEPP was used to simulate total runoff and total sediment yield from a single event for 58 different rainfall experiments. The use of ordinary Monte Carlo-based random sampling was not feasible because of the number of potential parameter sets that had to be generated to explore the complex parameter space of the WEPP model. DREAM_(ZS) is adaptive and efficient for finding "acceptable" parameter sets in complex inverse modeling problems (Vrugt et al. 2009). The resulting posterior parameter distributions from WEPP calibration using DREAM_(ZS) were used to explore the uncertainty associated with model parameters and model predictions, as well as parameter identifiability/sensitivity, and overall prediction performance.

The main advantage of using DREAM_(ZS) to derive posterior distributions of model parameters is that the sampling procedure learns from experience (i.e., model performance in predicting runoff and sediment yield) and provides denser sampling in the model parameter ranges that are associated with acceptable model runs. Consequently, fewer model runs (and less computer processing time) are necessary to adequately sample the model parameter space. DREAM_(ZS) was used to generate 10,000 unique sets of model parameter values for each rainfall experiment. The number of model runs was selected based on analysis of chain convergence for each of the model parameters.

During model calibration, the range of parameter values for effective hydraulic conductivity was 0.1 to 10 mm $hour^{-1}$ (Table 1), based on previous field estimates of hydraulic conductivity at the stream-crossing approaches (Brown et al. 2014). WEPP model runs were performed with hydraulic conductivity held constant, meaning that WEPP did not internally adjust hydraulic conductivity during event simulations. The range for initial ridge roughness was 0–0.08 m, with the lower values representing road conditions immediately after bulldozer trafficking and the higher val-

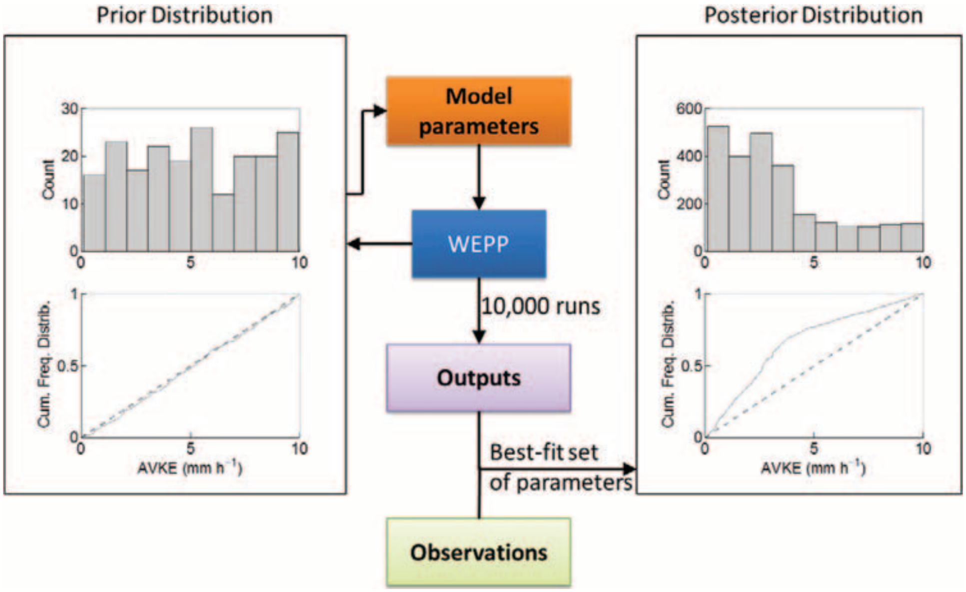

Figure 5. Idealized schematic diagram depicting the relationship between the probability density function and cumulative distribution function of the model parameter values (e.g., effective hydraulic conductivity) before and after MCMC sampling of the parameter ranges. The dashed gray lines represent the cumulative distribution function of a uniform distribution. The solid gray lines indicate the cumulative distribution function of the actual range of parameter values, before and after MCMC sampling.

ues associated with maximum gravel cover on the stream approaches.

