MEASURING AND MODELING IMPACTS OF GRAVEL ROAD DESIGN ON SEDIMENT GENERATION IN THE SOUTHEASTERN U.S.

William J. Elliot[1,*,†], Sarah A. Lewis[1], Chelsea L. Cannard[2]

[1] Rocky Mountain Research Station, USDA Forest Service, Moscow, Idaho, USA.
[2] Environmental Pubic Health, Washington State Department of Health, Olympia, Washington, USA.
[*] Correspondence: welliot@moscow.com
[†] Retired.

HIGHLIGHTS

- The erodibility of heavily trafficked gravel roads can be much greater than that of low volume forest roads.
- Improved designs of heavily trafficked gravel roads can decrease sediment generation by more than 90 percent.
- The WEPP Model can be successfully parameterized for high traffic gravel roads to reflect the effects of weather, road design, and topography.

ABSTRACT. *The purposes of this study were to support a watershed modeling analysis by evaluating the ability to the Water Erosion Prediction Project (WEPP) model to estimate sediment generated by high traffic gravel roads, and to determine the erodibility of two designs of high-traffic gravel roads. In many watersheds, the road network can be a major source of sediment. The ability to predict erosion from roads, evaluate the effects of design and management on road sedimentation, and compare sediment from roads to other sources of sediment in the watershed is an ongoing need by watershed managers. The Water Erosion Prediction Project (WEPP) model is a widely used model for predicting sediment from forest roads. There has, however, been little information published on erosion from high traffic gravel roads and WEPP applications to such roads. To evaluate road erosion predictions, a study was conducted incorporating two road designs at Fort Benning, Georgia, U.S. One design followed a common practice of starting with a native material road and adding gravel and grading as required. Erosion and rutting on the road surface were common occurrences on this type of road. The improved design was a "graded aggregate base" design, built with compacted aggregate layers. To evaluate erosion risks for these two road designs, runoff and sediment delivery were measured from ten plots ranging in size from 63 to 150 m^2. Runoff depths up to 50 mm occurred from daily rainfall amounts up to nearly 60 mm, with least square mean event runoff values of 6.5 mm from unimproved plots and 14.9 mm from improved road plots. Delivered sediment ranged from zero to 18 Mg ha^{-1} from individual storms with least square mean amounts of 2.27 Mg ha^{-1} of sediment delivered from unimproved road plots compared to only 0.026 Mg ha^{-1} delivered from improved road design plots for a given runoff event. Hydraulic conductivity was found by calibration to be 3.0 mm h^{-1} for unimproved roads and 1.3 mm h^{-1} for improved road segments. Rill erodibility was 0.09 s m^{-1} for unimproved roads and 0.0008 s m^{-1} for improved roads, values that were greater than had been measured on road erosion studies elsewhere that were typically less than 0.0004 s m^{-1}. The critical shear for the unimproved roads was the minimum that the WEPP model would accept, 0.0001 Pa, but was a more typical value of 1.5 Pa for the improved road segments. When applying the calibrated erodibility values to a validation data set, the Willmott indices of agreement were 0.62 and 0.82 for runoff for unimproved and improved roads, respectively, and 0.67 and 0.66 for sediment delivery from unimproved and improved roads, respectively, indicating good agreement between observed and WEPP-estimated runoff and erosion rates. A sensitivity analysis and calibration analysis found that the WEPP model was not sensitive to interrill erosion for this application. A sensitivity analysis coupled with a WEPP validation analysis showed that WEPP could incorporate weather, topography, soil, and road design features to predict sediment delivery from highly erodible road segments. The study suggests that there is a need for a simulated runoff study to determine high values of rill erodibility more precisely on unimproved high traffic roads, and that there is a need to incorporate more erodible road erodibility values into the online WEPP:Road interface for the WEPP model. The road erosion rates and effectiveness of improved road designs for reducing off-road sediment reported in this study will be useful to managers seeking to quantify and reduce road erosion rates from high-traffic gravel roads in sensitive watersheds.*

Keywords. *Erodibility, Gravel Roads, Soil Erosion, WEPP.*

Submitted for review on 16 January 2023 as manuscript number NRES 15539; approved for publication as a Research Article by Associate Editor Dr. Prem Parajuli and Community Editor Dr. Kati Migliaccio of the Natural Resources & Environmental Systems Community of ASABE on 3 July 2023.

Mention of company or trade names is for description only and does not imply endorsement by the USDA. The USDA is an equal opportunity provider and employer.

The purposes of this study were to support a watershed modeling analysis by evaluating the ability to the Water Erosion Prediction Project (WEPP) model (Laflen et al., 1997) to estimate sediment generated by high traffic roads, and to determine the erodibility of high-traffic gravel roads for two road designs. To achieve this purpose, this article provides an overview of road erosion processes and prediction, a description of the key equations in WEPP that influence erosion prediction, a sensitivity analysis of the WEPP model focusing on a road segment, a field experiment measuring runoff and sediment delivery from two road designs, and a calibration exercise to determine soil erodibility values for the WEPP model based on sediment delivery measured from two types of road design.

BACKGROUND

In many non-agricultural watersheds, onsite erosion and offsite sediment delivery are dominated by the road network (Donigian, 2013; Elliot, 2013; Grace, 2017). Road sediment delivery rates are dependent on topography, road surfacing, soil properties, management practices, and climate (Elliot, 2013; Gucinski et al., 2001; Luce and Black, 1999; Robichaud et al., 2010). Compared to cropland and forest soils, infiltration rates on roads are low due to compaction during construction and subsequent traffic (Foltz et al., 2009; Grace, 2017; Huffman et al., 2013; Reid and Dunne, 1984), resulting in higher runoff rates and sediment delivery compared to most other land uses (Elliot, 2013).

ROAD EROSION MODELLING

Predictive models can assist in determining the relative sources of sediment within forested watersheds, whether sediment is coming from undisturbed, thinned, harvested, or burned hillslopes, access networks such as roads, skid trails, or recreational trails, or the drainage and stream system. Models can also aid in quantifying the benefits and impacts of the design and management of sediment-generating features, such as roads, on sediment delivery. In the 1980s, cumulative watershed effects models were developed that predicted overall sediment contributions from the entire road network within a watershed as a function of a soil factor, a location factor, and the age of the road (USDA Forest Service, 1990). The interface to the 1990 model was like an accounting sheet where users entered road segments within each category of management or age for several years, and the tool tallied the annual sediment from the road network, along with sediment from other forest management practices for each year. The model was customized for specific forests in the northern Rocky Mountains, U.S., with names like "Boised" for application in the Boise National Forest and "Nez Sed" for application in the Nez Perce National Forest. These models included factors to account for landscape topography, road segment design and use categories, and time since construction, but did not focus on the variation of sediment delivery among individual road segments. Similar models were developed for eastern U.S. forests, including roads of different ages and uses in a tally sheet format for estimating the overall contribution of the road network to sediment generated within a watershed using a Universal Soil Loss Equation (USLE)-based erosion model (Marion and Clingenpeel, 2012; Wischmeier and Smith, 1978). In these watershed sediment accounting models, the location of a road segment, whether it was located on the ridge top, parallel to a stream, or crossing a stream, was not easily incorporated.

The Washington Forest Practices Board (Dubé et al., 2004) built on this factor approach by allowing users to address the conditions of individual road segments, but it was up to the user to build the tally sheet. The Washington Forest Practices approach was then linked to a GIS (SEDMODL) to aid in determining road topography and identifying stream crossings (Dubé et al., 2004).

Researchers with the USDA Forest Service developed an approach to modeling road network erosion based on detailed road surveys called the Geomorphic Road Analysis and Inventory Package (GRAIP) Model (Black et al., 2012; Cissel et al., 2012). With GRAIP technology, a detailed road survey was carried out, noting road surface conditions and details of intersections between roads and streams. Concurrent with many of the GRAIP surveys, sediment boxes were installed to measure runoff and sediment delivery from different lengths and gradients of road segments for several years (Black and Luce, 2013). Empirical relationships derived from the sediment boxes were used to estimate runoff and erosion from individual road segments for the entire road network (Cissel et al., 2012). GIS tools were developed as an alternative to the field survey to determine road segment topography and subsequent sediment delivery from a road network (Prasad, 2007). Applications of GRAIP were limited to locations where road monitoring studies were completed, or where soil and climate properties were similar to those of sites that had been monitored.

The above road erosion models provided average annual erosion rates. They were not well suited for estimating erosion from individual storms, runoff generation, downslope sediment delivery, or for addressing road management or use scenarios or soils that were outside of the database that had been developed from monitoring projects. This suite of models often required several years of annual erosion data for calibration for climates or soils that were not already in the database.

Another road erosion technology that was under development since the late 1980s was the application of the USDA Water Erosion Prediction Project (WEPP) Model to forest roads (Elliot, 2004; Elliot et al., 1994, 1995; Laflen et al., 1997). WEPP was a complex FORTRAN model that was run for single storms, a useful feature for calibration and validation, or continuously with daily weather input in yearly increments from 1 to 999 years (Flanagan and Livingston, 1995). For predictive modeling, the WEPP model was usually run for 30 to 100 years using a stochastic climate (Baffaut et al., 1996). For calibration or validation using observed data, WEPP was run either for several individual observed runoff events from natural or simulated rainfall in a single storm mode (Foltz et al., 2009, 2011), or for a limited number of years in continuous mode using an observed

climate (Elliot et al., 1994; Laflen et al., 2004; Robichaud et al., 2016).

Lang et al. (2017) compared estimating road erosion with the WEPP model and the Revised Universal Soil Loss Equation (RUSLE, Renard et al., 1997). They found that variations of both the RUSLE and WEPP models did not perform well, likely due to the magnitude of the variability in the observed erosion rates, which is typical of erosion studies (Elliot and Flanagan, 2023; Laflen et al., 2004; Robichaud et al., 2007). Overall, a carefully parameterized application of the WEPP model performed best in the Lang et al. (2017) study.

All the above models estimated annual sediment delivery. The WEPP model was the only tool among those reviewed that was developed to predict sediment generated by individual storms (Flanagan and Nearing, 1995). That feature made it particularly suited to complement the Hydrologic Simulation Program – Fortran (HSPF) watershed tool that also ran on a daily time step for evaluating watershed scale sediment sources in a recent watershed analysis for the Fort Benning, GA, military base (Donigian, 2013).

The Water Erosion Prediction Project (WEPP) Model

The WEPP model mathematically described the processes that caused erosion, including sub-daily precipitation, infiltration and runoff, a daily soil water balance, and, if there was runoff, sediment detachment, transport, and deposition distributed along a hillslope, and sediment delivery from the hillslope (Laflen et al., 1997, 2004). For describing precipitation, users could either specify a precipitation amount, duration, time to peak intensity as a fraction of duration, and peak intensity as a multiple of average intensity (average intensity = rainfall depth / storm duration), or format precipitation as a "break point" file with pairs of time of day and cumulative precipitation for as many points as the user desired (Flanagan and Livingston, 1995). Generally, the amount-duration format was used for long term modeling with a stochastic weather file (Baffaut et al., 1996) and the breakpoint format for model calibration and validation from simulated rainfall and single storm or limited annual research data sets (Foltz et al., 2009, 2011). In both formats, WEPP internally used precipitation information to construct a double exponential hyetograph for subsequent internal runoff and erosion calculations. Nicks et al. (1995) described the development of a rainfall hyetograph from either breakpoint or amount/duration input. The method assumed an exponential curve for rainfall intensity that rose from the beginning of the storm to the peak intensity occurring when the time to peak intensity was either input or determined from the breakpoint data, and then exponentially declining until the final duration time was achieved. The equations describing these two exponential curves were (Nicks et al., 1995):

$$i(t) = \begin{cases} i_p e^{b(t-t_p)} & 0 \le t \le t_p \\ i_p e^{d(t_p-t)} & t_p < t < 1.0 \end{cases} \quad (1)$$

where

$i(t)$ = normalized rainfall intensity at the normalized time t

i_p = peak intensity as a multiple of the average intensity

t_p = normalized time to peak intensity.

The variables b and d were functions of i_p and t_p for a given storm (Nicks et al., 1995). Once the values for b and d were calculated, the normalized value for $i(t)$ was determined for the desired number of time steps. With breakpoint data, there could be multiple peaks in the resulting hyetographs. If there was a pause in precipitation within a 24-hr day WEPP merged the hyetographs for each sub daily event within a day to generate a multi peak hyetograph to simulate the day's rainfall pattern. An example comparing of how WEPP processed breakpoint data into a simulated hyetograph is presented later in this article.

On hillslopes, WEPP predicted erosion from rain drop splash and shallow overland flow ("interrill erosion") and from concentrated flow ("rill erosion"). Interrill erosion was predicted from rainfall and runoff by Elliot and Flanagan (2023) and Flanagan and Nearing (1995):

$$D_i = K_i S_f i q \quad (2)$$

where

D_i = interrill detachment in kg m^{-2} s^{-1}

K_i = interrill erodibility, a soil property in kg s m^{-4}

S_f = slope factor (dimensionless)

i = rainfall intensity (m s^{-1})

q = runoff rate (m s^{-1}).

Additional factors that addressed rill width and spacing, ground and canopy cover, and surface roughness were also considered when predicting interrill erosion (Flanagan and Nearing, 1995). Interrill erosion was usually measured with rainfall simulation studies on small, bordered plots where rainfall intensities, runoff rates, and sediment concentration in the runoff were measured, along with topography and, in some cases, ground cover. Interrill erodibility was then calculated by solving equation 2 for K_i (Al-Hamdam et al., 2022; Elliot and Flanagan, 2023). Alternatively, Foltz et al. (2009, 2011) used an iterative solution with the WEPP model within Mathcad to minimize the differences between a WEPP-predicted sediment delivery from interrill plots and observed values in replicated plot studies.

Rill erosion was predicted within WEPP by a hydraulic shear model (Elliot and Flanagan, 2023; Flanagan and Nearing, 1995):

$$D_r = K_r (\tau - \tau_c)\left(1 - \frac{G}{T_c}\right) \quad (3)$$

where

D_r = detachment rate in a rill after accounting for entrained sediment (kg m^{-2} s^{-1})

K_r = rill erodibility (s m^{-1})

τ = net hydraulic shear of flow in rill having accounted for shear lost through surface plant residue and roughness (Pa).

τ_c = critical shear, a soil property that must be exceeded by τ before rill erosion occurs (Pa)

G = transport rate of sediment in the rill from upstream rill and interrill erosion (kg s^{-1} m^{-1})

T_c = transport capacity of the rill flow, a function of sediment size distribution and density, and hydraulic shear τ (kg s^{-1} m^{-1})

$$= \gamma r s \tag{4}$$

where

γ= Specific weight of water (N m^{-3})

r = hydraulic radius of rill flow (m)

s = slope of rill, channel (m m^{-1}).

Rill erodibility and critical shear were frequently estimated in runoff simulation studies where a range of simulated flows were applied to an eroding rill (Al-Hamdan et al., 2022; Elliot, 1988; Elliot and Flanagan, 2023; Foltz et al., 2008; Wagenbrenner et al., 2010). Hydraulic shear for each flow rate was determined from the rill cross sectional shape or width and depth, and rill flow velocity. From those values, the hydraulic radius and shear (eq. 4) for a range of simulated rill flow rates were estimated. The measured sediment delivery was divided by the rill area (rill length x width) to calculate a rill erosion rate for each flow rate. The rill erosion rates for the range of flows were then plotted against the hydraulic shear for each respective flow rate. The slope of that plot was the rill erodibility, K_r, and the intercept with the horizontal (shear) axis was the critical shear, τ_c. Advanced analytical methods for incorporating limitations to rill detachment due to sediment already in transport were necessary on highly erodible cropland soils (Elliot, 1988; Elliot and Flanagan, 2023). Foltz et al. (2008) used runoff simulation on forest roads with bordered plots to measure rill erodibility, reducing variability by having a constant rill width. They estimated K_r and τ_c by assuming that the limitations to detachment in equation 3 was equal to 1 ($(1 - G / T_c) \cong 1$). Wagenbrenner et al. (2010) applied a series of overland flow rates to naturally formed rills following wildfire. They measured the rill width and depth with a ruler for each flow rate to estimate cross sectional shape and flow depth, and to subsequently calculate r and τ. Such analytical approaches, however, did not work with natural rainfall, as it was not possible to measure rill geometry during every rainfall event. For natural rainfall applications, researchers have used optimizing methods to vary K_r and τ_c values to match observed sediment delivery. Grace (2017) manually calibrated a subset of observed sediment delivery rates from natural rainfall on forest road plots to determine optimal values for K_r and τ_c. Srivastava et al. (2020) used the PEST model (Doherty, 2005) in combination with WEPP to minimize the least-square error between observed and simulated annual sediment yields at in a nested watershed study. They found that in a 25 km^2 forested watershed, sediment delivery was dominated by channel processes, and that observed sediment delivery was only sensitive to the critical shear of the channels. Srivastava et al. (2020) reported that the WEPP-calibrated critical shear values were related to the sediment size distribution for a given channel reach.

Infiltration rate in the WEPP model was estimated by a form of the Green-Ampt Mein-Larson model at a sub-hourly time step as a function of soil properties, soil water content, and infiltration depth from previous time increments (Flanagan and Nearing, 1995; Mein and Larson, 1973; Risse et al., 1994). The WEPP model internally estimated soil infiltration rates from the user's inputs of soil profile textural properties and a baseline saturated hydraulic conductivity value (K_b) (Flanagan and Livingston, 1995). The K_b value was then adjusted internally in WEPP to account for seasonal weathering, consolidation, and other temporal effects to estimate an effective hydraulic conductivity K_e for a given storm. This K_e was then used in the Green-Ampt equation to estimate subhourly infiltration using the equation (Risse et al., 1994):

$$f = K_e\left(1+\frac{N_s}{F}\right) \tag{5}$$

where

f = infiltration rate (L/T)

K_e = effective hydraulic conductivity (L/T)

N_s = effective matric potential (L)

F = cumulative infiltration (L).

