Journal of Hydrology 672 (2026) 135340

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Evaluating the influence of climate inputs on WEPP predictions for a forest road in the Florida Panhandle, USA

Jingqiu Chen [a,1,*], Shuyuan Wang [a,b,1], William J. Elliot [c], Bernard A. Engel [d], Feng Pan [d], Yaoze Liu [e], Johnny M. Grace III [f]

[a] Biological Systems Engineering, College of Agriculture and Food Sciences, Florida A&M University, Tallahassee, FL 32307, USA
[b] Department of Statistics and Actuarial Science, The University of Iowa, Iowa City, IA 52242, USA
[c] Department of Soil and Water Systems, University of Idaho, Moscow, ID 83844, USA
[d] Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907, USA
[e] Department of Environmental and Sustainable Engineering, University at Albany, State University of New York, Albany, NY 12222, USA
[f] USDA-Forest Service, Southern Research Station, Center for Forest Watershed Research, Tallahassee, FL 32307, USA

A R T I C L E I N F O

This manuscript was handled by Dan Lu, with the assistance of Yang Gao, Associate Editor

Keywords:
Forest road
WEPP
Climate input formats
Machine learning
Mediation analysis

A B S T R A C T

Rainfall temporal resolution and climate input formats can influence runoff and soil erosion predictions on forest roads. This study evaluated their effects using the Water Erosion Prediction Project (WEPP) simulations for a forest road in Florida's Chipola Experimental Forest. High-resolution NOAA-ASOS precipitation data (2000–2024) were used to generate breakpoint, fixed-interval (5–60 min), and time to peak/peak intensity (TPIP) climate inputs. WEPP predictions were generally stable across input types, with modest differences in average annual runoff and soil loss. Finer temporal resolutions (5–10 min) for breakpoint inputs produced slightly smaller deviations, while TPIP showed greater sensitivity to storm irregularity, particularly for high-intensity, highly peaked, or long-duration events. Machine learning analysis identified total precipitation and average intensity as the main drivers of daily runoff differences, with ratio of peak intensity to average intensity (ip), ratio of duration where peak intensity occurs to total storm duration (tp), and duration having minor influence. Across soil erodibility conditions, average treatment effects (ATEs) were near zero for small storms and increased moderately for larger events, especially under higher infiltration capacities. TPIP inputs exhibited greater variability, suggesting potential need for runoff adjustments. Mediation analysis indicated that natural indirect effects (NIEs) could be reduced by improving runoff predictions, whereas direct effects (NDEs) were mainly influenced by critical shear stress (τ_c). These results highlight the importance of rainfall temporal structure and climate input selection in hydrologic and erosion modeling, providing guidance for improving model predictions and informing future applications.

1. Introduction

Soil erosion is a critical environmental challenge that threatens soil productivity, water quality, and ecosystem stability worldwide. In forested landscapes, road networks constructed for forestry operations introduce substantial disturbances to soil systems. The soils of forest roads differ markedly from adjacent undisturbed topsoil, exhibiting altered physical and hydrological properties that increase their susceptibility to erosion (Zemke, 2016). Unpaved roads are particularly problematic, as they are widely recognized as dominant sediment sources within forested watersheds (Grace III, 2017; Yu et al., 2024). During rainfall events, components of the road prism, including cut slopes, fill slopes, and road surfaces, are highly vulnerable to splash erosion and runoff scouring, resulting in the detachment and mobilization of soil particles (Wang et al., 2023; Elliot et al., 2023; Chen et al., 2025). Moreover, forest roads act as efficient conduits for runoff and sediment, altering natural hydrologic pathways and enhancing sediment transport efficiency within watersheds (Eisenbies et al., 2007; Yousefi et al., 2016; Farias et al., 2021). Understanding and quantifying these erosion processes is therefore essential for sustainable watershed

* Corresponding author.
E-mail address: jingqiu.chen@famu.edu (J. Chen).
[1] Co-author both contributed equally to this work (Co-first authorship).

https://doi.org/10.1016/j.jhydrol.2026.135340
Received 16 December 2025; Received in revised form 27 February 2026; Accepted 16 March 2026
Available online 19 March 2026

management and the development of effective mitigation strategies.

The importance of understanding road-related erosion is particularly pronounced in Florida, where climatic and landscape characteristics heighten hydrologic sensitivity. The state's humid subtropical climate is marked by frequent, high-intensity rainfall events, including tropical storms and hurricanes, which generate substantial runoff and accelerate sediment mobilization from disturbed surfaces (Chen et al., 2025, Chen et al., 2026). Florida's Panhandle, in particular, is dominated by extensive pine forests and unpaved road networks that support silvicultural operations, recreation, and rural access. These roads often intersect with shallow water tables, poorly drained soils, and dense drainage networks, creating direct hydrologic connectivity between road surfaces and nearby streams and wetlands. As a result, even localized erosion from forest roads can have disproportionately large impacts on water quality, aquatic habitats, and downstream ecosystems. Given Florida's reliance on its forest resources and the ecological value of its wetlands and streams, quantifying and managing road-induced erosion is essential for sustainable land management and watershed protection in the region.

The Water Erosion Prediction Project (WEPP) model is based on fundamentals of infiltration, surface runoff, plant growth, residue decomposition, hydraulics, tillage management, soil consolidation, and erosion mechanics (Flanagan and Livingston, 1995). This model combines physically based erosion and hydrology models with a stochastic climate generator to estimate soil loss and deposition and thus facilitate the selection of management practices to minimize soil erosion. The WEPP model has been widely used in forest runoff and erosion applications (Elliot, 2013; Cao et al., 2021; Dun, et al., 2009; Grace III, 2017). Elliot and Hall (2010) have published several online interfaces for WEPP applications to forest and rangeland hillslopes that require minimal user inputs to direct the interface to access online databases for soils, vegetation and climate inputs (Elliot, 2004). These have been very popular with the USDA Forest Service and other land management agencies, completing tens of thousands of WEPP runs annually.

When the WEPP model was developed in the late 1980 s, a national climate database was required to support its simulations. Nicks et al. (1995) initiated such a database for daily precipitation and subsequently enhanced it by incorporating sub-daily rainfall duration and intensity, using data from a limited number of recording rain gages. Since then, recording rain gages have become widely available, highlighting the need to evaluate whether this enhanced dataset can improve WEPP model performance and to identify the most effective way to integrate the expanding rain gage data into WEPP input files. Precipitation intensity is one of the most influential climate drivers of infiltration, runoff, and soil erosion, and both detachment and transport processes are highly sensitive to intensity inputs (Flanagan et al., 2020; Nearing et al., 1990). The WEPP model accommodates this sensitivity through two primary options for climate input: (1) stochastic climate data generated by CLImate GENerator (CLIGEN), which provides daily storm characteristics including storm depth, storm duration, ratio of peak intensity to average intensity (ip), and ratio of duration where peak intensity occurs to total storm duration (tp), and (2) direct use of observed breakpoint precipitation data that capture the true variability of rainfall events (Flanagan and Nearing, 1995; Flanagan et al., 2001; Nicks et al., 1995). CLIGEN-generated files only allow a single rainfall with a single peak intensity in a day, with the event starting at midnight. While CLIGEN is widely used for conservation planning and long-term simulations, its reliance on statistical summaries can introduce uncertainty when representing storm intensity and distribution. In contrast, breakpoint data preserve rainfall variability more accurately. A breakpoint input file has pairs of points providing the cumulative precipitation for a given time of day. The format allows users to specify the starting time of an event, and the format has the ability to describe hyetographs with more than one peak, or precipitation patterns that may have multiple storms within a 24hr period. Breakpoint data are often used for validation studies where observed runoff and erosion rates can be linked to precipitation patterns (Elliot et al, 2023). However, breakpoint availability is often limited in spatial and temporal coverage. Fixed-interval precipitation records (e.g., 5-, 15-, or 60-min data) offer broader coverage but tend to underestimate peak intensities, potentially biasing runoff and soil loss predictions (Hollinger et al., 2002). These differences in climate inputs are critical, as underrepresentation of rainfall intensity can lead to systematic underestimation of sediment detachment and transport in WEPP, underscoring the importance of input resolution for reliable hydrologic and erosion predictions.

Runoff within the WEPP model is a two-step process following the Green-Ampt Mein Larson approach (Flanagan and Nearing, 1995). In the first step, the "time to ponding" is estimated:

$$t_p = \frac{\varphi_f \theta}{i - K_e} \tag{1}$$

where t_p is time to ponding (s), φ_f is the average soil capillary potential (m), θ is the soil moisture deficit (m m^{-1}), i is the rainfall rate (m s^{-1}) and K_e is the effective saturated hydraulic conductivity (m s^{-1} internally, mm h^{-1} as a WEPP input variable). Once ponding has occurred ($t > t_p$), the infiltration rate $f(t)$ for a given time t is:

$$f(t) = K_e\left(1 + \frac{\varphi_f \theta}{F(t)}\right) \tag{2}$$

where $F(t)$ is the cumulative infiltration. K_e is a WEPP soil file input and other variables are calculated internally by WEPP as functions of other input soil properties.

Despite widespread application of the WEPP model in forested environments, erosion assessments have primarily relied on CLIGEN-generated climate inputs, or adjusted datasets parameterized for CLIGEN but still based on daily storm characteristics (tp and ip) (Elliot, 2013; Dun et al., 2009), including applications in the southeastern United States (Dun et al., 2013; Elliot et al., 2023). In humid subtropical landscapes, where unpaved forest roads are hydraulically connected to streams and erosion is driven by short-duration, high-intensity rainfall, this approach may inadequately represent dominant runoff and sediment transport processes. Given the sensitivity of WEPP runoff and erosion predictions to precipitation intensity, differences in the choice and temporal resolution of climate inputs may affect simulation accuracy, yet their influence on forest road erosion estimates derived from high-resolution observed precipitation data remains insufficiently documented in the region.

Although erosion processes in forest ecosystems have been widely studied, model applications have primarily focused on agricultural settings. Forest roads, however, remain less explored. While WEPP predictions are known to be sensitive to precipitation intensity, the influence of climate input temporal resolution on runoff and erosion estimates for forest roads has received limited attention, particularly in subtropical regions such as Florida where short-duration, high-intensity rainfall dominates.

To address these gaps, this study evaluates the effects of multiple climate input types, including breakpoint data, fixed-interval records, and CLIGEN TPIP data, on WEPP predictions for a forest road in the Florida Panhandle. By comparing model outcomes across different climate inputs, we aim to (1) assess the influence of climate input types on the WEPP-predicted runoff and erosion for the three road surface types including bare, compacted, and gravel surfaces; (2) identify the most influential precipitation events; (3) estimate the average treatment effects (ATE) of climate inputs on predicted runoff under varying hydraulic conductivity (K_e) values; and (4) evaluate the natural direct effects (NDE) and natural indirect effects (NIE) of climate inputs on soil erosion via runoff, considering the influence of three key soil erodibility parameters (interrill erodibility K_i, rill erodibility K_r, and rill critical shear τ_c). Collectively, these objectives aim to advance methodological understanding of WEPP applications in forest road contexts, improve insight into how climate inputs influence runoff and erosion predictions across precipitation patterns, hydraulic conductivity, and soil erodibility, and provide guidance for potential model adjustments.