No distinct rills (concentrated overland flows) were observed during the field rainfall experiments. However, we used a parameter range of 0–0.2 m for rill width to reflect the likelihood that erosion by concentrated runoff had at least a small effect on observed sediment yields. The parameter range for rill erodibility (0.0001–0.01 s m^{-1}) was set to reflect the wide range of values reported in the peer-reviewed literature (e.g., Foltz et al. 2008). The range of parameter values used for interrill erodibility (2×10^6 to 11×10^6 kg s m^{-4}) was based on the range of values reported in the WEPP technical documentation (NSERL 1995). The range of parameter values used for baseline critical shear was 0.4 to 2.6 N m^{-2} (Foltz et al. 2008).

Evaluation of Model Predictions

Of the 10,000 model runs for each rainfall experiment, the last 25% were chosen for estimating the posterior distributions of parameters (Figure 5). Posterior distributions were estimated from the last samples in the Markov chains when convergence of the individual chains was consistently below the threshold of 1.2 for the Gelman and Rubin statistic (Gelman and Rubin 1996). It has been suggested that the last 25% of the samples in each chain are an appropriate characterization of the posterior distribution (Vrugt et al. 2009). In all simulations for this study, convergence was reached within 1,500 samples; thus, the last 25% was a conservative estimate. To account for variability in applied rainfall amounts and intensities, model performance was based on the comparison of observed and predicted runoff coefficients (runoff depth/rainfall depth) and sediment yield per unit rainfall (mg m^{-2} mm^{-1}).

The 95% confidence intervals of model predictions resulting from the posterior parameter distributions were used to evaluate model performance in comparison to event-based runoff and sediment yield for a succession of applied rainfall events, as well as a succession of gravel treatments that represented increasing intensities of BMP implementation. Posterior distributions of the model parameter values were expressed as empirical cumulative distribution function (ECDF) plots to identify regions of the model parameter space (i.e., specific ranges of values for each of the model parameters) that were associated with acceptable model runs. For a continuous variable, the gradient of an ECDF plot is equal to the probability density at that point. This means that the steepest slopes on the ECDF plot indicate the highest relative frequencies on a histogram of the posterior distribution (Figure 5). Therefore, we can use ECDF plots to identify the best range of model parameter values (as indicated by the steepest slopes on the ECDF plots) to be used for different road surface treatments (e.g., no gravel, low gravel, and high gravel).

Results and Discussion

Model Performance in Predicting Event-Based Surface Runoff

For many of the rainfall experiments, WEPP predictions matched the observed runoff coefficient (Figure 6). This means that for a given rainfall experiment, there was at least one parameter set that resulted in a prediction that matched observed runoff. WEPP also predicted higher runoff coefficients for the no gravel treatment, which is similar to findings from the field rainfall experiments (Brown et al. 2014). However, the ranges of predicted runoff coefficients were often very wide (Figure 6), reflecting the substantial uncertainty associated with model parameter values related to runoff generation (e.g., effective hydraulic conductivity). Variability in runoff is also influenced by the water content of the road on the day of the event and the duration and intensity of the event. Initial water content, rainfall duration, and intensity were fixed (i.e., held at the field-measured values) for each rainfall event. However, initial water content of the road surface is an important factor controlling runoff for a given event (Flanagan et al. 2012).

Model Performance in Predicting Event-Based Sediment Yield

WEPP performed well in predicting reductions in sediment yield for successive rainfall events within a given treatment and

