Journal of Environmental Management 397 (2026) 128207

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Estimating runoff and erosion on forest roads under various climate projections in Florida Panhandle region, USA

Jingqiu Chen [a,1,*], Shuyuan Wang [a,b,1], William J. Elliot [c], Zhihan Chen [b], 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

ARTICLE INFO

Keywords:
Forest road
Storm runoff
Soil erosion
WEPP
Climate models
Florida Panhandle region

ABSTRACT

Forest roads directly influence storm runoff, soil erosion, and sediment transport, making soil and water conservation essential. This study estimated runoff and erosion using the Water Erosion Prediction Project (WEPP) model for a typical unimproved forest road in the Chipola Experimental Forest, considering two alternative designs: a compacted "graded aggregate base" and a gravel surface. Both historical and future performance were evaluated using five widely used climate models in their Coupled Model Intercomparison Project (CMIP) fifth phase (CMIP5) and sixth phase (CMIP6) versions under emission scenarios 4.5 and 8.5. Model reliability during the historical period (2015–2023) was assessed using the ADF test for stationarity and ARMA modeling of precipitation biases. Extreme events were extracted to evaluate model performance under high-impact conditions. For 2024–2099, uncertainty was quantified using bootstrapped confidence intervals for projected precipitation, and estimated runoff and erosion. Results indicate that all selected models were stationary and provided reasonable annual and monthly precipitation estimates. CMIP6 models outperformed CMIP5, exhibiting lower ARMA intercepts for precipitation differences and more consistent projections of precipitation, runoff, and erosion across all road designs. CMIP5 models showed a tendency toward more extreme events. CMIP6 bootstrapped confidence intervals were more precise while maintaining comparable accuracy. The fractions of acceptable monthly precipitation, runoff, and erosion estimates exceeded 79 %, 64 %, and 45 %, respectively, with better performance for compacted and gravel roads than unimproved roads. These findings establish a robust framework for estimating runoff and erosion using climate models, underscore the benefits of improved road designs, and elucidate the differences between CMIP5 and CMIP6 models. The results can inform the design and maintenance of forest road networks under changing climate conditions and support long-term infrastructure resilience planning.

1. Introduction

Forest roads enable access for timber, recreation, and conservation but disturb watersheds and contribute to runoff and erosion (Foltz et al., 2009). Sediment-laden runoff degrades water quality and harms ecosystems (Dissmeyer, 2000), while erosion undermines road stability and accelerates land degradation. As increasing pressures from changing climate conditions intensify rainfall and alter hydrology, understanding runoff and erosion is key to reducing environmental impacts and supporting sustainable forest management. Elliot et al. (2023) evaluated various road designs using the WEPP model (Flanagan and Livingston, 1995), which effectively simulates road erosion by accounting for detachment and delivery processes. Improving roads is essential for managing runoff and erosion. Compacted roads limit water pooling and channelization, while gravel surfaces reduce runoff velocity and shield soil from rainfall and traffic. These methods reduce sediment transport

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

https://doi.org/10.1016/j.jenvman.2025.128207
Received 6 May 2025; Received in revised form 13 November 2025; Accepted 1 December 2025

and enhance road resilience in high-rainfall areas (Chen et al., 2025). WEPP can be effectively parameterized for high-traffic gravel roads, incorporating weather, design, and topographic effects.

The impacts of changing climate on soil erosion have been observed since the 1940s (Bryan and Albritton, 1943) and include both direct and indirect impacts (Li and Fang, 2016). The direct impacts are driven by the changes in rainfall amount, rainfall intensity, and rainfall spatio-temporal distributional patterns (Bangash et al., 2013; Pruski and Nearing, 2002; Tang et al., 2015; Zhang, 2012; Maeda et al., 2010; Lenderink et al., 2025). The indirect impacts are related to the rising temperature, which exacerbates soil desiccation and vegetation stress, reducing the soil's capacity to resist erosion. By changing the vegetation cover and soil moisture, a warming climate influences soil erosion (Nearing et al., 2004). Moreover, the indirect impacts are brought by socio-economic factors, such as the changes in crop management, planting and harvesting dates, type of cultivars, and price of crops to accommodate the new climate regime (Cao et al., 2024; Paroissien et al., 2015; Mullan, 2013; Parajuli et al., 2016; Garbrecht and Zhang, 2015; Garbrecht et al., 2007). Therefore, accurate estimation of runoff and erosion under changing climate conditions is critical for developing adaptive management strategies. Researchers can simulate future scenarios, identify vulnerable areas, and design effective mitigation measures by integrating climate models with hydrological and erosion models (Ma et al., 2025; Liu et al., 2024; Takhellambam et al., 2023; Anache et al., 2018; Zhang and Nearing, 2005). Most of the previous future climate studies were focused on the runoff and erosion estimation on farmland, however, the study of forest roads under climate change scenarios is needed to inform sustainable forest road management practices.

Florida is among the most climate-vulnerable U.S. regions, facing more frequent and intense hurricanes (https://www.weather.gov/key/storms). Rising sea surface temperatures fuel stronger, wetter, and slower storms, causing extreme rainfall, storm surge, and flooding. Florida's low elevation and long coastline heighten its susceptibility to these climate-driven impacts. Intensifying storms increase forest road runoff and erosion, as heavy rainfall boosts runoff volume and speed, overwhelming drainage and damaging road surfaces and slopes. Hurricanes also strip vegetation, uproot trees, and destabilize soil, leaving forests vulnerable to further erosion. Hurricane Michael in 2018, for example, destroyed nearly 1.2 million hectares of forest in Florida's Panhandle. Estimating forest road runoff and erosion in Florida must account for these growing climate impacts.

Climate models, including global (GCMs) and regional (RCMs) models, are essential for predicting how future climate may impact runoff and erosion. While GCMs provide broad-scale projections of climate variables, RCMs operate at a much finer spatial resolution, typically ranging from 10 to 50 km. This higher resolution enables RCMs to capture local climate features more accurately, such as topographic influences on precipitation patterns and temperature variations across heterogeneous landscapes. Additionally, RCMs are better equipped to simulate extreme weather events, such as heavy rainfall or intense storms, which are critical drivers of runoff and erosion. For regional or local studies about runoff, erosion, and related parameters, RCMs are widely used (Zhu et al., 2024; Takhellambam et al., 2023; Gaur et al., 2023; Huang et al., 2013). The widely used models included CanESM, CNRM-CM, GFDL-ESM, HadGEM, and MPI-ESM (Takhellambam et al., 2023; Kumar et al., 2021; Wang et al., 2023; Rajbanshi and Bhattacharya, 2022).

The Coupled Model Intercomparison Project (CMIP) improves climate change understanding through model comparisons. CMIP6, the latest phase, enhances CMIP5 with more models, better atmospheric simulations, and improved feedback processes. Chen et al. (2021) found CMIP6 outperformed CMIP5 in simulating precipitation extremes. CMIP6 also incorporates new socioeconomic pathways (SSPs) for future climate scenarios (Eyring et al., 2016). Despite these advances, CMIP5 remains in use, making comparisons between the two critical (Takhellambam et al., 2023). Climate projections, whether derived from GCMs or RCMs, CMIP5 or CMIP6, inherently contain uncertainties due to the complexity of climate systems and assumptions used in model development. Ignoring these uncertainties can lead to overconfidence in predictions. By systematically analyzing and quantifying uncertainties, researchers can better understand the range of possible outcomes, identify key sources of variability, and improve model robustness.

The use of ARMA models for comparing climate model performances has been well-established in the literature (Dimri et al., 2020; Dastorani et al., 2016; Sarhadi et al., 2014). For example, Wilks (2006) discussed how ARMA models, combined with residual analysis, can help identify model biases and quantify the goodness of fit for various climate models. The intercept of ARMA models can provide a direct measure of model bias. In the context of climate models, this can help determine whether a model systematically overestimates or underestimates precipitation variability. Giorgi and Mearns (1999) demonstrated how statistical models like ARMA can capture biases in simulated precipitation and provide a basis for comparing the model's performance across different time scales. Mulla et al. (2024) demonstrated the effectiveness of ARMA-type modes for climate-related time series forecasting and a SARIMAX model was successfully applied to predict monthly rainfall in coastal Phaltan.

The bootstrap method is a powerful statistical tool for analyzing uncertainties in climate change data for runoff and erosion estimation, which does not rely on strict assumptions about the underlying data distribution, making it well-suited for complex climate data and models. By incorporating bootstrap analyses, researchers can enhance the reliability of their results, providing more comprehensive insights into the uncertainties associated with climate change impacts on hydrological and erosion processes (Kim et al., 2014; Zhang et al., 2018; Kropp and Schellnhuber, 2010; Wu et al., 2024).

Although climate change impacts on runoff and erosion have been widely studied, limited attention has been given to forest road networks, particularly in regions like the Florida Panhandle where high rainfall intensity and sandy soils heighten erosion risk. Existing studies often use a narrow set of climate models without rigorously validating historical performance or comparing CMIP5 and CMIP6 frameworks across emission scenarios. Additionally, few studies focus on model behavior during extreme precipitation events or incorporate uncertainty in a systematic way. This study addresses these gaps by evaluating multiple climate models using time-series and resampling techniques to assess historical performance, analyze extremes, and project future runoff and erosion under different road design scenarios. The findings provide site-specific insights and support climate-resilient planning for forest infrastructure.

Focusing on a representative forest road within the Chipola Experimental Forest in the Florida Panhandle near Clarksville, FL, this study evaluates precipitation, runoff, and erosion across multiple climate models. The objectives are to: (1) assess projected precipitation from selected climate models under various CMIP frameworks and emission scenarios by comparing them to observational data; (2) analyze extreme precipitation events within defined time windows while accounting for model uncertainties in short-term applications; (3) evaluate the performance of selected climate models in predicting runoff and erosion for the forest road under three distinct design scenarios; and (4) project future precipitation, runoff, and erosion for the study site while incorporating uncertainty analysis. By providing a comprehensive evaluation of climate model performance across different CMIP frameworks and emission scenarios, this study offers robust future projections of precipitation, runoff, and erosion while addressing associated uncertainties.

2. Materials and methods

This study utilized multiple climate models to simulate runoff and erosion using the WEPP model. To ensure the reliability of the selected models, their performance during the historical baseline period

(2015–2023) was evaluated. The Augmented Dickey-Fuller (ADF) test was applied to the time series of differences between modeled and observed daily precipitation to confirm the stationarity of model biases. As high-intensity precipitation events are more relevant to erosion processes, large events were identified based on a defined criterion and used to assess model performance under extreme conditions. An ARMA framework was employed to model residuals and rank the climate models based on their temporal fidelity. For the future period (2024–2099), a non-parametric bootstrap approach was applied to quantify uncertainty and generate confidence intervals for projected precipitation, and estimated runoff and erosion, providing a robust basis for climate-informed road management planning.

2.1. Study area

This study focused on estimating runoff and erosion for a forest road (30.4178°N, −85.2628°E) within the Chipola Experimental Forest in Florida, U.S. (Fig. 1). The predominant soil type was Fuquay loamy sand, composed of 86 % sand, 10.2 % silt, and 3.8 % clay. The analyzed road segment was 100 m in length, 4 m in width, and had a 1 % gradient. Runoff and erosion estimates were obtained using the Water Erosion Prediction Project (WEPP) model (Elliot et al., 1999). Three road scenarios were evaluated: (1) an untreated bare-surface road representing current conditions, (2) an improved design featuring a graded aggregate base as described in Elliot et al. (2023), and (3) a gravel-covered surface. The erodibility parameters for all scenarios are presented in Table 1a.

2.2. Climate models

To assess the impacts of climate change, road runoff and erosion were modeled under various climate scenarios. This study considered five widely used climate models from CMIP5 and CMIP6 across different frameworks, including CanESM, CNRM-CM, GFDL-ESM, HadGEM, and MPI-ESM (Table 1b).

