

A Climate-Compatible Machine-Learning Framework for Future Rainfall Erosivity Projections Using Standardized Climate Model Outputs

Kenan İnce

kananince@gmail.com

Ankara University

Günay Erpul

Ankara University

Research Article

Keywords: Rainfall erosivity (R30), Climate-compatible machine learning, Machine-learning emulator, Random Forest, Climate projections, Soil erosion

Posted Date: August 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10525957/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Rainfall erosivity (R_{30}) is a key driver of soil erosion, yet projecting its future evolution remains challenging because standard climate-model outputs generally lack the sub-daily rainfall information required by conventional erosivity estimation methods. This study develops a climate-compatible machine-learning framework that estimates rainfall erosivity directly from predictor variables routinely available in climate-model outputs, providing a transferable approach for large-scale climate impact assessments. Historical rainfall erosivity observations from 209 meteorological stations across Türkiye were used to train a Random Forest emulator using strict station-wise GroupKFold spatial cross-validation. The validated model was then driven by annual precipitation projections from three global climate models (GFDL, HadGEM, and MPI) under the RCP4.5 and RCP8.5 scenarios to project rainfall erosivity for 2025–2095. The emulator explained approximately 54% of the observed variability under spatial validation ($R^2 = 0.54$). National projections indicate a moderate decline in mean rainfall erosivity across all climate model–scenario combinations, while regional and station-level analyses reveal pronounced spatial heterogeneity, including localized increases in several coastal and transitional climatic regions. Multi-model ensemble analyses further demonstrate increasing divergence among climate projections toward the end of the century, emphasizing the importance of accounting for climate-model uncertainty. By relying exclusively on predictors consistently available from climate projections, the proposed framework provides a practical and computationally efficient bridge between standardized climate-model outputs and rainfall erosivity assessment, supporting large-scale soil erosion risk evaluation and climate adaptation planning.

1. Introduction

Rainfall erosivity is a key driver of soil erosion, representing the capacity of rainfall to detach and transport soil particles. Because erosivity is strongly governed by rainfall intensity and temporal distribution, projected changes in the global hydrological cycle under climate change are expected to substantially modify erosion processes. Rising temperatures increase the atmosphere's moisture-holding capacity, enhancing the frequency and magnitude of extreme precipitation events (Trenberth, 2011; IPCC, 2023). Consequently, intensified runoff, accelerated soil degradation, nutrient depletion, and increased sediment transport are anticipated, making reliable projections of future rainfall erosivity essential for sustainable land and water management (Almagro et al., 2017; Nearing et al., 2004; Kundzewicz et al., 2005).

Long-term observations already reveal a global increase in extreme precipitation events (Papalexiou and Montanari, 2019), while climate projections indicate that this trend will continue throughout the twenty-first century (Sun et al., 2007). Accordingly, numerous studies have investigated the implications of climate change for rainfall erosivity and soil erosion (Panagos et al., 2022; Takhellambam et al., 2023; Zhao et al., 2024). Global assessments suggest that climate change alone could increase soil erosion by 9–56%, although responses are expected to vary considerably among regions depending on future

precipitation regimes, topography, and land surface characteristics (Yang et al., 2003; Borrelli et al., 2020).

In Türkiye, climate projections indicate pronounced regional differences in future rainfall regimes. Although annual precipitation is projected to decrease across much of western and southern Türkiye, the frequency of short-duration, high-intensity rainfall events and associated flooding is expected to increase (Dalfes et al., 2007; Demir et al., 2008; Önol and Semazzi, 2009; Türkeş, 1999, 2003). These changes are likely to alter the spatial distribution of rainfall erosivity, with important implications for soil conservation, watershed management, and agricultural sustainability (Kılıç and Gunal, 2021).

Considerable progress has been achieved in projecting future rainfall erosivity under climate change. Dynamical and statistical downscaling, bias-corrected climate products, stochastic weather generators, and machine-learning-based rainfall reconstruction techniques have substantially improved the representation of precipitation extremes and enabled future erosivity assessments across a wide range of spatial and temporal scales. These developments have significantly expanded the applicability of climate-driven soil erosion modelling.

Despite these advances, projecting future rainfall erosivity remains methodologically challenging. Conventional erosivity estimation relies heavily on rainfall intensity characteristics derived from sub-daily observations, whereas standard climate model outputs generally provide only aggregated precipitation variables. Consequently, future erosivity assessments often require multiple preprocessing steps, including rainfall disaggregation, bias correction, or statistical reconstruction, increasing computational complexity and limiting transferability across climate models and emission scenarios. Developing approaches capable of translating routinely available climate projection variables directly into reliable rainfall erosivity estimates therefore represents an important research priority for large-scale climate change impact assessment.