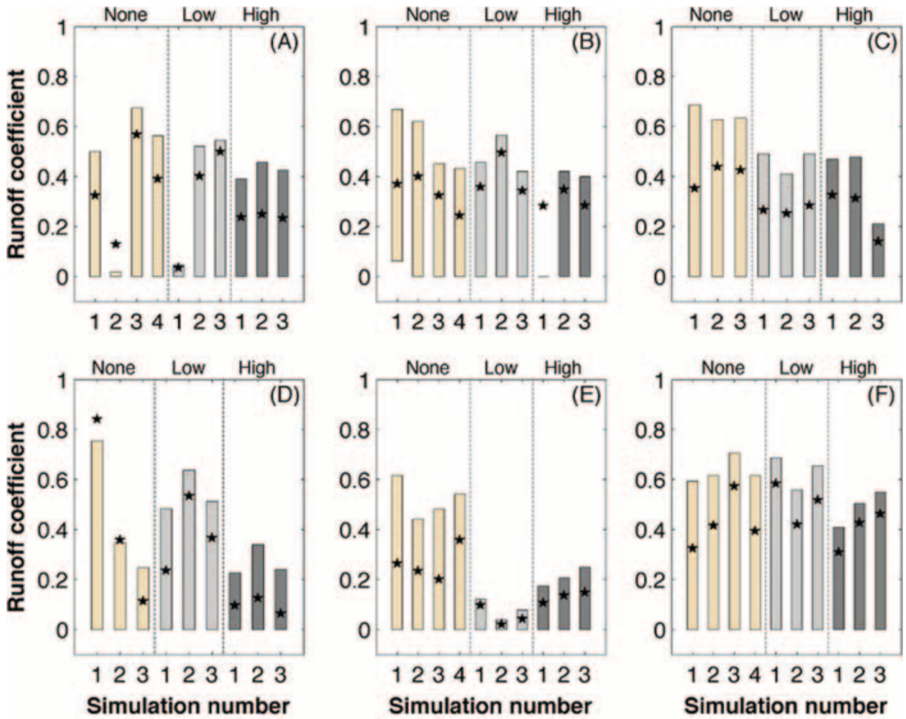

Figure 6. Predicted (bars) versus observed (stars) runoff coefficients for the six stream-crossing approaches used in this study (sites 1–6 are shown as A–F) and by treatment type (none = no gravel, low gravel, and high gravel). Bars represent the 95% confidence intervals for the model predictions for each rainfall experiment. Simulation number specifies the order in which rainfall experiments were conducted within each treatment.

reductions in sediment yield associated with increasing gravel surface cover on the stream-crossing approaches (Figure 7). However, the prediction intervals were often wide, reflecting the substantial uncertainty associated with model parameters related to sediment yield. For example, observed sediment yield at site 3 (Figure 7C) decreased with successive rainfall events for the no gravel treatment (event numbers 1, 2, and 3) as a result of decreased soil erodibility. WEPP predictions also reflect the effect of decreased soil erodibility as a function of successive rainfall events within a given treatment.

WEPP predictions of sediment yield clearly show differences (i.e., sediment reductions) among the different road surface treatments that represent increasing intensities of BMP implementation. This capability is important for evaluating the effectiveness of different BMP implementations to reduce sediment delivery for a wide range of road conditions (i.e., climate, soil, topography, surface cover, and traffic). However, such wide prediction intervals for runoff and sediment yield underscore the importance of explicitly accounting for the uncertainty associated with model parameterization by utilizing a range of erosion predictions (Hession et al. 1996), as opposed to a single erosion prediction, to aid forestland managers in prescribing site-specific BMP implementations to reduce sediment delivery to water bodies. The complexity of data requirements for many physically based erosion models make them best suited for academicians and state and federal agencies that have the resources to couple field monitoring with evaluations of model performance. Forestry practitioners can reduce sediment delivery from major sources (e.g., roads, trails, and associated stream crossings) through a careful emphasis on preharvest planning and the use of erosion models with readily attainable parameter values (e.g., the Universal Soil Loss Equation modified for forestland) to estimate the sediment-reduction efficacy of site-specific BMP implementations (Dissmeyer and Foster 1984).

Therefore, despite the ability of WEPP to predict relative differences in event-based sediment yield among different types of BMP, such wide prediction intervals suggest limited applicability for scenarios that demand a high level of prediction accuracy, such as total maximum daily load development. This issue is not specific to the WEPP model. Because of the inherent variability in measuring soil erosion rates, it follows that model predictions are also highly variable (Brooks et al. 2006). Therefore, it is recommended that prediction intervals be used to show the substantial variability in sediment yield predictions that can result from measurement error and parameter uncertainty, among other sources (i.e., model structure). Commonly, evaluations of model performance have compared an optimal model run with observations of runoff and sediment yield (Croke and Nethery 2006, Sawyers et al. 2012, Wade et al. 2012, Brown et al. 2013). Our study findings show that although a single optimal model run may be useful for comparing relative differences among treatments, it is less meaningful if a large subset of model runs (with unique combinations of values for model parameter sets) can yield equally acceptable model predictions as defined by an objective function such as the least squares of the model residuals.