To estimate the baseline saturated hydraulic conductivity K_b, some users calculated the difference between runoff and rainfall from prolonged rainfall simulation studies (Al-Hamdam et al., 2022; Rawls et al., 1989; Risse et al., 1994). More recently, researchers have estimated K_b through an iterative process using WEPP itself to determine an optimal baseline hydraulic conductivity value to best match predicted runoff hydrograph or amount for a given event with observed rainfall, runoff, and the soil water content prior to rainfall (Foltz et al., 2009, 2011). For watershed analyses, Srivastava et al. (2017, 2020) applied the PEST model (Doherty, 2005) in conjunction with WEPP to minimize the least square error between observed and predicted daily streamflow, and the logarithm of daily streamflow to estimate saturated hydraulic conductivity of the upland soils and other soil hydrologic properties that were important for watershed modeling.

Sensitivity analyses can be used to better quantify or qualify model performance and prioritize model calibration. The value of the sensitivity parameter S was calculated by Nearing et al. (1990) and Ascough et al. (2013):

$$S = \frac{\left[\dfrac{(O_2 - O_1)}{\dfrac{(O_2 + O_1)}{2}}\right]}{\left[\dfrac{(I_2 - I_1)}{\dfrac{(I_2 + I_1)}{2}}\right]} \tag{6}$$

where O_2 and O_1 were adjacent model outputs estimated by the input values of I_2 and I_1. In WEPP studies, the input values were usually associated with weather, topography, soil, and vegetation (Ascough et al., 2013; Miller et al., 2011; Nearing et al., 1990). In erosion modeling, S was dependent on the range of the specific input values and other site conditions. Both Nearing et al. (1990) and Ascough et al. (2013) analyses were done using midwestern U.S. cropland soils

and management. Ascough et al. (2013) found in their analysis that a natural break in sensitivity tended to occur around S = 0.2. Input variables with S<0.2 were considered "not very sensitive." Sensitivity has also been evaluated graphically, showing the response of a given model output for a range of input conditions (Miller et al., 2011). The steeper the slope of a graph, positive or negative, of output vs. input, the greater the model sensitivity. Others have used higher level statistics to evaluate the impact of the variability of WEPP input parameters on the variability of output values (Brazier et al., 2000; Tiscareno-Lopez et al., 1993).

Nearing et al. (1990) reported that runoff predicted by WEPP was most sensitive to the total rainfall depth and duration of a given storm, while Ascough et al. (2013) reported that the soil hydraulic conductivity K_b was the most sensitive input variable for runoff. Nearing et al. (1990) reported that the sediment delivery was most sensitive to rill erodibility K_r and hydraulic conductivity K_b (eqs. 3 and 5). Ascough et al. (2013) reported that the sensitivity of sediment delivery depended on cropland management system but was most sensitive to slope length and steepness, K_r (eq. 3), K_b (eq. 5), and in some cases, critical shear τ_c (eq. 3). It is noteworthy that neither investigator reported sensitivity to interrill erodibility K_i (eq. 2) for their conditions. Tiscarno-Lopez et al. (1993) found that runoff depth was most sensitive to precipitation depth and duration, and K_b (eq. 5) when modeling rangelands with variable rainfall characteristics within a given storm. Miller et al. (2011) reported a study focused on post wildfire erosion where they found that WEPP-estimated sediment delivery was sensitive to slope steepness for both wet and dry climates and was sensitive to slope length for a wet climate. Miller et al. (2011) found that WEPP was less sensitive to slope lengths above 200 m for a drier climate. Zhang (2016) found that when comparing WEPP predictions with ^{137}Cs erosion estimates, the erosion distribution predicted by WEPP was sensitive to downslope distance, slope steepness, and rill spacing on cropland plots. Zhang (2016) found that a 10-m rill spacing near the top of a 200 m plot, dropping to 0.5 m midslope, and decreasing to sheet flow at the lower gradient toe of the slope resulted in an erosion distribution that was similar to the distribution determined from ^{137}Cs observations. Nearing et al. (1990), however, reported that the WEPP model was not sensitive to rill spacing with S = -0.1 for a 22-m long slope with a 9% gradient. In the context of road management, Elliot et al. (1999a) showed that WEPP was sensitive to both road segment slope and segment length, particularly for steeper segments (8%-12%) shorter than 50 m. The Ascough et al. (2013) study demonstrated that WEPP sensitivity varies with baseline conditions that were modeled and emphasized the effect of the variability of erodibility properties on erosion prediction. Both Ascough et al. (2013) and Nearing et al. (1990) studies showed that to better understand the importance of the main WEPP input variables for any unique situation, a sensitivity analysis specifically addressing that condition was warranted.

Low volume road erosion may be sensitive to road design. Low volume roads have generally been grouped into three designs (fig. 1; Cao et al., 2021; Elliot, 2004; Elliot and Hall, 1997): (1) insloping toward a ditch that is bounded by

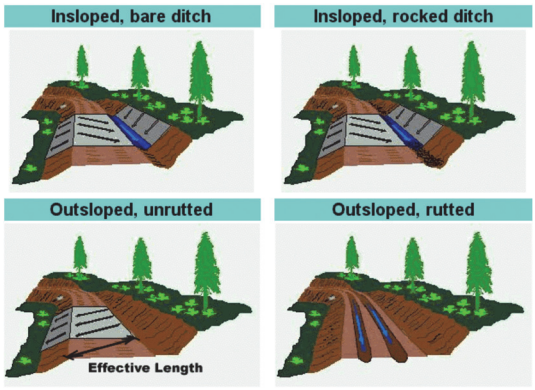

Figure 1. Three road designs for modeling erosion in the WEPP:Road interface plus the insloped design with a vegetated or rocked ditch (Elliot, 2004).

an uphill bank, (2) outsloping, where surfaces are sloped toward a fill slope with no ditch, allowing the runoff to continue downslope, and (3) roads with ruts where the ruts contain road runoff until a cross drain feature intercepts the flow or the rut depth decreases, allowing water to flow to a ditch, over a fillslope, and on down the hill.

Runoff flow paths on roads are likely to influence road erosion rates. Foltz and Elliot (1997) noted that ruts on 60-m road plots increased flow path lengths to the length of the plot (60 m), whereas if rut formation was limited because of reduced tire pressure, flow paths tended to be the resolution of the road grade and the degree of side slope and were 4-5 m long (fig. 1; table 1). The plots with longer wheel ruts delivered more sediment than plots with rills with an outsloped flow path (fig. 1).

Elliot and Hall (1997) described how to model the three road designs (fig. 1) with the WEPP model, including road ditch practices to reduce erosion (adding gravel or allowing revegetation). Tysdal et al. (1999) expanded on insloping road erosion processes, reporting that erosion from the cutslope on the other side of an insloped ditch was not a significant source of sediment. The sediment was either detached from the road surface or the ditch.

To complement road design, the measured erodibility values (eqs. 2, 3, and 5) of road soils in the western U.S. were determined from simulated and natural rainfall (Foltz, 1996; Luce and Black, 1999) and incorporated into an online interface to the WEPP model called WEPP:Road (Elliot, 2004, 2014). Elliot and Foltz (2001) and Laflen et al. (2004) found that WEPP:Road predicted reasonable road erosion rates for several Northwestern U.S. forest road studies. Welsh (2008) found that WEPP:Road underpredicted road erosion by a factor of 20 on a loamy coarse sand in Colorado, USA, while Stafford (2011) found that WEPP:Road overpredicted road erosion by a factor of 3 on low traffic forest roads with a sandy loam soil in the Sierra Nevada mountains in California. In another study in the Sierra Nevada mountains, Foltz et al. (2011) found that hydraulic conductivity on coarse-textured soils was greater than the WEPP:Road database and was likely one reason for the overprediction in the Stafford (2011) study. Foltz et al. (2011) also reported that interrill erodibility was lower than values in the WEPP:Road interface. In a study in Northern Georgia, U.S. Grace (2017)

Table 1. Features of the three road designs in WEPP:Road plus the insloped rocked or vegetated ditch design (Elliot, 2004; Elliot and Hall, 1997). A run was completed for each road design using the same inputs: Climate: 100-yr stochastic climate using Talbotton, GA climate data (1351 mm average annual precipitation); Graveled sandy loam soil; High traffic; 8% Road Gradient, 15 m length 5.3 m wide; Fillslope 12% slope, 0.3 m length; Forest buffer 12% slope, 0.3 m length.

Feature	Insloped, Bare Ditch	Insloped, Rocked Ditch	Outslope, Unrutted	Outslope, Rutted
WEPP input slope	Road Gradient	Road Gradient	$\sqrt{\text{Road Gradient}^2 + 4^2}$	Road Gradient
WEPP input slope length	Road segment length 15 m	Road segment length 15 m	$\text{Segment width} \times \frac{\text{Input slope}}{4\%} = 11.85$ m	Road segment length 15 m
Rill spacing	4 m	4 m	1 m	2 m
Critical shear τ_c	2 Pa	10 Pa	2 Pa	2 Pa
Width of plot	Width of road segment 5.3 m	Width of road segment 5.3 m	$\frac{\text{Segment length} \times \text{width}}{\text{WEPP input slope length}} = 6.7$ m	Width of road segment 5.3 m
Rain + snow runoff	267 mm	190 mm	245 mm	232 mm
Sediment leaving road	32.5 Mg ha^{-1}	22.8 Mg ha^{-1}	29.2 Mg ha^{-1}	30.8 Mg ha^{-1}

found similar saturated hydraulic conductivity values to the WEPP:Road database, but estimated rill erodibility to be about a fourth of the WEPP:Road value on a road with limited traffic. These studies all tend to confirm what Foltz et al. (2011) concluded:

> Future erosion modeling studies on unpaved forest roads should focus on determining better proxies for erosion parameters. On road surfaces, parent material and soil texture are marginal indicators of effective hydraulic conductivity. Levels of traffic, time since road grading, and wetting and drying cycles may be better indicators. We suggest that these indicators be included in future rainfall simulation studies of water erosion parameters.

The roads in the above studies were in forested watersheds where road traffic was light. Only during intense thinning or timber harvest operations were roads subjected to high traffic, and were subsequently highly erodible (Foltz, 1996; Guciniski et al., 2001; Luce and Black, 1999; Reid and Dunne, 1984).

The WEPP:Road interface not only had 3 road designs (fig. 1), but it also allowed the user to access the WEPP climate database with 2600 climate stations, select from four soil textures, specify the traffic level (high, low, none) and specify the road surface (native, gravel, or paved). WEPP:Road assumed an outslope grade of 4% for the outslope design (fig. 1). Users provided the length, width, and gradient of the road surface, as well as the lengths and steepness values of the road's fill slope and any downslope forested buffer. Table 1 describes the features of the three road designs in figure 1, plus the insloped design with a rocked or vegetated ditch. Additional details of WEPP:Road inputs can be found in Elliot (2004) and Elliot et al. (1999b).

Road Erosion on Military Bases

One federal agency concerned about road erosion was the U.S. Department of Defense (DoD). The DoD conducts training and testing activities on more than 1200 km^2 of forest, grassland, boreal, and urban landscapes. Many roads on military lands receive heavier traffic than most forest roads and are thus likely susceptible to elevated erosion rates. Aqua Terra Consultants (2012) determined that the road network in Fort Benning, Georgia, U.S. (fig. 2) only accounted for 3% of the surface area yet was the source of 28% of the sediment within that installation (Donigian, 2013). To predict the relative role of road erosion on the site, they developed a method for linking the WEPP model outputs to their watershed tool, the Hydrologic Simulation Program Fortran (HSPF), to synthesize sediment from all sources (roads, unburned and burned forests, developed areas). One of the limitations to the Donigian (2013) study, however, was the same as that noted by Foltz et al. (2011), in that the erodibility of the highly trafficked roads on the military base was not known. To fill this knowledge gap, the U.S. DoD Environmental Security Technology Certification Program (ESTCP) funded a study to measure the erodibility of roads in Fort Benning and determine if the WEPP technology could be applied to heavily trafficked roads as part of a larger watershed modeling project (Donigian et al., 2018).

Historically, Fort Benning site managers added gravel and graded roads as necessary to maintain trafficability. Recently, the managers upgraded some of the more heavily used roads by rebuilding them in compacted layers of aggregate to improve trafficability and reduce erosion. Improved roads were designated as "graded aggregate base" or GAB roads (Donigian et al., 2018). Watershed managers and modelers, however, did not know the amount of sediment that was delivered from the original unimproved gravel roads, nor the effects of the GAB road design on sediment delivery. Managers also needed tools to estimate the amount of sediment delivered from these high traffic road networks within Fort Benning, and from road networks elsewhere (Donigian et al., 2018). They were especially interested in tools that could account for differences in climate, topography, and road design (Elliot et al., 2015). Earlier estimates of sediment generated by roads used WEPP, but the application of this model to Fort Benning conditions had not been validated (Donigian et al., 2018; Elliot et al., 2015).

The objectives of this article are:

1. To report on a study to determine erosion rates and WEPP soil erodibility values for two designs of high-traffic gravel roads, and
2. To demonstrate the ability of WEPP to account for the effects of weather, topography, and management on road sediment delivery.

The hypotheses are that an improved road design will result in reduced sediment delivery and that WEPP will be able

Figure 2. Location of Fort Benning, Georgia, in Southeastern U.S., and location of the two sets of research plots, the Digital Multipurpose Range Complex (DMPRC) with the improved road design and Hourglass Road with the unimproved road design, for the Fort Benning road erosion study. In the bottom map, note the location of Talbotton, GA, in the upper right corner of the map, the location of the weather station that provided the long term weather statistics for the sensitivity analyses.

to predict the effects of weather, topography, and management on sediment delivery. To achieve these objectives, we carried out a field study to measure runoff and sediment delivery from roads at Fort Benning. We completed a sensitivity analysis with the WEPP model to identify the sensitivity of the model to road design, topography, soil hydraulic conductivity, and WEPP soil erodibility. From the field study results, we calibrated the WEPP model for the two road designs with half the field data set. We validated the application of the WEPP model for this site with the other half of the field data set. The results of this study are useful for planning road erosion studies on other sites, for evaluating road erosion risks from high-traffic gravel roads on similar soils, and for quantifying the watershed benefits of an improved road design.

Materials and Methods

The study was carried out at Fort Benning, in Eastern Georgia, U.S. (fig. 2). Fort Benning covers 740 km^2, the majority of which is forested. The dominant undisturbed soil was Troup loamy sand with 86% sand, 3% clay, and 11% silt (table 2; USDA, 2023). The average high temperature was 24.6°C, and the low was 12.8°C. The average annual rainfall was 1190 mm, and snowfall was 5 mm. March had the highest monthly precipitation, averaging 129 mm (https://www.milbases.com/georgia/fort-benning-army-base/weather). Precipitation often occurred as high intensity localized thunderstorms (Trewartha and Horn, 1980). The Fort Benning road network included 100 km of surfaced roads. Most roads were either graveled and frequently graded or had improved GAB construction that required less frequent maintenance (Donigian et al., 2018).

Field Data Collection

Runoff plots were initially installed on improved (GAB) and unimproved (gravel with grading) road segments during the last week of February and first week of March, 2014. Data collection took place in the historically wettest months of March, June, and July of 2014 and 2015. Plots were sufficiently long to ensure that both interrill and rill erosion processes were occurring (table 3), but not too long as to overwhelm the monitoring equipment during a heavy thunderstorm (Grace, 2017; Grace et al., 1998). The plots had four straight sides but were not rectangular (fig. 3a). The lengths of the sides and diagonals of all plots were measured to calculate the plot area. The elevations of the corners of the plots were surveyed with a laser level and combined with the plot

Table 2. Soil textural distribution by horizon for Troup loamy sand (USDA, 2023). The rock content (> 2mm) was trace for all horizons. The Troup loamy sand is the dominant soil series associated with the Fort Benning road erosion study.

			% of <2mm Mineral Soil									
			Total			Silt		Sand				
Depth (cm)	Horz	Lab Texture	Clay <.002	Silt	Sand .05-2	Fine 002. -.02	Coarse .02-.05	VF .05-.10	F .10-.25	M .25-.50	C .5-1	VC 1-2
0-8	A	ls	3.3	10.7	86.0	6.6	4.1	9.3	24.4	30.2	18.6	3.5
8-20	AE	s	3.6	9.5	86.9	5.5	4.0	9.4	20.3	34.1	18.3	4.8
20-53	E1	cos	3.2	8.5	88.3	6.0	2.5	7.0	22.7	33.6	19.8	5.2
53-100	E2	ls	3.3	10.2	86.5	5.9	4.3	9.8	26.2	29.8	16.6	4.1

Table 3. Locations and descriptions of plots within site locations for the Fort Benning road erosion study.

Site	Plot Name	Plot Sampling Type	Plot Area (m^2)	Plot Length (m)	Plot Road Gradient (%)
GAB – improved site (DMPRC – digital multipurpose range complex)	DMRP1	Automatic	97.93	18.37	11.04
	DMRP2	Automatic	102.71	18.35	11.87
	DM4P1	Manual	58.63	12.30	10.76
	DM4P2	Manual	62.51	12.44	3.66
	DM5P1	Manual	62.91	12.10	7.94
	DM5P2	Manual	62.91	12.10	7.04
Gravel – unimproved road (HG– Hourglass Road)	HGP1	Manual	65.10	12.23	6.32
	HGP2	Manual	67.48	]12.27	6.02
	HGRP3[a]	Automatic	~~342.15~~; 146.68	~~54.41~~; 24.45	~~4.52~~; 3.97
	HGRP4[a]	Automatic	~~213.32~~; 148.56	~~35.66~~; 24.32	~~3.54~~; 3.38

[a] HGRP3 and HGRP4 plots were shortened after the first runoff event overwhelmed the monitoring equipment. The strike through values were before shortening, and the normal values were after shortening.