2. Materials and methods

This study was conducted at the Chipola Experimental Forest in Florida, U.S. (Fig. 1), an area significantly altered by Hurricane Michael

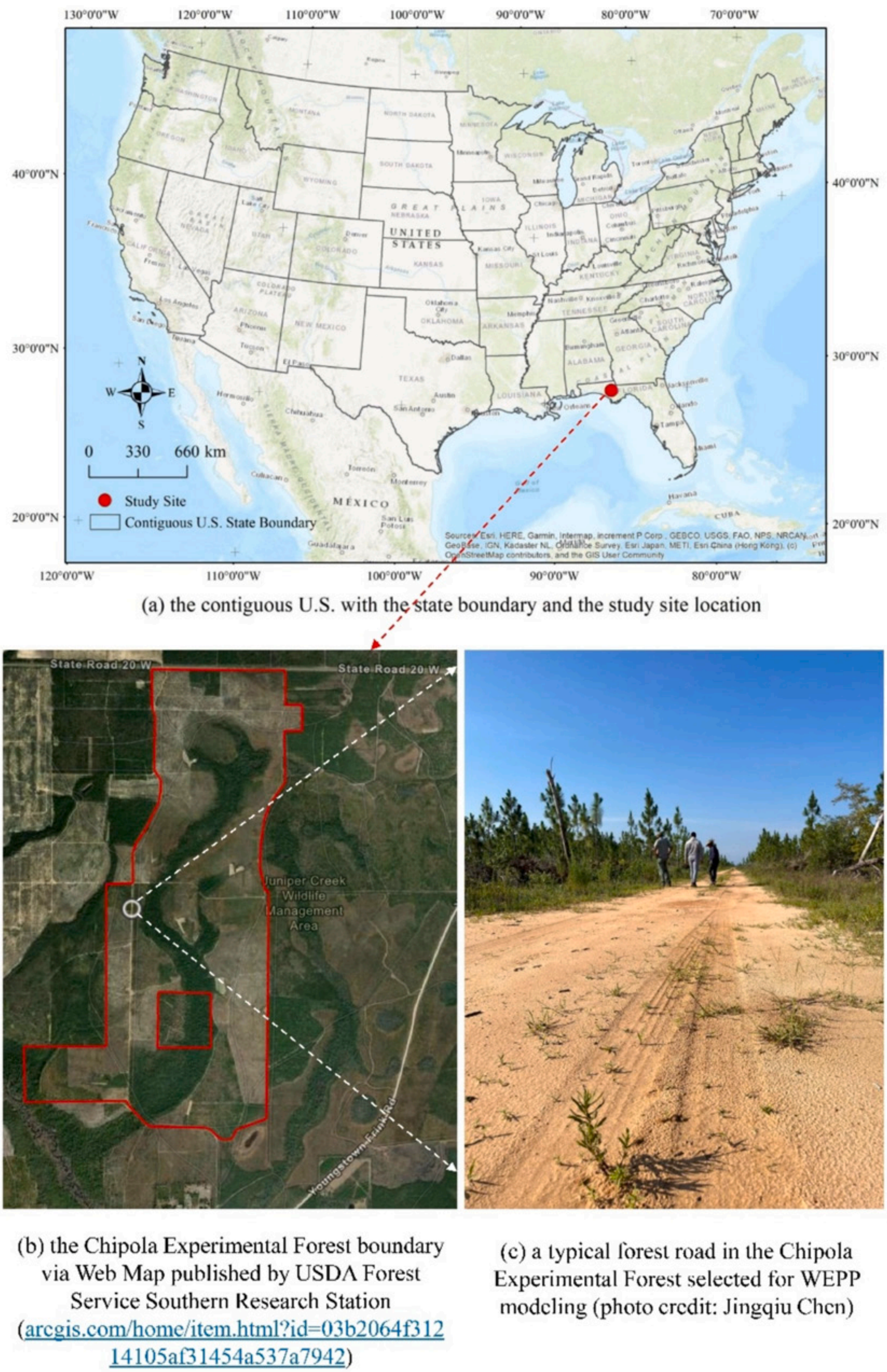

Fig. 1. Location of the typical forest road study site in the Chipola Experimental Forest, FL.

in 2018. The frequent reforestation activities increased the disturbance of the road surfaces in the study area, which led to increased runoff and erosion in this region.

A typical forest road in the Chipola Experimental Forest was selected for modeling (30.4178° N, −85.2628° E). The climate data was obtained from the closest NOAA-ASOS (the National Oceanic and Atmospheric Administration's (NOAA) Automated Surface Observation System (ASOS) network) station, Mariana (30.8356° N, −85.1839° E) for 2000 – 2024. The average annual precipitation was 1244.62 mm, with June to August as the wettest season. However, precipitation is generally distributed throughout the year with May, September, October, and November the relatively dry months. The average maximum temperature was 26.39 °C, and the minimum was 13.89 °C. The road segment studied was 100 m long, 4 m wide, with 1% gradient. The segment was crowned with the surface sloping to either side. The "inslope" template in WEPP was applied, which assumes all runoff is diverted to the inside ditch, rill spacing equals the road width (4 m) (Elliot et al., 1999). Field observations of minimal rill development on the study road support the applicability of these assumptions. Management and vegetation inputs were assumed unchanged, reflecting the absence of active disturbances such as weeding, ditch cleaning, or surface operations.

An overview of the modeling and analytical workflow used in this study is provided in Fig. 2.

2.1. Climate input scenarios and WEPP simulation setup

The WEPP model is a FORTRAN program with two versions, a hillslope version and a watershed version that links hillslopes with channels and in-channel impoundments. A windows interface allows users to customize WEPP's four input files (soil, topography, vegetation and climate) (Flanagan et al., 2001). Online interfaces have been developed including WEPP: Road for road erosion modeling that limits user access to WEPP's input files while allowing users to select from a limited menu one of four soil textural categories, level of traffic, and whether the road is graveled. Both interfaces have been used to model road erosion.

WEPP: Road (https://forest.moscowfsl.wsu.edu/fswepp) is an interface to the WEPP model that enables road erosion simulations and is linked exclusively to the Rock: Clime online interface to the CLIGEN model, with a database of 2600 weather stations in the U.S. However, this study aims to simulate road runoff and erosion at multiple climate temporal scales, which is not supported by WEPP: Road, as it is designed for average annual scale assessments. Therefore, the latest version of the WEPP Windows hillslope version (September 2024) was used for this study.

When using the WEPP windows hillslope model, the road surface can be simulated as a single overland flow element (OFE). This study focused on simulating the forest road, which was one OFE. To utilize the WEPP Windows hillslope model, four input files are required: climate, soil, slope, and management input files (Flanagan and Livingston, 1995).

For the climate input, WEPP requires multiple climate variables, including precipitation, maximum, minimum and dew point temperatures, wind speed and direction, and solar radiation. All but precipitation is required as daily inputs to WEPP, and the resolution of the climate inputs can vary for observed data. For this study, we used the "breakpoint" format for the climate file.

The climate inputs were generated using WEPP Climate File Formatter (WEPPCLIFF; McGehee et al., 2020) including the breakpoint, 5-, 10-, 15-, 30-, 60-min from the 5-min ASOS gauge data, which is free to the public. The breakpoint climate input was created from 5-min gauge data using a novel disaggregation 'spreading' procedure for intensities that were less than the precision limitation of the ASOS gauges. This method was summarized in Fullhart et al. (2020).

The 'TPIP' format as described above was also evaluated in this study, which was generated by summarizing the 'breakpoint' format data. The breakpoint and TPIP formats provide concise yet informative descriptions of storm structure and dynamics, allowing for systematic comparisons across events and facilitating the analysis of precipitation patterns in hydrological and climatological studies.

The dominant undisturbed soil was mainly Fuquay loamy sand, which contains 86% sand, 10.2% silt, and 3.8% clay. The soil input file contains soil texture and erodibility parameters. The selected study road is bare surface (Fig. 1). Two additional road mitigation scenarios were considered in the study: (1) changing bare soil road surfaces to a gravel-covered surface, and (2) adopting the improved "graded aggregate base" design as described in Elliot et al. (2023). The erodibility parameter values for these two scenarios are detailed in Table 1. Parameters for vegetated or gravel-covered road surfaces were derived from the road WEPP soil database, while those for the improved compacted roads were sourced from Elliot et al. (2023). Elliot et al. (2023) studied Troup loamy sand soil, which consists of 86% sand, 11% silt, and 3% clay, closely matching the Fuquay loamy sand in this study. Given the close similarity in texture between Fuquay and Troup loamy sands and the reduction of natural soil structure differences by road compaction, the Troup-derived erodibility parameters are considered representative for the study site. While soil erodibility can be measured directly through flume experiments (Wang et al., 2020), the similarity of the soils and the time-intensive nature of such measurements justified adopting the parameter values directly.

The slope file contains topographic information. The studied road is an insloped road with limited or no vegetation or gravel cover in the ditches (Fig. 1). The road is 100 m long and 4 m wide, with a 1% gradient.

The management file describes the vegetation cover and management details. This study includes one bare road segment without road fills or forest buffer. The management file named "road-insloped surface with bare ditches" was selected from the WEPP windows management database.

Climate Input Scenarios + Soil Parameters
↓
WEPP Model Simulations
(Runoff → Erosion)
↓
Runoff-Focused Analyses
├─ Machine Learning
│ (Identify influential rainfall events)
└─ Causal Inference
(Quantify effects of rainfall characteristics and hydraulic conductivity on runoff)
↓
Erosion-Focused Analysis
└─ Mediation Analysis
(Rainfall → Runoff → Erosion, conditioned on critical shear stress, rill and interrill erodibility)

Fig. 2. Overview of the modeling and analytical workflow used in this study.

Table 1
Soil erodibility parameter values used in WEPP for different surface conditions.

Parameters	Bare road surface	Gravel road surface	Compacted road surface (Elliot et al., 2023)
K_e (mm/h)	3.8	3.8	1.3
K_i (kg·s/m^4)	2,000,000	2,000,000	1,000,000
K_r (s/m)	0.0004	0.0004	0.0008
τ_c (N/m^2)	0.04	10	1.5

Note: Values are shown for the default WEPP sandy loam insloped bare road surface, vegetated rock ditch, gravel road surface and the improved loamy sand road surface developed by Elliot et al. (2023) for Fort Benning, GA.

2.2. Machine learning analysis of runoff differences

The runoff and erosion were estimated daily by the WEPP model. This study summarized the predictions to monthly, annual, and average annual for each of the climate scenarios. All data analysis and machine learning were for the original daily scale. In this study, we considered 7 climate scenarios: breakpoint, fixed intervals including 5-, 10-, 15-, 30-, 60-min, and TPIP data.

Breakpoint precipitation data were used as the reference in this study to compare other climate variables. Breakpoint data preserve the temporal variability of rainfall intensity, which is critical for accurately simulating infiltration-excess runoff and erosion processes in WEPP. Although the breakpoint data employed were derived from 5-min precipitation records, this format retains essential storm structure information and is the recommended default input for WEPP (Flanagan and Livingston, 1995; Flanagan et al., 2020). Breakpoint rainfall data were generated by aggregating consecutive intervals with measurable precipitation into single periods of constant intensity. For intervals with alternating zero and small non-zero values ($\leq$ instrument precision), the total accumulation over the duration was summed and uniformly redistributed to represent low-intensity rainfall (Fullhart et al., 2020).

To further investigate how precipitation characteristics influence discrepancies in WEPP-simulated runoff among different climate input formats, we employed machine learning models to predict the runoff difference between each climate input and the breakpoint data. For each simulation, runoff predictions generated using breakpoint data were treated as the baseline, and differences in runoff from other climate inputs were calculated relative to this reference. We hypothesized that these runoff differences were driven by variations in key storm characteristics. Accordingly, five precipitation variables were selected as explanatory covariates: total storm depth (prcp, mm), storm duration (dur, h), average storm intensity (intensity, mm h^{-1}), tp, and ip.

Four machine learning algorithms were used to model the relationship between these precipitation characteristics and the runoff differences: random forest, least absolute shrinkage and selection operator regression (LASSO), generalized additive models (GAM), and extreme gradient boosting (XGBoost). LASSO allows sparse, linear selection of relevant covariates; GAM captures nonlinear and additive relationships; RF and XGBoost capture complex nonlinear interactions, thresholds, and storm-event-specific effects. For each model, the dataset was randomly split into 70% training and 30% validation sets. Model performance was evaluated using the coefficient of determination (R2), root mean square error (RMSE) and mean absolute error (MAE) to quantify each algorithm's ability to explain the variability in runoff differences relative to the breakpoint reference. Let y_i denote the observed runoff difference between the runoff predicted using the WEPP breakpoint climate input and the runoff predicted using one of the six alternative climate inputs (5-min, 10-min, 15-min, 30-min, 60-min, and TPIP). Let $\hat{y}_i$ be the corresponding model-predicted runoff difference and let n be the number of rainfall events. The three metrics are defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

where $\bar{y}$ is the sample mean of the observed runoff differences.

This modeling framework allowed us to identify the rainfall conditions, such as storm size, intensity structure, and timing, under which runoff predictions from alternative climate input formats deviated most from those generated with breakpoint data.