It is possible that in this research, the substantial uncertainty associated with WEPP predictions of event-based runoff and sediment yield is partly a function of the relatively small quantities of runoff and sediment yield observed during the field rainfall experiments (Brown et al. 2014). The rainfall simulator used in this study has a designed rainfall application rate of 50.8 mm hour^{-1} (Dillaha et al. 1988). At 50.8 mm hour^{-1}, the Rain Jet 78C nozzles provide about 40% of the kinetic energy of natural rainfall (Renard 1989). In addition, after road reopening by bulldozer blading, traffic was limited to light-vehicle use to complete the rainfall experiments (i.e., one to two passes per week), as well as two passes by a dump truck to spread gravel on the approaches. As a result, in a few cases we are attempting to predict runoff depth as low as 0.3 mm and sediment yield as low as 0.04 kg. A runoff prediction of 1 mm in comparison to an observed runoff depth of 0.3 mm represents an overprediction by 233%. Prediction accuracies would probably improve in the case of much greater observations of runoff and sediment yield (i.e., for very large storm events or for annual runoff amounts and rates of sediment delivery). For example, another study suggested that WEPP predictions of erosion could be assumed to be within ±50% of observations for erosion predictions over longer timescales, such as in annual sediment budget analyses (Brooks et al. 2006).

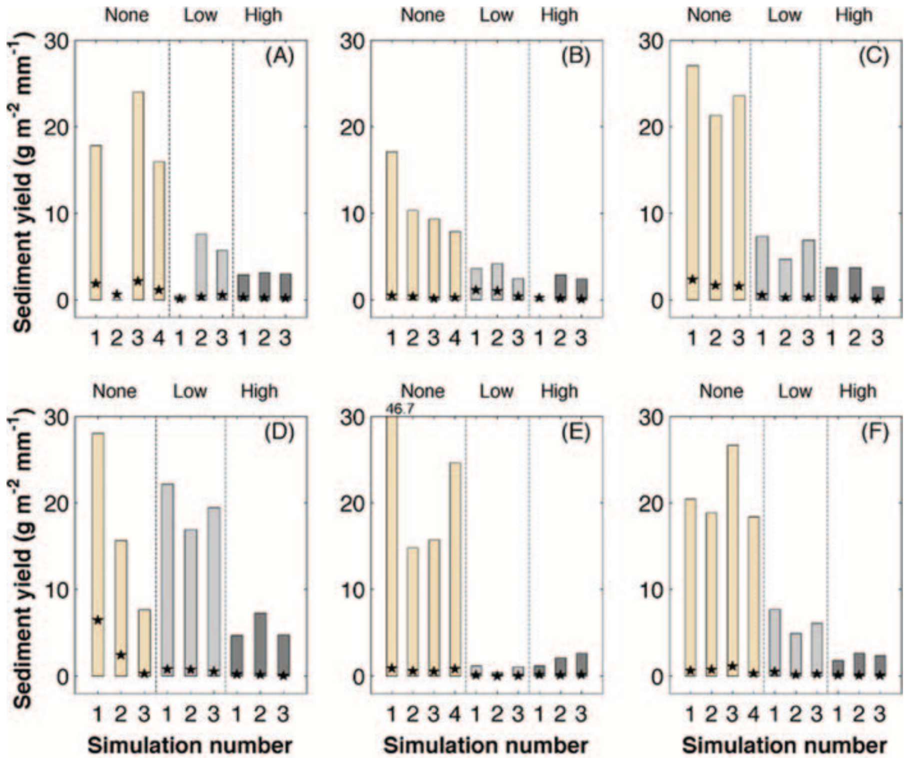

Figure 7. Predicted (bars) versus observed (stars) sediment yields for the six stream-crossing approaches used in this study (sites 1–6 are shown as A–F) and by treatment type (none = no gravel, low gravel, and high gravel). Bars represent the 95% confidence intervals of predicted sediment yield for each rainfall experiment. For instances in which the prediction limit exceeded the *y*-axis limit (30 g m^{-2} mm^{-1}), the value is labeled at the top of the figures. Simulation number specifies the order in which rainfall experiments were conducted within each treatment.