Figure 3. In the Fort Benning road erosion study: (a) Grab sample plot on the improved site. Note the natural slope break serving as the top of the plot, the metal side gutter, the shallow trench diverting water from the bottom of the plot to the collection point, and the gray color of the compacted aggregate; and (b) military vehicles on the unimproved plot. Note the deposition in the road ditch leading to the inlet to the collection pipe (fig. 5), the rills delivering sediment to the ditch, the ruts forming on the road from the traffic, and the red hue of the weathered loamy sand soil.

dimensions to determine the direction vector for runoff, and the length and slope steepness of each plot's flow path assuming an outsloped road (figs. 1 and 3a, and table 1; Elliot, 2004; Elliot and Hall, 1997). The "effective width" of the plot was calculated from the formulae in table 1. The sensitivity analysis to evaluate whether this diagonal vector with 1-m rill spacing resulted in higher indices of agreement d_r (Willmott et al., 2012) than a vector following the gradient of the road with a 2-m rill spacing in the case of rutting showed that there was minimal difference between the two methods (table 1), so the diagonal flow path directions and slope values were used for all subsequent analyses for all plots (table 1 and fig. 3).

The improved site was on the Digital Multipurpose Range Complex (DMPRC), a site heavily used for military tank training (figs. 2 and 3a). The DMPRC plots were graded in March 2014 and June 2015. The nearby unimproved road site was on Hourglass Road (HG) (figs. 2 and 3b), a heavily trafficked road used by a wide range of vehicles, including trucks of all sizes and tracked vehicles (fig. 3b). Hourglass Road was graveled as needed and frequently graded to prevent excessive rutting, including several times during the study. The two sites selected were within about 3.5 km of each other to facilitate plot management (fig. 2).

The study had two sampling methods: "manual sampling plots" that were monitored by manual grab sampling during runoff events (fig. 3a; table 3), and "automatic sampling plots" that were continuously monitored by recording equipment (figs. 4 and 5, table 3). Runoffs from improved plots were collected from roadside sheet metal gutters (fig. 3a). On the unimproved plots, runoff from the road was routed down the ditches (figs. 3b and 5a) before diverting to the manual or automatic sampling equipment (figs. 4 and 5b). The improved site contained two automatic sampling plots and four manual sampling plots, for a total of six plots. The unimproved gravel site contained two automatic sampling

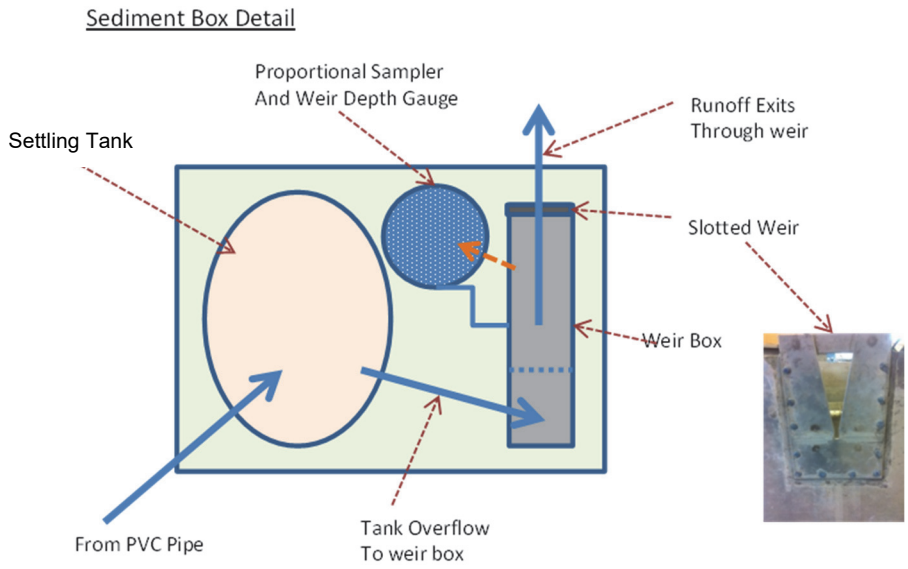

Figure 4. Diagram of components within the sediment box for automatic sampling plots installed at the Fort Benning road erosion study.

(a)

(b)

Figure 5. Automatic sampling installation (a) unimproved site with pipe leading from road ditch to sediment basin, and (b) emptying sediment basin after first major event showing tank used for sediment basin, settling basin above the weir and buckets full of sediment to be weighed or transported to the laboratory to determine water content and dry sediment delivery for the Fort Benning road erosion study.

plots, and two manual sampling plots, for a total of four plots. (table 3).

There were no additions or alterations to the existing road to impede traffic other than a small diversion cross drain at the bottom of each plot (fig. 3a). Occasional plot repair was carried out following damage to the cross drains or the gutters from road traffic or maintenance. No sampling or access was allowed on the improved site when it was in use for training activities, but access was available early mornings, evenings, and on weekends for servicing the automatic samplers and manual sampling if there was a rainfall event.

At manual sampling plots, timed 1-L grab samples of runoff were taken at 5-minute intervals from each plot during a runoff event. Manual sampling plots were installed in pairs so that, for a given storm, a runoff sample could be collected from two plots sequentially.

For automatic sampling plots, runoff was diverted from the road plot to a box containing a 180-L sediment settling tank, a depth data logger monitoring a slotted weir for measuring runoff rate, and a proportional sampler for collecting runoff samples to determine the suspended sediment load leaving the sediment tank (figs. 4 and 5b). Each weir was individually calibrated to determine its flow rate versus depth rating curve. The sediment boxes at each automatic sampling plot were supported by 150 mm diameter, 450 mm deep concrete piers. The automatic sampling plots on the improved site used the natural break in slope as the upper plot boundary, and a 25 mm deep, 100 mm wide diversion trench to define the lower plot boundary. A series of 16-gauge sheet metal gutters on the side of the road section were installed on all improved plots (fig. 3a). Runoff was diverted from the gutters through 200-mm diameter plastic culverts to the sampling boxes that were located 5 to 6 m from the road. The two unimproved automatic sampling plots used grader-incised earthen ditches to route runoff from the road surface to road ditch “turnouts” (fig. 5a). Natural breaks in the slope were used as the upper plot boundaries. A diversion in the turnout was installed to guide the runoff Into a 200-mm diameter plastic culvert that conveyed the runoff to the sampling box located 18 to 20 m from the road (fig. 3). During the first storm event of the study on 28 March 2014, one of the sediment boxes on the automatic plots on the unimproved roads was unable to capture all the sediment generated from a 60-mm rainfall event. Two automatic sampling plots on the unimproved road (HGRP3 and HGRP4) were

shortened for the remainder of the study to prevent overwhelming of the monitoring equipment for all subsequent events (table 3). The sediment yield from the 18 March 2014 event from the unimproved plots was not included in subsequent statistical analyses. The runoff data from this first event, however, were retained as the weir functioned for the higher flow rates as intended. The automatic samplers were removed at the end of March and reinstalled for the June and July collection period. The samplers were removed after the July 2014 sampling period and reinstalled in February 2015, to minimize the risk of equipment damage.

Weather Station and Rain Gages

A weather station was installed in a cleared area in the south corner of the DMRPC site (fig. 2). The station recorded precipitation, temperature, humidity, solar radiation, wind direction, and wind speed. Four additional tipping bucket rain gages were installed near each of the four automatic sampling plot locations to record precipitation. On the manual plots, a portable tipping bucket recording rain gage was set up adjacent to each pair of plots when sampling.

Sediment Concentration and Sediment Water Content

Sediment concentrations in 1-L bottles from grab samples, the settling tank for the automatic samplers (figs. 4 and 5b), and runoff collected by the automatic samplers were analyzed at the USDA Forest Service Rocky Mountain Research Station (RMRS) engineering laboratory in Moscow, ID, U.S. For all but the largest event, the dry mass and the water content of the sediment collected in the sediment basins were also determined in the RMRS laboratory. For the large event on 28 March 2014, on the unimproved site, the wet mass of buckets of wet sediment were measured onsite (fig. 5b). A wet sediment sample was taken from each bucket and stored in a sealed 1-L bottle. The bottles were taken to the laboratory to determine the water content of each bucket full of sediment. All sediment collected in the sediment basins for other events was packed into 20-L buckets for storage and transport to the lab. At the lab, the buckets were emptied onto large metal trays, dried in ovens at 105°C, and the dry mass measured. Runoff collected in bottles from the automatic samplers and the manual grab samples was tightly capped to restrict evaporation. They, along with the sediment samples from the buckets from the 28 March event, were transported to the laboratory and emptied into 1-L beakers, weighed, dried at 105°C and reweighed. The tare weights of all beakers were also determined. All data were entered into spreadsheets for processing to determine runoff rates for the grab samples and sediment flux for all samples.

Data Analysis

Precipitation

Figure 6 is a simplified example of the nature of rainfall data that were observed, along with the runoff response from two plots, and the WEPP-predicted rainfall hyetograph and runoff hydrograph for a rainfall event that occurred on 17 June 2015. Breakpoint format weather files were prepared for all rainfall events, capturing the within-storm cumulative precipitation. The timings of precipitation and runoff were adjusted in figure 6 to account for differences in the clocks in the rain gage and the runoff data loggers. In some cases, with the automatic samplers, rainfall and runoff may have started before midnight on one day and finished after midnight the following day. For such storms, the date assigned to the storm was the day the storm ended.

Runoff and Sediment Delivery

For both manual and automatic plots, instantaneous flow rates, as shown in figure 6a, were integrated in spreadsheets using Simpson's Rule (Whyte, 1976) to determine total runoff. If there were fewer than three observations, then the integration was carried out using the Trapezoidal Rule (Whyte, 1976). Care was needed as there were periods within a single

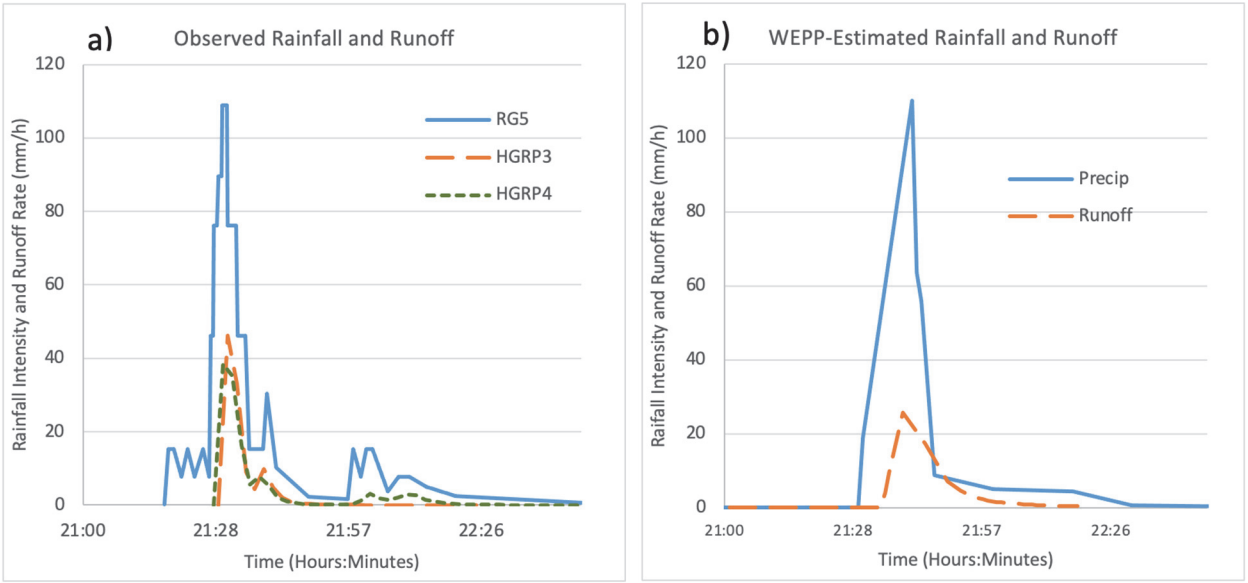

Figure 6. (a) Hyetograph of rainfall during the 17 June 2015 event as measured by rain gage 5 (RG5), the runoff hydrographs observed on unimproved plots HGRP3 and HGRP4; and (b) the internal rainfall hyetograph and runoff hydrograph generated by the WEPP model from the breakpoimt climate file based on the recording rain gage data for the Fort Benning road erosion study.

storm event when precipitation and runoff rates were near or at zero. An example of this is shown for plot HGRP4 in figure 6, where observed runoff stopped at approximately 21:40, but with an increase in rainfall intensity, runoff began again at 22:00. Figure 6b shows that internally, WEPP was limited in its ability to consider such fine detail in the breakpoint rainfall file and transformed the breakpoint data into a double exponential curve for internal analysis. The runoff coefficient *RC* (Observed Runoff ÷ Precipitation) was calculated for each runoff event.

The sediment concentrations from the proportional samplers were multiplied by the runoff rates for each sampling time increment to determine sediment flux, generating a sediment delivery curve similar to the runoff curve in figure 6a. The sediment flux curve was then integrated to obtain the total sediment exiting through the weir. The exiting sediment was added to the sediment that was suspended in the settling tank and the sediment that was deposited in the settling tank and weir box (fig. 4) to determine sediment delivery from each runoff event. Calculated sediment delivery was divided by plot area (table 3) to determine erosion rates (tables 5 and 6). There were no significant differences due to sampling between the manual and automatic samplers for runoff (p = 0.29) or sediment (p = 0.68) measurements (SAS proc GLM; SAS Institute, 2003); therefore, observations from automatic and manual plots were combined for all analyses.

Model Sensitivity, Calibration, and Validation

Sensitivity

A WEPP sensitivity analysis based on road design, topography, and erodibility was carried out for this study's site conditions as recommended by Ascough et al. (2013) (table 4). A 100-year stochastic climate for Talbotton, GA, located 33 km NE of the study area, was used for all analyses (fig. 2). The WEPP:Road batch interface was used for the topographic and road design analyses because of the simplicity of building all of the input files for a single batch run (https://forest.moscowfsl.wsu.edu/cgi-bin/fswepp/wr/wepproadbat.pl). For the road design analysis (inslope, outslope, rutted) the WEPP:Road batch interface was used with the sandy loam, graveled, and high traffic soil properties in the WEPP:Road database (table 4) for all segment designs. The average of the study plot dimensions 15 m long, 5.3 m wide, with a uniform grade of 8%, were entered into the WEPP:Road batch interface. To determine sensitivity to rill spacing, the results for an insloping plot with an eroding ditch segment (rill spacing = 4 m), an outsloping plot with no ruts (rill spacing = 1 m), and a rutted plot (rill spacing = 2 m) were compared (fig. 1 and table 1). The WEPP:Road batch interface was also used for determining the sensitivity to road segment slope length from 5 to 30 m and gradients from 2% to 12%. The WEPP:Road interface does not allow users to alter soil erodibility values, so the WEPP Windows interface (Flanagan et al., 1998) was used for evaluating sensitivity to the WEPP erodibility parameters (K_b, K_i, K_r, and τ_c). An insloping road with an eroding ditch template was used, setting the rill spacing to 4 m, a length of 15 m, and a width of 5.3 m. The baseline erodibility values are in table 4. Baseline erodibility values were multiplied by 0.25, 0.5, 1, 2, 4, and 8 for the study. The multiple of 8 for the interrill erodibility value was not evaluated as it was much greater than had ever been reported (Elliot and Flanagan, 2023; Elliot et al., 1999a; Foltz et al., 2011).

Table 4. Baseline values for a sensitivity analysis of WEPP when applied to road segments. Erodibility values are those for a high traffic graveled sandy loam soil in the WEPP:Road interface. The climate was a 100-yr stochastic weather sequence based on Talbotton, GA, climate data. Topographic values are the mean plot slope length and steepness values from the Fort Benning road erosion study.

Property	Value
Length, width, and gradient	15 m, 5.3 m, and 8%
Base hydraulic conductivity K_b	10.2 mm h^{-1}
Interrill erodibility K_i	2,000,000 kg s m^{-4}
Rill erodibility K_r	0.0004 s m^{-1}
Critical shear τ_c	2 Pa

The sensitivity S (eq. 6) was calculated for the road design (table 1; 1, 2, or 4 m rill spacing). For topographic and erodibility analyses, the WEPP input values and sediment delivery amounts were compared graphically. For some graphic analyses, it was useful to "normalize" the results with the normalized input defined as the input value divided by the baseline input value (table 4), and for the normalized runoff or sediment delivery, the estimated value for a given set of inputs was divided by the output value for the baseline conditions.

Calibration

The potential data set for calibration included all the automatic and manual runoff events, plus the four largest rainfall events that had no observed runoff for both treatments to aid in determining the threshold hydraulic conductivity (tables 5 and 6). Calibration of the WEPP model for runoff amounts was carried out with approximately half of the plots at each site randomly assigned to "calibration" with the other half saved for "validation" (tables 5 and 6). Plots that measured runoff greater than precipitation were not used for calibration but were included in the validation data set. Potential reasons for observed runoff greater than precipitation are in the discussion section. Single storm "breakpoint" precipitation files were built for each storm from the recording rain gages nearest each plot (Flanagan and Livingston, 1995). Other weather information on the day of each event required for the WEPP weather file (maximum and minimum temperatures, wind speed and direction, humidity, and solar radiation) were obtained from the main site weather station.

The WEPP Windows interface (Flanagan et al., 1998) has a feature that allows the user to complete multiple runs as a batch. A batch run was set up for the 13 unimproved calibration plot events (table 6), and a second batch was set up for the 15 improved calibration plot events (table 5). Plot dimensions (table 3) and breakpoint weather files (tables 5 and 6; fig. 6) were defined for each of the calibration plot events. The soil input file was common for all plots for a given treatment (13 unimproved or 15 improved). With a common soil file, all the calibration runs were completed as a single batch run for a given treatment, and the runoff and sediment delivery from every event were generated as a single table. The predicted runoff and sediment yield amounts were copied from the WEPP Windows batch output, pasted into a spreadsheet to be compared to observed values for each event.

Table 5. Summary of events for the Improved (GAB) plots. Model designation is calibration (C) or validation (V); *RC* is runoff coefficient (*RC* = observed runoff ÷ precipitation) for the Fort Benning road erosion study. The sources of the predicted runoff and sediment delivery values are described in the methods and results sections.