These models were chosen based on their consistent availability across both CMIP5 and CMIP6 generations, structural diversity in climate sensitivity and Earth system representation, and demonstrated capability in simulating precipitation patterns critical to hydrological assessments. GFDL-ESM and HadGEM incorporate advanced atmospheric and land-surface parameterizations that are particularly effective in representing extreme rainfall events (Dunne et al., 2020; Rabiee et al., 2023). MPI-ESM and CNRM-CM offer well-balanced performance across key climate variables, including realistic simulations of subtropical precipitation regimes that are relevant for the southeastern United States (Gutjahr et al., 2019; Voldoire et al., 2013; Decharme and Colin, 2024; Eekhout et al., 2021). CanESM features a high equilibrium climate sensitivity, includes fully coupled land and ocean carbon cycle modules, and supports large ensemble simulations, making it particularly suitable for bounding upper-end projections in erosion stress-testing scenarios (Swart et al., 2019; Ahmadi et al., 2019; Güven et al., 2024). Moreover, these models are widely used as boundary forcings for RCMs. Their use ensures consistency with established downscaling protocols and supports improved spatial resolution and topographic representation, which are critical for hydrological modeling in complex forest road systems.

CMIP6 represents the latest phase of the CMIP, officially launched in 2016, with most simulations conducted between 2017 and 2020. The corresponding RCMs have been released more recently, such as the downscaled climate projections for future climates as represented in the Localized Constructed Analogs (LOCA) version 2 dataset at a 6 km resolution for the North American domain, which was completed in late 2022. Previous runoff and erosion studies incorporating climate change scenarios primarily relied on CMIP5. By evaluating and comparing the performance of both CMIP5 and CMIP6, this study establishes a

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

Table 1
WEPP model simulation parameters.

a. Soil erodibility parameters

Parameters	Bare road surface	Gravel road ditch	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

b. Climate models selected

		Acronym	GCM	RCM
CMIP5	RCP 4.5 and RCP 8.5	CanESM2_CRCM5	Canadian Earth System Model, version 2	Canadian Regional Climate model, version 5
		CNRM-CM5_CRCM5	Centre National de Recherches Météorologiques, version 5	Canadian Regional Climate model, version 5
		GFDL-ESM2M_WRF	Earth System Model – Geophysical Fluid Dynamics Laboratory	Weather Research and Forecasting
		HadGEM2-ES_WRF	Hadley Centre Global Environment Model version 2 Earth system model	Weather Research and Forecasting
		MPI-ESM-LR_WRF	Max Planck Institute for Meteorology Earth System Model LR	Weather Research and Forecasting
CMIP6	SSP2-4.5 and SSP5-8.5	CanESM5_LOCA2	Canadian Earth System Model, version 5	LOCA version 2 for North America
		CNRM-CM6-1_LOCA2	Centre National de Recherches Météorologiques, version 6	LOCA version 2 for North America
		GFDL-ESM4_LOCA2	Earth System Model – Geophysical Fluid Dynamics Laboratory	LOCA version 2 for North America
		HadGEM3-GC31-LL_LOCA2	Hadley Centre Global Environment Model version 3 Earth system model	LOCA version 2 for North America
		MPI-ESM1-2-LR_LOCA2	Max Planck Institute for Meteorology Earth System Model LR	LOCA version 2 for North America

Data reference: CMIP5: https://esgf-ui.ceda.ac.uk/cog/search/cordex-ceda/. CMIP6: https://loca.ucsd.edu/loca-version-2-for-north-america-ca-jan-2023/.

connection between prior research and the latest advancements in climate modeling across different CMIP phases.

RCMs are published with a single realization due to the high cost and time constraints; therefore, to ensure a fair comparison among all climate models, this study considered a single realization for each model. In practice, ensemble methods are commonly used in RCM applications by integrating multiple simulations from different global climate models (GCMs) and exploring various emission scenarios to account for uncertainty. This approach was adopted in this study. Two emission scenarios, 4.5 and 8.5, were considered for all climate models over an 85-year period (2015–2099).

2.3. Climate inputs in WEPP

The CLImate GENerator (CLIGEN 5.3) was utilized to generate stochastic daily weather inputs for long-term and average annual modeling in the WEPP model (https://www.ars.usda.gov/midwest-area/west-lafayette-in/national-soil-erosion-research/docs/wepp/cligen). Climate data from all selected models were converted into WEPP-compatible inputs using the ratio of maximum rainfall intensity to average rainfall intensity (IP) and the ratio of time to rainfall peak to rainfall duration (TP).

In CLIGEN, the PAR file (GLIGEN water station static input file) was calibrated using observational data from the nearest Automated Surface Observation System (ASOS) station, Marianna (30.8356°N, −85.1839°E). Observations from 2000 to 2023, including daily maximum and minimum air temperature, 5-min precipitation, wind speed, wind direction, solar radiation, latitude, longitude, and elevation, were utilized for this calibration. The PRN files were generated based on both the observational data and all selected climate models, containing the required inputs of daily precipitation, maximum and minimum air temperatures. Given that the study was conducted in late 2024, simulated runoff and erosion were assessed for the historical period of 2015–2023. Furthermore, model performance and associated uncertainties were evaluated for future projections spanning 2023–2099.

2.4. Augmented Dickey-Fuller test

The Augmented Dickey-Fuller (ADF) test is a widely used statistical test for assessing the presence of a unit root in a time series, which is a key indicator of stationarity. In the time series analysis, stationarity means the statistical properties of the series such as mean and variance remain constant over time. The ADF test is widely used in climate studies (Chandio et al., 2020; Abdi et al., 2022). In this study, the historical climate parameters (precipitation, maximum temperature, and minimum temperature) were observed from the local ASOS weather station. Before using the selected climate models for either historical or future runoff and erosion estimations, the differences between the climate models and the observed climate parameters were evaluated. The ADF tests were conducted to check the stationary of the precipitation differences. The null hypothesis is the difference data has a unit root, meaning non-stationary, which suggests that trends, long-term dependencies, or structural shifts may exist. The alternative hypothesis is the difference data is stationary, suggesting the statistical properties (mean, variance, and autocorrelation) remain constant over time Running the ADF on daily model-observation differences is the first check before prediction. Divergent stationarity signals structural mismatches (e.g., trends, cycles) that may call for data transformation or revised assumptions. A significant ADF result (stationarity confirmed) builds confidence in the models and any downstream forecasting or anomaly detection.

Given that most RCMs have only one realization available, this study selected a single realization from each model, CMIP, and emission scenario (Table 1b). The climate models are mathematical representations of the Earth's climate system that simulate atmospheric, oceanic, and land surface processes, which are used to project future climate conditions based on physical principles and historical data. Due to natural climate variability, local weather fluctuations, and model uncertainties, their short-term outputs may differ from the observed conditions, and they are more widely used and more reliable for capturing long-term climate trends and responses to external forcings. While short-term mismatches between model outputs and observations are expected, the precipitation differences were evaluated based on the moving average values, and various moving average window sizes were considered in this study. By applying the ADF test to the moving average of climate data, we can reduce the short-term noise, which makes it easier to assess long-term stationarity and enhance the detection of long-term trends.

The "adf.test(.)" function in the R package of "tseries" was used in this study, which offers a robust and convenient approach for assessing stationarity. This function provides a test statistic and p-value, enabling yo an informed decision about stationarity. However, for this function, while the p-value might be smaller than 0.01, it is rounded to 0.01. For this study, 0.01 is strong enough for stationarity and there is no need for a more precise p-value. Other functions could be utilized if the exact p-value is required.

2.5. Identification of relevant large events

In hydrology and soil erosion studies, extreme precipitation events play a crucial role in shaping runoff dynamics and soil erosion processes (Elliot et al., 2023), particularly in Florida. Table 2 presents the number of events and their contributions to runoff and erosion across different rainfall intensity levels. Between 2015 and 2023, 38 extreme rainfall events with daily precipitation exceeding 50 mm accounted for 48.27 % of total runoff and 46.73 % of total erosion. In contrast, rainfall events below 10 mm, which constituted 65.79 % of the 988 recorded events during this period, contributed less than 0.1 % to total runoff and erosion. Therefore, beyond assessing climate model performance across all events, greater emphasis is needed for predicting extreme precipitation events. However, given the expected short-term discrepancies between climate model outputs and observations, accurately identifying relevant extreme events remains a significant challenge.

Given that climate models are generally more reliable for long-term projections, a 91-day window (spanning 45 days before and 45 days after each observed extreme event) was used to identify corresponding large events within the climate model outputs. The selection of a 91-day window ensured that matched events occurred within the same season. Two comparative analyses were conducted between the observed climate and likely climate scenarios generated for the same period using the climate models described in Table 1b: (1) extreme rainfall events with daily precipitation exceeding 50 mm and (2) rainfall events exceeding 20 mm. As shown in Table 2, the first comparison accounted for 48.27 % of total runoff and 46.73 % of total erosion, while the second encompassed 94.51 % of total runoff and 93.98 % of total erosion. Given the greater influence of larger rainfall events, event selection was performed in descending order, starting with the largest observed rainfall event and ensuring a one-to-one match. This approach enhances the robustness of climate model evaluations by capturing the temporal dynamics and lag effects associated with extreme precipitation events.

Paired t-tests were conducted to compare climate models across different CMIP phases and emission scenarios to the historic climate. However, caution was required when interpreting these comparisons, as climate models are primarily designed for long-term projections rather than short-term event matching. Therefore, an insignificant statistic only suggested the study model may have some benefits in capturing the short-term events. The significant statistic may have been due to some typical short-term events, which may not support a good long-term estimation.

2.6. Autoregressive moving-average model

To compare the model performances for precipitation to the observed data, the best fit ARMA models were generated for daily precipitation differences for all considered climate models (Table 1b) and the estimated intercepts (c in equation (1)) were utilized for estimation. Given the ARMA model is particularly well-suited for stationary time series, the ADF tests should be conducted first for stationarity checks. In ARMA models, the notation "ARx_MAy" refers to the order of autoregressive (AR) and moving average (MA) components, where "x" gives the number of AR terms, meaning how many past values of the series influence the current value and "y" gives the number of MA terms, meaning how many past error terms influence the current value. For example, AR1_MA2 model is shown below.

$$Y_t = c + \phi_1 Y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \quad (1)$$

where Y_t is the time series at time t; c is the intercept; ϕ_1 is the autoregessive coefficient for lag 1 (AR1), meaning the current value depends on the previous value; ϵ_t is white noise; θ_1 and θ_2 are the moving average coefficients, meaning the current value is influenced by the past two random shocks (ϵ_{t-1} and ϵ_{t-2}). Various ARMA models were generated for a given climate model in the differences of precipitation, and the best model was selected based on the Akaike Information Criterion (AIC).

Table 2
Predicted runoff and erosion contributions for historical (2015–2023) rainfall events from a road 100 m long with a 1 % grade with a bare surface (Table 1a).

Daily rainfall (mm)	Number of rainfall events	Contributed events (%)	Contributed runoff (%)	Contributed erosion (%)
<10	660	65.79	0.09	0.08
10–20	148	14.98	5.40	5.94
20–30	83	8.40	14.76	15.12
30–50	69	6.98	31.48	32.13
>50	38	3.85	48.27	46.73

2.7. Bootstrap

RCMs often provide only one realization per emission scenario due to computational limitations. To address the resulting limitations in uncertainty characterization, the bootstrap resampling method offers a powerful solution (Efron and Tibshirani, 1993). This method allows researchers to assess the uncertainty of climate projections, even with a limited set of model runs. Specifically, the bootstrap is applied to the outputs from all selected climate models, repeatedly resampling the model predictions with replacement to generate an empirical distribution of outcomes. This approach allowed us to derive robust estimates of the confidence intervals for the climate projections, effectively capturing the variability introduced by model selection. By leveraging the bootstrap, we account for model-based uncertainty without making strict parametric assumptions, enhancing the reliability of the uncertainty quantification in the climate change assessments.