Recent studies in Türkiye have shown that annual precipitation, together with topographic and regional variables, explains a substantial proportion of rainfall erosivity variability (İnce, 2025), suggesting that climate-compatible predictor variables may provide sufficient information to estimate future erosivity without explicitly reconstructing rainfall intensity. Building on this concept, the present study develops a climate-compatible machine-learning emulator that estimates rainfall erosivity directly from predictor variables consistently available in climate model outputs. The emulator is trained using historical observations from meteorological stations across Türkiye, rigorously evaluated through strict station-wise spatial cross-validation, and subsequently coupled with precipitation projections from three global climate models under moderate- and high-emission scenarios to estimate future rainfall erosivity through 2095.

The main contributions of this study are threefold. First, a climate-compatible machine-learning framework is developed to project future rainfall erosivity directly from predictor variables routinely available in climate model outputs, eliminating the need for rainfall disaggregation or rainfall-intensity reconstruction. Second, the framework provides spatially explicit national-, regional-, and station-level

projections of future rainfall erosivity under multiple climate models and emission scenarios. Third, projection uncertainty is explicitly quantified using a multi-model ensemble approach, providing a more comprehensive basis for climate-informed soil erosion risk assessment and adaptation planning.

Accordingly, the objectives of this study are to:

- Develop and validate a climate-compatible machine-learning emulator for rainfall erosivity using strict station-wise spatial cross-validation;
- Project future rainfall erosivity across Türkiye using precipitation projections from multiple global climate models under RCP4.5 and RCP8.5 scenarios;
- Quantify projected national, regional, and station-level changes in rainfall erosivity; and
- Evaluate inter-model uncertainty associated with future rainfall erosivity projections.

2. Materials and Methods

2.1 Study Area

Türkiye was selected as the study area because its pronounced climatic and physiographic diversity provides an ideal natural laboratory for developing and evaluating climate-compatible models of future rainfall erosivity. Located at the junction of Southeastern Europe and Western Asia (36°–42° N, 26°–45° E), the country encompasses a wide spectrum of climatic conditions, ranging from the humid Black Sea climate in the north to Mediterranean climates along the western and southern coasts and semi-arid continental environments across Central and Eastern Anatolia (Fig. 1).

The country's complex topography further amplifies this climatic diversity. Elevation ranges from sea level to mountain systems exceeding 3000 m above mean sea level, generating strong spatial gradients in precipitation amount, seasonality, and rainfall erosivity. Consequently, Türkiye exhibits substantial regional variability in erosion-driving climatic processes, making it particularly suitable for evaluating the spatial transferability of machine-learning models under contrasting environmental conditions.

The study utilizes rainfall erosivity observations from 209 meteorological stations distributed across the country's seven geographical regions (Marmara, Aegean, Mediterranean, Central Anatolia, Black Sea, Eastern Anatolia, and Southeastern Anatolia). The broad spatial coverage of the monitoring network captures the major climatic and physiographic gradients influencing rainfall erosivity and provides a robust foundation for developing a nationally representative projection framework.

Beyond its geographical diversity, Türkiye represents an appropriate test case because high-quality historical rainfall erosivity observations are available together with dynamically downscaled climate projections from multiple global climate models. This combination enables the development of a climate-compatible machine-learning emulator that can be trained using observed erosivity while remaining directly applicable to routinely available climate-model outputs. Although demonstrated here

for Türkiye, the proposed framework is designed to be transferable to other regions where comparable rainfall erosivity observations and climate projections are available.

The overall methodological framework adopted in this study is summarized in Fig. 2.

Figure 2. Overall methodological framework of the proposed climate-compatible machine-learning emulator for projecting future rainfall erosivity. Historical rainfall erosivity observations from 209 meteorological stations were combined with climate-compatible predictor variables to train and evaluate multiple machine-learning models using station-wise GroupKFold cross-validation. Following model selection, annual precipitation projections from three global climate models (GFDL, HadGEM, and MPI) under the RCP4.5 and RCP8.5 scenarios were used to estimate rainfall erosivity for the period 2025–2095. The resulting projections were analysed at national, regional, and station scales, while projection uncertainty was quantified using a multi-model ensemble framework.

2.2 Data Sources and Overall Projection Framework

Historical rainfall erosivity observations were obtained from the Dynamic Erosion Modeling and Monitoring System (DEMMS), which provides annual RUSLE rainfall erosivity (R_{30}) estimates derived from minute-resolution rainfall records collected at 209 meteorological stations distributed across Türkiye. The historical database covers the period 2004–2021 and includes annual rainfall erosivity together with station latitude, longitude, elevation, annual precipitation, and regional climatic classifications.