Plot/ Sampling	Year	Month	Day	Model	Precipitation (mm)	Observed Runoff (mm)	Predicted Runoff (mm)	Observed Sediment (Mg ha^{-1})	Predicted Sediment (Mg ha^{-1})	*RC*
DMRP1 Automatic	2014	3	6	C	11.18	0	1.92	0	0.01	0.00
	2014	3	16	V	58.93	53.98	41.05	3.22	5.04	0.92
	2014	3	28	C	49.02	26.59	25.07	0.34	1.56	0.54
	2014	7	19	V	29.21	22.46	17.08	2.93	2.04	0.77
	2014	7	28	V	9.91	0	4.72	0	0.39	0.00
	2015	3	5	C	9.65	4.66	5.76	1.53	0.75	0.48
	2015	3	12	V	21.84	7.49	8.15	0.30	0.50	0.34
	2015	6	3	C	10.41	0	6.89	0	0.86	0.00
	2015	6	8	C	9.91	0	3.83	0	0.19	0.00
	2015	6	9	C	12.19	5.56	4.31	1.37	0.29	0.46
	2015	6	28	V	18.03	6.52	6.05	2.12	0.74	0.36
	2015	7	4	C	5.33	1.33	2.08	0	0.21	0.25
	2015	7	4	C	9.14	3.90	5.11	0.15	0.70	0.43
	2015	7	15	V	28.19	12.67	22.01	7.48	3.09	0.45
DMRP2 Automatic	2014	3	16	C	58.93	53.45	41.05	5.81	5.66	0.91
	2014	3	28	V	49.01	23.01	25.07	0.46	1.97	0.47
	2014	7	19	C	29.21	11.07	17.08	0.91	2.32	0.38
	2015	3	5	V	9.65	5.16	5.76	2.37	0.84	0.53
	2015	3	12	V	21.84	10.75	8.15	0.53	0.63	0.49
	2015	6	9	C	12.19	3.72	4.31	0.32	0.36	0.30
	2015	6	28	V	18.03	3.72	6.05	0.39	0.83	0.21
DM4P1 Manual	2014	3	16	V	26.16	23.87	14.28	2.33	1.00	0.91
	2014	3	29	V	2.54	1.25	0	0.01	0	0.49
DM4P2 Manual	2014	3	16	V	26.16	21.15	13.23	1.18	0.08	0.81
	2014	3	29	V	2.54	0.54	0	0.01	0	0.21
DM5P1 Manual	2014	3	12	C	6.60	1.25	1.53	0.18	0.09	0.19
	2014	6	30	C	25.40	18.66	18.54	2.40	2.00	0.73
	2014	7	20	V	4.06	0.81	0	0.01	0	0.20
DM5P2 Manual	2014	3	12	V	6.60	1.19	1.25	0.44	0.05	0.18
	2014	6	30	C	25.40	19.40	18.27	2.63	1.74	0.76
	2014	7	19	V	16.26	21.16[a]	10.40	1.42	0.69	1.30
	2014	7	20	C	4.06	1.19	0	0.02	0	0.29
Mean of non-zero runoff events					18.15	10.46	9.76	1.42	1.01	0.49

[a] More runoff than precipitation, and not used for calibration.

Willmott et al.'s (2012) index of agreement (eq. 7) was calculated from the runs. The soil file was altered in subsequent runs by either increasing or decreasing the specific erodibility value under calibration to obtain a maximum value for the index of agreement. The index of agreement d_r was defined as (Willmott et al., 2012):

$$d_r = \begin{cases} 1 - \dfrac{\sum_{i=1}^{n} \left| P_i - O_i \right|}{\sum_{i=1}^{n} \left| O_i - \overline{O} \right|}, \text{ when } \sum_{i=1}^{n} \left| P_i - O_i \right| \le 2 \sum_{i=1}^{n} \left| O_i - \overline{O} \right| \\ \dfrac{\sum_{i=1}^{n} \left| O_i - \overline{O} \right|}{\sum_{i=1}^{n} \left| P_i - O_i \right|} - 1, \text{ when } \sum_{i=1}^{n} \left| P_i - O_i \right| > 2 \sum_{i=1}^{n} \left| O_i - \overline{O} \right| \end{cases} \quad (7)$$

where

d_r = index of agreement

P_i = predicted runoff or sediment delivery for the i^{th} storm

O_i = observed runoff or sediment delivery for the i^{th} storm

$\overline{O}$ = mean of observed values

n = number of observations

i = summation index.

The range of d_r is from -1 to +1. A value of zero indicates the sum of the magnitudes of the model errors and the observed deviation magnitudes are equal (Willmott et al., 2012). Values greater than zero suggest a good fit, and values less than zero suggest a poor fit. "A value for d_r of +0.5, for example, indicates that the sum of the error-magnitudes is one half of the sum of the perfect-model-deviation and observed-deviation magnitudes." (Willmott et al., 2012). This index was selected because it considers the linear rather than the square of differences between observed and predicted values, so the selection of input variables will not be disproportionately influenced by very large events, as is the case with most goodness-of-fit statistics.

Runoff was first optimized for a given road design set (unimproved or improved). For runoff predictions, the baseline hydraulic conductivity K_b in the WEPP:Road sandy loam soil file (10 mm h^{-1}) was manually adjusted in increments of 0.1 mm h^{-1} or more to increase or decrease estimated runoff, and WEPP was run for each plot and storm to maximize d_r (eq. 7). The 0.1 mm h^{-1} incremental minimum was considered sufficiently precise because of the inherent variability in soil erodibility properties; any greater precision was not necessary (Bosch and West, 1998; Morgan and Nearing, 2011). K_b was the only input variable in the soil file that affects runoff (Ascough et al., 2013; Nearing et al., 1990), so there was no chance of more than one optimal solution for hydraulic conductivity. Once the optimal value for K_b was estimated, the first soil erodibility calibration step

Table 6. Summary of events for the unimproved (gravel + grading) plots. Model designation is calibration (C) or validation (V); *RC* is runoff coefficient (*RC* = observed runoff ÷ precipitation) for the Fort Benning road erosion study. The sources of the predicted runoff and sediment delivery values are described in the methods and results sections.

Plot/ Sampling	Year	Month	Day	Model	Precipitation (mm)	Observed Runoff (mm)	Predicted Runoff (mm)	Observed Sediment (Mg ha^{-1})	Predicted Sediment (Mg ha^{-1})	*RC*
HGP1	2014	3	6	V	8.13	12.23[a]	0	1.57	0	1.50
Manual	2014	7	11	C	16.51	10.62	8.49	4.38	7.27	0.64
	2015	3	5	V	4.83	1.04	0	0.22	0	0.22
HGP2	2014	3	6	C	8.13	6.62	0	0.63	0	0.81
Manual	2014	7	11	V	16.51	10.23	7.64	3.78	5.13	0.62
	2015	3	5	V	4.83	0.93	0	0.13	0	0.19
HGRP3	2014	3	28	V	46.23	36.29	9.20	7.10	3.50	0.78
Automatic	2014	6	30	C	10.41	0.71	0.65	1.55	0.25	0.07
	2014	7	11	C	16.00	6.04	6.06	8.09	2.26	0.38
	2014	7	19	C	31.50	17.01	9.90	15.80	3.75	0.54
	2015	3	5	V	7.87	4.05	0.43	7.16	0.16	0.51
	2015	3	12	V	22.86	0	2.76	2.86	1.05	0.00
	2015	6	9	V	33.53	17.64	19.74	13.70	7.32	0.53
	2015	6	17	V	17.53	3.98	3.82	3.41	1.45	0.23
HGRP4	2014	3	28	V	46.23	47.33[a]	9.06	5.54	2.87	1.02
Automatic	2014	6	6	C	11.18	0	0	0	0	0.00
	2014	6	8	C	14.73	0	4.68	0	1.49	0.00
	2014	6	23	C	39.37	13.90	22.35	11.03	6.80	0.35
	2014	6	30	V	10.41	0.71	0	1.53	0	0.07
	2014	7	11	C	13.72	3.23	3.41	3.67	1.09	0.24
	2014	7	19	C	30.99	11.02	8.19	7.32	2.60	0.36
	2014	7	21	C	33.02	18.86	20.57	10.92	6.30	0.57
	2015	3	5	C	7.87	2.09	0.29	4.66	0.09	0.26
	2015	3	12	V	22.86	9.80	2.62	3.38	0.84	0.43
	2015	6	9	V	33.53	13.75	19.59	10.03	6.12	0.41
	2015	6	17	V	17.53	4.63	3.68	1.60	1.18	0.26
	2015	6	28	V	10.41	0.12	0	0.50	0	0.01
	2015	7	2	V	7.87	0	0	0	0	0.00
	2015	7	4	C	7.87	0	0	0	0	0.00
	2015	7	15	V	36.07	11.42	27.16	18.07	8.19	0.32
Mean of non-zero events					24.48	9.74	8,77	7.11	2.88	0.39

[a] More runoff than precipitation, and not used for calibration or statistical analysis.

was to estimate rill erodibility K_r, as this was the most sensitive soil erodibility variable affecting sediment delivery (Ascough et al., 2013; Nearing et al., 1990). For the K_r calibration, critical shear was initially set to 2 Pa, and interrill erodibility was initially estimated to be 1,000,000 kg s m^{-4}. Once the rill erodibility was determined for both treatments, optimal values were determined for critical shear and then for interrill erodibility. Rill erodibility was checked a final time to ensure that it was still the value that resulted in the maximum index of agreement.

STATISTICAL ANALYSES

The runoff depths and sediment yields were log-transformed to homogenize the variance of the residuals (Helsel and Hirsch, 2002). Small values equivalent to 1/10th of the minimum non-zero observed runoff and sediment yield values (runoff, 0.01 mm; sediment yield, 0.001 Mg ha^{-1}) were added to all runoff and sediment yield data so that the zero-value data could be log-transformed. The statistical models for the transformed runoff or sediment yield used plot design (unimproved or improved) as the fixed effect, and the precipitation depth or the 30-minute rainfall intensity (I_{30}) for each event were used as covariates (Littell et al., 2006). The covariance structure of the repeated measures on each plot was modeled using a compound symmetry function and the Julian date. The plot within each site was a random effect in the model. Differences in the log-transformed runoff and sediment yields were compared using the least squares mean estimates for each road design. A Tukey–Kramer adjustment was used for comparisons of least-squares means. The significance level was 0.05 for all comparisons.

The model performance was tested with observations for the validation events, with 17 events on the unimproved plots and 16 events on the improved plots (tables 5 and 6). The goodness of fit for runoff and sediment delivery for both road treatments was evaluated by linear regression, noting the coefficient of determination (R^2), the slope of the regression line (a), and the y-axis intercept of the regression line (b). A second evaluation of the validation data set was with the Nash-Sutcliff Efficiency (*NSE*) coefficient, which determines how the modeling error compares to the experimental error. (Nash and Sutcliffe, 1970) *NSE* is calculated from the relationship:

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(O_i - P_i\right)^2}{\sum_{t=1}^{n}\left(O_i - \bar{O}\right)^2} \quad (8)$$

where

O_i = Observed runoff or sediment delivery for event i

P_i = Predicted runoff or sediment delivery for event i

$\bar{O}$ = Mean of the observed runoff or sediment delivery for the treatment

n = Total number of observations.

The goodness of fit was also evaluated by comparing the root mean square error (*RMSE*) to the mean values for the runoff and sediment delivery for both treatments (Elliot et al., 1991). *RMSE* is a measure of the difference between each estimated and observed value, and is calculated from:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(P_i - O_i\right)^2}{n}} \quad (9)$$

where the variables are the same as equation 8. The *RMSE* for each treatment was divided by the mean to normalize the statistic to allow comparisons within this study and with other studies.

RESULTS

FIELD OBSERVATIONS

During March, June, and July 2014, there were 40 days when precipitation was measurable by at least one of the rain gages. Runoff was collected on 7 of those days on the improved road plots and 8 days on the unimproved plots (tables 5 and 6). In 2015, there were 49 days when precipitation was measured by at least one rain gage, and runoff was only collected on 6 days on both the improved and unimproved plots. There was likely one additional runoff event from an improved plot on 4 July 2015, as sediment data were collected, but the flow recorder malfunctioned. The precipitation depths in tables 5 and 6 were from the nearest functioning rain gage to the plot where the runoff was observed. On a given day, the rain gages did not all record the same amounts. For example, on 28 March 2014, the rain gages measured 49.02 and 49.01 mm on the improved plots, and 46.23 mm on the unimproved plots. Such variation was observed throughout the study due to the nature of the small, intense thunderstorms that were typical of this climate (Trewartha and Horn, 1980). The spatial variability in rainfall resulted in dates when runoff was recorded at one site, and not at the other. Plot design (table 3) and local microtopography such as ruts and rills (fig. 3) may also have contributed to the variation in runoff response, particularly from smaller events. For example, on 11 July 2014, runoff was collected from both automatic plots and a manual plot on the unimproved site (table 6), but there was no runoff from improved plots (table 5). The 21 July 2014 runoff event recorded on plot HGRP4 was not recorded on the other three plots as there was insufficient time to clean out the accumulated sediment from the 19 July storms on all four sites before 21 July. On 4 July 2015, there were two separate small rainfall and runoff events recorded on plot DMRP1, but no runoff was measured on DMRP2 (table 5) and no runoff from the unimproved plots (table 6) from either rainfall event on that day.

On the improved plots (table 5), average storm runoff was 10.5 mm from 28 observations, averaging 18.2 mm of rainfall per event, while average runoff from unimproved plots was 9.7 mm from 25 observations with an average precipitation depth of 24.5 mm (table 6). The averaged runoff coefficients were 0.39 for the unimproved roads and 0.49 for the improved roads (tables 5 and 6). Even though the improved roads generated more runoff, the sediment yields were less than the unimproved plots (tables 5 and 6; fig. 7). The average observed storm sediment deliveries were 1.42 Mg ha^{-1} from the improved site, compared to 7.11 Mg ha^{-1} from the unimproved/graveled plots during the study period (tables 5 and 6).

The relationships between runoff and precipitation, and sediment yield and runoff are shown in figure 7. The coefficients of determination (R^2s) were greater than 0.7 for runoff versus precipitation for both road designs. For sediment yield versus runoff, the R^2 values were 0.69 for the unimproved road and 0.38 for the improved road design.

There were three instances where more runoff was collected than precipitation (tables 5 and 6). We did not use these data in the statistical or calibration analyses but have reported the values to give the reader a sense of some of the challenges associated with collecting road erosion data. The three events, however, were used for the validation analysis.

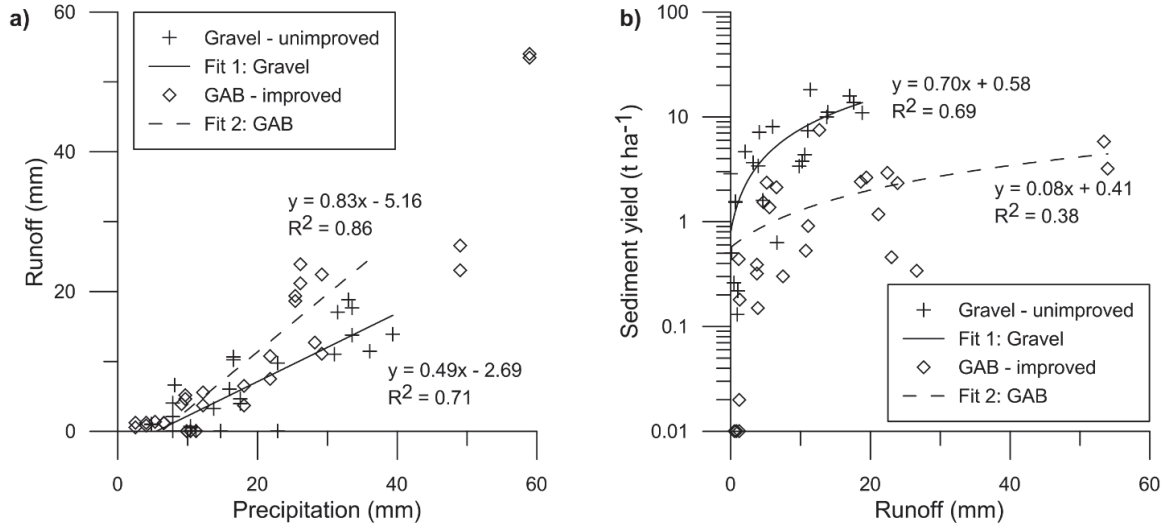

Figure 7. Relationships between (a) precipitation and runoff and (b) runoff and sediment yield by treatment for the road erosion study on Fort Benning. Note the log scale on the sediment yield axes. The two largest events on the unimproved plots on 28 March 2014 were not included in the sediment yield analysis.

Analysis of the sediment collected in the sediment box system (figs. 4 and 5) revealed that, on average, 79% of the collected sediment was deposited in the settling tank, with less than 1% remaining in suspension in the tank. Runoff leaving the sediment box system transported the remaining 20% through the slotted weir. Larger runoff events tended to transport a greater percent of the sediment through the weir, ranging from less than 1% for rainfall events less than 12 mm to more than 30% of the measured sediment for events with greater than 50 mm of rainfall.

Table 7 shows the results of applying a Linear Mixed Model (LMM) to the data set. Precipitation depth was the most significant factor in determining runoff and sediment yield from the plots (p<0.0001) (table 7). Rainfall intensity (I_{30}) was also significant for predicting runoff (p = 0.001) and sediment yield (p = 0.005). However, precipitation depth and rainfall intensity were not independent of one another, and only one could be used in a final model; therefore, precipitation depth was used for subsequent analysis because it was more significant. There were no differences due to the sampling method (manual or automatic), so results of both methods were combined for comparing road designs. Road design alone did not have a significant effect on sediment yield (p = 0.07) or runoff (p = 0.31) (table 7); however, there was an interaction effect between precipitation depth and road design, necessitating the interaction term in the final model. The interactions for both runoff and sediment yield can also be seen in figures 7a and 7b, where the regression lines for the two road designs are not parallel.