In this study, five models were selected for each CMIP and emission scenario (Table 1b). A bootstrap resampling procedure was performed 1000 times for each CMIP and emission scenario using a fixed seed value of 135 in R to ensure reproducibility. The 95 % confidence intervals were generated. The accuracy of runoff and erosion estimates for each CMIP and emission scenario was assessed using a tolerance range of 0.5–2.0. This criterion considers model estimates acceptable if they exceed half of the observed values but remain within twice the observed data. The percentages of acceptable estimates were summarized for monthly and large event precipitation, runoff and erosion.

3. Results and discussion

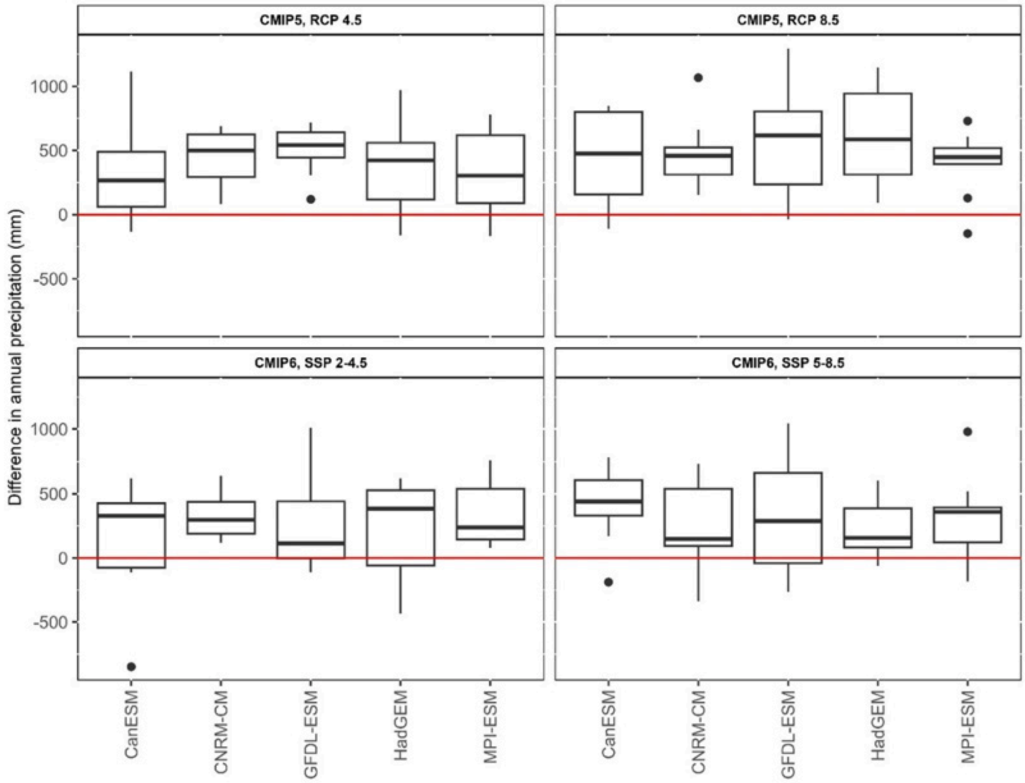
Given that precipitation is a key driver of hydrological processes, directly affecting runoff generation and soil erosion, it is essential to evaluate climate model precipitation before applying it in impact studies. Therefore, this study first assessed the baseline precipitation data before analyzing runoff and erosion. The evaluation began with an analysis of climate model outputs for the historical period (2015–2023), which is crucial for identifying biases, understanding model uncertainties, and ensuring the reliability of projections for future climate impact assessments. The predictions for the future period (2024–2099) were conducted and compared to the historical dataset.

3.1. Climate model evaluation based on precipitation

3.1.1. Annual and monthly evaluation

Box plots show annual precipitation differences across climate models (Fig. 2a), displaying the interquartile range (IQR), median, and outliers (values beyond 1.5 × IQR). During 2015–2023, most models overestimated precipitation under both 4.5 and 8.5 scenarios. This study focused on high-emission scenarios (RCP 8.5/SSP 5–8.5) because: (1) they better capture extreme precipitation events critical for runoff and erosion studies, and (2) they support risk-based planning, helping policymakers develop conservative adaptation strategies (IPCC, 2021).

Table 3a shows t-test p-values comparing CMIP5 and CMIP6 annual



a. Differences of annual precipitation between the climate models and the observational data for the historical period.

Fig. 2. Comparison of climate models and observational data: Annual precipitation (a, b), runoff (c), and erosion (d) differences (2015–2023), with monthly precipitation variability for three road designs.

precipitation projections against observations. CMIP6 models showed better agreement with historical data, with more insignificant p-values. While the 9-year sample size introduces uncertainty, most models performed better under the 4.5 scenario. The CNRM-CM model was an exception, performing better under the 8.5 scenario in CMIP6. These results highlight the importance of considering multiple models and scenarios for reliable precipitation estimates.

Fig. 2b illustrates the monthly precipitation across all considered models. For most months, the climate model data provided reasonable simulations, as the results fell within the observed minimum and maximum precipitation ranges. However, May exhibited a significant overestimation compared to other months. During the historical period considered, local minimum precipitation values were observed in May in several years, including 2016, 2019, 2021, 2022, and 2023. Since this overestimation occurred across all models, it may be attributed to local land surface interactions affecting soil moisture or vegetation dynamics in May. Given the study's focus on the impacts of large precipitation events and the corresponding wet months, no additional adjustments were made for May.

An interesting feature of Fig. 2b is that all but two of the outliers are greater than the band of observed precipitation depths, suggesting that the future climate algorithms may have more extreme wet periods, but are unlikely to have extreme droughts for this region.

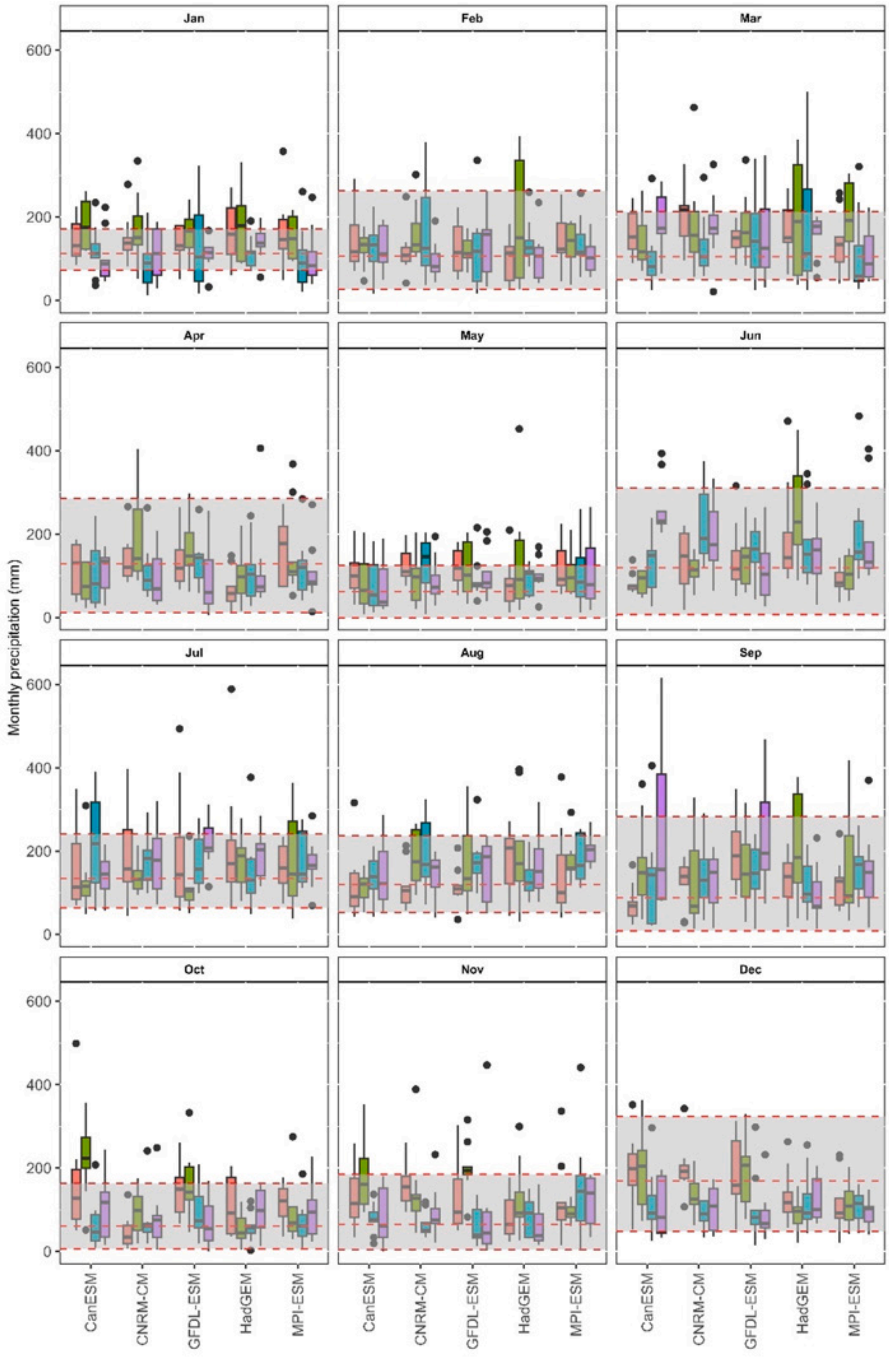
In this study, despite considering relatively high-emission scenarios, many models demonstrated comparable annual precipitation simulations. Additionally, the monthly precipitation projections for nearly all months were consistent with observed patterns, indicating reasonable model performance.

3.1.2. Model comparison based on the ARMA models

The ADF tests were conducted for all considered models for precipitation with various moving average window sizes. Table 3b presents the p-values for window sizes ranging from 1 day to 91 days (45 days before and 45 days after). Significant results across all window sizes rejected the null hypothesis, indicating that the data are likely stationary. For the 1-day window, the ADF test was applied to the daily precipitation differences. The results confirmed stationarity, suggesting no trends or seasonality in the original differences. The moving average was used to identify underlying trends and low-frequency variability in the climate models, with different window sizes capturing variations at various time scales. By removing short-term fluctuations, this approach provided a more robust assessment. The stationarity of the data indicates that the climate model precipitation outputs do not exhibit trends or systematic changes over time and the long-term predictions using these models may be more reliable, as future values are expected to statistically behave similarly to past values.

Given the stationarity observed for all considered models, the best-fit ARMA models were generated, with intercepts presented in Table 4. Initially, different ARMA models were fitted to each climate model. The primary objective of ARMA modeling is to capture the temporal dependence structure of precipitation differences as accurately as possible. Since climate models exhibit varying biases and time-dependent structures, it is appropriate to fit distinct ARMA models to each dataset (Wilks, 2006).