The annual Revised Universal Soil Loss Equation (RUSLE) using the sum of event-based erosivity values was determined, as expressed in Eq. (1), by summing the R_i values of all rainfall events within a given year that were capable of inducing erosion (RUSLE- R_{30}).

$$R_y = \sum_{i=1}^m R_i = \sum_{i=1}^m (E_i x I_{30})_i$$

1

Here, R_y denotes the annual total energy flux ($\text{MJ ha}^{-1} \text{ mm h}^{-1}$), and m represents the total number of rainfall events within a given year that satisfy the predefined calculation criteria. The long-term mean RUSLE- R_{30} values were subsequently computed using Eq. (2).

$$R = \frac{1}{N} \sum_{j=1}^N \left[\sum_{i=1}^m (E_i x I_{30})_i \right]$$

2

E_i is rainfall kinetic energy, I_{30} is the maximum 30-min rainfall intensity, and mmm is the number of erosive rainfall events during a given year. Where N denotes the number of observation years.

Future climate information was obtained from three global climate models (GFDL-ESM2M, HadGEM2-ES, and MPI-ESM-MR; hereafter referred to as GFDL, HadGEM, and MPI) under the RCP4.5 and RCP8.5 emission scenarios. The simulations were dynamically downscaled by the Turkish State Meteorological Service to a spatial resolution of 20 km × 20 km. Future annual precipitation totals were available at five-year intervals covering the period 2025–2095.

Projected precipitation grids were processed using ArcGIS 10.2. Meteorological station locations were spatially matched with the corresponding climate grid cells, and projected annual precipitation totals were extracted for every station and projection period. These climate-compatible precipitation variables constituted the inputs to the machine-learning emulator.

The overall methodological workflow is summarized in Fig. 1. Historical rainfall erosivity observations were first used to develop and spatially validate multiple machine-learning models. The optimal model was subsequently driven by projected annual precipitation from the three climate models to estimate future rainfall erosivity. The resulting projections were analysed at national, regional, and station scales, while inter-model uncertainty was quantified using a multi-model ensemble approach.

2.3 Machine-Learning Emulator Development

Five regression algorithms representing complementary machine-learning paradigms were evaluated: Ridge Regression, ElasticNet Regression, Random Forest, Gradient Boosting Regression, and Multilayer Perceptron (MLP). These algorithms were selected to represent linear, regularized, ensemble tree-based, and nonlinear learning approaches while providing a comprehensive comparison of predictive capability.

Model hyperparameters were optimized using RandomizedSearchCV, which enables efficient exploration of parameter space while substantially reducing computational cost relative to exhaustive grid search. Hyperparameters governing model complexity, regularization, tree structure, ensemble size, and feature selection were randomly sampled within predefined ranges. To ensure an unbiased comparison, all algorithms were optimized and evaluated using an identical spatial validation strategy.

Random Forest was ultimately selected as the climate emulator because it consistently achieved the highest predictive accuracy and demonstrated the strongest spatial generalization performance.

2.4 Spatial Cross-Validation and Model Evaluation

Model performance was assessed using a station-wise GroupKFold cross-validation strategy. All observations belonging to the same meteorological station were assigned exclusively to either the training or validation subsets, thereby preventing spatial information leakage that commonly affects conventional random cross-validation.

Compared with random partitioning, GroupKFold provides a substantially more realistic estimate of predictive performance because the model is evaluated only at previously unseen locations. This validation strategy therefore better represents the intended application of the emulator for projecting rainfall erosivity under future climate conditions.

Predictive performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) (Table 1).

Table 1 Model Performance Metrics and Mean Values			
Model Performance Metrics	Coefficient of determination (R^2)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)
Mean	0.537 ± 0.073	≈ 984 mm	≈ 461 mm

2.5 Future Rainfall Erosivity Projection

Following model validation, the selected Random Forest emulator was retrained using the complete historical dataset to maximize the information available for projection. The final model was subsequently coupled with annual precipitation projections from the three global climate models (GFDL, HadGEM, and MPI) under the RCP4.5 and RCP8.5 emission scenarios.

For each climate model–scenario combination, projected annual precipitation totals at five-year intervals (2025–2095) were extracted for every meteorological station. These climate-compatible precipitation variables were combined with static geographic predictors, including latitude, longitude, elevation, and regional classification, to construct predictor datasets consistent with those used during model development. The trained emulator was then applied to estimate annual rainfall erosivity (R_{30}) for each station and projection period.