Although validation performance provides an indication of model generalization, the primary objective of this analysis was explanatory rather than predictive. Because the goal was to identify stable relationships between precipitation characteristics and runoff differences, moderate validation R^2 values did not preclude interpretation. After evaluating model performance using the training–validation split, the best-performing model was refit using the full dataset to maximize the information available for interpretation. Partial dependence plots (PDP) were then generated from this final model to characterize precipitation conditions associated with the largest runoff differences.

Machine learning methods were selected for their ability to capture complex nonlinear relationships and to support post-hoc interpretation of marginal effects between precipitation characteristics and runoff differences through techniques such as partial dependence analysis. Machine learning was applied to predicted runoff rather than soil erosion because runoff is the primary driver of sediment detachment and transport in WEPP. Directly linking precipitation to erosion could conflate effects mediated through runoff, making interpretation unclear. Focusing on runoff allows for a more direct and interpretable identification of the most influential precipitation events. The framework developed for analyzing runoff via machine learning in this study can guide future studies applying similar approaches to soil erosion predictions.

2.3. Causal inference analysis of climate input effects on runoff

To quantify the influence of climate input formats on WEPP-predicted runoff, we employed a causal inference framework based on the average treatment effect (ATE). Each climate input format (5-, 10-, 15-, 30-, 60-min, and TPIP) was considered a "treatment," with the breakpoint data serving as the reference (control). The outcome variable was the WEPP-simulated runoff corresponding to each input format. Key precipitation characteristics, such as precipitation amount, duration, average intensity, tp, and ip, were included as confounders, as they influence both the climate input representation and the resulting runoff response.

The identification of causal effects rested on four standard assumptions: (1) no interference, meaning runoff under one climate input is unaffected by other inputs; (2) exchangeability, assuming that, conditional on the measured storm characteristics, the assignment of climate input formats is independent of potential runoff outcomes; (3) positivity, requiring that each format has a nonzero probability of occurrence across the observed range of storm characteristics; and (4) consistency, indicating that the observed runoff under a specific climate input corresponds to its potential outcome.

Under these assumptions, the ATE estimates the expected difference in runoff between each climate input format and the breakpoint reference, adjusting for confounding by storm characteristics. This framework isolates the effect of input resolution on runoff predictions, mitigating bias from differences in precipitation features. Fig. 3 presents the corresponding directed acyclic graph (DAG), in which climate input

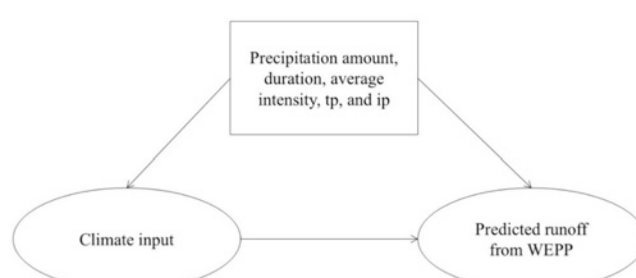

Fig. 3. Direct Acyclic Graph (DAG) showing the assumed causal relationships between precipitation input formats (treatment), runoff differences (outcome), and storm characteristics as confounders.

format is the treatment, runoff is the outcome, and precipitation characteristics act as confounders influencing both.

In WEPP, hydraulic conductivity (K_e) is a key parameter that governs infiltration and runoff generation. To assess how soil surface properties influence the causal effects of climate inputs, ATEs were calculated across six K_e scenarios (1.3, 1.8, 2.3, 2.8, 3.3, and 3.8 mm/h) including but not limited to the ones for the soil surfaces considered in this study: bare, compacted, and gravel. This approach allowed us to evaluate the sensitivity of climate input effects on runoff predictions to soil hydraulic properties, providing insight into the interactions among rainfall characteristics, soil infiltration capacity, and runoff response.

2.4. Mediation analysis of climate input effects on erosion via runoff

To elucidate the mechanisms through which climate input formats influence soil erosion, we applied a causal mediation analysis with WEPP-predicted runoff as the mediator. In this framework, climate input format was defined as the exposure, predicted erosion as the outcome, and runoff as the mediating variable transmitting part of the treatment effect. The same set of storm characteristics was included as confounders to adjust for factors influencing both runoff and erosion responses.

A counterfactual-based mediation framework was used to decompose the total effect of climate input format on predicted erosion into natural direct and indirect components. The natural direct effect (NDE) represents the expected change in predicted erosion attributable to differences in climate input format while holding runoff fixed at the level it would attain under the reference (breakpoint) condition. The natural indirect effect (NIE) quantifies the portion of the effect that operates through changes in runoff. Together, these components characterize the extent to which the influence of climate input format on erosion is mediated by hydrological processes.

The assumed causal structure is illustrated in Fig. 4, where climate input format serves as the treatment, predicted runoff as the mediator, and predicted erosion as the outcome. Arrows from climate input to erosion, both directly and indirectly through runoff, represent the hypothesized causal pathways underlying this analysis.

The mediation analysis was conducted in R using the mediation package. Two regression models were specified: (i) a *mediator model*, in which predicted runoff was regressed on climate input format and the confounders, and (ii) an *outcome model*, in which predicted erosion was regressed on climate input format, runoff, and the same confounders. The *mediate* function in R was then applied to estimate the NDE and NIE, along with their 95% confidence intervals, using nonparametric bootstrap procedures. This approach accounts for potential confounding between the treatment and mediator, assuming that all relevant confounders have been measured and included in the analysis. Under this framework, causal interpretation of the estimated natural direct and indirect effects requires several assumptions. First, there must be no unmeasured confounding of the treatment–outcome, treatment–mediator, or mediator–outcome relationships after adjusting for the included covariates. Second, the standard causal inference assumptions of consistency and positivity must also be held.

To further assess the soil property influences on the decomposition of treatment effects, we performed a sensitivity analysis of key WEPP erodibility parameters including the critical shear stress (τ_c), rill erodibility (K_r), and interrill erodibility (K_i). In this analysis, the hydraulic conductivity (K_e) was fixed at the bare soil surface value, and each of the other parameters, τ_c, K_r, and K_i, were varied one at a time, while all remaining parameters were kept at their bare-surface values. Specifically, τ_c, K_r, and K_i were each varied across physically meaningful ranges (Table 2) to isolate their respective influences on the estimated NDE and NIE. This approach allows the effects of each parameter to be examined independently, avoiding potential confounding from interactions among parameters, and substantially reduces computational demand compared to evaluating all possible parameter combinations.

3. Results

3.1. Influence of climate inputs on average annual runoff and erosion predictions

The average annual runoff and soil loss predicted by the WEPP model for the selected forest road varied substantially with climate input type and road surface condition (Fig. 5). For the bare surface, the predicted average annual runoff ranged from 405.2 mm (60-min input) to 428.3 mm (TPIP), with breakpoint and 5-min inputs yielding similar values around 411 mm (Fig. 5a). Fine-resolution inputs (BK, 5-min) produced slightly lower runoff than TPIP, whereas intermediate resolutions (10-, 15-, and 30-min) generated marginally higher estimates (416–418 mm). The coarsest 60-min input estimated lower runoff by approximately 5% relative to TPIP. Similar temporal patterns were observed for the compacted and gravel surfaces, though absolute runoff values differed considerably due to the influence of road properties.

Table 2
Soil erodibility parameter values considered to assess their impact on the estimated NDE and NIE.

Parameter	Values
K_i (kg·s/m^4)	1000000, 1200000, 1400000, 1600000, 1800000, 2,000,000
K_r (s/m)	0.0004, 0.0005, 0.0006, 0.0007, 0.0008
τ_c (N/m^2)	0.04, 0.1, 0.2, 0.5, 1.0, 1.5, 3.0, 5.0, 10.0

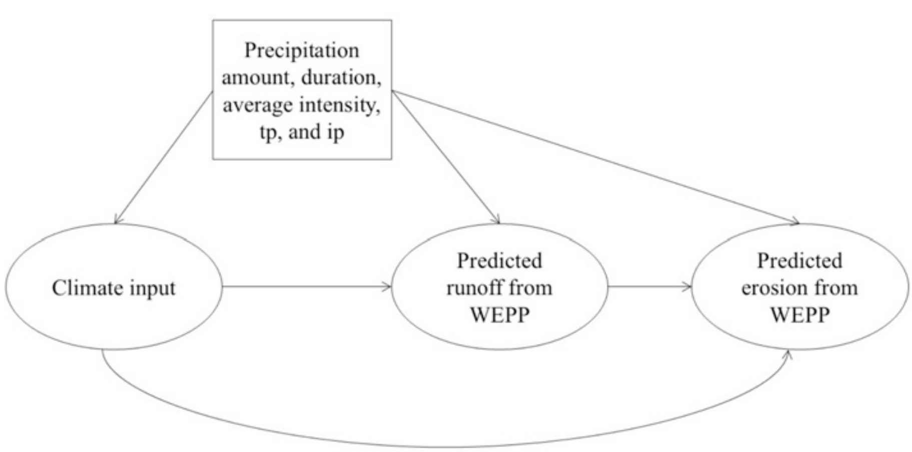

Fig. 4. DAG showing the assumed causal relationships between precipitation input formats (treatment), predicted runoff (mediator), predicted erosion (outcome), and storm characteristics as confounders.

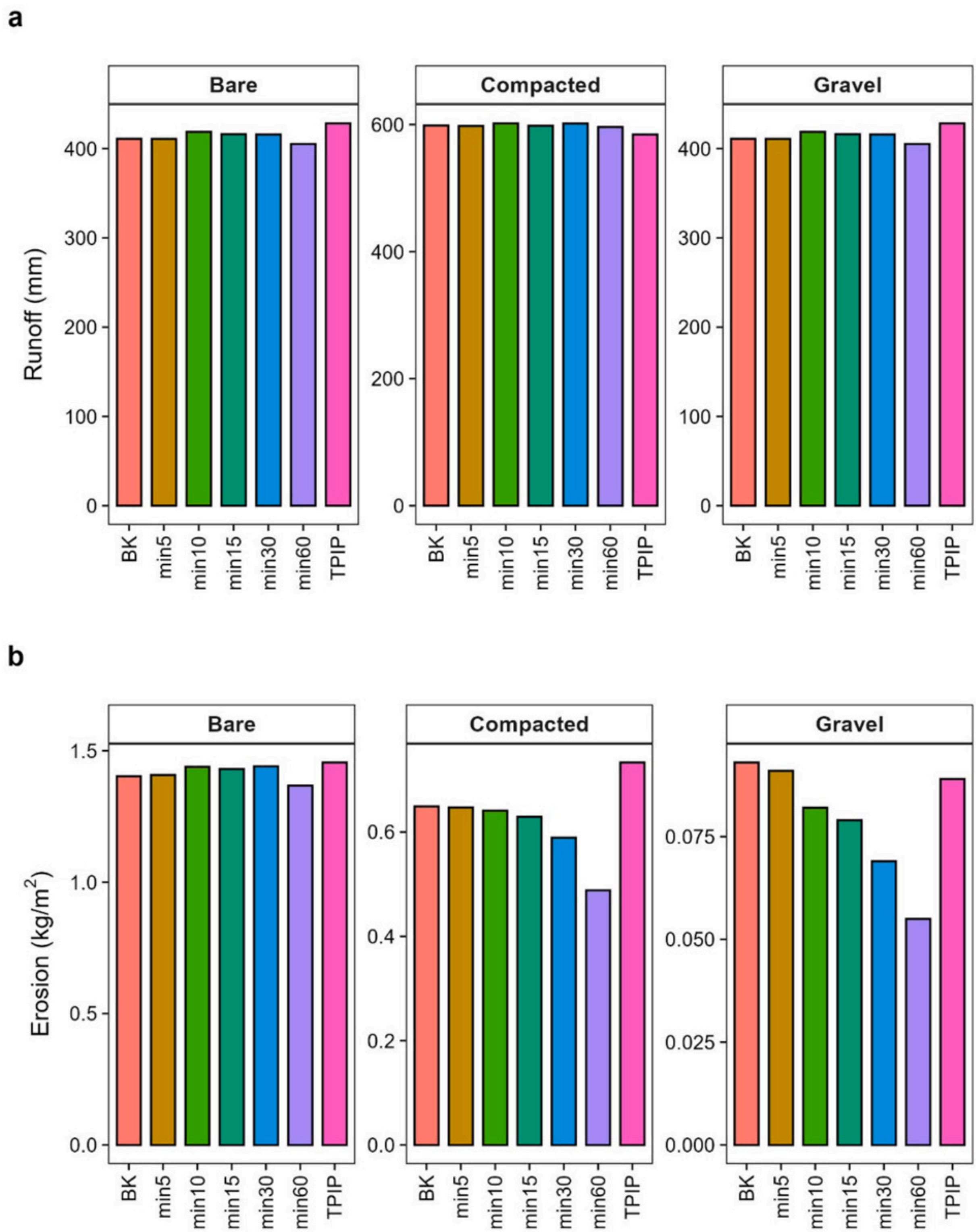

Fig. 5. WEPP-predicted average annual (a) runoff and (b) erosion under different climate inputs for bare, compacted, and gravel forest road surfaces.