Second, no negative intercept estimates were obtained for any of the considered climate models. Since the best-fit ARMA models were applied to precipitation differences, the point estimate of the intercept represents the mean bias in precipitation differences. Smaller intercepts

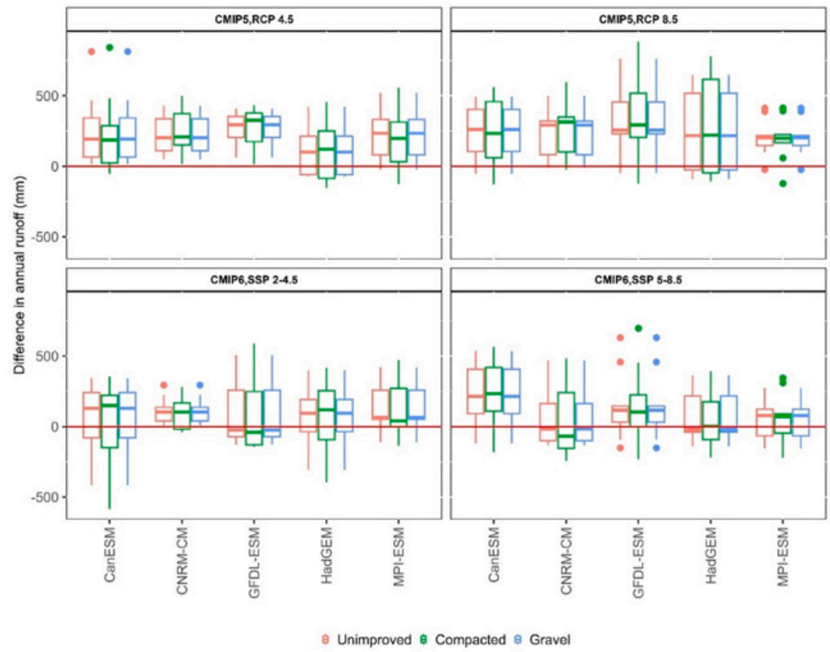


b. Average monthly precipitations for the climate models for the historical period (Notes: the three red dashed lines are the maximum, mean, and minimum of the observational monthly precipitation in the historical period).

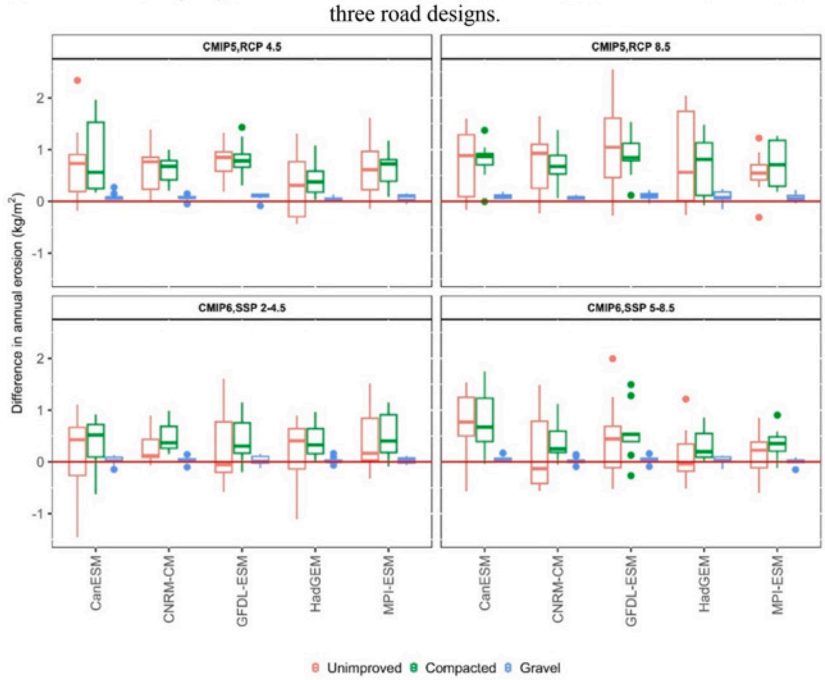
Fig. 2. (*continued*).

(closer to 0) suggest a more unbiased model in terms of precipitation differences, while larger positive intercepts indicate greater precipitation amounts in the future. The 95 % confidence interval provides a range within which the true mean bias is likely to fall with 95 % confidence. The estimated intercept for the CMIP6 SSP 2–4.5 CanESM model was 0, which also showed insignificant differences in the annual precipitation comparison in Table 3a. Based on the confidence intervals, 0 was included within the ranges for the CMIP6 SSP 2–4.5 HadGEM model and CMIP6 SSP 5–8.5 GFDL-ESM model, indicating insignificant overestimations in daily precipitation for these two models as well.

Third, Table 4 presents the rankings of the 20 selected models and scenarios based on the intercepts in their corresponding best-fit ARMA models. Overall, the performance of CMIP6 models was superior to that of CMIP5 models. Therefore, studies based on CMIP5 may anticipate a reduction in precipitation projections when transitioning to CMIP6, even with the same emission scenarios. The emission scenario of 8.5 tended to show larger overestimations compared to 4.5, though this was not always the case. For RCMs, simulated precipitation also depends on regional climate responses, model configurations, and feedback mechanisms, in addition to emission scenarios. While the overall



c. Differences of annual runoff (mm) between the climate models and the observed data for the historical period for three road designs.



d. Differences of annual erosion between the climate models and the observed data for the historical period for three road designs.

Fig. 2. (*continued*).

overestimation was higher for 8.5, extreme precipitation events were expected to be more frequent and intense under 8.5 than under 4.5. When considering the margin of error for the intercepts, most of the 8.5 models exhibited larger uncertainties than their corresponding 4.5 models, even with the best-fit ARMA models applied to each. Table 4 can also serve as a reference for selecting climate models in the study area when working with a limited number of models. However, when possible, it is recommended to use multiple climate models and apply bootstrapping techniques to achieve a more robust estimate of uncertainty.

3.1.3. *Model comparison based on relevant large events*

Given the significant impact of large events on runoff and erosion (Table 2), the performance of all considered climate models was evaluated based on extreme rainfall events (daily precipitation >50 mm) and relatively large events (daily precipitation >20 mm). These events were selected using a 91-day window size to address the mismatch problem. A total of 38 extreme events and 190 relatively large events were included in the analysis. The corresponding p-values from paired t-tests are presented in Table 3c.

For both CMIP5 and CMIP6, the models generally performed better for events with daily precipitation greater than 20 mm than for those exceeding 50 mm, as indicated by less significant statistical results for

Table 3
P-values from *t*-test (model vs. observation precipitation differences) and ADF test (stationarity across moving average windows).

a. P-values from the *t*-test assessing annual precipitation differences between the climate models and the observational data.

Models	CMIP5		CMIP6	
	RCP 4.5	RCP 8.5	SSP 2–4.5	SSP 5–8.5
CanESM	0.027*	0.005**	0.343	0.003**
CNRM-CM	0.000***	0.001***	0.000***	0.092
GFDL-ESM	0.000***	0.003**	0.053	0.079
HadGEM	0.023*	0.001***	0.142	0.015*
MPI-ESM	0.016*	0.002**	0.003**	0.024*

b. P-values from the ADF test assessing stationarity for different moving average window sizes.

Models	CMIP	Emission scenarios	Moving average window size							
			1-day	15-days	31-days	45-days	61-days	71-days	81-days	91-days
CanESM	CMIP5, CMIP6	RCP 4.5, RCP 8.5, SSP 2–4.5, and SSP 5–8.5	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
CNRM-CM	CMIP5, CMIP6	RCP 4.5, RCP 8.5, SSP 2–4.5, and SSP 5–8.5	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
GFDL-ESM	CMIP5, CMIP6	RCP 4.5, RCP 8.5, SSP 2–4.5, and SSP 5–8.5	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
HadGEM	CMIP5, CMIP6	RCP 4.5, RCP 8.5, SSP 2–4.5, and SSP 5–8.5	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
MPI-ESM	CMIP5, CMIP6	RCP 4.5, RCP 8.5, SSP 2–4.5, and SSP 5–8.5	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

c. P-values from the paired t-tests for precipitation of the relevant large events.

Model	Events with daily precipitation > 50 mm				Events with daily precipitation > 20 mm			
	CMIP5		CMIP6		CMIP5		CMIP6	
	RCP 4.5	RCP 8.5	SSP 2–4.5	SSP 5–8.5	RCP 4.5	RCP 8.5	SSP 2–4.5	SSP 5–8.5
CanESM	0.064	0.010**	0.000***	0.731	0.346	0.652	0.017*	0.757
CNRM-CM	0.000***	0.144	0.000***	0.000***	0.440	0.860	0.031*	0.001***
GFDL-ESM	0.055	0.455	0.000***	0.003**	0.047*	0.070	0.000***	0.162
HadGEM	0.002**	0.000***	0.001***	0.010**	0.017*	0.866	0.000***	0.001***
MPI-ESM	0.111	0.020*	0.000***	0.000***	0.475	0.447	0.021*	0.000***

Table 4
The best fit ARMA models of daily precipitation differences between the climate models and the observational data.

	CMIP	Emission scenarios	Model	Best fit ARMA model	Intercept (mm)		
					Estimate	Confidence interval (95 %)	
						lower	upper
1	CMIP6	SSP245	CanESM	AR1_MA2	0.00	0.00	0.00
2	CMIP6	SSP245	HadGEM	AR0_MA1	0.56	−0.02	1.14
3	CMIP6	SSP585	CNRM-CM	AR2_MA2	0.63	0.01	1.25
4	CMIP6	SSP585	HadGEM	AR0_MA2	0.65	0.06	1.24
5	CMIP6	SSP245	GFDL-ESM	AR0_MA1	0.79	0.22	1.35
6	CMIP6	SSP585	GFDL-ESM	AR1_MA2	0.81	−0.06	1.68
7	CMIP6	SSP585	MPI-ESM	AR0_MA2	0.84	0.26	1.43
8	CMIP5	RCP45	MPI-ESM	AR1_MA0	0.87	0.25	1.49
9	CMIP6	SSP245	CNRM-CM	AR1_MA0	0.87	0.24	1.50
10	CMIP5	RCP45	HadGEM	AR1_MA0	0.95	0.27	1.63
11	CMIP5	RCP45	CanESM	AR0_MA1	0.96	0.33	1.59
12	CMIP6	SSP245	MPI-ESM	AR1_MA0	0.97	0.34	1.61
13	CMIP5	RCP85	MPI-ESM	AR0_MA1	1.09	0.49	1.68
14	CMIP6	SSP585	CanESM	AR1_MA2	1.16	0.35	1.98
15	CMIP5	RCP85	CanESM	AR0_MA1	1.20	0.59	1.82
16	CMIP5	RCP45	CNRM-CM	AR0_MA1	1.21	0.63	1.79
17	CMIP5	RCP85	CNRM-CM	AR0_MA1	1.26	0.67	1.86
18	CMIP5	RCP45	GFDL-ESM	AR1_MA0	1.39	0.78	2.00
19	CMIP5	RCP85	GFDL-ESM	AR0_MA1	1.60	0.96	2.23
20	CMIP5	RCP85	HadGEM	AR1_MA2	1.74	0.91	2.58

the latter. Extreme rainfall events (daily precipitation >50 mm) are typically associated with highly localized phenomena, such as thunderstorms, cyclones, or tropical storms, which are challenging to capture even with a 91-day window size. Climate models tend to perform better for more frequent, less intense events, which are more representative of typical atmospheric conditions. At a significance level of 0.001, p-values from all CMIP5 models for events with daily precipitation greater than 20 mm were not significant, and no significant differences were detected for 5 out of 10 CMIP6 models. Even for the extreme events with daily precipitation exceeding 50 mm, several models from both CMIP5 and CMIP6 produced insignificant p-values.

Second, more significant estimates were obtained from CMIP6 than CMIP5 when considering large events. This difference can be attributed to the greater frequency of large events generated by the CMIP5 models. As a result, it was easier to identify a corresponding large event within the 91-day window for the CMIP5 models compared to CMIP6. Consistent with the observations in Table 4, which show larger overestimations from CMIP5 models, the CMIP5 models generally tend to overestimate precipitation across all event scales, including large events.

For both event levels, significant differences were observed in some models based on paired t-tests for either CMIP5 or CMIP6, under both emission scenarios (4.5 and 8.5). Given the challenges associated with simulating extreme events in climate models (Kharin et al., 2007), neither model considered demonstrated consistently good performance across all scenarios. It is therefore recommended to incorporate multiple models in the evaluation of large events, as this approach provides a more robust uncertainty analysis.