Model Parameter Identifiability

The posterior parameter distributions for interrill erodibility, rill erodibility, and critical shear did not differ substantially from their prior distribution (i.e., a uniform distribution) (Figure 8). This is indicated by the relatively constant slope steepness of the ECDF plots over the full range of model parameter values for interrill erodibility, rill erodibility, and critical shear. Consequently, there are no discernible differences among the road surface treatments. Therefore, for these parameters, we cannot recommend parameter ranges that are specific to the road surface treatments used in this study. This finding indicates that these parameters are insensitive to changes in soil erodibility associated with successive rainfall events, as well as surface cover associated with the different gravel surface treatments. Parameter identifiability may improve with further field experimentation to better define or narrow the initial ranges for model parameters that are integral to WEPP predictions of runoff and sediment yield. In the case of interrill erodibility, we used a wide range of potential values (2–11 × 10^6 kg s m^{-4}) (NSERL 1995), which was higher than that often observed on roads (see Foltz et al. 2009, 2011), because we had limited a priori knowledge of erodibility parameter ranges for soils at RHFRRC.

Despite limited parameter identifiability for interrill erodibility, rill erodibility, and critical shear, WEPP predictions showed decreases in sediment yield associated with successive rainfall events and increased gravel surface cover (Figure 7). WEPP predictions showed decreases in sediment yield because we manually changed the model parameter values for initial rill and interrill cover (corresponding to the succession of gravel treatments) and the cumulative rainfall amount since last disturbance for each rainfall experiment. For initial ridge roughness, better model runs for the high gravel treatment were associated with lower values, whereas better model runs for the no gravel treatment were associated with higher values (initial range = 0–0.08 m). For initial rill width, better model runs were associated with lower values (initial range = 0–0.2 m) for all treatments, and this finding was most pronounced for the no gravel treatment, followed by low gravel, and then high gravel. For effective hydraulic conductivity, better model runs for the low gravel and high gravel treatment were associated with lower values (initial range = 0.1–10 mm hour^{-1}).

Overall, despite using a MCMC algorithm to search the model parameter ranges, we found that it was difficult to identify parameter ranges that were associated with acceptable model runs, especially for interrill erodibility, rill erodibility, and critical shear. Brazier et al. (2000) also found that parameter identifiability was difficult for interrill erodibility. For physically based models such as WEPP that have many detailed mathematical equations and model parameters, there are complex interactions among model parameters that confound parameter identifiability (Beven 2008). Therefore, in this case, it is possible for predictions to match observed runoff and sediment, but it is difficult to know whether the model parameters adequately represent runoff and erosion processes that were observed in the field.

Conclusions

Sediment delivery from forest roads at stream crossings can be a major threat to water quality and aquatic habitat. Models are needed to evaluate the effectiveness of forestry BMPs to reduce sediment delivery over large spatial scales and to guide site-specific BMP implementations to protect water quality. In this study, WEPP model performance was evaluated for the prediction of event-based runoff and sediment yield at forest stream-crossing approaches and for different gravel surfacing treatments that represented increasing intensities of BMP implementation. WEPP was evaluated based on prediction performance for runoff and sediment yield, as well as its ability to distinguish between the different BMP treatments. The posterior parameter distributions that resulted from MCMC sampling were evaluated to determine whether we could recommend parameter ranges that are specific to the different road surface treatments used in this study.