Station-level projections constituted the basis for all subsequent analyses. Projected R_{30} values were aggregated to derive national and regional statistics, while spatial changes were quantified by calculating ΔR_{30} as the difference between projected late-century and baseline rainfall erosivity for each station. The overall projection workflow is summarized in Fig. 1.

2.6 Climate Impact Assessment and Uncertainty Analysis

Projected rainfall erosivity was evaluated at national, regional, and station scales. National-scale analyses were based on aggregated station averages for each climate model–scenario combination and projection period, whereas regional analyses were performed separately for Türkiye's seven geographical regions. Station-level projections were used to characterize localized spatial variability and to generate spatial distributions of projected change (ΔR_{30}).

To facilitate visualization of regional patterns, station-based ΔR_{30} values were spatially interpolated. These interpolated surfaces were intended solely to illustrate generalized spatial trends and should not be interpreted as exact local predictions.

Projection uncertainty was quantified using a multi-model ensemble approach across the six climate model–scenario combinations. National uncertainty was summarized using the ensemble median together with the minimum–maximum range and the 10th–90th percentile interval. Spatial uncertainty

was evaluated by calculating the standard deviation of projected ΔR_{30} among all model–scenario combinations at each monitoring station, thereby identifying regions exhibiting greater inter-model disagreement.

3. Results

3.1 Validation of the Climate-Compatible Machine-Learning Emulator

The predictive performance of the climate-compatible machine-learning emulator was evaluated using five-fold station-wise GroupKFold cross-validation, whereby all observations from an individual meteorological station were assigned exclusively to either the training or validation subset. This validation strategy prevents spatial information leakage and provides a realistic assessment of the model's ability to generalize to previously unseen locations.

The selected Random Forest emulator, developed exclusively from predictor variables routinely available in climate model outputs (annual precipitation, latitude, longitude, elevation, year, and regional classification), demonstrated satisfactory predictive performance under spatial validation (Table 2). The model explained 53.7% of the observed variability in annual rainfall erosivity (mean $R^2 = 0.537 \pm 0.073$), with a root mean square error (RMSE) of 984 MJ mm ha⁻¹ h⁻¹ yr⁻¹ and a mean absolute error (MAE) of 461 MJ mm ha⁻¹ h⁻¹ yr⁻¹.

Table 2
Station-wise cross-validated performance of the selected Random Forest emulator

Metric	Mean	Standard Deviation
R ²	0.537	0.073
RMSE (mm)	984	—
MAE (mm)	461	—

The relatively small variation in R^2 among validation folds indicates stable predictive performance across geographically independent station groups, suggesting that the emulator is not overly sensitive to the spatial composition of the training data. Although its predictive accuracy is lower than that of models incorporating rainfall-intensity descriptors, this reduction is expected because the emulator intentionally relies only on variables that are consistently available from climate model projections. The results therefore demonstrate that a reduced set of climate-compatible predictors captures a substantial proportion of the spatial and temporal variability in rainfall erosivity, providing a robust and transferable basis for subsequent climate-change projections.

3.2 Comparative Performance of Alternative Predictor Configurations

To examine the influence of predictor selection on model performance, three alternative feature configurations were evaluated using the same station-wise GroupKFold cross-validation framework:

- 1. Combined, including rainfall intensity indices, annual precipitation, and geographic variables;
- 2. Indices Only, including rainfall intensity indices and geographic variables; and
- 3. Totals Only, including annual precipitation and geographic variables.

Five regression algorithms representing linear, nonlinear, and ensemble-learning approaches were trained and evaluated for each feature configuration. Their cross-validated performance is summarized in Table 3.

Table 3
Performance comparison of machine-learning algorithms under different predictor configurations using station-wise spatial cross-validation

Feature Set	Model	Mean R ²	Std R ²	Mean RMSE (mm)	Mean MAE (mm)
Combined	GB	0.712	0.054	772	376
	RF	0.705	0.060	783	380
	Ridge	0.659	0.047	842	425
	MLP	0.553	0.294	936	412
	ElasticNet	0.225	0.100	1269	627
Indices Only	GB	0.707	0.058	780	376
	MLP	0.678	0.057	815	395
	Ridge	0.677	0.069	815	430
	RF	0.673	0.055	826	387
	ElasticNet	0.204	0.076	1288	633
Totals Only	GB	0.611	0.073	897	442
	RF	0.544	0.072	976	459
	MLP	0.492	0.224	996	465
	ElasticNet	0.091	0.123	1377	693
	Ridge	-13.806	18.977	3962	766

Marked differences in predictive performance were observed among the three predictor configurations. Models developed using the Combined and Indices Only feature sets consistently achieved the highest predictive accuracy. Gradient Boosting produced the best overall performance, with mean R^2 values of 0.712 and 0.707 for the Combined and Indices Only configurations, respectively. The small difference between these two configurations indicates that rainfall intensity indices account for most of the predictive information required to estimate rainfall erosivity.