The evaluation of climate models showed that (1) based on the exploratory data analysis of annual and monthly precipitation, CMIP6 outperformed CMIP5. For all climate models considered in the study area, it was not possible to determine which emission scenario, 4.5 or 8.5, provided better simulations. Simulations for most months were reasonable, although May was notably overestimated compared to the others. (2) The ADF test indicated stationary precipitation differences for all considered moving average window sizes, with no extreme large intercepts observed for the best-fit ARMA models. CMIP6 models yielded intercepts closer to 0, indicating less bias in the precipitation differences compared to CMIP5. The non-negative intercepts suggested that all climate models tended to overestimate precipitation. (3) The climate models performed better for less intense events, and CMIP5 models produced a higher frequency of large events than CMIP6 models.

3.2. Evaluation based on the estimated runoff and erosion

The WEPP model predicted forest road runoff and erosion using both climate model and observed data inputs. Fig. 2c shows annual runoff differences across models for three road designs. Like precipitation (Fig. 2a), future scenarios produced more runoff than observed. Despite differing hydraulic conductivities (3.8 mm h^{-1} for bare/gravel vs 1.3 mm h^{-1} for compacted), road designs showed no significant runoff differences. CMIP6 runoff predictions matched observations better than CMIP5, with similar historical standard errors across designs. Some CMIP6 models (GFDL-ESM SSP 2–4.5, MPI-ESM SSP 2–4.5, HadGEM SSP 5–8.5, CNRM-CM SSP 5–8.5) yielded runoff values closest to observations.

Fig. 2d shows annual erosion differences across climate models. Improved road designs (compacted surface, graveled ditch) reduced erosion and standard errors. The gravel ditch design showed negligible erosion differences compared to other designs. CMIP6 models performed better, with difference distributions closer to zero. Given CMIP6's superior precipitation and runoff predictions, their improved erosion estimates are expected, as erosion depends on these factors.

Bootstrapping of monthly precipitation, runoff, and erosion across multiple climate models (Table 5a) showed observed runoff generally within 95 % confidence intervals (Fig. 3a). CMIP5 had wider intervals than CMIP6, indicating greater uncertainty. Despite narrower intervals, CMIP6 accurately predicted the runoff across event magnitudes.

Table 5a shows acceptable estimation fractions (95 % CI within 0.5-2 × observed values). Average acceptance rates were: precipitation (0.813) > runoff (0.685) > erosion (0.601), reflecting their dependency chain. While road designs showed no differences for precipitation/runoff, unimproved roads had lower erosion accuracy due to higher variability. CMIP5/6 performed similarly, though CMIP6 SSP 5–8.5 yielded best runoff/erosion predictions across all designs.

Fig. 3b shows monthly erosion bootstrap results. Like runoff, CMIP5 exhibited wider confidence intervals than CMIP6. CMIP5's consistent overestimation across both scenarios better captured large erosions but missed smaller events. Compacted and gravel roads showed narrower intervals with more accurate erosion estimates than unimproved roads (Fig. 3a and b). For future monthly runoff/erosion studies, multi-scenario bootstrapping reduces model bias but increases variability, potentially widening confidence intervals.

Given the challenges associated with daily-scale comparisons, bootstrapping was performed for relevant large events across all considered climate models. The large events were selected based on daily precipitation amounts. The fractions of acceptable estimations for precipitation, runoff, and erosion were 0.961, 0.950, and 0.816 for events with daily precipitation exceeding 50 mm, and 0.992, 0.862, and

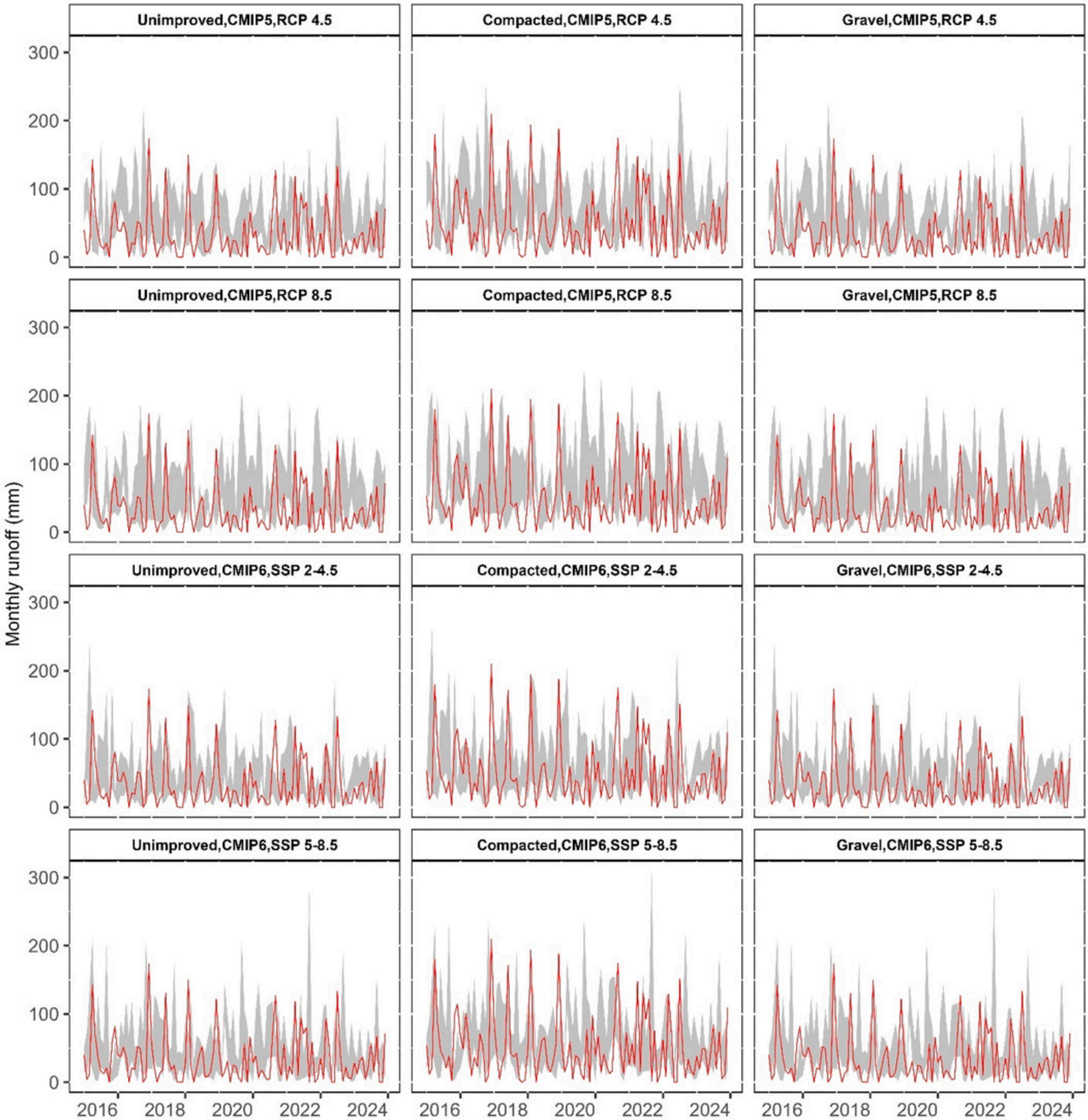
Table 5
Fractions of acceptable estimations for monthly and large-event precipitation, runoff, and erosion across climate models and road designs (vs. historic climate data).

a. Fractions of acceptable estimations of monthly precipitation, runoff, and erosion for the considered climate models with three road designs compared to using historic climate data.

Design	CMIP	Emission scenario	Fraction of acceptable estimations		
			Precipitation	Runoff	Erosion
Unimproved	CMIP5	RCP 4.5	0.833	0.731	0.454
		RCP 8.5	0.806	0.722	0.509
	CMIP6	SSP 2–4.5	0.806	0.676	0.519
		SSP 5–8.5	0.824	0.769	0.556
Compacted	CMIP5	RCP 4.5	0.833	0.667	0.611
		RCP 8.5	0.806	0.648	0.602
	CMIP6	SSP 2–4.5	0.806	0.648	0.630
		SSP 5–8.5	0.824	0.694	0.704
Gravel	CMIP5	RCP 4.5	0.833	0.667	0.639
		RCP 8.5	0.806	0.648	0.630
	CMIP6	SSP 2–4.5	0.806	0.648	0.630
		SSP 5–8.5	0.824	0.704	0.722

b. Fractions of acceptable estimations of large event precipitation, runoff, and erosion for the considered climate models with three road designs.

Design	CMIP	Emission scenario	Fraction of acceptable estimations for events with daily precipitation >50 mm			Fraction of acceptable estimations for events with daily precipitation >20 mm		
			Precipitation	Runoff	Erosion	Precipitation	Runoff	Erosion
Unimproved	CMIP5	RCP 4.5	0.974	0.974	0.632	0.995	0.947	0.542
		RCP 8.5	1.000	1.000	0.684	1.000	0.963	0.516
	CMIP6	SSP 2–4.5	0.921	0.895	0.789	0.984	0.942	0.663
		SSP 5–8.5	0.947	0.947	0.658	0.989	0.942	0.626
Compacted	CMIP5	RCP 4.5	0.974	0.947	0.842	0.995	0.832	0.647
		RCP 8.5	1.000	1.000	0.789	1.000	0.816	0.642
	CMIP6	SSP 2–4.5	0.921	0.895	0.842	0.984	0.811	0.705
		SSP 5–8.5	0.947	0.947	0.868	0.989	0.816	0.747
Gravel	CMIP5	RCP 4.5	0.974	0.947	0.947	0.995	0.837	0.816
		RCP 8.5	1.000	1.000	0.947	1.000	0.816	0.789
	CMIP6	SSP 2–4.5	0.921	0.895	0.868	0.984	0.805	0.795
		SSP 5–8.5	0.947	0.947	0.921	0.989	0.816	0.789



a. Bootstrapped monthly runoff for the historical period (Note: the red solid lines are the observed monthly runoff; the grey shaded areas are the bootstrapped 95% confidence intervals for monthly runoff).

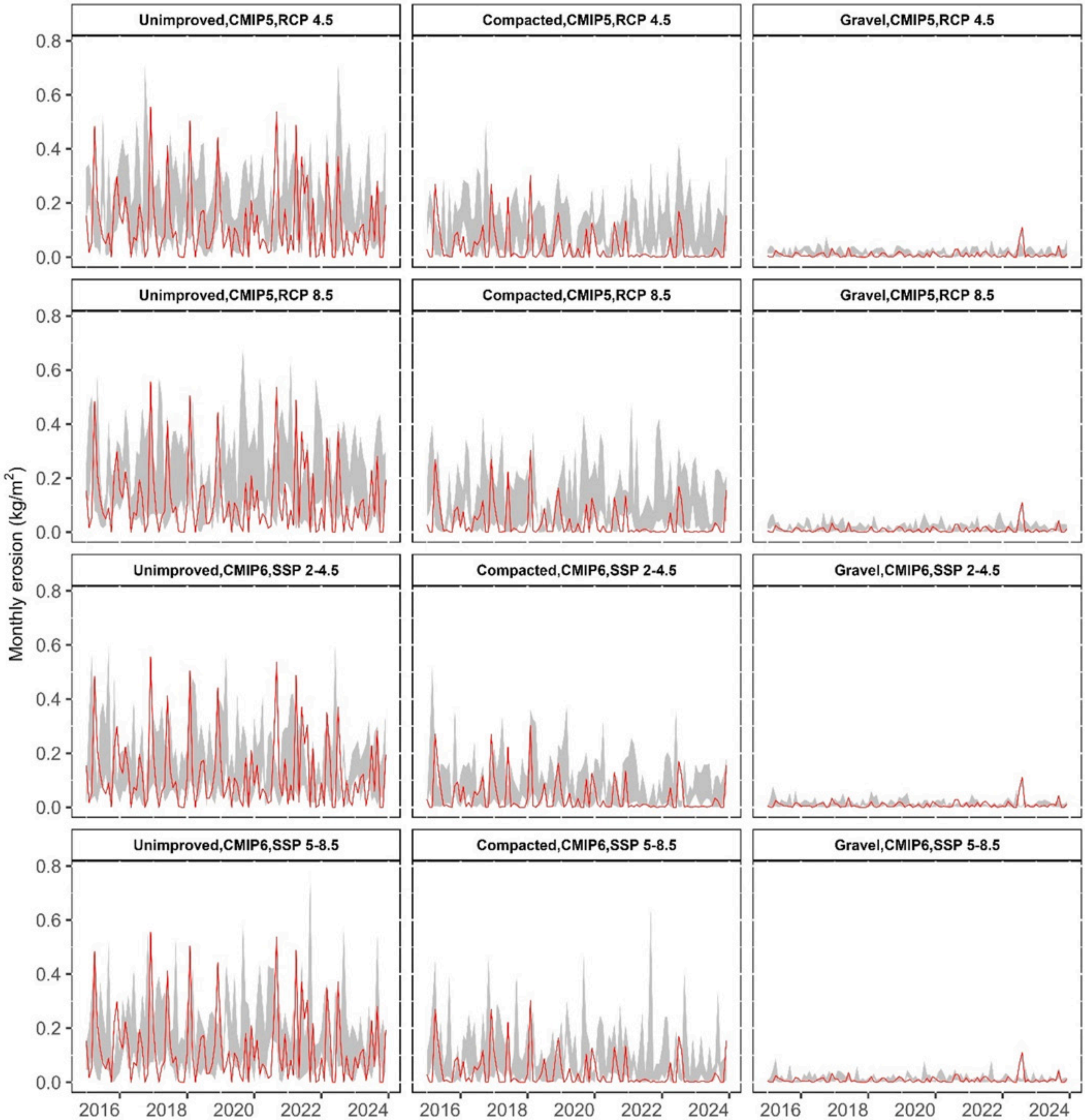
Fig. 3. Bootstrapped monthly runoff and erosion for the historical period with observed values and 95 % confidence intervals