Soil loss exhibited greater sensitivity to both precipitation resolution and road surface condition (Fig. 5b). For the bare road, the predicted average annual erosion ranged from 1.367 kg/m^2 (60-min) to 1.455 kg/m^2 (TPIP). Finer-resolution inputs (BK, 5-min) produced erosion estimates near 1.4 kg/m^2, while coarser inputs (30- and 60-min) estimated lower erosion by about 6–7% compared with TPIP. Erosion from compacted surfaces was substantially lower (0.488–0.733 kg/m^2), highlighting the dominant role of compaction in reducing detachment and sediment yield. Gravel surfaces exhibited minimal erosion (<0.1 kg/m^2), with finer-resolution inputs producing slightly higher estimates than coarser ones.

Across all surface types, TPIP produced the highest runoff and erosion estimates for bare and compacted roads, whereas the 60-min input systematically underestimated both. Intermediate resolutions (10-, 15-, and 30-min) yielded runoff comparable to BK and 5-min inputs but occasionally estimated lower erosion. Overall, the results show that temporal resolution systematically affected WEPP predictions of runoff and soil loss, particularly for less compacted surfaces.

3.2. Machine learning model performance and identification of influential precipitation events

To investigate how precipitation characteristics influence runoff differences at the daily scale, we modeled the day-to-day deviations in WEPP-simulated runoff relative to breakpoint data. Five storm variables, total precipitation amount, duration, average intensity, tp, and ip, were used as predictors. Four machine learning algorithms (LASSO, GAM, random forest, and XGBoost) were applied to capture both linear and nonlinear relationships, allowing us to identify the rainfall conditions that drive the largest daily deviations from the breakpoint reference.

Model performance was initially evaluated on the bare road surface using a 70% training and 30% validation split for algorithm selection. As shown in Table 3, XGBoost consistently achieved the highest predictive accuracy across all three surfaces (bare: $R^2 = 0.510$, RMSE = 0.372 mm, MAE = 0.090 mm; compacted and gravel surfaces in parentheses), indicating it reliably captures runoff deviations driven by precipitation characteristics.

Based on the training–validation evaluation, XGBoost was selected as the best-performing algorithm and subsequently refit using the full dataset to maximize information for interpretive analysis. On the full dataset, the final model achieved $R^2 = 0.938$, RMSE = 0.105 mm, and MAE = 0.042 mm, reflecting a close fit to observed runoff differences. Table 4 summarizes the importance of precipitation variables in predicting daily WEPP-simulated runoff differences as estimated by the XGBoost model based on the entire dataset. Three importance metrics are reported: Gain, Cover, and Frequency. Gain represents the relative contribution of each variable to improving the model's predictions, reflecting its importance in reducing prediction error, Cover indicates the proportion of observations affected by splits on a given variable, providing insights into how broadly the variable influences the dataset, and Frequency denotes how often a variable is used for splitting across all trees, indicating the consistency of its influence on model decisions. Among the five storm characteristics considered, total precipitation was the most influential, accounting for 54.3% of the model gain and affecting a majority of the splits (Cover = 0.607; Frequency = 0.460). Average intensity was the second most important variable (Gain = 0.223), suggesting that rainfall intensity also substantially influences model performance. In contrast, peak timing (tp), peak intensity (ip), and storm duration exhibited comparatively lower importance across all three metrics, indicating that these factors have secondary effects. These results indicate that daily runoff differences are primarily controlled by storm magnitude and intensity, with secondary influence from peak timing and duration, highlighting the importance of rainfall structure in driving deviations from breakpoint-based runoff simulations.

Table 3
Performance of four machine learning algorithms for predicting daily runoff differences. Values outside parentheses are for the bare surface; values in parentheses correspond to compacted and gravel surfaces, respectively.

Model	Type	R^2	RMSE	MAE
LASSO	Linear, penalized regression	0.209 (0.262, 0.201)	0.473 (0.153, 0.438)	0.100 (0.076, 0.092)
GAM	Additive, nonlinear smooths	0.464 (0.345, 0.469)	0.389 (0.144, 0.357)	0.098 (0.072, 0.077)
Random Forest	Nonlinear, ensemble trees	0.377 (0.328, 0.353)	0.419 (0.146, 0.395)	0.089 (0.060, 0.073)
XGBoost	Nonlinear, boosted trees	0.510 (0.427, 0.511)	0.372 (0.135, 0.343)	0.090 (0.064, 0.077)

Table 4
Feature importance of precipitation variables in the XGBoost model predicting daily runoff differences relative to breakpoint inputs for bare road surface.

Feature	Gain	Cover	Frequency
Total precipitation amount	0.543	0.607	0.460
Average intensity	0.223	0.149	0.177
ip	0.083	0.068	0.115
tp	0.078	0.092	0.125
Duration	0.073	0.084	0.124

Fig. 6 compares the observed and predicted daily differences in runoff between the stated climate format and the breakpoint format produced by the XGBoost model relative to the breakpoint data for the bare road surface. For the fixed-interval climate inputs (min5, min10, min15, min30, and min60), the daily runoff differences were generally small, with over 95% of the values falling within ± 2 mm and nearly all (≥99%) within ± 5 mm. The 5-min input yielded the smallest deviations, with 99.13% of the differences within ± 2 mm. In contrast, replacing the breakpoint input with the TPIP data led to slightly larger discrepancies, with 94.54% of the differences within ± 5 mm and a maximum deviation of approximately 20 mm; 88.07% of the TPIP-based differences remained within ± 2 mm.

The XGBoost model showed excellent predictive performance, with coefficients of determination (R^2) exceeding 0.96 and root mean square errors (RMSE) below 0.18 mm for the fixed-interval inputs, and an RMSE of 0.34 mm for the TPIP input. Although the overall R^2 for the combined dataset was 0.938, the subtype-specific R^2 values for each climate input were even higher (all above 0.95). This difference arises because the overall R^2 reflects both within-type and between-type variability, whereas the subtype-specific values are computed within relatively homogeneous climatic conditions, where predictor–response relationships are more consistent. As a result, the unexplained variance is smaller within each subtype, leading to higher R^2 values. Minor systematic biases between climate input types may slightly reduce the overall R^2 when data are pooled, even though the model performs strongly for each input category.

Fig. 7 presents the partial dependence plots (PDPs) generated from the XGBoost model for the bare ditch scenario, showing how individual precipitation characteristics influence the predicted daily runoff differences relative to the breakpoint data across all climate input formats. PDPs isolate the marginal contribution of each predictor by averaging over the effects of other variables, thereby providing an interpretable measure of variable importance within a complex nonlinear model. Positive PDP values indicate increased runoff differences associated with the corresponding predictor values. The observed nonlinear and interactive patterns highlight relationships that are not detectable through conventional regression or exploratory analysis, demonstrating the strength of machine learning approaches in capturing and visualizing complex, multidimensional dependencies in hydrologic simulations.

Among all precipitation characteristics, total precipitation was the most influential factor, producing the largest daily runoff differences, exceeding 6 mm for certain climate inputs. For total precipitation below 150 mm, the runoff differences across the fixed-interval inputs (5–60 min) were similar and generally within ± 3 mm. Larger positive differences appeared between 150 and 200 mm, with finer temporal resolutions (e.g., 5- and 10-min inputs) yielding smaller and more constrained differences. Negative runoff differences were observed for the 5-min input when total precipitation ranged between 100 and 150 mm.

For average intensity, runoff differences remained within ± 1 mm across all climate inputs. The finer-resolution inputs (5- and 10-min) maintained near-zero differences across the entire intensity range, whereas the coarser inputs (15–60 min) exhibited moderate fluctuations around 20–30 mm h^{-1} and tended toward stable negative differences

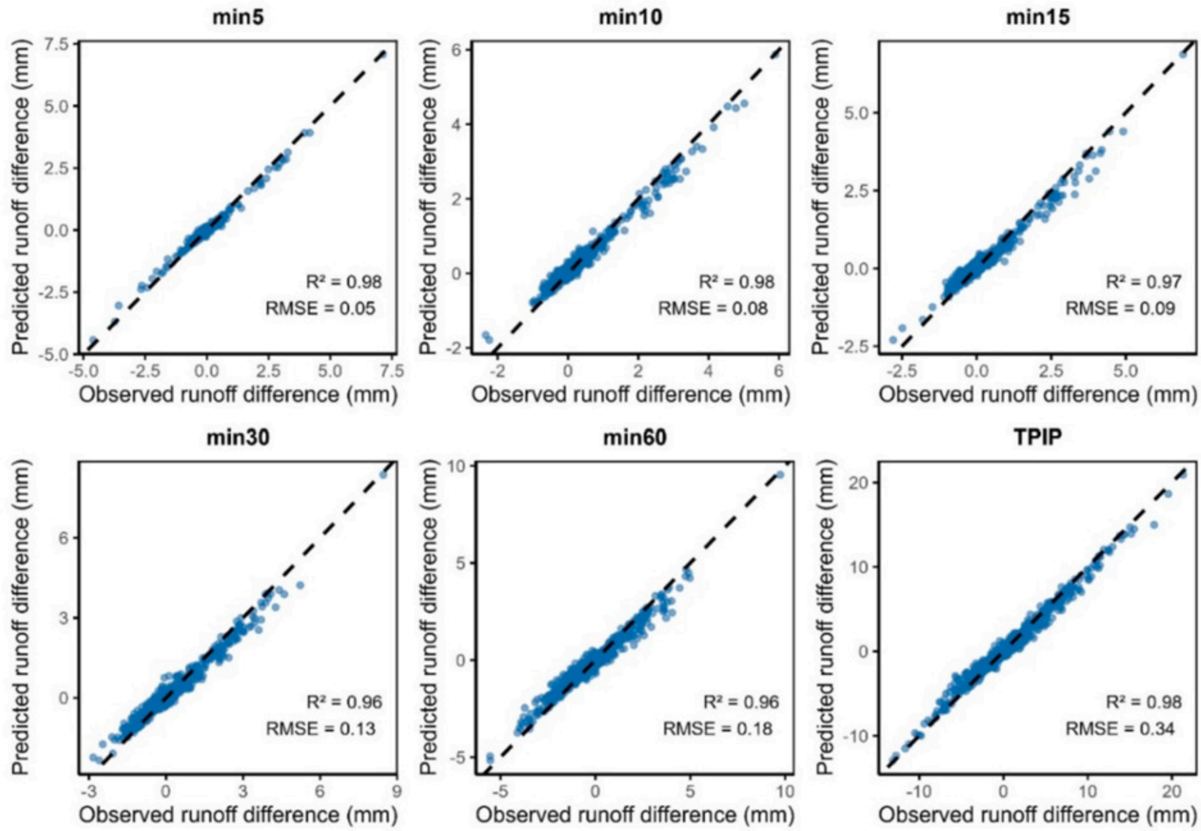

Fig. 6. Scatterplot comparing observed and predicted daily differences in runoff between stated climate format and breakpoint format from the XGBoost model relative to the breakpoint data for bare road surface. (Note: the dashed line is the 1:1 line).

beyond approximately 45 mm h^{-1}. For the fixed-interval inputs, runoff differences were only weakly affected by other precipitation attributes such as the peak intensity (ip), the timing of peaks (tp), or storm duration.