In contrast, models based exclusively on annual precipitation and geographic variables showed lower predictive performance. Within this Totals Only configuration, Gradient Boosting remained the highest-performing algorithm ($R^2 = 0.611$), followed by Random Forest ($R^2 = 0.544$), whereas the linear models exhibited substantially lower predictive skill. Ridge Regression performed particularly poorly under this configuration, indicating limited ability to capture the nonlinear relationship between annual precipitation and rainfall erosivity in the absence of rainfall intensity descriptors.

Overall, ensemble tree-based algorithms (Gradient Boosting and Random Forest) consistently outperformed the linear regression approaches across all predictor configurations, demonstrating their superior capacity to represent the complex nonlinear relationships governing rainfall erosivity. Because rainfall intensity indices are not routinely available from future climate model outputs, the Totals Only configuration was adopted for subsequent climate projections despite its lower predictive accuracy. Accordingly, the Totals Only predictor configuration was selected as the basis for the climate-compatible projection framework presented in the subsequent sections.

3.3 National-Scale Projections of Future Rainfall Erosivity

National-scale projections of rainfall erosivity generated by the climate-compatible Random Forest emulator are presented in Figs. 2 and 3, with the corresponding late-century changes summarized in Table 4. Across all climate model–scenario combinations, projected national mean rainfall erosivity (R_{30}) generally declines between 2025 and 2095, although the magnitude and temporal evolution vary among simulations.

Table 4
Projected national mean R_{30} change between 2025 and 2095

Model	Scenario	Mean R_{30} (2025)	Mean R_{30} (2095)	ΔR_{30} (2095–2025)
GFDL	4.5	812.01	695.28	-116.73
GFDL	8.5	735.95	508.85	-227.11
HadGEM	4.5	882.51	779.21	-103.30
HadGEM	8.5	796.17	675.66	-120.51
MPI	4.5	637.50	596.25	-41.25
MPI	8.5	674.22	552.18	-122.05

Figure 3 shows the temporal evolution of national mean rainfall erosivity for the six climate model–scenario combinations. Under the RCP4.5 scenarios, projected R_{30} generally decreases gradually throughout the projection period, with relatively small differences among climate models. In contrast, the RCP8.5 scenarios exhibit greater temporal variability and progressively increasing divergence after approximately 2070. Nevertheless, all simulations project lower national mean rainfall erosivity in 2095 than in 2025.

The magnitude of the projected decline differs among climate models. The largest reduction is simulated by the GFDL model under RCP8.5 ($\Delta R_{30} = -227.11 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$), whereas the smallest change occurs under the MPI RCP4.5 scenario ($\Delta R_{30} = -41.25 \text{ MJ mm ha}^{-1} \text{ h}^{-1} \text{ yr}^{-1}$) (Table 4). For all three climate models, the high-emission scenario generally produces larger reductions in rainfall erosivity than the corresponding moderate-emission scenario.

The temporal patterns summarized in Fig. 3 are further illustrated in the heatmap presented in Fig. 4, which facilitates comparison of projected rainfall erosivity among climate model–scenario combinations across successive projection periods. While projected values remain relatively similar during the early decades, divergence among simulations becomes increasingly evident toward the end of the century, particularly under the RCP8.5 scenarios.

Overall, the national-scale projections indicate a consistent long-term tendency toward declining rainfall erosivity across Türkiye, accompanied by progressively increasing differences among climate model projections. These patterns provide the basis for the uncertainty assessment presented in Section 3.5. The close agreement in the direction of change across all 209 monitoring stations, despite differences in magnitude among climate model–scenario combinations, suggests that the projected national tendency is robust, whereas the magnitude of future rainfall erosivity remains sensitive to the choice of climate model and emission scenario.

3.4 Spatial Patterns of Projected Rainfall Erosivity

Although the national-scale projections indicate an overall decline in rainfall erosivity, the spatial analyses reveal pronounced geographical heterogeneity across Türkiye (Fig. 5). Projected changes vary considerably among regions and climate model–scenario combinations, demonstrating that national averages conceal substantial local variability.

At the station scale, negative ΔR_{30} values predominate throughout most of the monitoring network, indicating a widespread reduction in projected rainfall erosivity by the end of the century (Fig. 5). Nevertheless, localized increases remain evident at several stations, particularly in southern and eastern Türkiye, illustrating that future rainfall erosivity is unlikely to evolve uniformly across the country.