0.690 for events with daily precipitation exceeding 20 mm, respectively (Table 5b). The performance for precipitation and runoff was consistent and similar across all considered CMIPs and emission scenarios for both the >50 mm and >20 mm precipitation events. However, the performance for erosion estimation was poorer for the unimproved road compared to the compacted and gravel roads.

For short-term climate modeling, bootstrapping reduces bias and improves reliability. Given poor erosion estimates for unimproved roads, additional models/scenarios may address uncertainties. While CMIP6 produced narrower confidence intervals, its acceptable estimation rates were similar to CMIP5, though with slightly better erosion predictions.

3.3. Estimations for the future period

Our analysis shows GCMs can reasonably predict historic



b. Bootstrapped monthly erosion for the historical period (Note: the red solid lines are the monthly erosion estimated from observed weather; the grey shaded areas are the bootstrapped 95% confidence intervals for monthly erosion).

Fig. 3. (*continued*).

precipitation, runoff, and erosion, enabling confident future projections using WEPP. Fig. 4 compares future (2024–2099) and historical periods across models. Most models projected increased average annual precipitation under higher emission scenarios (4.5 and 8.5), though CanESM showed decreases for 8.5. CMIP5 HadGEM (RCP 8.5) was an outlier with significantly lower precipitation, while its CMIP6 version projected increases. CMIP5 exhibited greater variability than CMIP6, where only CNRM-CM (SSP 2–4.5) and CanESM (SSP 5–8.5) projected decreases.

Fig. 4 reveals amplified runoff and erosion changes compared to precipitation, consistent across road designs. CMIP5 models under RCP 4.5 showed minimal changes from historical periods, though maintaining high event magnitudes. Under RCP 8.5, most CMIP5 models exhibited little change except MPI-ESM (increase) and HadGEM (decrease). CMIP6 models under SSP 4.5 showed runoff/erosion increases exceeding precipitation (CanESM, GFDL-ESM, HadGEM, MPI-ESM), suggesting more extreme rainfall, while CNRM-CM's decreases

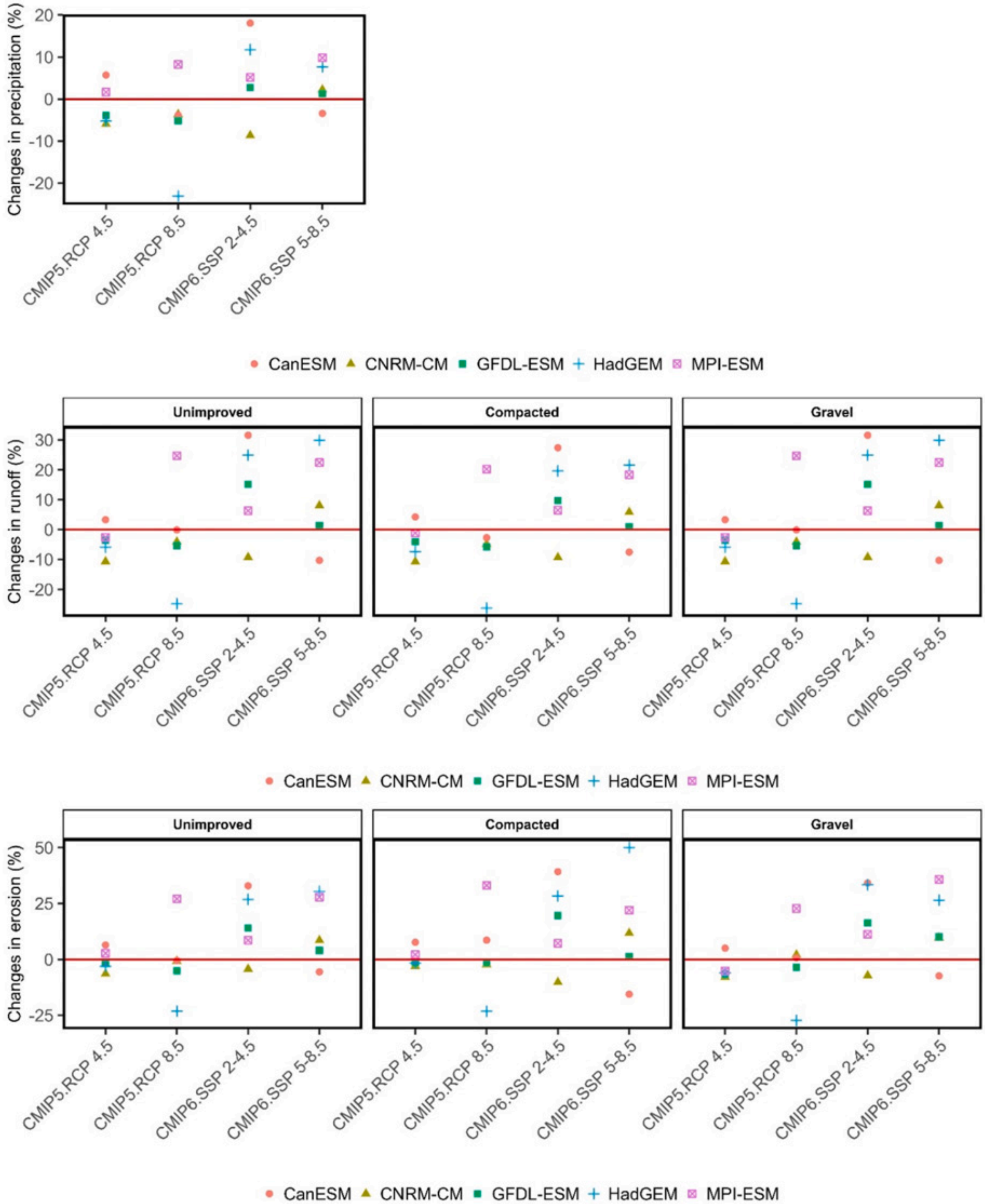
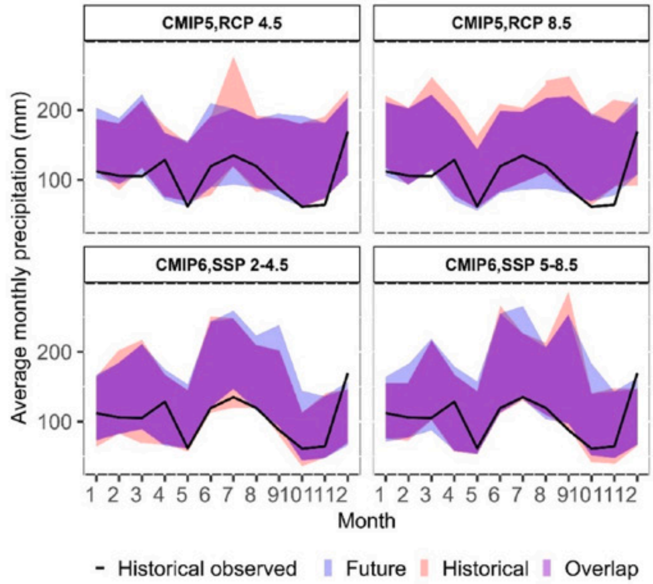


Fig. 4. Changes in average annual precipitation, runoff, and erosion for the future period compared to the historical period.

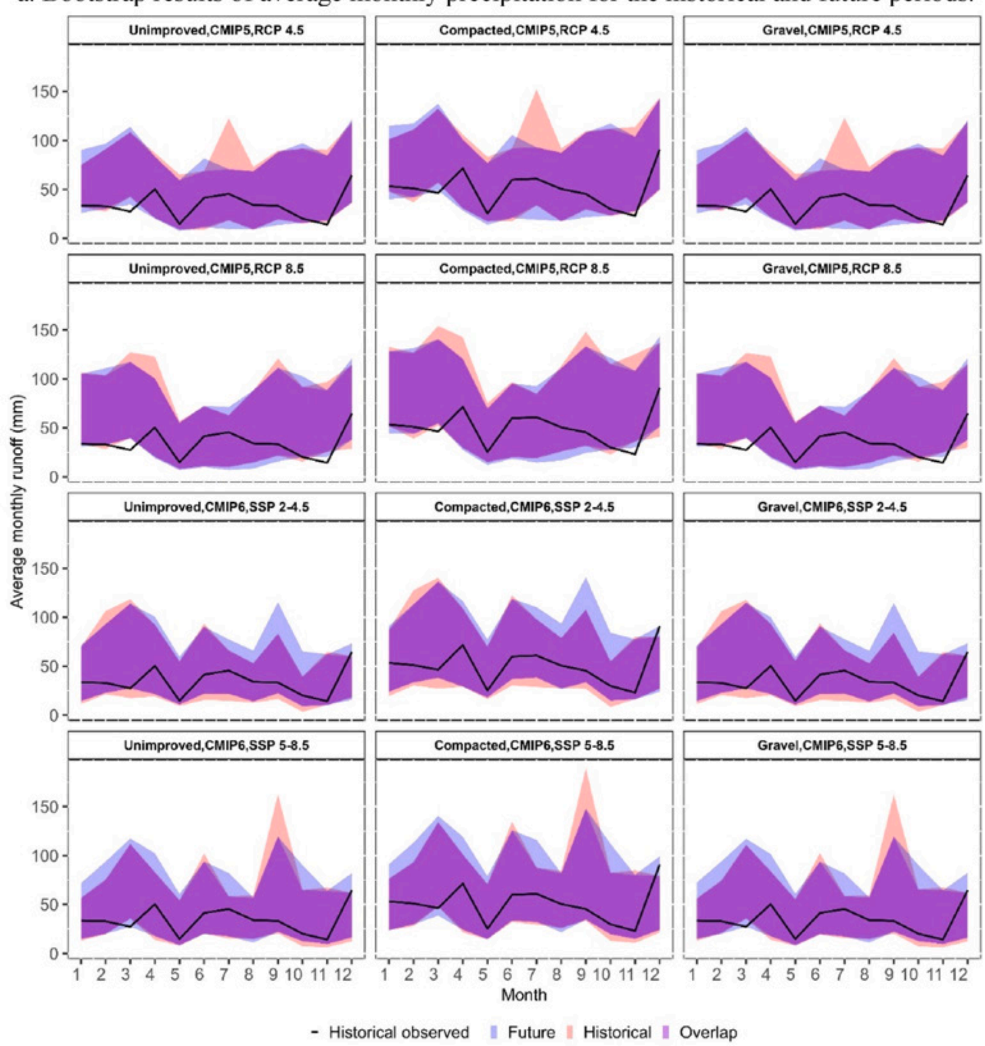
matched precipitation changes. Under SSP 8.5, all CMIP6 models displayed greater runoff/erosion changes than precipitation, indicating shifts in both averages and extremes.