WEPP was able to match observed runoff and sediment yield for many of the rainfall experiments. WEPP predicted reductions in sediment yield that were observed in the field resulting from decreased soil erodibility associated with successive rainfall

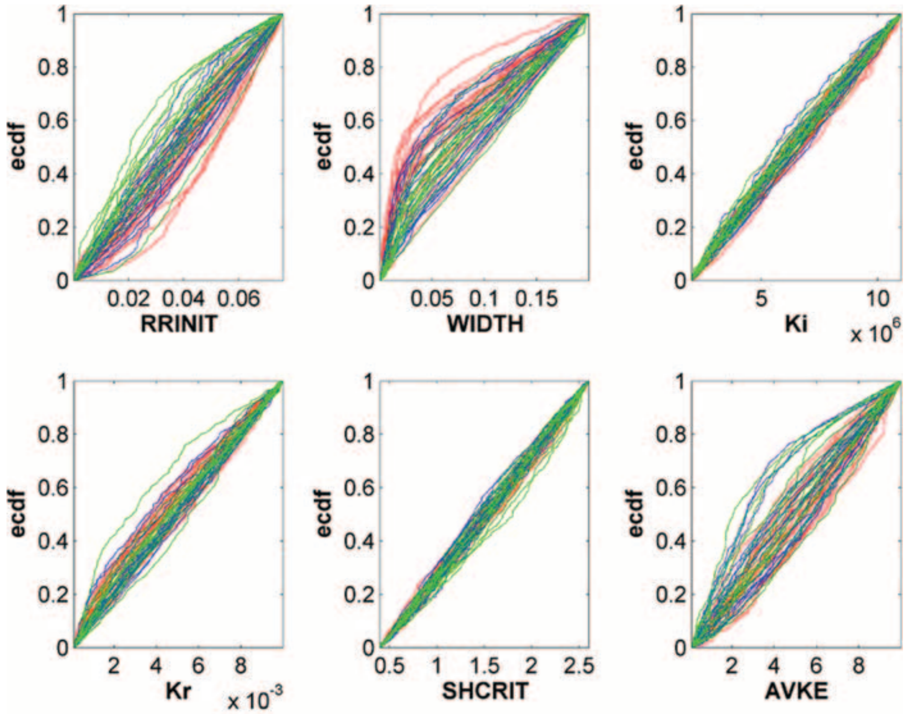

Figure 8. ECDF plots for model parameters that were varied by way of MCMC sampling. The ECDF plots shown here represent the best 2,500 model runs for each applied rainfall experiment. No gravel, low gravel, and high gravel treatments, respectively, are associated with red, blue, and green lines. From left to right on the top row: initial ridge roughness (RRINIT, m), initial rill width (WIDTH, m), and interrill erodibility (K_i, kg s m^{-4}). From left to right on the bottom row: rill erodibility (K_r, s m^{-1}), critical shear (SHCRIT, N m^{-2}), and effective hydraulic conductivity (AVKE, mm $hour^{-1}$).

events, as well as the treatment effect of increasing gravel cover on the stream-crossing approaches. However, 95% confidence intervals representing the range of predicted runoff and sediment yield for the best model runs were often very wide for each rainfall experiment. This result reflects the substantial uncertainty in model parameter values and model predictions. Based on analysis of the posterior distributions of model parameters, we could not recommend ranges of parameter values that were specific to the different road surface treatments for interrill erodibility, rill erodibility, or critical shear.

Overall, these results suggest that there is limited utility in estimating soil erosion or sediment delivery based on a single, optimized model run (i.e., one set of model parameters that result in an acceptable prediction for runoff and sediment yield). Rather, predictions should be made with a range of potential values for model parameters related to runoff generation and sediment yield to reflect the uncertainty associated with model parameterization. In this way, a range of erosion predictions associated with different intensities of BMP implementations can be compared to aid in watershed management efforts to protect water quality, while explicitly accounting for the uncertainty associated with model predictions. These results also suggest that watershed management decisions should not be based on model predictions of sediment yield alone but rather on a combined effort that includes field monitoring to determine BMP effectiveness in reducing sediment delivery, improve a priori estimates of model parameters, and evaluate the performance of models to estimate BMP efficacy.

Literature Cited

Anderson, C.J., and B.G. Lockaby. 2011a. Research gaps related to forest management and stream sediment in the United States. *Environ. Manage.* 47(2):303–313.