The strongest projected decreases are concentrated in western coastal areas, whereas much of the continental interior exhibits comparatively smaller changes. Despite differences in the magnitude of projected change among climate models, the principal spatial patterns remain consistent across the ensemble. Regions influenced by Mediterranean and transitional climatic conditions generally experience larger projected changes than the semi-arid interior, while isolated areas of increasing rainfall erosivity persist under several climate model–scenario combinations.

Despite differences in the magnitude of projected change, all six climate model–scenario combinations reproduce a broadly similar spatial organization of rainfall erosivity, indicating that the geographical distribution of projected change is more robust than its absolute magnitude.

Overall, the spatial analyses demonstrate that future rainfall erosivity across Türkiye is characterized by marked regional variability superimposed on a common national tendency toward declining mean R_{30} . These geographically explicit projections complement the national-scale assessment and provide the spatial context for the uncertainty analysis presented in the following section.

3.5 Uncertainty Analysis

Uncertainty in future rainfall erosivity projections was quantified using a multi-model ensemble comprising six climate model–scenario combinations (GFDL, HadGEM, and MPI under the RCP4.5 and RCP8.5 scenarios). Ensemble statistics were evaluated at both the national and station scales to characterize temporal and spatial variability in projected rainfall erosivity.

3.5.1 Temporal Evolution of Projection Uncertainty

The temporal evolution of national-scale uncertainty is presented in Fig. 6. Across the projection period, the ensemble median indicates a gradual decrease in nationally averaged rainfall erosivity (Fig. 6a). However, the spread among climate model projections increases progressively toward the end of the century.

The uncertainty envelope, represented by the 10th–90th percentile interval and the full minimum–maximum range, remains relatively narrow during the early projection period but expands during the

second half of the century. This widening indicates increasing divergence among climate model–scenario combinations in both the magnitude and timing of projected rainfall erosivity changes.

Figure 6b further quantifies the evolution of projection spread. Both uncertainty metrics exhibit a gradual increase throughout the century, with the largest spread occurring during the final projection decades. These results indicate that inter-model variability becomes progressively larger in late-century projections.

(a) National mean R_{30} projected by the six climate model–scenario combinations together with the ensemble median and the 10th–90th percentile uncertainty envelope. (b) Temporal evolution of projection spread expressed by the widths of the 10th–90th percentile interval and the full minimum–maximum range.

3.5.2 Spatial Distribution of Projection Uncertainty

The spatial distribution of projection uncertainty is illustrated in Fig. 7, which presents the standard deviation of projected late-century rainfall erosivity change ($\Delta R_{30} = 2095 - 2025$) across the six climate model–scenario combinations.

Standard deviation of projected ΔR_{30} (2095–2025) calculated from the six climate model–scenario combinations. Higher values indicate greater inter-model disagreement, whereas lower values indicate stronger agreement. The interpolated surface is provided solely to facilitate visualization of regional patterns.

Inter-model variability is spatially heterogeneous. Higher standard deviations occur predominantly in western and southwestern Türkiye, whereas lower values characterize much of Central Anatolia, indicating stronger agreement among climate models in these areas. Intermediate levels of uncertainty are observed across several transitional climatic regions.

Overall, the spatial pattern demonstrates that uncertainty is unevenly distributed across the country, with localized areas exhibiting substantially greater variability among climate projections than others.

4. Discussion

4.1 National and Regional Responses of Rainfall Erosivity to Climate Change

The climate-compatible machine-learning emulator projects an overall decline in national mean rainfall erosivity across Türkiye during the twenty-first century. However, this national tendency masks pronounced regional heterogeneity, with several areas exhibiting localized increases despite the overall downward trend. These findings demonstrate that future rainfall erosivity is unlikely to respond uniformly

to climate change and emphasize the importance of spatially explicit analyses when evaluating future erosion risk.

The projected decline contrasts with several recent studies that reported increasing rainfall erosivity under future climate scenarios. For example, Chen et al. (2024) projected widespread increases in global rainfall erosivity using CMIP5 and CMIP6 climate simulations, while Das and Jain (2024) reported substantial increases across India, particularly within semi-arid environments. Similarly, Weng et al. (2024) identified increasing rainfall erosivity across southern China using Bayesian multi-model ensembles. Rather than contradicting these studies, the present results illustrate the strong regional dependence of future rainfall erosivity. The projected decline over Türkiye most likely reflects the precipitation changes simulated for the eastern Mediterranean climate system, where several climate models consistently project reductions in annual precipitation under future warming.