Fig. 5a shows bootstrap results for monthly precipitation. CMIP6's historical estimates captured observations within confidence intervals, while CMIP5 overestimated March and November. CMIP5 under both scenarios showed lower precipitation in several months, with July (4.5 scenario) as an outlier due to extreme events. The 77-year future period should reduce extreme event impacts versus the 9-year historical period. Under RCP 8.5, most CMIP5 months showed decreased future precipitation, largely influenced by HadGEM's notably lower projections.

For CMIP6 models, the patterns of average monthly precipitation remained relatively consistent from the historical to the future period, with the highest precipitation occurring from June to September. For both emission scenarios (4.5 and 8.5), average monthly precipitation increased for most months. In the historical period, which only considered 9 years, the bootstrap results varied between emission scenarios for each month. However, in the future period, which spans 77 years, the bootstrap results for each month became similar across the two emission scenarios. This suggests that the future emission scenarios will not typically increase the overall precipitation but change the probabilities of extreme events.

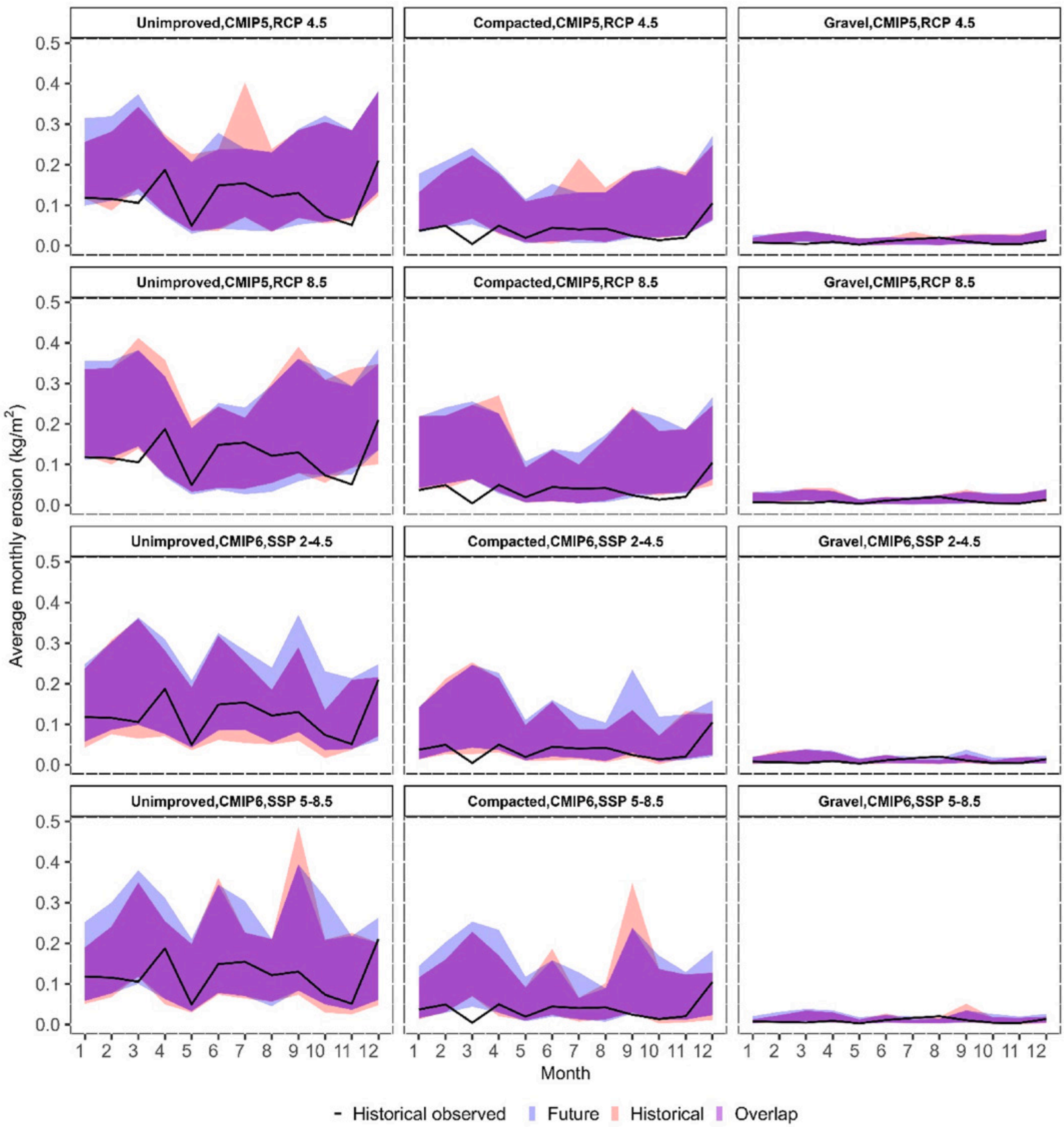


a. Bootstrap results of average monthly precipitation for the historical and future periods.



b. Bootstrap results of average monthly runoff for the historical and future periods as estimated with the WEPP model.

Fig. 5. Bootstrapped average monthly precipitation, runoff (WEPP), and erosion (WEPP) for historical and future periods



c. Bootstrap results of average monthly erosion for the historical and future periods estimated with the WEPP model.

Fig. 5. (*continued*).

Fig. 5b presents bootstrap results for estimated average monthly runoff during the historical and future periods. No significant differences in monthly runoff were observed across the three road designs. The bootstrap confidence intervals for the CMIP6 models accurately captured the observed runoff, whereas the CMIP5 models showed several discrepancies. The future runoff projections based on CMIP5 were generally similar to those in the historical period for most months. The runoff decreased dramatically in July for CMIP5 RCP 4.5 and in March and April for CMIP5 RCP 8.5. In contrast, the CMIP6 models provided consistent runoff estimates for both emission scenarios, with increased runoff observed for most months, peaking in March, June, and September.

Fig. 5c presents the bootstrap results for average monthly erosion. The highest estimated erosion was observed in the unimproved roads, followed by the compacted roads, with the least erosion occurring in the gravel roads. The CMIP6 models provided better estimations, but they showed higher estimated erosion rates for March in the case of compacted roads. Several months had higher estimated erosion rates with the CMIP5 models. The CMIP6 models displayed clear increases in erosion, with similar characteristics observed between the historical and future datasets. For the CMIP5 models, estimated erosion increased for certain months, including January, February, June, and August, while it decreased in May, with opposite trends for the remaining months across the two emission scenarios.

This study recommends using the CMIP6 models for estimating future scenarios for hydrologic modeling. A comparison of historical and future datasets indicates that the CMIP6 models provided consistent estimations. This is supported by the absence of outliers in the average annual estimates, confirmed stationarity in precipitation differences (ADF test), superior rankings in ARMA model assessments, and higher fractions of acceptable estimates for large-event precipitation, runoff, and erosion. When applying bootstrap to account for uncertainties, the characteristics of average monthly precipitation, runoff, and erosion remained relatively consistent between the historical and future periods. The bootstrap confidence intervals from the CMIP6 models captured nearly all observed historical values for both emission scenarios, with only one exception for March erosion on the compacted roads. Based on the evaluation of historical datasets in Table 5a and the bootstrap confidence intervals for each month (totaling 1032 months in the future period) for the CMIP6 models, the expected fractions of acceptable estimations for monthly precipitation, runoff, and erosion are projected to exceed 79 %, 64 %, and 51 %, respectively, across all considered road designs.

4. Conclusions

This study evaluated the projected precipitation from multiple climate models and investigated the corresponding runoff and erosion estimated by the WEPP model for a typical forest road in the Chipola Experimental Forest in the Florida Panhandle region. The analysis provided a comprehensive evaluation of the selected climate models under different CMIPs and emission scenarios, considering uncertainties and future predictions for precipitation, runoff, and erosion.

The climate models CanESM, CNRM-CM, GFDL-ESM, HadGEM, and MPI-ESM were analyzed in both their CMIP5 and CMIP6 versions under emission scenarios 4.5 and 8.5. A comparison of the historical period indicated that all models provided reasonable long-term precipitation projections. The best-fit ARMA models established the rank of projected precipitation for all models based on stationarity achieved in both short-term and long-term analyses. The CMIP6 models demonstrated superior performance, as evidenced by lower intercepts in the ARMA models for precipitation differences, along with more consistent projected precipitation and estimated runoff and erosion across both historical and future periods. When incorporating bootstrapped confidence intervals, both CMIP5 and CMIP6 models demonstrated strong predictive performance, with the fractions of acceptable estimated monthly precipitation, runoff, and erosion exceeding 79 %, 64 %, and 45 %, respectively, across all road designs. For large events within a 91-day window, these fractions increased to over 98 %, 80 %, and 51 % for events with daily precipitation exceeding 20 mm, with better performances for events exceeding 50 mm. This study generated bootstrapped confidence intervals for each CMIP and emission scenario. Future research could extend this approach by constructing confidence intervals across multiple emission scenarios, which may help reduce bias but at the cost of decreased precision.

This study establishes a framework for estimating runoff and erosion on forest roads using climate models while accounting for uncertainties. The comprehensive evaluation of the selected models provides a valuable reference for future model selection. By highlighting the differences between CMIP5 and CMIP6, this study bridges research conducted with these two CMIPs. Incorporating three road designs under both historical and future scenarios, it supports informed decision-making in forest management, road construction, and climate adaptation strategies within the study area.

Future research should consider expanding the spatial scope to include diverse topographic and soil conditions across the southeastern U.S. and beyond, to test the generalizability of the approach. Integrating economic assessments of road maintenance and erosion mitigation under projected climate scenarios would also support more comprehensive, cost-effective management planning. Finally, long-term field monitoring of forest roads, where feasible, would provide critical data for validating model projections and enhancing the accuracy of future erosion prediction models.

CRediT authorship contribution statement

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

Funding was provided by the USDA Forest Service Southern Research Station under the Cooperative Agreement (Number 23-CA-11330180-066) with Florida A&M University. We thank Dr. Nathan Wikle (Assistant Professor, Department of Statistics and Actuarial Science, The University of Iowa) for his valuable input.

Data availability

Data will be made available on request.

References

Abdi, A.H., Warsame, A. Al, Sheik-Ali, I.A., 2022. Modelling the impacts of climate change on cereal crop production in East Africa: evidence from heterogeneous panel cointegration analysis. Environ. Sci. Pollut. Control Ser. 30, 35246–35257.

Ahmadi, M., Motamedvaziri, B., Ahmadi, H., Moeini, A., Zehtabiyan, G.R., 2019. Assessment of climate change impact on surface runoff, statistical downscaling and hydrological modeling. Phys. Chem. Earth, Parts A/B/C 114, 102800.

Anache, J.A.A., Flanagan, D.C., Srivastava, A., Wendland, E.C., 2018. Land use and climate change impacts on runoff and soil erosion at the hillslope scale in the Brazilian Cerrado. Sci. Total Environ. 622–623, 140–151. https://doi.org/10.1016/j.scitotenv.2017.11.257.

Bangash, R.F., Passuello, A., Sanchez-Canales, M., Terrado, M., López, A., Elorza, F.J., Ziv, G., Acuña, V., Schuhmacher, M., 2013. Ecosystem services in Mediterranean river basin: climate change impact on water provisioning and erosion control. Sci. Total Environ. 458–460, 246–255.