The least square (LS) means in table 7 for runoff and sediment delivery are not the same as the arithmetic means in tables 5 and 6 due to the imbalanced statistical design and the necessity to use precipitation as a covariate but are a more appropriate representation of the mean values for this study (https://webpages.uidaho.edu/cals-statprog/sas/workshops/glm/lsmeans.htm). The LS Means in table 7, however, show the same relationships as the arithmetic means in tables 5 and 6, with more runoff (14.9 mm vs. 6.5 mm) for a given event, but much less sediment (0.026 Mg ha^{-1} vs. 2.27 Mg ha^{-1}) from the improved roads. Table 7 shows that there was no significant difference in runoff due to road design, but that the sediment generated by the improved road design was significantly less than the unimproved road.

Sensitivity Analysis

The sensitivity of runoff and sediment delivery to road design, which includes rill spacings, is presented in table 1. The runoff was not sensitive to road design, with a sensitivity parameter value (eq. 6) of S = -0.02 between 1 and 2 m rill spacing (outslope and rutted), and +0.05 between 2 and 4 m spacing (rutted and inslope). The outslope design estimated the least runoff of 232 mm and the inslope to a bare ditch the greatest runoff of 267 mm from an average annual precipitation of 1351 mm (table 1). The runoff from the insloped design to a rocked or vegetated ditch generated even less runoff (190 mm). The sediment delivery was also similar for the three designs, with sediment delivery estimates for outslope, inslope, and rutted designs of 29, 32, and 31 Mg ha^{-1}, respectively, with S = 0.02. Adding rock or vegetation to the ditch decreased sediment delivery from 32 to 23 Mg ha^{-1} for the inslope scenario. The outslope scenario with the shortest flow path length (12 m) and smallest rill spacing (1 m) generated the lowest runoff and least sediment delivery, whereas the inslope scenario with the widest rill spacing (4 m) generated the most runoff and sediment (table 1).

The mean S values for the other variables are summarized in table 8. Averaged over the entire range of input values, road runoff was sensitive to base hydraulic conductivity (K_b), and sediment delivery sensitive to road segment length and gradient, hydraulic conductivity (K_b), rill erodibility (K_r), and critical shear (τ_c). Sediment delivery was not sensitive to interrill erodibility (K_i).

The results of the sensitivity analysis of road segment length and steepness on sediment delivery are graphed in figure 8 for a range of slope lengths (5–30 m) and gradients (2%–12%). The 2% gradient appears nearly constant at 12-13 Mg ha^{-1}, with an average sensitivity to slope length of S = 0.08. The low slope suggests that sediment delivery is

Table 8. Summary of sensitivity parameter S (eq. 6) for the WEPP model for a road with the base conditions of a sandy loam, high traffic graveled road (table 4) with a base gradient of 12% and length of 15 m. Only variables with $S \geq 0.2$ or $S \leq -0.2$ are listed.

Road Design (fig. 1)	Input Variable	Mean Sensitivity	
		Runoff	Sediment
Insloped eroding ditch	Segment length		0.37
	Segment Gradient		0.75
Outsloped Unrutted	K_b	-0.88	-0.66
	K_r		0.2
	τ_c		-0.2

Table 7. Linear modeling results for runoff and sediment yield; covariate testing (top) and final model (bottom). P-values are significant at 0.05 for the Fort Benning road erosion study. Least square mean estimates of the runoff and sediment yields with different letters in the same row are significantly different.

	Response Variable				
	Runoff		Sediment Yield		
	F-value	p-value	F-value		p-value
Model covariates					
Precipitation	30.41	<0.0001	24.31		<0.0001
I30	12.55	0.0011	14.39		0.0005
Sampling Method	1.05	0.31	0.29		0.59
Road Design	1.03	0.31	3.45		0.07
Precipitation*Design	15.33	<0.0001	15.95		<0.0001
	Least square mean estimates		p-value		
Final model	Improved	Unimproved	Precipitation	Design	Precipitation*design
Runoff = Precipitation | Design (mm)	14.9 a	6.5 a	<0.0001	0.29	0.34
Sediment = Precipitation | Design (Mg ha^{-1})	0.026 a	2.27 b	<0.0001	0.63	0.19

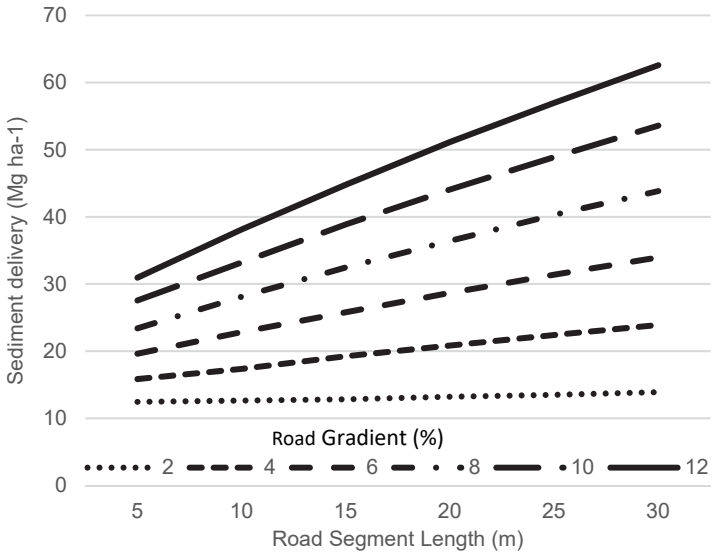

Figure 8. Sensitivity analysis showing sediment delivery estimated by WEPP:Road (Elliot, 2004) from a high traffic graveled sandy loam soil road segment for different segment lengths and road gradients, assuming an insloping road segment with an eroding ditch and a 100-y stochastic climate based on Talbotton, GA weather statistics. Fill slope and forest buffer elements were both set to a 0.3-m length and 12% slope.

due to interrill erosion only because the hydraulic shear does not likely exceed the critical shear in the ditch (eq. 3), so the erosion rate does not increase with increased segment length. The 4% gradient line has an inflection point at 10 m, with interrill erosion likely in shorter lengths, but rill erosion becoming dominant in segments longer than 10 m. S ranged from 0.27 for the 4% gradient to 0.44 for the 12% gradient.

The sensitivities of the WEPP model to base saturated hydraulic conductivity on runoff and sediment delivery are shown in figure 9a, and to interrill and rill erodibility and critical shear in figure 9b. The most striking feature of figure 9a is the major impact of base hydraulic conductivity K_b on not only runoff but also sediment delivery. The normalized curves are nearly identical for the two outputs. The mean sensitivities of runoff and sediment delivery to K_b were -0.88 and -0.66, respectively (table 8), greater than for any of the other parameters that were evaluated. The greatest sensitivity of K_b with S = -1.59 was for the estimated runoff that dropped from 78 to 24 mm by increasing K_b from 40 to 80 mm h^{-1} (fig. 9a). The same was true for sediment with S = -1.42 when estimated sediment delivery decreased from 9.8 to 3.5 Mg ha^{-1} following an increase in hydraulic conductivity from 40 to 80 mm h^{-1}.

Changes in the three soil erodibility values showed different responses in sediment yield (fig. 9b). Sediment delivery showed the greatest sensitivity to rill erosion from N = 1 to N = 2 (K_r = 0.0004 to 0.0008 s m^{-1}) with S = 0.45. There was little additional increase in sediment yield from N = 2 to N = 8, where S decreased from 0.07 to 0.0. Interrill erodibility K_i showed a similar trend, with a large response in sediment delivery from N = 0.25 to N = 0.5 (fig. 9b; S = 0.37), but with no response from additional increases in sediment delivery for larger values of K_i. Critical shear τ_c showed the greatest response in decreasing sediment delivery as values increased from N = 0.25 (0.5 Pa) to N = 1 (2 Pa), and for τ_c values greater than 2N (4 Pa).

CALIBRATION

Table 9 presents calibrated erodibility values for unimproved and improved plots, and the Willmott et al. (2012) indices of agreement d_r (eq. 7) for both calibration and validation. The hydraulic conductivity was greater on the unimproved road segments (3.0 versus 1.3 mm h^{-1}), and the rill erodibility was higher for the unimproved segments (0.09 versus 0.0008 s m^{-1}). The critical shear was greater on the improved plots (1.5 Pa) than on the unimproved plots (0.0001 Pa). We could not improve the index of agreement by adjusting the interrill erodibility value for either road design.

Figure 10 shows the relationships between the predicted and observed runoff amounts and sediment yields for the combined calibration and validation data sets for both road designs. Predicted versus observed runoff on the improved plots had a stronger relationship (R^2 = 0.89) than on the unimproved gravel site (R^2 = 0.25), and the slope of the regression line was greater for the improved plots than the unimproved plots (0.73 vs. 0.36). The sediment delivery prediction for the unimproved plots had an R^2 value of 0.58 and a regression line slope of 0.6 compared to the improved plot estimates (R^2 = 0.64 and slope = 0.43). The only intercept that was significantly different from zero was the intercept for the runoff from the improved plot (b = 2.19; (P($b \neq 0$) = 0.014).

Table 10 shows the results of an analysis of variance comparing the logs of observed and predicted runoff depths and sediment delivery. There were no differences between the logs of observed and predicted runoff and erosion rates. The analyses of the interactions were carried out to see if any observed differences between runoff and erosion estimates within the data set could be attributed to plot length, slope, treatment (unimproved versus improved), or precipitation. The only significant interaction ($P < 0.05$) was with the plot slope for sediment yield. Figure 11 shows why there was a significant interaction between observed and predicted

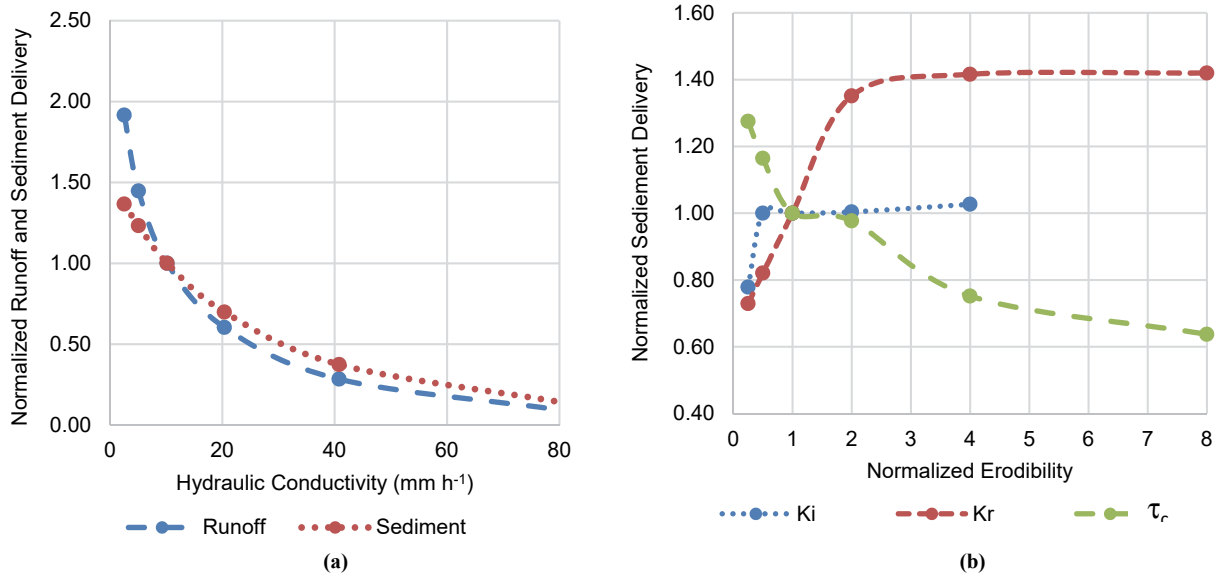

Figure 9. (a) Normalized runoff and sediment delivery vs hydraulic conductivity as predicted by the WEPP model for a sandy loam road soil with gravel and high traffic (table 4). Road width = 5.3 m, length = 15 m, insloping design to an eroding ditch. "Normal" values (N = 1) were 275 mm for average annual runoff from 1351 mm precipitation and 26 Mg ha^{-1} for sediment yield; and (b) Normalized sediment delivery vs. normalized values for interrill erodibility K_i, rill erodibility K_r and soil critical shear τ_c. "Normal input values" (N = 1) were K_i = 2,000,000 kg s m^{-4}, K_r = 0.0004 s m^{-1}, and τ_c = 2 Pa.

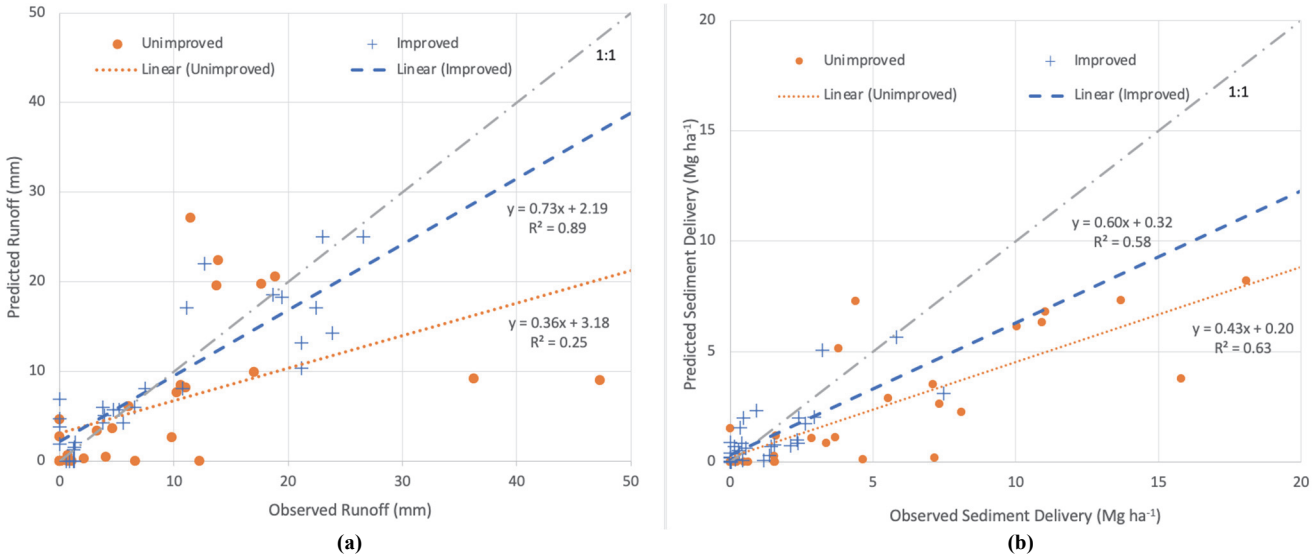

Figure 10. Predicted versus observed (a) runoff and (b) sediment yield for both calibration and validation data sets for the unimproved and improved road segments for the Fort Benning road erosion study.

Table 9. Calibrated soil erodibility values and Willmott et al. (2012) Indices of Agreement d_r for calibration and validation of unimproved and improved plots on the Fort Benning road erosion study.

Variable	Improved	Unimproved
Hydraulic Conductivity (mm h^{-1})	1.3	3.0
d_r for runoff for Calibration	0.866	0.770
and Validation	0.818	0.624
Rill erodiblity (s m^{-1})	0.0008	0.09
Critical shear (Pa)	1.5	0.0001
Interrill erodibility (kg s m^{-4})	1,000,000	1,000,000
d_r for sediment delivery for Calibration	0.790	0.584
and Validation	0.664	0.670

Table 10. Statistics of observed versus predicted runoff and sediment delivery logartihm analyses based on Type III SSE for the Fort Benning road erodibility study.

	Response Variable			
	logRunoff		logSediment Yield	
Model	F-value	p-value	F-value	p-value
ObsPred	0.21	0.65	1.28	0.26
ObsPred * Length	2.02	0.16	0.05	0.83
ObsPred * Slope	1.56	0.21	4.26	0.04
ObsPred * Treat	0.85	0.36	3.25	0.07
ObsPred * Precip	1.56	0.21	1.14	0.29

erosion rates and the plot slope in table 10. From the trend lines in figure 11, the observed erosion rates decreased with plot slope at a greater rate than the predicted erosion rates. This interaction and the negative relationship observed between sediment delivery and slope steepness will be discussed further in the next section.

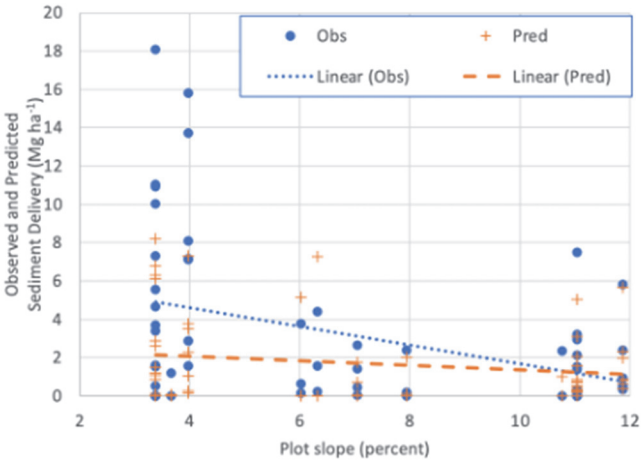

Figure 11. Observed and predicted sediment delivery versus road segment slope for the Fort Benning road erodibility study.