Although the national signal indicates declining rainfall erosivity, regional analyses reveal that the Mediterranean and Aegean regions generally experience larger projected changes than the continental interior. These regional contrasts are consistent with previous studies demonstrating that rainfall erosivity is controlled not only by precipitation amount but also by regional atmospheric circulation, topography, and storm climatology (Panagos et al., 2022; Takhellambam et al., 2023). Consequently, national averages alone are insufficient for evaluating future erosion hazards because they may obscure climatically sensitive regions where localized increases remain possible.

The projected spatial heterogeneity therefore has important implications for climate adaptation. Soil conservation strategies based exclusively on national statistics may underestimate future erosion risk in regional hotspots, whereas spatially explicit projections allow adaptation measures to be prioritized according to regional climatic sensitivity.

4.2 Methodological Contributions of the Climate-Compatible Machine-Learning Emulator

The principal contribution of this study is methodological rather than algorithmic. Most recent rainfall erosivity projection studies rely on rainfall downscaling, bias correction, stochastic rainfall generation, or reconstruction of sub-daily rainfall intensity before erosivity estimation (Chen et al., 2024; Bagarello et al., 2025). Although these approaches provide high physical realism, they require climate variables that are often unavailable or inconsistent among global climate models and usually involve substantial computational preprocessing.

The framework proposed here follows a different philosophy by directly estimating rainfall erosivity from predictor variables routinely available in climate-model outputs. Using only annual precipitation together with static geographical variables, the proposed emulator provides a transferable workflow that can be applied consistently across multiple climate models and emission scenarios without rainfall disaggregation or temporal reconstruction.

This simplification inevitably involves a trade-off between predictive accuracy and transferability. Under strict station-wise spatial validation, the reduced-feature Random Forest emulator explained approximately 54% of the observed rainfall erosivity variability. Although this performance is lower than that achieved by models incorporating rainfall-intensity indices, those variables are generally unavailable for future climate simulations. The reduced-feature emulator therefore prioritizes operational applicability for climate impact assessment rather than maximizing historical prediction accuracy.

Recent studies published in CATENA have highlighted the increasing role of machine learning in soil erosion assessment because of its ability to capture nonlinear environmental relationships while maintaining computational efficiency (Senanayake et al., 2024). The present study extends this direction by demonstrating that machine learning can function not only as a predictive model but also as a climate-compatible emulator, translating standardized climate projections directly into rainfall erosivity estimates. Rather than replacing physically based erosion models, the proposed framework complements existing approaches by providing a scalable interface between climate projections and erosion assessment.

An additional methodological strength is the adoption of GroupKFold spatial cross-validation. Conventional random cross-validation frequently overestimates predictive performance because neighbouring observations are shared between calibration and validation datasets. By assigning entire meteorological stations exclusively to either training or validation subsets, the present framework provides a more realistic assessment of model transferability, consistent with current recommendations for environmental machine-learning applications (Roberts et al., 2017).

4.3 Projection Uncertainty and Interpretation

Future rainfall erosivity projections inevitably inherit uncertainties associated with climate models, emission scenarios, and the statistical representation of climate–erosivity relationships. Rather than selecting a single climate projection, this study explicitly quantified uncertainty using a multi-model ensemble framework.

The ensemble analysis demonstrates that although the overall direction of change remains relatively consistent, disagreement among climate models increases progressively toward the end of the century, particularly under the higher-emission scenarios. Similar behaviour has recently been reported in global rainfall erosivity assessments, where inter-model agreement decreases with increasing radiative forcing and projection horizon (Chen et al., 2024). These findings reinforce the importance of ensemble-based analyses instead of deterministic interpretations based on individual climate models.

Spatial uncertainty also exhibits a coherent geographical structure. Higher inter-model disagreement is concentrated mainly in western and southwestern Türkiye, whereas projections across central Anatolia remain comparatively consistent. These patterns indicate that uncertainty primarily reflects regional differences in projected precipitation responses rather than random model variability. Consequently,

regions exhibiting both substantial projected changes and elevated uncertainty should be interpreted with greater caution when supporting adaptation planning.

4.4 Limitations and Future Perspectives

Several limitations should be considered when interpreting the present findings.

First, the proposed emulator estimates rainfall erosivity using annual precipitation together with static geographical variables. Consequently, future changes in rainfall intensity, storm duration, and intra-event precipitation structure are not explicitly represented, although these processes strongly influence rainfall erosivity.

Second, projected rainfall erosivity directly inherits uncertainties associated with global climate models and emission scenarios. Although the ensemble framework quantifies inter-model variability, uncertainties arising from climate sensitivity, internal climate variability, and future socioeconomic pathways remain unavoidable.