Anderson, C.J., and B.G. Lockaby. 2011b. The effectiveness of forestry best management practices for sediment control in the southeastern United States: A literature review. *South. J. Appl. For.* 35(4):170–177.

Appelboom, T.W., G.M. Chescheir, R.W. Skaggs, and D.L. Hesterberg. 2002. Management practices for sediment reduction from forest roads in the coastal plains. *Trans. ASABE* 45(2):337–344.

Ascough, J.C. II, D.C. Flanagan, M.A. Nearing, and B.A. Engel. 2013. Sensitivity and first-order/Monte Carlo uncertainty analysis of the WEPP hillslope erosion model. *Trans. ASABE* 56(2):437–452.

Beven, K. 2008. *Environmental modelling: An uncertain future?* Routledge, London. 328 p.

Boston, K. 2012. Impact of the Ninth Circuit court ruling (*Northwest Environmental Defense Center v. Brown*) regarding forest roads and the clean water act. *J. For.* 110(6):344–346.

Brazier, R.E., K.J. Beven, J. Freer, and J.S. Rowan. 2000. Equifinality and uncertainty in physically based soil erosion models: Application of the glue methodology to WEPP-the water erosion prediction project-for sites in the UK and USA. *Earth Surf. Process. Landf.* 25(8):825–845.

Brooks, E.S., J. Boll, W.J. Elliot, and T. Dechert. 2006. Global positioning system/GIS-based approach for modeling erosion from large road networks. *J. Hydrol. Eng.* 11(5):418–426.

Brown, K.R., W.M. Aust, and K.J. McGuire. 2013. Sediment delivery from bare and graveled forest road stream crossing approaches in the Virginia Piedmont. *For. Ecol. Manage.* 310:836–846.

Brown, K.R., K.J. McGuire, W.M. Aust, W.C. Hession, and C.A. Dolloff. 2014. The effect of increasing gravel cover on forest roads for reduced sediment delivery to stream crossings. *Hydrol. Process.* doi: 10.1002/hyp.10232.

Croke, J., and M. Nethery. 2006. Modelling runoff and soil erosion in logged forests: Scope and application of some existing models. *Catena* 67(1):35–49.

Dillaha, T.A., B.B. Ross, S. Mostaghimi, C.D. Heatwole, and V.O. Shanholtz. 1988. Rainfall simulation: A tool for best management practice education. *J. Soil Water Conserv.* 43:288–290.

Dissmeyer, G.E., and G.R. Foster. 1984. A guide for predicting sheet and rill erosion on forest land. USDA For. Serv., Southern Region, Tech. Publ. R8-TP 6, Atlanta, Georgia. 39 p.

Elliot, W.J. 2013. Erosion processes and prediction with WEPP technology in forests in the northwestern US. *Trans. ASABE* 56(2):563–579.

Flanagan, D.C., J.R. Frankenberger, and J.C. Ascough II. 2012. WEPP: Model use, calibration, and validation. *Trans. ASABE* 55(4):1463–1477.

Foltz, R.B., N.S. Copeland, and W.J. Elliot. 2009. Reopening abandoned forest roads in northern Idaho, USA: Quantification of runoff, sediment concentration, infiltration, and interrill erosion parameters. *J. Environ. Manage.* 90(8):2542–2550.

Foltz, R.B., W.J. Elliot, and N.S. Wagenbrenner. 2011. Soil erosion model predictions using parent material/soil texture-based parameters compared to using site-specific parameters. *Trans. ASABE* 54(4):1347–1356.

Foltz, R.B., H.J. Rhee, and W.J. Elliot. 2008. Modeling changes in rill erodibility and critical shear stress on native surface roads. *Hydrol. Process.* 22(24):4783–4788.

Fu, B.H., L.T.H. Newham, and C.E. Ramos-Scharron. 2010. A review of surface erosion and sediment delivery models for unsealed roads. *Environ. Model. Softw.* 25(1):1–14.

Gelman, A., and D.B. Rubin. 1996. Markov chain Monte Carlo methods in biostatistics. *Stat. Methods Med. Res.* 5(4):339–355.