VALIDATION

The goodness-of-fit for both the calibration and validation data sets for the Willmot et al. (2012) index of agreement d_r were presented table 9. The values for d_r were generally greater for the improved plots, with the lowest d_r (0.62) for validating runoff from the unimproved plots, and the greatest d_r (0.82) for validating runoff from the improved plots. The summary of another goodness-of-fit analysis was presented in figures 10a and 10b with linear regression statistics for the combined calibration and validation plots from tables 5 and 6. Figure 10a shows that the lowest R^2 value (0.25) was for the unimproved road runoff predictions and the best R^2 value (0.89) was for runoff estimates for the improved road runoff events. The slope a of the regression lines that was closest to 1.0 was 0.73 for runoff estimates for the improved road plots (fig. 10a). The a value that was furthest from 1.0 was 0.36 for the estimated runoff from the unimproved plots. Table 11 presents goodness-of-fit statistics for the validation events only from tables 5 and 6. There were 17 validation events on unimproved plots and 16 events on improved plots. Estimating runoff from the unimproved plots had the poorest result with a low R^2 (0.18), the lowest *NSE* (0.09), and the highest *RMSE*:*Mean* ratio (1.2). Runoff from the improved plots had the best validation results, with $R^2 = 0.87$, *NSE* = 0.84 and lowest *RMSE*:*Mean* ratio (0.45). Sediment prediction success was similar for both road designs with *NSE* = 0.41, and other goodness-of-fit statistics were similar for both treatments (table 11). In all cases, the slope of the regression line a was less than 1, indicating that the largest values were under predicted, while none of the regression intercepts b were significantly different from zero. It is notable that even though the estimated runoff from the unimproved validation plots was poorly predicted, the estimates for sediment delivery were reasonable.

Table 11. Goodness-of-fit analyses for the validation data sets from tables 5 and 6 for the Fort Benning road erosion study. *NSE* is the Nash-Sutcliff Efficiency (eq. 8; Nash and Sutcliffe, 1970), *RMSE* is the root mean square error (eq. 9), and d_r is the Willmott et al. (2012) index of agreement (eq. 7).

	Runoff		Sediment Delivery	
Analysis	Unimproved	Improved	Unimproved	Improved
Mean	10.24 mm	12.16 mm	4.72 Mg ha^{-1}	1.49 Mg ha^{-1}
n	17	16	17	16
Linear Regression[a]				
a	0.27	0.73	0.49	0.48
b	3.50	1.87	-0.09	0.36
R^2	0.18	0.87	0.79	0.47
NSE	0.09	0.84	0.41	0.41
RMSE	12.29	5.48	3.76	1.45
RMSE:Mean	1.20	0.45	0.79	0.97
d_r	0.62	0.82	0.67	0.66

[a] $y = ax + b$

DISCUSSION

The purpose of this study was to support a watershed modeling analysis to evaluate the ability of the WEPP model to estimate sediment generated by high traffic roads, and to determine the erodibility of high-traffic gravel roads. We compared runoff and sediment delivery from two gravel road designs with high-traffic and evaluated the ability of WEPP to account for the effects of weather, topography, and management on road sediment delivery. The hypotheses were that improved road designs will result in less erosion and that WEPP can predict the effects of weather, topography, and management on sediment delivery. The results showed that the improved road design resulted in a reduction in sediment delivery by more than 98% (table 7). WEPP was able to model differences due to weather, plot length, and management, but the data set could not account for differences due to plot steepness (table 9). The validation study, however, clearly showed that WEPP was able to address the effects of road segment length and gradient on sediment delivery (fig. 8). The reasons that the observed data did not demonstrate WEPP's topographic capabilities are discussed below.

DATA COLLECTION

Precipitation was the single most important factor influencing both runoff and sediment delivery, as it was an overwhelming covariate (table 7). Measuring precipitation proved to be a challenge. The local climate was dominated by small thunderstorms that could result in significant precipitation at one site but not another. For example, on 9 June 2015, recorded precipitation was 12 mm on the improved plots (table 5) but 34 mm on the unimproved plots (table 6). This highly variable but locally intense rainfall pattern made collecting data manually challenging, and manual data were only collected from one event in the three-month sampling period in 2015 (table 6). We found that there could be large differences in measured precipitation between nearby gauges but were not able to determine if these differences reflected spatial variability in precipitation, or a malfunction in a rain gauge as the differences were not consistent. Elliot and Rhee (2022) reported that numerous rain gages failed during the 3-yr study, but they had installed six rain gages per treatment to offset anticipated challenges in measuring precipitation beneath a forest canopy. Ciach (2002) described numerous reasons for erroneous readings from tipping bucket rain gauges as well as problems with using a point measurement for the surrounding area. When determining which rain gauge to use for the WEPP calibration/validation analysis, we considered the timing of the rainfall compared to the timing of the runoff event as well as the rainfall depth (fig. 6a), particularly if one rain gauge was

measuring less precipitation than the observed runoff (tables 5 and 6).

WEPP's internal smoothing of the breakpoint file (fig. 6b) was likely a source of error for the calibration/validation analysis. We were able to describe the variability of rainfall intensity within a storm using WEPP's breakpoint input format, but internally, the WEPP model smoothed over some of that variability to create a uniform rainfall hyetograph as shown in figure 6b. Despite these known sources of experimental error, we felt that the erodibility values we calculated were reasonable estimates and the indices of agreement for both calibration and validation were acceptable (tables 9 and 11). Nicks et al. (1995) described similar challenges to modeling storms with the WEPP climate format and arrived at the same conclusion. Since the earliest days of the development of WEPP technology, storm variability has been identified as a challenge to erosion modeling (Nicks et al., 1995). Another challenge associated with single-storm calibration is that one of the important inputs to the single storm version of the WEPP model is the soil water content prior to the rainfall event. Some sampling of the water content of road soils was carried out prior to some rainfall events. Because of the highly variable and unpredictable nature of precipitation, it was not possible to collect a soil water content value before every storm, so it was necessary to estimate this value for the WEPP input soil file for each event based on recent weather when soil water content had not been measured. Such known sources of error in measuring and modeling precipitation, runoff, and erosion are not unique to this study (Elliot and Rhee, 2022; Grace, 2017; Robichaud et al., 2007), but contribute to the overall error associated with any erosion study or predictive technology. This inherent variability underscores the importance of interpreting any estimate of erosion as representing a range of potential outcomes (Brazier et al., 2000; Nearing, 1998; Robichaud et al., 2007).

There were three events when observed runoff was greater than precipitation (tables 5 and 6), one event on the improved plots, and two events on the unimproved plots. This may be due to a rain gauge error, as previously discussed, with the rain gauge measuring less precipitation than occurred, or because of high variability, the precipitation depth on the plot was greater than at the nearby rain gauge. More likely, however, rutting on the roads had increased the area of the plot, with ruts directing runoff across the road crown. Assuming the road crown was a plot boundary was one of the weaker assumptions in the experimental design but overcoming this weakness by either reshaping the plot surface to have no crown but sloping only one way, or collecting runoff from both sides of the road and installing two sediment trap devices, or routing the runoff under the road to a single trap (Foltz and Elliot, 1997) would have been costly and likely not acceptable to the road managers on the heavily trafficked roads. This confirms what Elliot (2013) reported: that ruts tend to keep water on the road surface, resulting in increased sediment delivery, although the sensitivity analysis for plot design suggests this may not necessarily be the case on the relatively short road segments in this study (table 1). Another potential source of excessive runoff on the unimproved plots may have been surface runoff or lateral flow from the forest (Gucinski et al., 2001), as observed from a road ditch in Elliot and Glaza (2008). This external runoff was possibly observed by Grace (2017), where forested hillslopes could potentially have increased runoff on some of the unimproved plots where runoff was routed down incised ditches between the forest and the road (fig. 3b). Lateral flow would not have been a source of excessive runoff on the improved plots, where runoff was collected from the road shoulders on all plots using sheet metal gutters (fig. 3a).

The main reason for the managers to upgrade road networks at Fort Benning was to improve trafficability, particularly in this highly erosive climate. Anecdotal observation suggested that they had achieved their goal. The improved road required minimum maintenance despite heavy use from tracked military vehicles (fig. 3b). The unimproved road was often rutted and required frequent grading as well as additional gravel during the study period. The reduced sediment generation we observed from the improved roads can be an additional justification for upgrading roads if the managers are under pressure to reduce offsite pollution from sediments (Donigian, 2013). The unimproved plots were generating more than 80 times the sediment measured from the improved plots (2.27 vs. 0.026 Mg ha^{-1} per event) (table 7). Onsite observations suggested that the unimproved ditches were areas of deposition of material eroded from the road surface (fig. 3b). Had it been possible to install gutters on the unimproved plots as was done on the improved plots, even more sediment may have been captured from some events.

Sensitivity

Forest road studies involving traffic by Foltz (1996) and Luce and Black (1999) indicated that a single pass by a logging truck each day was sufficient to maintain a "traffic" condition for several weeks unless a large runoff event removed readily available sediment. Traffic volume was identified as one of the major factors contributing to increased erosion on unpaved roads by others as well (Reid and Dunne, 1984; Ziegler et al., 2001; Sheridan et al., 2006; Fu et al., 2010). Thus, we assumed that both sites in this study would meet the "traffic" condition for road soils in the WEPP:Road soil database (Elliot et al., 1999b). We selected the high traffic graveled sandy loam soil as a starting point for the sensitivity study and the calibration exercise (table 4).

Table 1 showed that the model was not particularly sensitive to road design and the associated flow path lengths and rill spacings, with the differences in estimated sediment delivery less than 10% for the three road designs and sensitivity values well below the 0.2 threshold for the sensitivity parameter *S* among the treatments (Ascough et al., 2013). The addition of rock or vegetation to the ditch did, however, reduce the estimated sediment delivery in table 1. The insensitivity to road design in this study was likely because the insloping and rutted designs had a rill length of 15 m, a relatively short length for road segments, whereas the relatively wide segments (5.3 m) resulted in a flow path length (fig. 1; table 1) of nearly 12 m for the outsloping design that was not much shorter than the 15-m segment length for the inslope or rutted designs. Ascough et al. (2013) stated that sensitivity analyses are "only true at the point where taken," so for other

conditions with longer road segments, the differences in design could result in a greater sensitivity in sediment delivery. As the flow path lengths were not greatly different, some of the difference in sediment delivery between the designs may have been due to rill spacing, as suggested by Zhang (2016). WEPP:Road estimated the greatest amount of sediment (32 Mg ha^{-1}) with the insloped design with the widest rill spacing (4 m). The WEPP:Road estimate for the outslope design with the shorter flow path (12 m) and closest rill spacing (1 m) delivered the least amount of sediment (29 Mg ha^{-1}). Zhang (2016) found that WEPP soil loss "increased substantially" when rill spacing was increased from 1 to 10 m, showing trends similar to those shown in table 1, where the 4-m spacing generated more sediment than the 1- or 2-m spacings. Nearing et al. (1990), however, found that sediment delivery slightly declined with increasing rill spacing for their conditions. For all three slope lengths in the Nearing et al. study (22, 50, and 200 m), S was greater than -0.2, and hence sediment delivery was not considered sensitive to rill spacing.

The sensitivities to slope length and steepness shown in figure 8 were similar to the graph presented in Elliot et al. (1999a) with sediment delivery increasing with road gradient and length. The relatively flat response to increasing segment length and a mean S value of only 0.08 for the 2 percent gradient suggested that interrill erosion was dominant. Zhang (2016) also credited a flat sediment delivery curve on shorter slopes, in his case near the top of the plot, to interrill erosion processes being dominant. Sediment delivery was sensitive to all other road gradients and segment lengths, with S values ranging from 0.27 to 0.9.

WEPP was not sensitive to high values of K_r (fig. 9b). This may be due to a sediment transport limiting erosion process within WEPP, as first proposed by Ellison (1947), who hypothesized that sediment delivery may be either detachment-limited, or transport-limited. Equation 3 showed that as the sediment in transport G approaches the transport capacity T_c, the rill detachment rate D_r approaches zero. Thus, the rate of sediment delivery G could be no greater than T_c, regardless of the value for K_r. The transport capacity was calculated internally by WEPP as a function of soil properties (Finker et al., 1989; Flanagan and Nearing, 1995). With field studies such as this that do not have details of rill shape, controlled runoff dynamics, and delivered sediment particle and aggregate size distributions, it is not possible to externally estimate T_c or solve equation 3 to estimate K_r and τ_c as was done by Elliot and Flanagan (2023), Foltz et al. (2008), and Wagenbrenner et al. (2010). Despite these concerns, however, the LS mean for sediment delivery was 2.27 Mg ha^{-1} for the unimproved plots, 87 times the delivery from the improved plots (0.026 Mg ha^{-1}; table 7). The K_r value for the unimproved plots was 0.09 s m^{-1}, 112 times the value of the improved plots (0.0008 s m^{-1}; table 9). The greater difference in K_r compared to the LS mean sediment delivery between treatments was likely the effect of G on estimating K_r and the lower runoff rates from the unimproved plots. Elliot et al. (1989) found that for K_r values less than 0.02, like the improved plots, the sediment in transport G had minimal impact on the estimation of K_r. Elliot et al. (1989) showed that for only a few high flow rates, G exceeded the estimated T_c on the most erodible soil, the Amarillo sandy loam (K_r = 0.045 s m^{-1}). None of the other soils with smaller K_r values in the data set had instances where G exceeded T_c.

We found that sediment delivery was not sensitive to interrill erodibility K_i (fig. 9b). This lack of sensitivity was confirmed in the calibration analysis when changes in K_i had no effect on the model performance for either road design. Neither Nearing et al. (1990) nor Ascough et al. (2013) found WEPP predictions to be sensitive to K_i. The sensitivity analysis supported the approach we took for calibration, starting with the variable with the greatest sensitivity, hydraulic conductivity, followed by rill erodibility, critical shear, and interrill erodibility (table 8; fig. 9).

The estimated sediment delivery in sensitivity results for the base conditions (N=1) for plots with dimensions and gradient similar to the study plots (table 3) was 26 Mg ha^{-1} per year (figs. 8 and 9). The observed sediment delivery rates for the three months in the study, selected because they were the wettest months, averaged 21 Mg ha^{-1} for the improved plots and 73 Mg ha^{-1} for the unimproved plots. This suggests that the current erodibility values in WEPP:Road were of similar magnitude but were likely too low for both designs. The under prediction of sediment delivery in the sensitivity study was likely because the baseline K_r (0.0004 s m^{-1}; table 4) was less than the calibrated values of 0.0008 and 0.09 s m^{-1} for the improved and unimproved plots (table 9).

CALIBRATION

Table 9 presents the estimated WEPP hydraulic conductivity and erodibility values from the calibration analyses. The unimproved plots had less runoff and lower runoff coefficients compared to the improved plots (tables 5 and 6; fig. 7a), resulting in a higher estimated hydraulic conductivity value of 3 mm h^{-1} compared to 1.3 mm h^{-1} for the improved plots (table 9). However, the greater sediment yields from the unimproved plots (tables 5 and 6, fig. 7b) despite lower runoff rates were due to a greater rill erodibility value (0.09 s m^{-1} for unimproved vs. 0.0008 s m^{-1} for improved) and lower critical shear (0.0001 Pa for unimproved vs. 1.5 Pa for improved, table 9).

With three variables to describe soil erodibility, there was a risk that we may not have found the optimal solution (Brazier et al., 2000). For the study conditions, the sensitivity analysis and subsequent calibration showed that interrill erosion was likely not a major source of erosion (fig. 9b), so the chances of missing an optimal solution for sediment delivery from a high estimate of interrill erosion offsetting a low erosion estimate from rill erosion was unlikely. The only other risk of missing an optimal solution was with a low value of critical shear, offsetting a high value of rill erodibility. However, for the unimproved plots, the critical shear was near zero (τ_c = 0.0001Pa) and could not have been any less, so in this instance, it was unlikely that a local optimum was missed. The only reason a value of zero was not used for critical shear was that WEPP was programmed to estimate an input value internally if there was a zero in the input file (Elliot and Flanagan, 2023; Flanagan and Livingston, 1995). On the improved plots, the solutions obtained for K_r (0.0008 s m^{-1}) and τ_c (1.5 Pa) were within the range of values that were observed elsewhere (Elliot et al., 1999b; Elliot and

Flanagan, 2023), and figure 9 shows that predicted erosion was not sensitive to values of τ_c around 2 Pa. The likelihood of missing a better, optimal solution for the improved road is unlikely.

The WEPP:Road database had a hydraulic conductivity for a sandy loam road soil of 3.8 mm h^{-1} for a native surface road and 10.2 mm h^{-1} for a graveled road; both values were higher than in this study (1.3 and 3.0 mm h^{-1} for improved and unimproved plots, respectively, table 9). Foltz et al. (2011) determined hydraulic conductivity input values on roads with sandy loam textures around 8 mm h^{-1}. The lower values for hydraulic conductivity in this study may be due to higher traffic levels continually compacting the soils on the unimproved road segments and compaction during construction on the improved road segments compared to the lower traffic in the Foltz et al. (2011) study. The higher very fine (*VFS*) and fine (*FS*) sand contents (37%) in this study (table 2) were more easily compacted and had lower hydraulic conductivities than the more coarse-grained sands in the Foltz et al. (2011) study (*VFS* + *FS* = 25%; Onur, 2014). Foltz and Elliot (1997) reported on high traffic forest roads estimated hydraulic conductivity values to be 1.1 mm h^{-1} for road ruts with low tire pressure logging trucks and 0.4 mm h^{-1} for road ruts with high tire pressure logging trucks on a soil with a clay content greater than 20%. They estimated the hydraulic conductivity of the road shoulder to be 2 mm h^{-1}, conductivity values more in line with this study. Similar lower hydraulic conductivity values (2.9 mm h^{-1}), however, were reported by Elliot et al. (1994) on newly constructed roads in the Western U.S., where time had been too short for macropores to form as likely occur on lower traffic forest roads. In Northern Georgia, USA, 300 km northeast of this study, Grace's (2017) study on a low use forest road determined that the hydraulic conductivity was only 0.025 mm h^{-1}. His soil, however, had a clay content of 30% and a sand content of only 30%, compared to 3% clay and 86% sand on our sites (table 2). likely accounting for a much lower hydraulic conductivity value on the Grace (2017) site.