In contrast, the TPIP input exhibited distinct behavior relative to the fixed-interval data. Substantial runoff differences (up to ~ 6 mm) occurred for total precipitation between 140 and 200 mm, and stronger fluctuations were observed for 50–100 mm precipitation compared with other inputs. For average intensity, TPIP produced negative runoff differences of approximately 1 mm for intensities between 10–15 mm/h and around 0.7 mm for intensities above 25 mm/h. The TPIP data were also more sensitive to the temporal structure of rainfall, particularly to ip and tp. For events with ip greater than about 50, runoff differences approached −2 mm, suggesting that TPIP inputs are less representative of highly irregular rainfall events. Large differences were also associated with both extremely small and large tp values, indicating difficulties in reproducing events where rainfall peaks occur near the beginning or end of the storm. The effect of duration was comparatively minor, with a maximum runoff difference of about 0.9 mm for events lasting 15–20 h.

The PDPs of runoff differences for the bare surface exhibited patterns similar to those of the gravel surface when the same hydraulic conductivity (K_e, 3.8 mm h^{-1}) values were used in simulations. Fig. 8 presents the PDPs for the compacted surface with the lower hydraulic conductivity (1.3 mm h^{-1}). Overall, the compacted surface showed comparable trends to the bare surface for variables such as average intensity, ip, tp, and duration, though the magnitudes differed. The runoff differences among the climate input types were generally smaller for the compacted surface than for the bare surface.

For the fixed-interval climate inputs, the trends were largely parallel across all considered variables, with smaller runoff differences for finer temporal resolutions (5- and 10-min inputs). The TPIP input remained distinct from the others. When total precipitation was considered, TPIP produced mostly positive runoff differences for the bare surface but more negative differences for the compacted surface.

When examining average intensity, ip, tp, and duration, similar TPIP-related patterns emerged across both surfaces. The largest negative runoff differences occurred at rainfall intensities between 15–20 mm/h and stabilized beyond 20 mm/h. Similarly, ip values greater than 100 corresponded to the largest negative runoff differences. Runoff differences tended to be strongly positive or negative when tp approached 1 or 0, respectively, and the greatest fluctuations were observed for durations between 15–20 h.

3.3. Average treatment effects of climate inputs on predicted runoff

Fig. 9 illustrates the average treatment effects (ATEs) of different climate inputs on predicted daily runoff across precipitation bins for varying soil hydraulic conductivity values (K_e). Positive ATE values indicate that a climate input increases runoff relative to the breakpoint scenario, whereas negative values indicate a reduction. For example, an ATE of 1 mm means that, on average, changing from the breakpoint to a given climate input increases predicted daily runoff by 1 mm, after accounting for precipitation characteristics and K_e. The results demonstrate how the influence of climate input on predicted runoff varies with precipitation amount and K_e, highlighting the sensitivity of runoff predictions to climate input choices under different environmental conditions.

For all considered climate inputs and soil hydraulic conductivity (K_e), the ATE values ranged from −1 mm to 2.3 mm, indicating that the WEPP model remained stable across different climate input formats. For events with total precipitation below 20 mm, the ATEs were near zero, suggesting negligible sensitivity to input type. The 5-min climate input

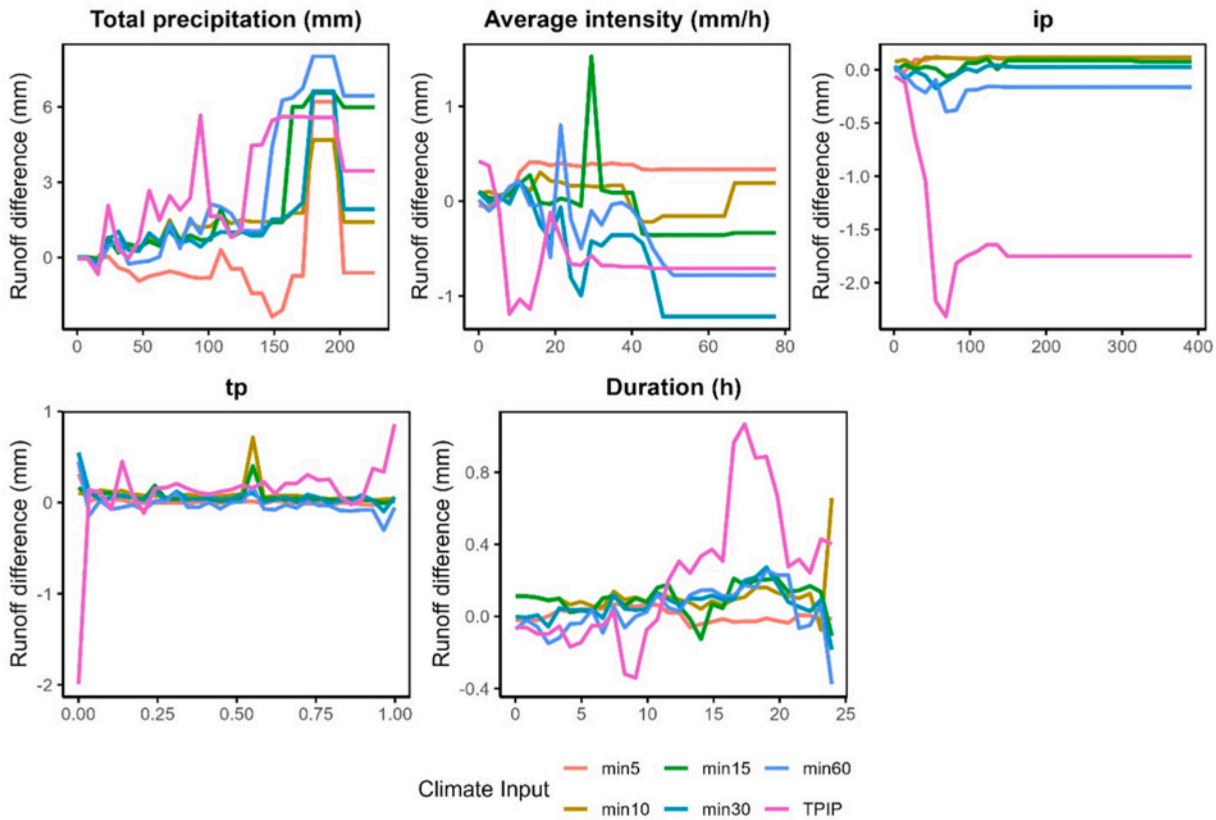

Fig. 7. Partial dependence plots (PDPs) illustrate the marginal effects of key precipitation characteristics on the predicted daily differences in runoff between the stated climate format and the breakpoint format for bare road surface.

consistently produced ATEs close to zero across all precipitation amounts and K_e values. For larger K_e values, representing more permeable roads with greater infiltration capacity, ATEs increased for the 30–40 mm and > 50 mm precipitation bins, reaching approximately 0.5–1.0 mm for most fixed-interval inputs except the 5-min data. The TPIP inputs exhibited greater variability than the fixed-interval inputs, with larger absolute ATE values for higher-intensity events. For rainfall between 20 and 30 mm, ATEs were positive across all K_e values. In contrast, for events exceeding 30 mm, ATEs were negative under lower K_e ($\leq$1.8 mm/h) and positive under higher K_e ($\geq$2.3 mm/h), consistent with Fig. 5, which showed decreased predicted runoff for TPIP on compacted surfaces but increased runoff on bare surfaces relative to the breakpoint input. These results suggest that when using TPIP data in WEPP simulations, adjustments to runoff predictions may be warranted for low ($\approx$1.3 mm/h) or high ($\geq$2.8 mm/h) K_e conditions.

3.4. Direct and indirect effects of climate inputs on erosion via runoff

The total effect of climate input on predicted soil erosion can be partitioned into a natural direct effect (NDE) and a natural indirect effect (NIE) via runoff. The NDE represents the effect of climate input on erosion independent of changes in runoff, whereas the NIE captures the effect mediated through runoff. For example, an NIE of 0.01 kg/m^2 indicates that changes in runoff due to the climate input increase daily erosion by 0.01 kg/m^2 on average. Similarly, an NDE of 0.01 kg m^{-2} indicates that the direct influence of the climate input increases erosion by the same amount, independent of runoff changes. This decomposition clarifies the pathways through which climate inputs affect erosion.

Fig. 10 shows the estimated NDE and NIE of climate input type on predicted daily soil erosion across precipitation bins and road surface designs. Points represent the effect estimates, and vertical bars indicate 95% confidence intervals (CIs), with values above or below zero (dashed line) corresponding to increases or decreases in erosion relative to the reference breakpoint input. For most climate inputs, the CIs overlapped the zero-effect line, indicating that differences in climate input generally did not lead to significant changes in predicted soil erosion.

Effect responses were generally smaller for the gravel road surface, reflecting the limited erosion potential on this treatment. Under light rainfall conditions (<20 mm), predicted erosion was nearly zero across all climate inputs and surfaces. In contrast, larger rainfall events and coarser temporal resolution inputs tended to produce more pronounced effects. The patterns of NDE and NIE were broadly similar across fixed climate inputs, though absolute effect sizes increased with coarser temporal resolution.

NIE results indicate that differences in erosion were primarily mediated through runoff variability, which could be mitigated with improved runoff predictions. Significant NIEs were observed for TPIP inputs under 30–40 mm and > 50 mm rainfall on bare surfaces, and > 50 mm rainfall on compacted and gravel surfaces. Significant NDEs were detected for 60-min inputs on compacted surfaces for events > 30 mm, and for inputs coarser or equal to 30-min on gravel surfaces for events > 20 mm. TPIP inputs also yielded significant NDEs for 20–40 mm rainfall on bare surfaces, >40 mm on compacted surfaces, and 20–40 mm on gravel surfaces.

Fig. 11 illustrates the natural direct effects (NDE) of climate inputs on predicted daily soil erosion across precipitation bins and varying soil erodibility parameters. Among the three considered parameters, K_i had the least influence, with NDE showing little change across its range for all climate inputs. In contrast, τ_c exhibited the largest impact, with NDE varying nonlinearly as τ_c increased. Larger absolute NDEs were generally associated with higher precipitation events and coarser temporal resolution for fixed-interval climate inputs.

For τ_c, clear negative NDEs were observed in the range 0.5–5 N/m^2 for nearly all fixed-interval climate inputs, except for the 5-min interval.

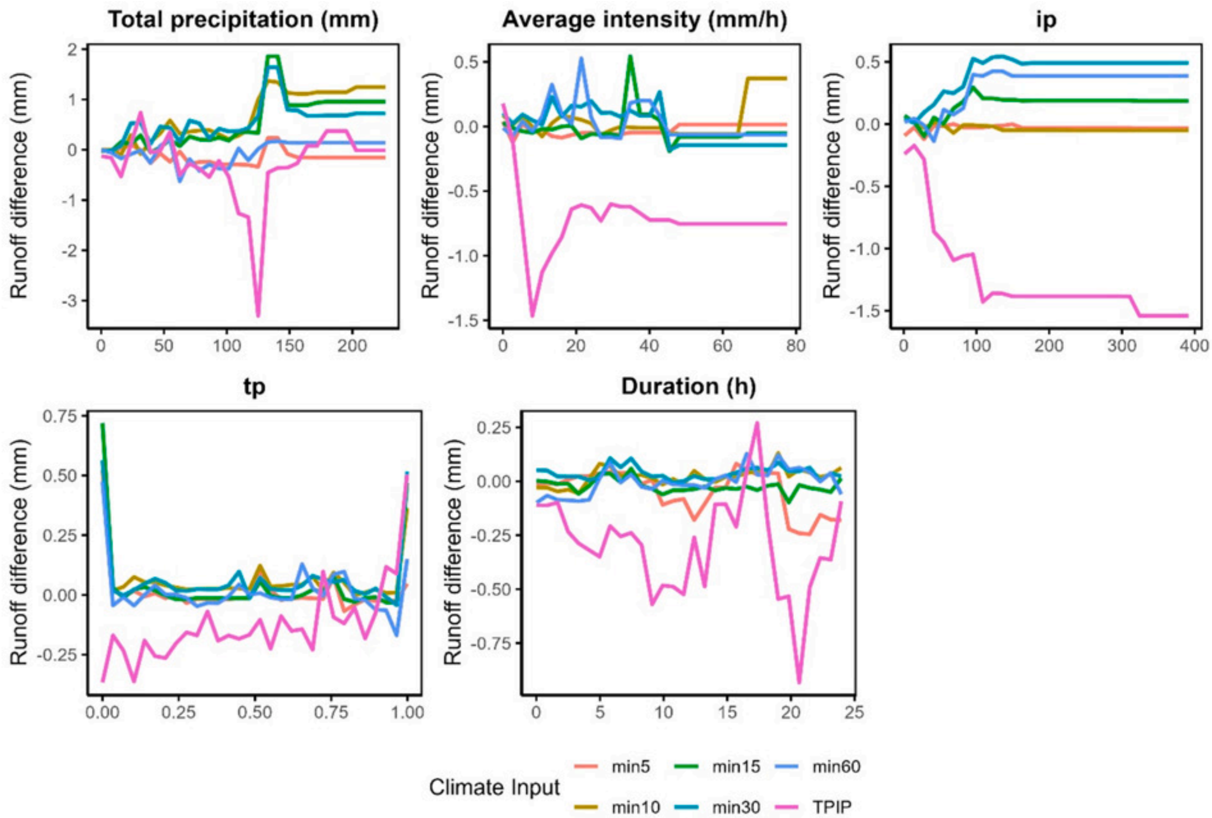

Fig. 8. Partial dependence plots (PDPs) illustrate the marginal effects of key precipitation characteristics on the predicted daily runoff differences relative to the breakpoint data for compacted road surface.