Harris, R., J. Gerstein, and P. Cafferata. 2008. Changes in stream channel morphology caused by replacing road-stream crossings on timber harvesting plans in northwestern California. *West. J. Appl. For.* 23(2):69–77.

Hession, W.C., and D.E. Storm. 2000. Watershed-level uncertainties: Implications for phosphorus management and eutrophication. *J. Environ. Qual.* 29(4):1172–1179.

Hession, W.C., D.E. Storm, and C.T. Haan. 1996. Two-phase uncertainty analysis: An example using the Universal Soil Loss Equation. *Trans. ASABE* 39(4):1309–1319.

Ice, G.G. 2011. Assessing best management practices effectiveness at the watershed scale. *Appl. Eng. Agric.* 27(6):925–931.

Jackson, C.R. 2014. RESPONSE: Forestry best management practices: A mitigated water pollution success story. *J. For.* 112(1):47–49.

Laflen, J.M., D.C. Flanagan, and B.A. Engel. 2004. Soil erosion and sediment yield prediction accuracy using WEPP. *J. Am. Water Resour. Assoc.* 40(2):289–297.

Laloy, E., and J.A. Vrugt. 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM_(ZS) and high-performance computing. *Water Resour. Res.* 48:18.

Lane, P.N.J., and G.J. Sheridan. 2002. Impact of an unsealed forest road stream crossing: Water quality and sediment sources. *Hydrol. Process.* 16(13):2599–2612.

Loehle, C., T.B. Wigley Jr., A. Lucier Jr., E. Schilling, R.J. Danehy, and G. Ice. 2014. Toward improved water quality in forestry: Opportunities and challenges in a changing regulatory environment discussion. *J. For.* 112(1):41–47.

MacDonald, L., and D. Coe. 2014. RESPONSE: A suggested tiered monitoring strategy for maximizing best management practice effectiveness and protecting water quality. *J. For.* 112(1): 49–50.

Natural Resources Conservation Service. 2013. *Web soil survey.* Available online at websoilsurvey.sc.egov.usda.gov/App/HomePage.htm; last accessed Apr. 23, 2014.

National Soil Erosion Research Laboratory. 1995. *USDA-Water Erosion Prediction Project hillslope profile and watershed model.* Doc. Rep. 10, National Soil Erosion Research Laboratory, West Lafayette, IN. 141 p.

Renard, K.G. 1989. What tools for BMP education? [Commentary in front matter]. *J. Soil Water Conserv.* 44:5–7.

Sawyers, B.C., M.C. Bolding, W.M. Aust, and W.A. Lakel. 2012. Effectiveness and implementation costs of overland skid trail closure techniques in the Virginia Piedmont. *J. Soil Water Conserv.* 67(4):300–310.

Soil Science Society of America. 1986. *Methods of soil analysis. Part 1. Physical and mineralogical methods.* Soil Science Society of America, Madison, WI. 1358 p.

ter Braak, C.J.F., and J.A. Vrugt. 2008. Differential evolution Markov chain with snooker updater and fewer chains. *Stat. Comput.* 18(4): 435–446.

Turton, D.J., M.D. Smolen, and E. Stebler. 2009. Effectiveness of BMPS in reducing sediment from unpaved roads in the Stillwater Creek, Oklahoma watershed. *J. Am. Water Resour. Assoc.* 45(6):1343–1351.

Vrugt, J.A., C.J.F. ter Braak, M.P. Clark, J.M. Hyman, and B.A. Robinson. 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resour. Res.* 44:15.

Vrugt, J.A., C.J.F. ter Braak, C.G.H. Diks, B.A. Robinson, J.M. Hyman, and D. Higdon. 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlinear Sci. Numer. Simul.* 10(3): 273–290.

Wade, C.R., M.C. Bolding, W.M. Aust, W.A. Lakel, and E.B. Schilling. 2012. Comparing sediment trap data with the USLE-forest, RUSLE2, and WEPP-Road erosion models for evaluation of bladed skid trail BMPs. *Trans. ASABE* 55(2):403–414.