Rill erodibility values in this study were 0.09 s m^{-1} for the unimproved road and 0.0008 s m^{-1} for the improved road segments (table 8); both values were greater than the WEPP:Road database value of 0.0004 s m^{-1} for high traffic sandy loam forest roads (Elliot, 2004). The constant heavy traffic on both treatments, coupled with regular maintenance on the unimproved plots, likely increased sediment availability and hence erodibility. Grace and Elliot (2008) determined that the online WEPP:Road predictions were reasonable when compared to observed sediment delivery from lower traffic forest roads in the Conecuh National Forest, 230 km southwest of the study site. Welsh (2008) found that WEPP:Road underpredicted observed sediment delivery by a factor of 20 on forest roads with sandy loam soils in Colorado, U.S., suggesting that the WEPP:Road rill erodibility values may be too low or hydraulic conductivity too high for some soils. Conversely, Stafford (2011) found that WEPP:Road overpredicted sediment by about a factor of three on the low use native roads in the Kings River Experimental Watershed in California, U.S., indicating that the rill erodibility values may be too high or the hydraulic conductivity too low in WEPP:Road. Grace (2017) estimated rill erodibility to be 0.0001 s m^{-1} on the high clay soils, lower than both our treatments. There was likely less traffic on the Stafford (2011) and Grace (2017) studies than in this study, confirming that, as with hydraulic conductivity, rill erodibility is influenced by both soil properties and traffic levels (Elliot, 2004; Foltz et al., 2011; Gucinski et al., 2001). This study and others suggest that the WEPP:Road database needs to be expanded to include a larger range of soils, including those of lower erodibility in the Sierra Nevada mountains in California, U.S. (Stafford, 2011; Foltz et al., 2011), and soils with higher erodibility such as the Colorado Rocky Mountains, U.S. (Welsh, 2008) and the roads in this study. The WEPP:Road database may also benefit from adding an option for improved road designs based on the GAB design in this study or similar practices on higher traffic gravel roads.

The estimated K_r value of 0.09 s m^{-1} for the unimproved road (table 9) was more than the greatest values published for croplands (K_r = 0.045 for Amarillo sandy loam), rangelands (K_r = 0.003 for Pratt sand) (Laflen et al., 1991), or roads (K_r = 0.0004 for sandy loam road soil; Elliot et al., 1999b). The cropland soils where K_r had been measured by rainfall simulation that were most like the Troup loamy sand on the unimproved plots (table 2) were Bonifay sand and Tifton loamy sand, located at Tifton, GA, 300 km southeast of the study site (Elliot and Flanagan, 2023). The texture of Troup loamy sand on the unimproved road in this study was nearly identical to that of Tifton loamy sand, with K_r equal to 0.013 s m^{-1}, and Bonifay sand with 91% sand had K_r equal to 0.018 s m^{-1} (Elliot and Flanagan, 2023). Both cropland soils had lower K_r values than the unimproved soil (K_r = 0.09 s m^{-1}), suggesting that the heavy traffic contributed to the elevated rill erodibility. There was also a possibility that the K_r values for the two cropland soils were underestimated as there were problems generating runoff from some of the rill plots and high variability within the plot data, resulting in incomplete field data sets and difficulties in calculating K_r (Elliot et al., 1989).

Figure 11 showed that observed and predicted erosion rates decreased as the slope increased. This was likely because the two longest plots in the study were the unimproved automatic plots (table 3). These plots had the highest erosion rates due to their management and their length, but also had the lowest gradients. These are the two plots with slopes under 4% in figure 11. This unexpected result of decreasing erosion with road gradient was also reported by Brake et al. (1997), who estimated sediment delivery from road sediment plume depositions within the forest. They found a similar result, with steeper plots delivering less sediment. In their case, like this study, the steeper road segments tended to be shorter. Anecdotal GIS erosion analyses of landscapes have shown similar relationships between slope length and steepness. Such results confirm the importance of considering the interaction between slope steepness and length with any erosion analysis. Both slope length and steepness are inputs for WEPP and other commonly used erosion models and thereby account for this landscape feature, but it is not uncommon for watershed managers to make landscape decisions based on slope steepness alone (Elliot et al., 2008).

VALIDATION

The Willmott et al. (2012) indices of agreement (d_r) were similar for calibration and validation (table 9), suggesting that the calibration and validation data sets as delineated in tables 5 and 6 likely had similar distributions of observed runoff and sediment delivery amounts. The index is the ratio of predicted error to observed error. Because of the high variability within the observed data, this error ratio was reduced, and subsequently, d_r was high (eq. 7). Grace (2017) reported similar high variabilities in observed runoff and sediment delivery, resulting in high values for d_r.

The linear regressions of predicted vs. observed runoffs and sediment delivery amounts on the entire data set (fig. 10) suggested that generally, larger events were underestimated (regression slopes < 1) while smaller events were generally overestimated (regression intercepts > 0). This is a common finding in erosion modeling and is due to "limitations in representing the random component of measured data within treatments (i.e., between replicates) with a deterministic model" (Nearing, 1998).

The decision to use the Willmott et al. (2012) index of agreement was made to prevent large events from dominating the validation analysis, as in natural erosion studies, including this one. There were few large events, and the largest event overwhelmed one of the automatic samplers on the unimproved site, limiting the usefulness of that event. The other common goodness-of-fit statistics all contained squared terms for evaluation (Coefficient of Determination, Nash-Sutcliffe and Root Mean Square Error). Table 11 showed that regardless of which statistic was selected for the validation data set, the conclusions drawn were the same: runoff from the unimproved plots was poorly predicted, whereas runoff from the improved plots and sediment delivery from both plots were reasonably well predicted.

The variety of goodness-of-fit statistics in table 11 does little to improve the model validation beyond what was presented in table 9 and figure 10, but the statistics do lend itself to better comparing results of this study to those of other studies. Elliot and Flanagan (2023) reported on an analysis that estimated rill erodibility and critical shear for 36 soils. The Elliot and Flanagan (2023) analysis used erosion mechanics to estimate erodibility, as the WEPP code was not yet complete when the study was done in the late 1980s. Elliot et al. (1991) used those erodibility values in an early version of WEPP to estimate sediment delivery from six rills on each of five soils that were in the original erodibility study described in Elliot and Flanagan (2023). Elliot et al. (1991) reported that the *RMSE:Mean* ratio for WEPP-estimated sediment delivery compared to observed sediment delivery for those five soils varied from 0.28 to 0.77, averaging 0.52. The ratios reported in this study (table 11) of 0.79 and 0.97 for sediment delivery are a little greater than Elliot et al. (1991). The lower ratios in Elliot et al. (1991) were likely due to less experimental error in a rainfall/runoff simulation experiment with fixed rainfall rates, inflow rates, plot boundaries, and six replicates (Elliot and Flanagan, 2023), compared to this study relying on natural rainfall and road topography for plot delineation with two to four replicates. Covert et al. (2005) used the WEPP Watershed tool to estimate runoff from three small, forested watersheds that had been harvested and burned. They calculated *RMSE:Mean* ratios of 0.48, 0.97, and 1.86 that were similar to the ratios of 1.2 and 0.45 found for runoff in this study. Srivastava et al. (2013) compared daily runoff estimated with the WEPP watershed model for a small, forested watershed in Idaho and found annual values for *NSE* ranging from 0.53-0.89, averaging 0.67. This study estimated the *NSE* to be 0.09 for runoff from the unimproved plots and 0.84 for the improved road plots. Reasons for the poor prediction on the unimproved plots were discussed previously. Moriasi et al. (2007) concluded that *NSE* values should be greater than 0.50 for a model simulation to be judged as satisfactory. Their evaluations, however, were focused on watershed studies with continuous daily or monthly flow observations, and they cited no studies that had applied the *NSE* statistic to sediment delivery. Laflen et al. (2004), however, did apply the *NSE* to erosion data when using 20 cropland studies originally designed to support the development of the USLE to validate the WEPP model. They reported that the *NSE* for WEPP estimates for sediment delivery ranged from -0.38 to 0.94 and exceeded 0.4 on only 6 of those 20 sites. They compared that to the RUSLE model that was developed from those plots that had an *NSE* value greater than 0.4 on 11 of the 20 sites. This suggests that the *NSE* values for sediment delivery for this study of 0.41 are reasonable. Grace (2017) measured and modeled runoff and sediment delivery from nine forest road segments, in a study similar to this study. He reported that *NSE* values ranged from -0.22 to 0.49, averaging 0.15 for runoff from the nine plots compared to 0.09 and 0.84 for this study. For sediment delivery, Grace (2017) reported *NSE* values ranging from -1.09 to 0.89, averaging 0.76 after deleting the lowest value on his road erosion study. In the Grace (2017) study, the Willmott index of agreement d_r for WEPP-predicted runoff ranged from 0.49 to 0.72, averaging 0.62 compared to 0.62 and 0.82 for the validation plots in this study (table 9). The d_r value for sediment delivery in the Grace (2017) study ranged from 0.37 to 0.73 for the nine plots, averaging 0.56 compared to the d_r values for sediment delivery for the validation plots in this study of 0.66 and 0.67. The validation analyses presented in tables 9 and 11 all lead to the same conclusion: that the WEPP model with the input erodibility values in table 9 was able to make valid estimates of sediment delivery from unimproved and improved roads in Fort Benning, GA.

In both the Grace (2017) study and this study, the d_r statistic gave a more optimistic indication of the model performance than did the *NSE* statistic. The *NSE* statistic was proposed to support the development of river flow predictions to evaluate how changing input watershed parameters or their values to a generic river flow model would impact the output modeling seasonal to multiyear hydrographs (Nash and Sutcliff, 1970). Examples of this application to river flows were reported in Moriasi et al. (2007) and have been applied to WEPP Watershed optimization analyses (Srivastava et al., 2013, 2017). The d_r statistic was developed to aid in evaluating models by comparing predicted values with observed values for a limited number of events (Willmott et al., 2012). The results of this validation analysis tend to confirm that the Willmott et al. (2012) index of agreement d_r

was the preferred statistic for calibrating and validating event-based soil erosion modeling.

Summary

The purpose of this study was to support a watershed modeling analysis in Fort Benning, GA, U.S. We evaluated the ability to the Water Erosion Prediction Project (WEPP) model to estimate sediment generated by high traffic graveled roads, and to determine the erodibility of such roads. We collected runoff and sediment delivery data from 10 road segments, including two different road designs, completed a sensitivity analysis of the WEPP model based on a road erosion plot, calibrated the results of the field study to determine the soil erodibility for both road designs, and validated the erodibility values. The study showed that the improved road design more than doubled runoff but decreased sediment delivery by 99%. The WEPP model did not perform well in predicting runoff from unimproved road segments but made reasonable predictions for runoff from the improved segments and sediment delivery from both road designs.

Conclusions

The study found that there was less runoff but significantly more sediment delivered from unimproved gravel roads with management focused on frequent grading and gravel addition as needed compared to roads constructed from compacted layers of gravel. This confirms what others have reported for forested watersheds: that roads are a major source of sediment in sensitive watersheds, and enhanced road designs may be needed to minimize downstream sedimentation. Hydraulic conductivity values were lower and rill erodibility values were greater than currently assumed for forest roads in the online WEPP:Road interface for predicting road erosion, underscoring the need to expand the online soil database to incorporate a wider range of soil properties and management practices. In the sensitivity, calibration, and validation analyses, the WEPP technology was able to consistently reflect the effects of precipitation amount, topography, and road management on sediment delivery. The study focused on a single site only, so the observed erosion rates and calculated erodibility values may not be applicable to other sites with different soils and climates. There is a need for additional studies on high traffic roads, whether they are public, private, or military, because in noncropland watersheds, high traffic roads are likely to be a significant source of sediment. There is also a need for more comprehensive runoff simulation studies on highly erodible roads to better quantify the unusually high rill erodibility measured in this study that may likely be found at other high traffic sites.

Acknowledgments

The authors wish to acknowledge the Department of Defense for funding the field work in this study through the Environmental Security Technology Certification Program (ESTCP), Contract number 201307. We acknowledge AQUA TERRA Consultants for administering the funding, and John Imhoff, PE, for contributing to an earlier manuscript, and providing review comments. We are indebted to Mr. Ben Kopyscianski, a retired hydrologist with Rocky Mountain Research Station (RMRS), who spent many long, hot days installing and maintaining plots, and collecting grab samples during storms that could occur any time during the day or night. Dr. Randy Foltz, a retired research civil engineer with RMRS, was the original principal investigator on the project, and with Mr. Kopyscianski developed the study plan. Ms. Sue Miller, RMRS Hydrologist, assisted with weather station and rain gage installation. Dr. Emily Carter, Research Soil Scientist with the Southern Research Station, Auburn Forestry Sciences Laboratory loaned some proportional water samplers and provided storage space for equipment between seasons. We greatly appreciated the assistance provided by Mr. Carlton Cribb, Site Supervisor of the Fort Benning DMPRC site where the improved road plots were installed. Ms. Victoria Nystrom, a summer employee with RMRS, assisted in building breakpoint climate files for every storm and slope files for each plot to support WEPP modeling.

References

Al-Hamdan, O. Z., Pierson, F. B., Robichaud, P., Elliot, W. J., & Williams, C. J. (2022). New erodibility parameterization for applying WEPP on rangelands using ERMiT. *J. ASABE, 65*(2), 251-264. https://doi.org/10.13031/ja.14564

Aqua Terra Consultants. (2012). Development of a watershed modeling system for Fort Benning using the USEPA BASINS Framework, Final Report for SERDP Project RC-1547. Alexandria, VA. Retrieved from https://serdp-estcp.org/projects/details/8d1e0869-049a-44ff-9003-558c765bf8f1

Ascough II, J. C., Flanagan, D. C., Nearing, M. A., & Engel, B. A. (2013). Sensitivity and first-order/Monte Carlo uncertainty analysis of the WEPP hillslope erosion model. *Trans. ASABE, 56*(2), 437-452. https://doi.org/10.13031/2013.42693

Baffaut, C., Nearing, M. A., & Nicks, A. D. (1996). Impact of CLIGEN parameters on WEPP-predicted average annual soil loss. *Trans. ASAE, 39*(2), 447-457. https://doi.org/10.13031/2013.27522

Black, T. A., & Luce, C. H. (2013). Measuring water and sediment discharge from a road plot with a settling basin and tipping bucket. Gen. Tech. Rep. RMRS-GTR-28. Fort Collins, CO: USDA, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/rmrs-gtr-287

Black, T. A., Cissel, R. M., & Luce, C. H. (2012). The Geomorphic Road Analysis and Inventory Package (GRAIP) Volume 1: Data collection method. Gen. Tech. Rep. RMRS-GTR- 280WWW. Fort Collins, CO: USDA, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/rmrs-gtr-280

Bosch, D. D., & West, L. T. (1998). Hydraulic conductivity variability for two sandy soils. *Soil Sci. Soc. Am. J., 62*(1), 90-98. https://doi.org/10.2136/sssaj1998.03615995006200010012x

Brake, D., Molnau, M., & King, J. G. (1997). Sediment transport distances and culvert spacings on logging roads within the Oregon Coast Mountain Range, Paper No. 975018. *Proc. 1997 ASAE Annu. Int. Meeting.* St. Joseph, MI: ASAE.

Brazier, R. E., Beven, K. J., Freer, J., & Rowan, J. S. (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.