For the TPIP input, the influential τ_c range was similar (0.5–5 N/m^2), but negative NDEs occurred for events below 40 mm, while positive NDEs appeared for events exceeding 40 mm. For K_r, NDE exhibited an approximately linear relationship only for large precipitation events (>50 mm for fixed-interval inputs and > 30 mm for TPIP), while it had little influence for smaller events. Across most scenarios, TPIP produced distinct or larger NDEs relative to fixed-interval inputs.

4. Discussions

4.1. *Interpretation of machine learning outputs under predictor dependence*

The machine learning analysis provides a quantitative framework for examining how rainfall characteristics shape runoff differences under alternative climate input formats. XGBoost was chosen for this study because of its ability to model complex nonlinear relationships and its robustness to multicollinearity, which is common among rainfall characteristics (e.g., precipitation amount, duration, and intensity).

In our analysis, concerns that strong multicollinearity might distort partial dependence plots (PDPs) are mitigated for several reasons. First, we used 25 years of data (2,876 rainfall observations), offering a large and varied representation of storm conditions, so most combinations of storm features occur in the training set. Second, because XGBoost is a tree-based ensemble method, it is relatively insensitive to correlated predictors: decision trees naturally select among correlated features during splits, reducing redundancy. As some have noted, tree-based models like boosted decision trees are much less affected by multicollinearity than regression-based models (Romero Martínez et al., 2025; Wang et al., 2025). Third, the pairwise correlations among our rainfall variables are moderate rather than severe; the largest observed Pearson correlation in Table 5 is 0.51 (between ip and duration), while correlations between duration and total precipitation, and between precipitation and average intensity, are 0.38 and 0.35, respectively. Given the dataset size, moderate predictor dependence, and model characteristics, the PDPs are considered reliable representations of how storm characteristics influence WEPP runoff predictions under varying climate input resolutions.

The machine learning results indicate that daily runoff differences among climate input formats were generally small, implying that the WEPP model produces numerically stable runoff predictions across a broad range of rainfall representations. The magnitude and direction of runoff differences varied systematically with road surface type, reflecting contrasts in hydraulic conductivity (K_e) and associated infiltration capacities. Surfaces with different hydraulic properties responded differently to identical precipitation characteristics, indicating that the effects of climate input resolution cannot be interpreted independently of surface conditions.

Collectively, these findings indicate that interpretable machine learning approaches have the potential to serve as diagnostic tools for evaluating process-based hydrologic models by highlighting dominant precipitation controls and surface-dependent sensitivities.

4.2. *Influence of storm intensity structure and climate input format on runoff*

Fig. 12 illustrates hyetographs for two representative storms (September 16, 2020, and April 24, 2021), characterized by high precipitation, long duration, and peak intensities (ip) of 11.87 and 17.62, respectively. Event (a) exhibits a multimodal intensity pattern, whereas event (b) is predominantly unimodal. Under TPIP input, WEPP applies a double-exponential function to approximate rainfall intensity, resulting in a smoother single peak.

Although the total precipitation is conserved across the three input

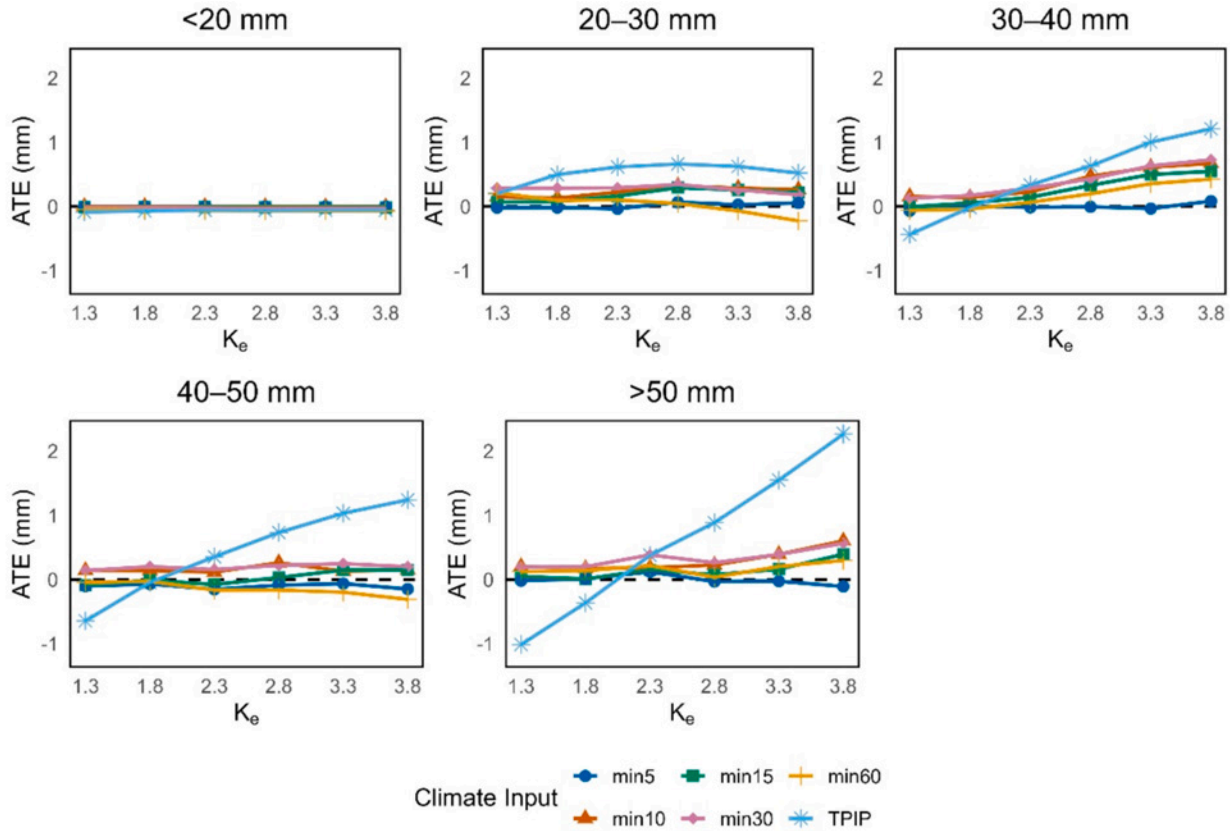

Fig. 9. Average treatment effect (ATE) of different climate scenarios on predicted daily runoff relative to breakpoint predictions across precipitation bins, shown for various values of the soil hydraulic conductivity (K_e). (Note: the dashed line is for ATE = 0 mm).

formats (breakpoint, 30 min, and TPIP), intensity redistribution alters runoff generation. When a decaying line representing the declining Green-Ampt hydraulic conductivity (like starting at 3.8 mm h^{-1}) is applied, the hyetograph area above the line corresponds to surface runoff. Breakpoint inputs capture short-duration, high-intensity bursts, while 30 min inputs preserve extended peaks but smooth rapid fluctuations. TPIP typically smooths or exaggerates maximum intensities.

Differences among temporal resolutions arise because infiltration responds rapidly to short-duration, high-intensity rainfall. Finer temporal resolution better preserves the timing and magnitude of peak intensities, allowing the Green–Ampt routine in WEPP to more accurately represent critical infiltration periods. In contrast, coarser inputs smooth short, intense bursts and may slightly over- or under-predict runoff during these intervals. This explains why finer fixed-interval inputs (5–10 min) generally produce smaller deviations from breakpoint simulations.

The contrasting ATE patterns observed in Section 3.3 can be interpreted in the context of the Green–Ampt infiltration routine in WEPP, which simulates declining infiltration capacity as soil water content increases with cumulative infiltration. Under this framework, large and long-duration storms with late-occurring intensity peaks may generate proportionally greater runoff, particularly on higher-conductivity soils where infiltration capacity decreases during the event. For both storms, rapid decay of the TPIP curve causes higher runoff contributions at higher hydraulic conductivity values, consistent with results in Fig. 9. The response of 30 min inputs varies depending on the number of short-term peaks: in event (a), runoff exceeded that of breakpoint, whereas for event (b), it was lower when $K_e = 1.3$ mm h^{-1}.

Table 6 summarizes storm characteristics by precipitation bin. Large storms in the study area, with high peak intensities and extended durations, are especially sensitive to the representation of rainfall in WEPP simulations. Climate input formats such as TPIP or coarser interval data can redistribute rainfall intensity, modifying both the timing and magnitude of peak runoff relative to breakpoint inputs. Accurate representation of short-duration, high-intensity bursts is therefore important for simulating runoff during large events.

For large rainfall events ($\geq$40 mm), TPIP-derived runoff exhibited a near-linear relationship with soil hydraulic conductivity (K_e) in Fig. 9. This indicates that TPIP inputs can systematically over- or under-predict runoff depending on surface infiltration properties. Modelers may consider applying a K_e-dependent correction or conducting sensitivity analyses across realistic K_e values to improve runoff estimates for extreme events.

4.3. Erosion prediction in WEPP

Erosion predictions in WEPP were jointly influenced by hydrologic representation and soil detachment resistance. Mediation analysis showed that improving runoff accuracy substantially reduces indirect effects on erosion, whereas τ_c, followed by K_r and K_i, primarily drives direct effects. The nonlinear response of τ_c, with reduced sensitivity above ~ 5 N/m^2, suggests that erosion predictions are most responsive within a limited parameter range, and that targeted adjustment of τ_c and K_r can improve model performance under varying rainfall inputs.