Bryan, K., Albritton, C.C., 1943. Soil phenomena as evidence of climatic changes. Am. J. Sci. 241 (8), 469–490.

Cao, Y., Hua, L., Peng, D., Liu, Y., Jiang, L., Tang, Q., Cai, C., 2024. Decoupling the effects of air temperature change on soil erosion in Northeast China. J. Environ. Manag. 351, 119626. https://doi.org/10.1016/j.jenvman.2023.119626.

Chen, C., Hsu, H., Liang, H., 2021. Evaluation and comparison of CMIP6 and CMIP5 model performance in simulating the seasonal extreme precipitation in the Western

North Pacific and East Asia. Weather Clim. Extrem. 31, 100303. https://doi.org/10.1016/j.wace.2021.100303.

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, USA. Int. Soil Water Conserv. Res. https://doi.org/10.1016/j.iswcr.2025.09.007.

Chandio, A.A., Jiang, Y., Rehman, A., Rauf, A., 2020. Short and long-run impacts of climate change on agriculture: an empirical evidence from China. International Journal of Climate Change Strategies and Management 12 (2), 201–221.

Dastorani, M., Mirzavand, M., Dastorani, M.T., Sadatinejad, S.J., 2016. Comparative study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition. Nat. Hazards 81, 1811–1827.

Decharme, B., Colin, J., 2024. Influence of Floodplains and Groundwater Dynamics on the present-day Climate Simulated by the CNRM Model. EGUsphere. https://doi.org/10.5194/egusphere-2024-3091 [preprint].

Dimri, T., Ahmad, S., Sharif, M., 2020. Time series analysis of climate variables using seasonal ARIMA approach. J. Earth Syst. Sci. 129, 149.

Drinking water from forests and grasslands: a synthesis of the scientific literature. In: Dissmeyer, G.E. (Ed.), 2000. Gen. Tech. Rep. SRS-39. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, p. 246.

Dunne, J.P., Horowitz, L.W., Adcroft, A., Ginoux, P., Held, I.M., Johns, J.C.H., Krasting, J.P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C.A., Zadeh, N., Balaji, V., Blanton, C., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P.P.G., Griffies, S.M., Guo, H., Hallberg, R.W., Harrison, M.J., He, J., Hurlin, W., McHugh, C.W., Menzel, R., Milly, P.C.D., Nikonov, S., Paynter, D., Ploshay, J.J., Radhakrishnan, A., Rand, K., Reichel, B., Robinson, T.C., Schwarzkopf, M.D., Sentman, L., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A.T., Wyman, B., Zeng, Y., Zhao, M., 2020. The GFDL Earth System Model version 4.1 (GFDL-ESM4.1): overall coupled model description and simulation characteristics. J. Adv. Model. Earth Syst. 12 e2019MS002015.

Eekhout, J.P.C., Millares-Valenzuela, A., Martínez-Salvador, A., García-Lorenzo, R., Pérez-Cutillas, P., Conesa-García, C., de Vente, J., 2021. A process-based soil erosion model ensemble to assess model uncertainty in climate-change impact assessments. Land Degrad. Dev. 32, 2409–2422.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall/CRC.

Elliot, W.J., Hall, D.E., Scheele, D.L., 1999. WEPP: road (Draft 12/1999) WEPP interference 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. https://doi.org/10.13031/ja.15539.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. (GMD) 9, 1937–1958.

Flanagan, D.C., Livingston, S.J., 1995. WEPP User Summary, NSERL Report No. 11. West Lafayette. USDA-ARS, National Soil Erosion Research Laboratory. https://www.ars.usda.gov/ARSUserFiles/50201000/WEPP/usersum.pdf.

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.

Garbrecht, J.D., Zhang, X.C., 2015. Soil erosion from winter wheat cropland under climate change in Central Oklahoma. Appl. Eng. Agric. 31 (3), 439–454.

Garbrecht, J.D., Steiner, J.L., Cox, C.A., 2007. The times they are changing: soil and water conservation in the 21st century. Hydrol. Process. 21 (19), 2677–2679.

Gaur, S., Singh, R., Bandyopadhyay, A., Singh, R., 2023. Diagnosis of GCM-RCM-driven rainfall patterns under changing climate through the robust selection of multi-model ensemble and sub-ensembles. Clim. Change 176, 13. https://doi.org/10.1007/s10584-022-03475-z.

Giorgi, F., Mearns, L.O., 1999. Introduction to special section: regional climate modeling revisited. J. Geophys. Res. 104 (D6), 6335–6352.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J.H., von Storch, J.-S., Brüggemann, N., Haak, H., Stössel, A., 2019. Max planck Institute Earth System model (MPI-ESM1.2) for the high-resolution model Intercomparison Project (HighResMIP). Geosci. Model Dev. (GMD) 12, 3241–3281.

Güven, A., Gün, M.V., Pala, A., 2024. Modeling and future projection of streamflow and sediment yield in a sub-basin of Euphrates River under the effect of climate change. Journal of Water & Climate Change 15 (6), 2628.

Huang, S., Hattermann, F.F., Krysanova, V., Bronstert, A., 2013. Projections of climate change impacts on river flood conditions in Germany by combining three different RCMs with a regional eco-hydrological model. Clim. Change 116, 631–663.

IPCC (Intergovernmental Panel on Climate Change), 2021. Climate Change 2021: The Physical Science Basis. In: Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Kharin, V.V., Zwiers, F.W., Zhang, X., Hegerl, G.C., 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. J. Clim. 20, 1419–1434.

Kim, K., Whelan, G., Purucker, T., Bohrmann, T.F., Cyterski, M.J., Molina, M., Gu, Y., Pachepsky, Y., Guber, A., Franklin, D.H., 2014. Rainfall-runoff model parameter estimation and uncertainty evaluation on small plots. Hydrol. Process. 28, 5220–5235.

Kropp, J.P., Schellnhuber, H.J., 2010. In: Extremis: Disruptive Events and Trends in Climate and Hydrology. Springer, Berlin, p. 400.

Kumar, N., Singh, S.K., Dubey, A.K., Ray, R.L., Mustak, Sk, Rawat, K.S., 2021. Prediction of soil erosion risk using earth observation data under recent emission scenarios of CMIP6. Geocarto Int. 37 (24), 7041–7064.

Lenderink, G., de Vries, H., van Meijgaard, E., de Rooy, W., van Ulft, L., Thompson, V., Qiu, X., Fowler, H.J., 2025. A pseudo global warming based system to study how climate change affects high impact rainfall events. Weather Clim. Extrem. 49, 100781.

Li, Z., Fang, H., 2016. Impacts of climate change on water erosion: a review. Earth Sci. Rev. 163, 94–117. https://doi.org/10.1016/j.earscirev.2016.10.004.

Liu, Z., Wang, M., Liu, X., Lyu, X., Wang, M., Wang, F., Ji, X., Li, X., 2024. Spatiotemporal simulation and projection of soil erosion as affected by climate change in Northeast China. Int. J. Appl. Earth Obs. Geoinf. 135, 104305.

Ma, D., Peng, S., Lin, Z., 2025. Towards a more comprehensive scenario analysis: response of soil erosion to future land use and climate change in the Central Yunnan Urban Agglomeration, China. J. Environ. Manag. 376, 124523.

Maeda, E.E., Pellikka, P.K.E., Siljander, M., Clark, B.J.F., 2010. Potential impacts of agricultural expansion and climate change on soil erosion in the Eastern Arc Mountains of Kenya. Geomorphology 123 (3–4), 279–289.

Mulla, S., Pande, C.B., Singh, S.K., 2024. Time series forecasting of monthly rainfall using Seasonal Auto Regressive Integrated Moving Average with exogenous variables (SARIMAX) model. Water Resour. Manag. 38, 1825–1846.

Mullan, D., 2013. Soil erosion under the impacts of future climate change: assessing the statistical significance of future changes and the potential on-site and off-site problems. Catena 109, 234–246.

Nearing, M.A., Pruski, F.F., O'Neal, M.R., 2004. Expected climate change impacts on soil erosion rates: a review. J. Soil Water Conserv. 59 (1), 43–50.

Parajuli, P.B., Jayakody, P., Sassenrath, G.F., Ouyang, Y., 2016. Assessing the impacts of climate change and tillage practices on stream flow, crop and sediment yields from the Mississippi River Basin. Agric. Water Manag. 168, 112–124.

Paroissien, J.B., Darboux, F., Couturier, A., Devillers, B., Mouillot, F., Raclot, D., Bissonnais, Y. Le, 2015. A method for modeling the effects of climate and land use changes on erosion and sustainability of soil in a Mediterranean watershed (Languedoc, France). J. Environ. Manag. 150, 57–68. https://doi.org/10.1016/j.jenvman.2014.10.034.

Pruski, F.F., Nearing, M.A., 2002. Runoff and soil-loss responses to changes in precipitation: a computer simulation study. J. Soil Water Conserv. 57 (1), 7–16.

Rabiee, M.R.S., Peyrowan, H.R., Arasteh, P.D., Akbary, M., Motamedvaziri, B., 2023. Investigating the consequences of climate change in runoff and sediment rate. Arabian J. Geosci. 16, 304.

Rajbanshi, J., Bhattacharya, S., 2022. Modelling the impact of climate change on soil erosion and sediment yield: a case study in a sub-tropical catchment, India. Modeling Earth Systems and Environment 8, 689–711.

Sarhadi, A., Kelly, R., Modarres, R., 2014. Snow water equivalent time-series forecasting in Ontario, Canada, in link to large atmospheric circulations. Hydrol. Process. 28, 4640–4653.

Swart, N.C., Cole, J.N.S., Kharin, V.V., Lazare, M., Scinocca, J.F., Gillett, N.P., Anstey, J., Arora, V., Christian, J.R., Hanna, S., Jiao, Y., Lee, W.G., Majaess, F., Saenko, O.A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., Winter, B., 2019. The Canadian Earth System Model version 5 (CanESM5.0.3). Geosci. Model Dev. (GMD) 12, 4823–4873.

Takhellambam, B.S., Srivastava, P., Lamba, J., McGehee, R.P., Kumar, H., Tian, D., 2023. Projected mid-century rainfall erosivity under climate change over the southeastern United States. Sci. Total Environ. 865, 161119.

Tang, J.L., Cheng, X.Q., Zhu, B., Gao, M.R., Wang, T., Zhang, X.F., Zhao, P., You, X., 2015. Rainfall and tillage impacts on soil erosion of sloping cropland with subtropical monsoon climate — a case study in hilly purple soil area, China. J. Mt. Sci. 12 (1), 134–144.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., Chauvin, F., 2013. The CNRM-CM5.1 global climate model: description and basic evaluation. Clim. Dyn. 40, 2091–2121.

Wang, W., Yin, S., He, Z., Chen, D., Wang, H., Klik, A., 2023. Projections of rainfall erosivity in climate change scenarios for mainland China. Catena 232, 107391.

Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences, second ed. Academic Press, Amsterdam, Boston.

Wu, P., Li, R., Xiao, L., He, M., Chen, H., 2024. Mediation effects of vegetation, land use and cover change and karst characteristics on hydrologic processes in heterogeneous karst basins. Hydrol. Process. 38, e15093.

Zhang, A., Shi, H., Li, T., Fu, X., 2018. Analysis of the influence of rainfall spatial uncertainty on hydrological simulations using the bootstrap method. Atmosphere 9, 71.

Zhang, X.C., 2012. Cropping and tillage systems effects on soil erosion under climate change in Oklahoma. Soil Sci. Soc. Am. J. 76 (5), 1789–1797.

Zhang, X.C., Nearing, M.A., 2005. Impact of climate change on soil erosion, runoff, and wheat productivity in central Oklahoma. Catena 61, 185–195.

Zhu, Q., Yang, X., Ji, F., Du, Z., 2024. Rainfall erosivity projection in South-East Australia using the improved regional climate simulations. Int. J. Climatol. 45, e8702.