https://doi.org/10.1002/1096-9837(200008)25:8<825::AID-ESP101>3.0.CO;2-3
Cao, L., Elliot, W., & Long, J. W. (2021). Spatial simulation of forest road effects on hydrology and soil erosion after a wildfire. *Hydrol. Process., 35*(6), e14139. https://doi.org/10.1002/hyp.14139
Ciach, G. J. (2002). Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol., 20*(5), 752-759. https://doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2
Cissel, R. M., Black, T. A., Schreuders, K. A., Prasad, A., Luce, C. H., Tarboton, D. G., & Nelson, N. A. (2012). The Geomorphic Road Analysis and Inventory Package (GRAIP) Volume 2: Office procedures. Gen. Tech. Rep. RMRS-GTR-281WWW. Fort Collins, CO: USDA, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/rmrs-gtr-281
Covert, A., Robichaud, P. R., Elliot, W. J., & Link, T. E. (2005). Evaluation of runoff prediction from WEPP-based erosion models for harvested and burned forest watersheds. *Trans. ASAE, 48*(3), 1091-1100. https://doi.org/10.13031/2013.18519
Doherty, J. (2005). PEST model independent parameter estimation user manual. *5th*. Watermark Numerical Computing. Retrieved from https://www.nrc.gov/docs/ML0923/ML092360221.pdf
Donigian, A. (2013). A watershed modeling system for Fort Benning, GA using the US EPA BASINS framework. SERDP Project RC-1547. AquaTerra Consultants. Retrieved from https://serdp-estcp-storage.s3.us-gov-west-1.amazonaws.com/s3fs-public/project_documents/RC-1547-FR.pdf?VersionId=VtSdlXt0UbsjRVx7Q6jDm4odZFQ33KrF
Donigian, A., Imhoff, J. C., Mishra, A., & Duda, P. B. (2018). Demonstration and validation of BASINS watershed modeling system enhanced for military installations - Fort Benning. ESTCP Project RC-201307. RESPEC Consulting & Services. Retrieved from https://serdp-estcp-storage.s3.us-gov-west-1.amazonaws.com/s3fs-public/project_documents/RC-201307%2BFinal%2BReport%2B-%2BFort%2BBenning.pdf?VersionId=TIc3FQLZetNUQYLYFXoPgl.dKVDspIpY
Dubé, K., Megahan, W., & McCalmon, M. (2004). Washington road surface erosion model. Olympia, WA: State of WA Department of Natural Resources.
Elliot, B. (2014). WEPP: Road input screen version 2014.09.05. Retrieved from http://forest.moscowfsl.wsu.edu/fswepp/
Elliot, W. J. (1988). A process-based rill erosion model. PhD diss. Ames, IA: Iowa State University, Dep. Agric. Eng. Retrieved from https://dr.lib.iastate.edu/server/api/core/bitstreams/cd5cccac-451e-4587-a32f-2842d3e629b7/content
Elliot, W. J. (2004). WEPP internet interfaces for forest erosion prediction. *JAWRA, 40*(2), 299-309. https://doi.org/10.1111/j.1752-1688.2004.tb01030.x
Elliot, W. J. (2013). Erosion processes and prediction with WEPP technology in forests in the Northwestern U.S. *Trans. ASABE, 56*(2), 563-579. https://doi.org/10.13031/2013.42680
Elliot, W. J., & Flanagan, D. C. (2023). Estimating WEPP cropland erodibility values from soil properties. *J. ASABE, 66*(2), 329-351. https://doi.org/10.13031/ja.15218
Elliot, W. J., & Foltz, M. (2001). Validation of the FS WEPP interfaces for forest roads and disturbances. *Proc. 2001 ASAE Annual Meeting.* St. Joseph, MI: ASAE. https://doi.org/10.13031/2013.5553
Elliot, W. J., & Glaza, B. D. (2008). Impacts of forest management on runoff and erosion. In R. M. Webb, & D. J. Semmens (Eds.), *Planning for an Uncertain Future - Monitoring, Integration, and Adaptation: Proc. of the Third Interagency Conference on Research in the Watersheds. USGS Scientific Investigations Report 2009-5049* (pp. 117-127). Estes Park, CO: USDOI & USGS. Retrieved from https://www.fs.usda.gov/research/treesearch/34683
Elliot, W. J., & Hall, D. E. (1997). Water erosion prediction project (WEPP) forest applications. Gen. Tech. Rep. INT-GTR-365. Ogden, UT: USDA, Forest Service, Intermountain Research Station. Retrieved from https://forest.moscowfsl.wsu.edu/cgi-bin/engr/library/searchpub.pl?pub=1997f
Elliot, W. J., & Rhee, H. (2022). Impacts of forest biomass operations on forest hydrologic and soil erosion processes. *Trees, Forests and People, 7*, 100186. https://doi.org/10.1016/j.tfp.2021.100186
Elliot, W. J., Elliot, A. V., Wu, Q., & Laflen, J. M. (1991). Validation of the WEPP model with rill erosion plot data. Paper no. 91-2557. *Proc. 1991 Int. Winter Meeting of the ASAE.* St. Joseph, MI: ASAE.
Elliot, W. J., Foltz, R. B., & Luce, C. H. (1995). Validation of the Water Erosion Prediction Project (WEPP) model for low-volume forest roads. *Proc. of the Sixth Int. Conf. on Low-Volume Roads* (pp. 178-186). Washington, DC: National Academy Press. Retrieved from https://www.fs.usda.gov/rm/pubs_journals/1995/rmrs_1995_elliot_w001.pdf
Elliot, W. J., Foltz, R. B., & Luce, C. H. (1999a). Modeling low-volume road erosion. *Transp. Res. Rec., 1652*(1), 244-249. https://doi.org/10.3141/1652-64
Elliot, W. J., Foltz, R. B., & Remboldt, M. D. (1994). Predicting sedimentation from roads at stream crossings with the WEPP model, Paper No. 947511. *Proc. ASAE Annu. Int. Meeting.* St. Joseph, MI: ASAE.
Elliot, W. J., Hall, D. E., & Scheele, D. L. (1999b). WEPP: Road (Draft 12/1999) WEPP interface for predicting forest road runoff, erosion and sediment delivery, technical documentation.
Elliot, W. J., Liebenow, A. M., Laflen, J. M., & Kohl, K. D. (1989). Compendium of soil erodibility data from WEPP cropland field erodibility experiments 1987 & 1988. NSERL Report No. 3. West Lafayette, IN: USDA-ARS, National Soil Erosion Research Laboratory. Retrieved from https://milford.nserl.purdue.edu/weppdocs/comperod/
Elliot, W. J., Wei, L., Imhoff, J. C., Foltz, R. B., & Nystrom, V. E. (2015). Impacts of road design on sediment generation. Paper No. 152181402. *2015 ASABE Ann. Int. Meeting.* St. Joseph, MI: ASABE. https://doi.org/10.13031/aim.20152181402
Elliot, W., Miller, W., Hartsough, B., & Stephens, S. (2008). Vegetation management in sensitive areas of the Lake Tahoe Basin, A workshop to evaluate risks and advance existing strategies and practices. Independent review panel report. Incline Village, NV: Tahoe Center for Environmental Studies. Retrieved from https://www.fs.usda.gov/research/treesearch/34120
Ellison, W. D. (1947). Soil erosion studies, part I. *Agric. Eng., 28*(4), 145-146. 145-146
Finkner, S. C., Hearing, M. A., Foster, G. R., & Gilley, J. E. (1989). A simplified equation for modeling sediment transport capacity. *Trans. ASAE, 32*(5), 1545-1550. https://doi.org/10.13031/2013.31187
Flanagan, D. C., & Livingston, S. J. (1995). WEPP User Summary, NSERL Report No. 11. West Lafayette, IN: USDA-ARS, National Soil Erosion Research Laboratory. https://www.ars.usda.gov/ARSUserFiles/50201000/WEPP/usersum.pdf
Flanagan, D. C., & Nearing, M. A. (1995). USDA-Water Erosion Prediction Project (WEPP) Hillslope Profile and Watershed Model Documentation. NSERL Report No. 10. West Lafayette, IN: USDA-ARS, National Soil Erosion Research Laboratory. Retrieved from https://www.ars.usda.gov/midwest-area/west-

lafayette-in/national-soil-erosion-research/docs/wepp/wepp-model-documentation/
Flanagan, D. C., Fu, J., Frankenberger, J. R., Livingston, S. J., & Meyer, C. R. (1998). A Windows® interface for the WEPP erosion model. Paper No. 98-2135. *Proc. 1998 ASAE Annu. Int. Meeting.* St. Joseph, MI: ASABE.
Foltz, R. B. (1996). Traffic and no-traffic on an aggregate surfaced road: Sediment production differences. *Proc. Seminar on Environmentally Sound Forest Roads and Wood Transport.* Rome, Italy: United Nations FAO.
Foltz, R. B., & Elliot, W. J. (1997). Effect of lowered tire pressures on road erosion. *Transp. Res. Rec.: J. Transp. Res. Board, 1589*(1), 19-25. https://doi.org/10.3141/1589-03
Foltz, R. B., Copeland, N. S., & Elliot, W. J. (2009). Reopening abandoned forest roads in northern Idaho, USA: Quantification of runoff, sediment concentration, infiltration, and interrill erosion parameters. *J. Environ. Manag., 90*(8), 2542-2550. https://doi.org/10.1016/j.jenvman.2009.01.014
Foltz, R. B., Elliot, W. J., & Wagenbrenner, N. S. (2011). Soil erosion model predictions using parent material/soil texture-based parameters compared to using site-specific parameters. *Trans. ASABE, 54*(4), 1347-1356. https://doi.org/10.13031/2013.39036
Foltz, R. B., Rhee, H., & Elliot, W. J. (2008). Modeling changes in rill erodibility and critical shear stress on native surface roads. *Hydrol. Process., 22*(24), 4783-4788. https://doi.org/10.1002/hyp.7092
Fu, B., Newham, L. T., & Ramos-Scharrón, C. E. (2010). A review of surface erosion and sediment delivery models for unsealed roads. *Environ. Model. Softw., 25*(1), 1-14. https://doi.org/10.1016/j.envsoft.2009.07.013
Grace III, J. M. (2017). Predicting forest road surface erosion and storm runoff from high-elevation sites. *Trans. ASABE, 60*(3), 705-719. https://doi.org/10.13031/trans.11646
Grace III, J. M., & Elliot, W. J. (2008). Determining soil erosion from roads in coastal plain of Alabama. *Environmental Connection 08, Proceedings of Conference 39.* International Erosion Control Association.
Grace III, J. M., Rummer, B., Stokes, B. J., & Wilhoit, J. (1998). Evaluation of erosion control techniques on forest roads. *Trans. ASAE, 41*(2), 383-391. https://doi.org/10.13031/2013.17188
Gucinski, H., Furniss, M. J., Ziemer, R. R., & Brookes, M. H. (2001). Forest roads: A synthesis of scientific information. Gen. Tech. Rep. PNW-GTR-509. Portland, OR: USDA, Forest Service, Pacific Northwest Research Station. https://doi.org/10.2737/pnw-gtr-509
Helsel, D. R., & Hirsch, R. M. (2002). Statistical methods in water resources: USGS Techniques of Water Resources Investigations, Book 4, Chapter A3. Reston, VA: USGS. https://doi.org/10.3133/twri04A3
Huffman, R. L., Fangmeier, D. D., Elliot, W. J., & Workman, S. R. (2013). Chapter 7: Soil erosion by water. In *Soil and water conservation engineering* (7th ed., pp. 145-170). St. Joseph, MI: ASABE. https://doi.org/10.13031/swce.2013.7
Laflen, J. M., Elliot, W. J., Flanagan, D. C., Meyer, C. R., & Nearing, M. A. (1997). WEPP-Predicting water erosion using a process-based model. *J. Soil Water Conserv., 52*(2), 96-102.
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 Water Conserv., 46*(1), 39-44. Retrieved from https://www.jswconline.org/content/jswc/46/1/39.full.pdf
Laflen, J. M., Flanagan, D. C., & Engel, B. A. (2004). Soil erosion and sediment yield prediction accuracy using WEPP. *JAWRA, 40*(2), 289-297. https://doi.org/10.1111/j.1752-1688.2004.tb01029.x
Lang, A. J., Aust, W. M., Bolding, M. C., McGuire, K. J., & Schilling, E. B. (2017). Comparing sediment trap data with erosion models for evaluation of forest haul road stream crossing approaches. *Trans. ASABE, 60*(2), 393-408. https://doi.org/10.13031/trans.11859
Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS® for Mixed Models* (2nd ed.). Cary, NC: SAS Institute.
Luce, C. H., & Black, T. A. (1999). Sediment production from forest roads in western Oregon. *Water Resour. Res., 35*(8), 2561-2570. https://doi.org/10.1029/1999WR900135
Marion, D. A., & Clingenpeel, J. A. (2012). Methods used for analyzing the cumulative watershed effects of fuel management on sediment in the Eastern United States. Gen. Tech. Rep. SRS-161. In R. LaFayette, M. T. Brooks, J. P. Potyondy, L. Audin, S. L. Krieger, & C. C. Trettin (Eds.), *Cumulative watershed effects of fuel management in the Eastern United States* (pp. 308-326). Asheville, NC: USDA Forest Service, Southern Research Station.
Mein, R. G., & Larson, C. L. (1973). Modeling infiltration during a steady rain. *Water Resour. Res., 9*(2), 384-394. https://doi.org/10.1029/WR009i002p00384
Miller, M. E., MacDonald, L. H., Robichaud, P. R., & Elliot, W. J. (2011). Predicting post-fire hillslope erosion in forest lands of the western United States. *Int. J. Wildland Fire, 20*(8), 982-999. https://doi.org/10.1071/WF09142
Morgan, R. P., & Nearing, M. A. (2011). *Handbook of erosion modelling.* Chichester, UK: Blackwell Publishing. https://doi.org/10.1002/9781444328455
Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE, 50*(3), 885-900. https://doi.org/10.13031/2013.23153
Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: Part I — A discussion of principles. *J. Hydrol., 10*(3), 282-290. https://doi.org/10.1016/0022-1694(70)90255-6
Nearing, M. A. (1998). Why soil erosion models over-predict small soil losses and under-predict large soil losses. *CATENA, 32*(1), 15-22. https://doi.org/10.1016/S0341-8162(97)00052-0
Nearing, M. A., Deer-Ascough, L., & Laflen, J. M. (1990). Sensitivity analysis of the WEPP hillslope profile erosion model. *Trans. ASAE, 33*(3), 839-849. https://doi.org/10.13031/2013.31409
Nicks, A. D., Lane, L. J., & Gander, G. A. (1995). Weather generator, Chapter 2. In D. C. Flanagan, & M. A. Nearing (Eds.), *USDA-Water Erosion Prediction Project (WEPP) hillslope profile and watershed model documentation. NSERL Report No. 10.* West Lafayette, IN: USDA-ARS, National Soil Erosion Research Laboratory.
Onur, E. M. (2014). Predicting the permeability of sandy soils from grain size distributions. MS thesis. Kent, OH: Kent State University. Retrieved from https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=kent1389550812&disposition=inline
Prasad, A. (2007). A tool to analyze environmental impacts of roads on forest watersheds. MS thesis. Logan, UT: Utah State University.
Rawls, W. J., Brakensiek, D. L., Simanton, J. R., & Kohl, K. D. (1989). Development of a crust factor for a Green Ampt model. *Trans. ASAE, 33*(4), 1224-1228. https://doi.org/10.13031/2013.31461
Reid, L. M., & Dunne, T. (1984). Sediment production from forest road surfaces. *Water Resour. Res., 20*(11), 1753-1761. https://doi.org/10.1029/WR020i011p01753

Renard, K. G., Foster, G. R., Weesies, G. A., McCool, D. K., & Yoder, D. C. (1997). Predicting soil erosion by water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). Handbook 703. Washington, DC: USDA-ARS.

Risse, L. M., Nearing, M. A., & Savabi, M. R. (1994). Determining the Green-Ampt effective hydraulic conductivity from rainfall-runoff data for the WEPP model. *Trans. ASAE, 37*(2), 411-418. https://doi.org/10.13031/2013.28092

Robichaud, P. R., Elliot, W. J., Lewis, S. A., & Miller, M. E. (2016). Validation of a probabilistic post-fire erosion model. *Int. J. Wildland Fire, 25*(3), 337-350. https://doi.org/10.1071/WF14171

Robichaud, P. R., Elliot, W. J., Pierson, F. B., Hall, D. E., & Moffet, C. A. (2007). Predicting postfire erosion and mitigation effectiveness with a web-based probabilistic erosion model. *CATENA, 71*(2), 229-241. https://doi.org/10.1016/j.catena.2007.03.003

Robichaud, P. R., MacDonald, L. H., & Foltz, R. B. (2010). Fuel management and erosion: Chapter 5. In W. J. Elliot, I. S. Miller, & L. Audin (Eds.), *Cumulative watershed effects of fuel management in the Western United States. Gen. Tech. Rep. RMRS-GTR-231* (pp. 79-100). Fort Collins, CO: USDA, Forest Service, Rocky Mountain Research Station. Retrieved from https://www.fs.usda.gov/treesearch/pubs/34310

SAS Institute. (2003). SAS System Software. Cary, NC: SAS Institute Inc.

Sheridan, G. J., Noske, P. J., Whipp, R. K., & Wijesinghe, N. (2006). The effect of truck traffic and road water content on sediment delivery from unpaved forest roads. *Hydrol. Process., 20*(8), 1683-1699. https://doi.org/10.1002/hyp.5966

Srivastava, A., Brooks, E. S., Dobre, M., Elliot, W. J., Wu, J. Q., Flanagan, D. C.,... Link, T. E. (2020). Modeling forest management effects on water and sediment yield from nested, paired watersheds in the interior Pacific Northwest, USA using WEPP. *Sci. Total Environ., 701*, 134877. https://doi.org/10.1016/j.scitotenv.2019.134877

Srivastava, A., Dobre, M., Wu, J. Q., Elliot, W. J., Bruner, E. A., Dun, S.,... Miller, I. S. (2013). Modifying WEPP to improve streamflow simulation in a Pacific Northwest Watershed. *Trans. ASABE, 56*(2), 603-611. https://doi.org/10.13031/2013.42691

Srivastava, A., Wu, J. Q., Elliot, W. J., Brooks, E. S., & Flanagan, D. C. (2017). Modeling streamflow in a snow-dominated forest watershed using the Water Erosion Prediction Project (WEPP) model. *Trans. ASABE, 60*(4), 1171-1187. https://doi.org/10.13031/trans.12035

Stafford, A. K. (2011). Sediment production and delivery from hillslopes and forest roads in the Southern Sierra Nevada, California. MS thesis. Fort Collins, CO: Colorado State University.

Tiscareno-Lopez, M., Lopes, V. L., Stone, J. J., & Lane, L. J. (1993). Sensitivity analysis of the WEPP watershed model for rangeland applications I: Hillslope processes. *Trans. ASAE, 36*(6), 1659-1672. https://doi.org/10.13031/2013.28509

Trewartha, G. T., & Horn, L. H. (1980). *An introduction to climate*. New York: McGraw-Hill Book Co.

Tysdal, L. M., Elliot, W. J., Luce, C. H., & Black, T. A. (1999). Modeling erosion from insloping low-volume roads with WEPP watershed model. *Transp. Res. Rec., 1652*(1), 250-255. https://doi.org/10.3141/1652-65

USDA. (2023). National cooperative soil survey soil characterization data (Lab data). Retrieved from https://ncsslabdatamart.sc.egov.usda.gov

USDA Forest Service. (1990). R1-WATSED Region 1 Water and Sediment Model. Missoula, MT: USDA Forest Service Region.

Wagenbrenner, J. W., Robichaud, P. R., & Elliot, W. J. (2010). Rill erosion in natural and disturbed forests: 2. Modeling Approaches. *Water Resour. Res., 46*(10). https://doi.org/10.1029/2009WR008315

Welsh, M. J. (2008). Sediment production and delivery from forest roads and off-highway vehicle trails in the Upper South Platte River watershed, Colorado. MS thesis. Fort Collins, CO: Department of Forest, Rangeland, and Watershed Stewardship, Colorado State University.

Whyte, W. S. (1976). *Basic metric surveying* (2nd ed.). London: Butterworth Scientific.

Willmott, C. J., Robeson, S. M., & Matsuura, K. (2012). A refined index of model performance. *Int. J. Climatol., 32*(13), 2088-2094. https://doi.org/10.1002/joc.2419

Wischmeier, W. H., & Smith, D. D. (1978). *Predicting rainfall erosion losses: A guide to conservation planning. Agriculture Handbook 537.* Washington, DC: USDA.

Zhang, X. C. (2016). Evaluating water erosion prediction project model using cesium-137-derived spatial soil redistribution data. *Soil Sci. Soc. Am. J., 81*(1), 179-188. https://doi.org/10.2136/sssaj2016.06.0172

Ziegler, A. D., Sutherland, R. A., & Giambelluca, T. W. (2001). Interstorm surface preparation and sediment detachment by vehicle traffic on unpaved mountain roads. *Earth Surf. Process Landf., 26*(3), 235-250. https://doi.org/10.1002/1096-9837(200103)26:3<235::AID-ESP171>3.0.CO;2-T