Surface condition strongly influenced mediation behavior. For the bare surface, most normalized direct and indirect effects (NDE and NIE) were nonsignificant but exhibited the widest confidence intervals, reflecting greater variability associated with higher erodibility. In contrast, compacted and gravel surfaces showed more consistent parameter sensitivities and narrower uncertainty bounds, indicating that increased surface stability enhances the detectability of parameter effects. Significant NIE values were observed for nearly all TPIP inputs during events exceeding 30 mm, suggesting that TPIP data have difficulty reproducing the temporal structure of larger storms. Errors in

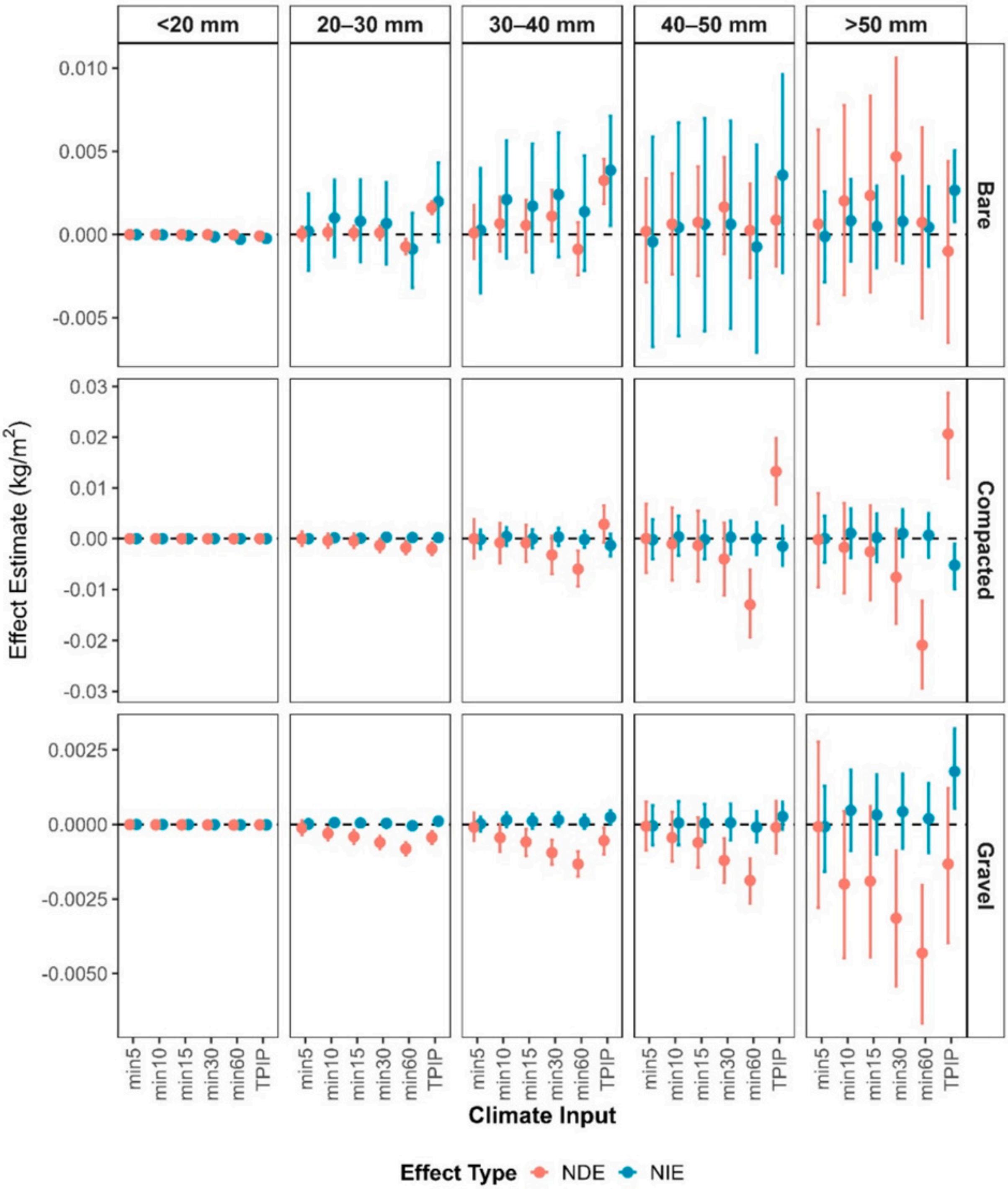

Fig. 10. Estimated natural direct effects (NDE) and natural indirect effects (NIE) of climate input type on predicted daily soil erosion (kg/m^2) across precipitation bins and surface designs. (Notes: Points indicate estimated effects, with vertical bars showing 95% confidence intervals. The dashed horizontal line at zero denotes no effect.).

rainfall sequencing propagate through runoff simulation and amplify indirect effects on erosion.

For storms larger than 30 mm, significant NDE patterns varied by road surfaces. The compacted surface showed strong direct effects under 60-min and TPIP inputs, whereas the gravel surface exhibited significant responses under 30-min, 60-min, and TPIP inputs. The influence of τ_c was greatest within 0.5–5 N/m^2, peaking near $\sim$ 3 N/m^2, which corresponds to the τ_c range of the compacted surface. For higher τ_c values (5–10 N/m^2), characteristic of the gravel surface, pronounced effects were mainly associated with events exceeding 50 mm.

These findings are consistent with previous WEPP sensitivity studies. Nearing et al. (1990) identified precipitation and rill erodibility as dominant factors influencing model response, with saturated hydraulic conductivity and interrill erodibility showing moderate sensitivity. Likewise, Elliot et al. (2023) confirmed hydraulic conductivity as the most sensitive calibration parameter for road erosion modeling, followed by rill erodibility and critical shear. Under high-intensity or extended-duration storms, erosion becomes more sensitive to detachment-resistance parameters and potential armoring of the surfaces as finer materials are detached, leaving coarser particles behind

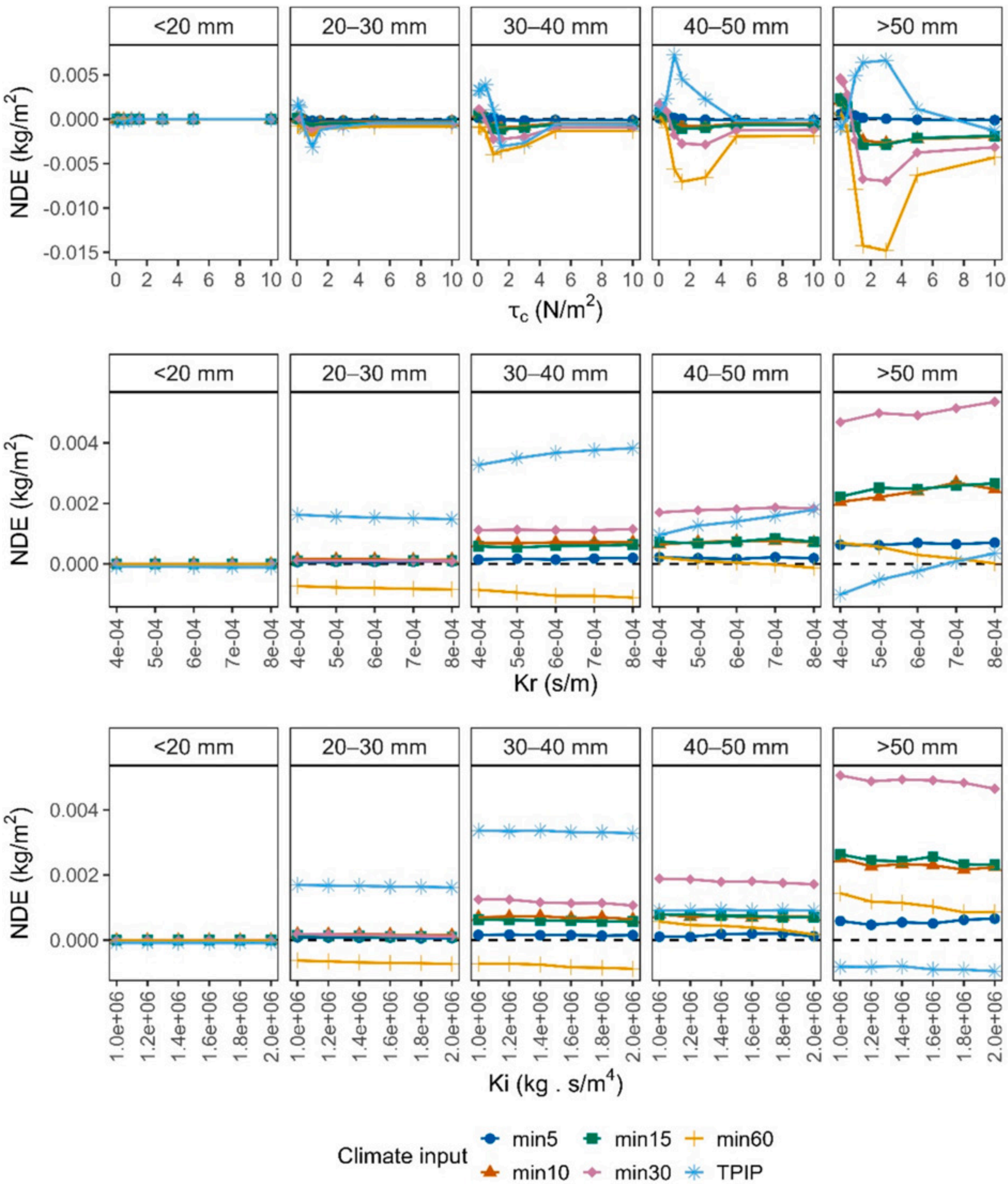

Fig. 11. Estimated natural direct effects (NDE) of climate input on predicted daily soil erosion (kg/m^2) across precipitation bins and soil erodibility parameters. (Notes: The dashed horizontal line at zero denotes no effect).

Table 5
Pairwise correlations among rainfall characteristics.

	Amount	Duration	tp	ip	Intensity
Amount	1.00	0.38	0.02	0.11	0.35
Duration		1.00	0.11	0.51	−0.23
tp			1.00	0.07	−0.08
ip				1.00	−0.14
Intensity					1.00

(Foltz et al., 2008). Accurate simulation of peak runoff is essential to support sediment detachment and transport estimates, after which strategic refinement of τ_c and K_r improves erosion prediction.

4.4. Comparison with previous studies and broader implications

Previous studies suggest that rainfall temporal resolution and structure can influence WEPP predictions, but the reported effects on runoff and erosion vary across locations, input types, and event magnitudes. Flanagan et al. (2020) found that simulations driven by breakpoint rainfall tended to better represent erosivity than coarser aggregated inputs, particularly for tilled cropland in Boston; however,

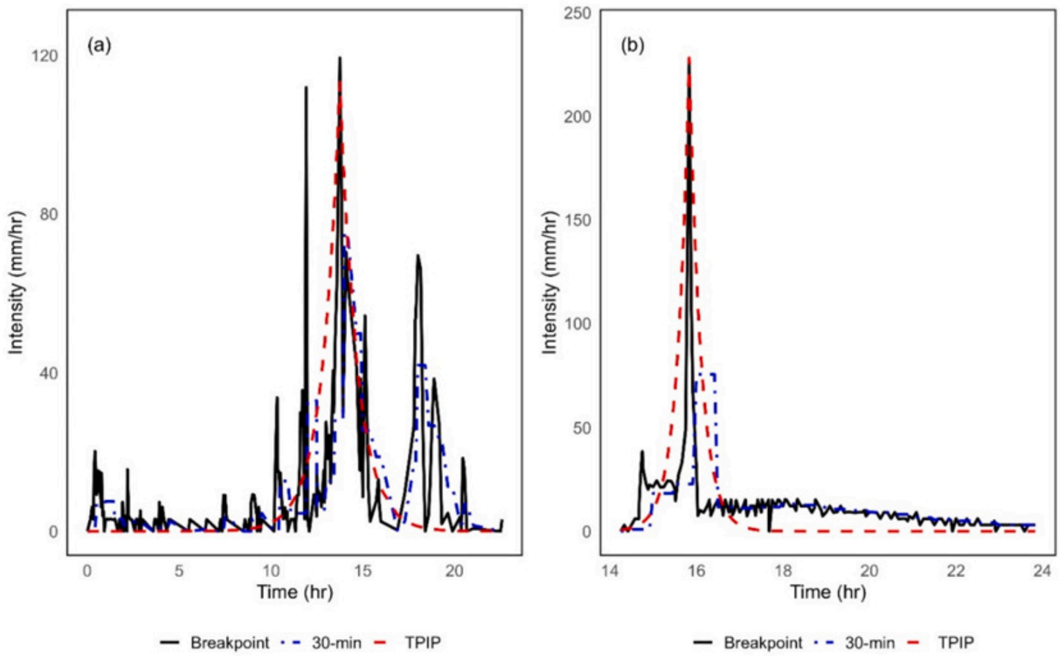

Fig. 12. Hyetographs for two representative high-intensity storms on September 16, 2020 (a) and April 24, 2021 (b).

Table 6
Storm summary statistics by precipitation bin.

Precipitation bins (mm)	Number of storms	Precipitation		Duration		tp		ip	
		Mean (mm)	SD (mm)	Mean (h)	SD (h)	Mean	SD	Mean	SD
< 20	2380	4.71	5.09	5.23	6.34	0.46	0.30	10.75	21.91
20–30	214	24.36	2.96	9.65	6.93	0.48	0.32	14.64	15.45
30–40	116	34.61	2.69	10.71	6.72	0.44	0.32	14.74	14.88
40–50	61	44.56	2.64	12.69	7.14	0.5	0.31	16.89	14.14
>50	105	75.72	31.41	13.27	6.87	0.49	0.31	15.4	16.25

their study included few extreme events, limiting relevance for high-intensity storms or forest roads in subtropical regions. Srivastava et al. (2019) examined differences between the original and updated CLIGEN databases (both TPIP-type inputs) and observed regional variations in runoff and erosion, but did not consider event-scale mechanisms, rainfall structure, or surface-specific responses. Fullhart et al. (2020) developed statistical and machine learning methods to downscale maximum 30-min intensities from coarser inputs, showing that fine-resolution corrections approximated breakpoint-derived intensities, whereas coarse inputs required more complex, climate-dependent adjustments; these methods were not directly evaluated in process-based runoff or erosion simulations.

The present study extends these efforts by examining event-scale runoff and erosion responses for forest roads in a humid subtropical setting, comparing multiple climate input formats (breakpoint, fixed-interval, and TPIP). Results suggest that runoff differences among inputs are generally modest for small to moderate storms but increase for large, high-intensity, or temporally uneven events, particularly on surfaces with contrasting infiltration properties. This partially resolves previous ambiguity regarding the influence of climate input resolution on erosion simulations: while input resolution may have limited effects under typical conditions, it could be more relevant for extreme events or sensitive surface types.

Overall, the findings suggest that, for forest roads in Florida, coarser or TPIP-type inputs may be sufficient for routine runoff estimation, whereas finer temporal resolution may improve representation of extreme or highly peaked storm events, providing a more nuanced understanding of when climate input resolution matters in hydrologic and erosion modeling. Although this study focuses on Florida, the key processes controlling runoff and erosion—such as high-intensity, short-duration storms and surface-dependent infiltration—are characteristic of humid subtropical regions worldwide. Thus, these findings can inform hydrologic and erosion modeling in other areas with similar climates, particularly regarding the need for finer temporal resolution during extreme storm events.

4.5. Limitations and future research

This study has several limitations that should be considered when interpreting the results. The analysis focused on a single forest road site in the Florida Panhandle, and the findings may not fully represent the range of soil, surface, and climatic conditions encountered in other regions. In addition, climate forcing was derived from a single meteorological station, which limits representation of spatial rainfall variability that can influence runoff and erosion responses.

The machine learning analysis was based on a finite sample of observed storm events, including a limited number of extreme cases, introducing uncertainty in inference under rare, high-magnitude conditions. Future research should extend this framework to diverse geographic settings, incorporate spatially distributed rainfall observations, and expand representation of extreme events.

5. Conclusions

This study evaluated the effects of rainfall temporal resolution and climate input formats on WEPP-simulated runoff and soil erosion for a forest road in Florida's Chipola Experimental Forest. Multiple climate inputs, including breakpoint, fixed-interval (5–60 min), and TPIP, were generated from NOAA-ASOS data (2000–2024) to drive WEPP hillslope simulations.

WEPP predictions were generally stable across input types, with modest differences in average annual runoff and soil loss. Bare surfaces showed the largest variability and uncertainty in runoff and erosion responses, compacted surfaces exhibited moderate sensitivity, and gravel surfaces were comparatively robust to changes in rainfall representation. Finer temporal resolutions (5–10 min) and breakpoint inputs produced slightly smaller deviations from the breakpoint file, while TPIP exhibited greater sensitivity to storm irregularity, particularly for high-intensity, highly peaked, or long-duration events. Machine learning analysis indicated that total precipitation and average intensity were the dominant factors controlling runoff differences, with storm structure variables (ip, tp, and duration) having smaller effects. Surface properties modulated these patterns.

Across all climate inputs and soil erodibility values, ATEs remained near zero for small storms (<20 mm) and increased modestly for larger events. TPIP inputs showed greater variability, suggesting that runoff adjustments may be warranted under low or high infiltration conditions. Mediation analysis indicated that natural indirect effects (NIEs) could be minimized by improving runoff predictions, while direct effects (NDE) were influenced primarily by τ_c, with minor contributions from K_r and K_i. Erosion responses diverged by surface type during storms > 30 mm: compacted and gravel roads showed significant τ_c-driven direct effects, whereas bare surfaces exhibited greater uncertainty associated with high erodibility. Maximum erosion sensitivity occurred within $\tau_c \approx$ 0.5–5 N/m^2 (peaking near 3 N/m^2), characteristic of compacted roads.

These results indicate that finer temporal resolutions (5–10 min) are recommended for accurately simulating both runoff and erosion, particularly for bare and compacted surfaces during high-intensity or short-duration rainfall events, while coarser resolutions (~30 min) are generally sufficient for gravel surfaces or low-intensity storms. TPIP data may be suitable for sensitivity analyses but require caution under irregular or extreme rainfall conditions. For practical applications, breakpoint or ≤ 10-min inputs should be prioritized for erosion assessment on bare and compacted roads subject to intense storms, whereas 30-min inputs are acceptable for routine prediction on gravel roads or in moderate rainfall regimes. TPIP formats are best reserved for exploratory or data-limited scenarios and should be supplemented with calibration when simulating large or complex storm events.

Overall, the findings underscore the relevance of climate input resolution and structure in hydrologic and erosion modeling. By integrating machine learning, average treatment effects, and mediation analysis, this study provides a practical framework for diagnosing how rainfall representation propagates through hydrologic and erosion processes and for selecting climate inputs that improve predictive reliability in road erosion modeling.

CRediT authorship contribution statement

Jingqiu Chen: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Shuyuan Wang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **William J. Elliot:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition. **Bernard A. Engel:** Writing – review & editing, Validation, Supervision, Resources, Funding acquisition. **Feng Pan:** Writing – review & editing, Validation, Resources, Methodology, Investigation. **Yaoze Liu:** Writing – review & editing, Validation, Resources, Investigation. **Johnny M. Grace:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by funding provided by the USDA Forest Service Southern Research Station under the Cooperative Agreement (Number 23-CA-11330180-066) with Florida A&M University.

Data availability

Data will be made available on request.

References

Cao, L., Elliot, W.J., Long, J.W., 2021. Spatial simulation of forest road effects on hydrology and soil erosion after a wildfire. Hydrol. Process. 35, e14139.

Chen, J., Wang, S., Elliot, W.J., Engel, B.A., Dobre, M., Srivastava, A., Pan, F., Liu, Y., Grace III, J.M., 2025. Modeling forest road runoff and erosion at multiple temporal scales in Florida Panhandle region. International Soil and Water Conservation Research, USA, 10.1016/j.iswcr.2025.09.007.

Chen, J., Wang, S., Elliot, W.J., Chen, Z., Pan, F., Liu, Y., Grace III, J.M., 2026. Estimating runoff and erosion on forest roads under various climate projections in Florida Panhandle region, USA. J. Environ. Manage. 397, 128207. https://doi.org/10.1016/j.jenvman.2025.128207.

Dun, S., Wu, J.Q., Elliot, W.J., Frankenberger, J.R., Flanagan, D.C., McCool, D.K., 2013. Applying online WEPP to assess forest watershed hydrology. Trans. ASABE 56 (2), 581–590.

Dun, S., Wu, J.Q., Elliot, W.J., Robichaud, P.R., Flanagan, D.C., Frankenberger, J.R., Brown, R.E., Cu, A.C., 2009. Adapting the Water erosion Prediction Project (WEPP) model for forest applications. J. Hydrol. 366, 46–54.

Eisenbies, M.H., Aust, W.M., Burger, J.A., Adams, M.B., 2007. Forest operations, extreme flooding events, and considerations for hydrologic modeling in the Appalachians – a review. For. Ecol. Manage. 242, 77–98.

Elliot, W.J., 2004. WEPP Internet interfaces for forest erosion prediction. J. Am. Water Resour. Assoc. 40, 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.

Elliot, W. J., Hall, D. E., 2010. Disturbed WEPP Model 2.0. Ver. 2014.04.14. Moscow, ID: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. Available at https://forest.moscowfsl.wsu.edu/fswepp/.

Elliot, W. J., Hall, D. E., Scheele, D. L., 1999. WEPP: Road (Draft 12/1999) WEPP interface for predicting forest road runoff, erosion and sediment delivery, technical documentation. https://forest.moscowfsl.wsu.edu/fswepp/docs/wepproaddoc.html.

Elliot, W.J., Lewis, S.A., Cannard, C.L., 2023. Measuring and modeling impacts of gravel road design on sediment generation in the southeastern U.S. Journal of the ASABE 66 (5), 1229–1254.

Farias, T.R.L., Medeiros, P.H.A., Araújo, J.C.D., Navarro-Hevia, J., 2021. The role of unpaved roads in the sediment budget of a semiarid mesoscale catchment. Land Degrad. Dev. 32, 5443–5454.

Flanagan, D.C., Ascough II, J.C., Nearing, M.A., Laflen, J.M., 2001. Chapter 7. the Water erosion Prediction Project (WEPP) model. In: Harmon, R.S., Doe, W.W. (Eds.), *Landscape Erosion and Evolution Modelling*. Kluwer Academic / Plenum Publishers, New York, pp. 145–199.

Flanagan, D.C., Livingston, S.J., 1995. WEPP User Summary, NSERL Report No. 11. USDA-ARS, National Soil Erosion Research Laboratory, West Lafayette, IN.

Flanagan, D.C., McGehee, R.P., Srivastava, A., 2020. Evaluation of different climate inputs to WEPP. ASABE 2020 Annual International Meeting 2000740.

Flanagan, D.C., Nearing, M.A. (Eds.), 1995. *USDA-Water Erosion Prediction Project (WEPP) Hillslope Profile and Watershed Model Documentation*. NSERL Report No. 10. USDA-ARS National Soil Erosion Research Laboratory, West Lafayette, Indiana.

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.

Fullhart, A.T., Nearing, M.A., McGehee, R.P., Weltz, M.A., 2020. Temporally downscaling a precipitation intensity factor for soil erosion modeling using the NOAA-ASOS weather station network. Catena 194, 104709.

Grace III, J.M., 2017. Predicting forest road surface erosion and storm runoff from high-elevation sites. Trans. ASABE 60 (3), 705–719.

Hollinger, S. E., Angel J. R., Palecki, M. A., 2002. Spatial distribution, variation, and trends in storm precipitation characteristics associated with soil erosion in the

United States. Contract report CR 2002-08. Illinois State Water Survey. Champaign, IL: Illinois Department of Natural Resources.

McGehee, R.P., Flanagan, D.C., Srivastava, P., 2020. WEPPCLIFF: a command-line tool to process climate inputs for soil loss models. Journal of Open Source Software 5, 1–7. https://doi.org/10.21105/joss.02029.

Nearing, M.A., Deer-Ascough, L., Laflen, J.M., 1990. Sensitivity analysis of the WEPP hillslope profile erosion model. Trans. ASABE 33 (3), 0839–0849.

Nicks, A. D., Lane, L. J., Gander, G. A., 1995. Chapter 2. Weather Generator. 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, Indiana: USDA-ARS National Soil Erosion Research Laboratory.

Romero Martínez, M., Carmona Ibáñez, P., Martínez Vargas, J., 2025. Predicting business failure with the XGBoost algorithm: the role of environmental risk. Sustainability 17 (11), 4948.

Srivastava, A., Flanagan, D.C., Frankenberger, J.R., Engel, B.A., 2019. Updated climate database and impacts on WEPP model predictions. J. Soil Water Conserv. 74 (4), 334–349.

Wang, R., Wang, X., Zhang, Z., Zhang, S., Li, K., 2025. Research on the nonlinear and interactive effects of multidimensional influencing factors on urban innovation cooperation: a method based on an explainable machine learning model. Systems 13 (3), 187.

Wang, S., Flanagan, D.C., Engel, B.A., Zhou, N., 2020. Impacts of subsurface hydrologic conditions on rill sediment transport capacity. J. Hydrol. 591, 125582.

Wang, S., McGehee, R.P., Guo, T., Flanagan, D.C., 2023. Calibration, validation, and evaluation of the Water erosion Prediction Project (WEPP) model for hillslopes with natural runoff plot data. International Soil and Water Conservation Research 11, 669–687.

Yousefi, S., Moradi, H., Boll, J., Schonbrodt-Stitt, S., 2016. Effects of road construction on soil degradation and nutrient transport in Caspian Hyrcanian mixed forests. Geoderma 284, 103–112.

Yu, J., Zhao, Q., Yu, Z., Liu, Y., Ding, S., 2024. A review of the sediment production and transport processes of forest road erosion. Forests 15, 454.

Zemke, J.J., 2016. Runoff and soil erosion assessment on forest roads using a small scale rainfall simulator. Hydrology 3, 25.