Third, the present framework evaluates only the climatic component of rainfall erosivity. Other important controls on future soil erosion—including land-use change, vegetation dynamics, soil degradation, conservation practices, and watershed management—were intentionally excluded. Therefore, the projections should be interpreted as estimates of future climatic erosivity rather than direct predictions of soil loss.

Future research should focus on integrating convection-permitting climate simulations, bias-corrected high-temporal-resolution precipitation datasets, and CMIP6 climate projections. Coupling climate-compatible emulators with physically based erosion and sediment transport models would enable more comprehensive assessments of future erosion risk. Furthermore, evaluating the proposed framework across different climatic regions would provide important evidence regarding its geographical transferability. As recently emphasized by Das et al. (2024) in the development of globally consistent rainfall erosivity datasets, transferable climate-compatible methodologies will become increasingly important as standardized climate projections continue to expand in both spatial resolution and geographical coverage.

5. Conclusions and Outlook

This study developed and demonstrated a climate-compatible machine-learning emulator for projecting future rainfall erosivity (R_{30}) across Türkiye under multiple climate models and emission scenarios. Unlike conventional approaches that rely on rainfall disaggregation, temporal downscaling, or reconstruction of sub-daily rainfall intensity, the proposed framework estimates rainfall erosivity directly from predictor variables routinely available in standardized climate-model outputs. This simplified workflow enables efficient integration of climate projections into rainfall erosivity assessments while maintaining compatibility across multiple climate models and emission scenarios.

The results indicate a general tendency toward declining nationally averaged rainfall erosivity during the twenty-first century across all six climate model–scenario combinations. However, this national-scale signal conceals pronounced spatial heterogeneity. Regional and station-level analyses revealed localized increases in rainfall erosivity despite declining national averages, demonstrating that climate-driven changes are geographically differentiated rather than spatially uniform. Consequently, future erosion assessments and adaptation strategies should be based on spatially explicit analyses rather than national averages alone.

The present findings should also be interpreted in the context of our previous national-scale projection study (İnce et al., 2026), which employed a parsimonious panel-data regression framework driven by downscaled climate projections. While that study identified broad regional tendencies—including increasing rainfall erosivity across Eastern Anatolia and generally declining tendencies in Southeastern Anatolia and parts of the Black Sea region—the present machine-learning framework captures comparable large-scale spatial variability using a completely different modelling philosophy. Rather than relying on empirical regression equations, the proposed emulator learns nonlinear relationships directly from historical observations while remaining compatible with standardized climate-model outputs. The consistency of the broad regional patterns between these two independent modelling approaches increases confidence that the projected spatial heterogeneity represents robust climate-driven tendencies rather than artefacts of a particular modelling technique.

The multi-model ensemble analysis further demonstrates that projection uncertainty increases progressively toward the late twenty-first century. Although the ensemble median consistently indicates a gradual long-term decline in rainfall erosivity, the widening spread among climate model projections highlights the importance of explicitly accounting for climate-model uncertainty when assessing future erosion hazards. Ensemble-based analyses therefore provide a more robust basis for climate adaptation planning than reliance on individual climate projections.

From a methodological perspective, the principal contribution of this study lies in demonstrating that meaningful rainfall erosivity projections can be obtained using only annual precipitation together with static geographic predictors. Similar to the operational projection model adopted by İnce et al. (2026), the proposed emulator intentionally prioritizes transferability over maximum historical predictive accuracy. Although the reduced-feature Random Forest model explains a smaller proportion of rainfall erosivity variability than models incorporating rainfall-intensity variables ($R^2 \approx 0.54$ under strict station-wise spatial cross-validation), it achieves direct compatibility with standardized climate projections while preserving satisfactory predictive skill for national-scale climate impact assessments. This represents a deliberate trade-off between predictive performance and operational applicability.

The proposed framework is therefore not intended to replace physically based erosion models, rainfall downscaling procedures, or rainfall-intensity reconstruction techniques. Instead, it complements these approaches by providing a computationally efficient and transferable alternative for translating

standardized climate-model outputs into erosion-relevant information, particularly in regions where high-temporal-resolution rainfall observations are unavailable.

Several limitations remain. Similar to the operational regression framework presented by Ince et al. (2026), the proposed emulator relies on annual precipitation rather than explicit sub-daily rainfall intensity. Consequently, the projections should be interpreted as regional-scale climate scenarios of future rainfall erosivity rather than deterministic predictions of individual erosive storms or local event-scale responses. Moreover, uncertainties associated with climate-model structure, emission scenarios, and unresolved spatial processes remain unavoidable and become increasingly important toward the end of the projection period.

Future research should focus on integrating bias-corrected CMIP6 and convection-permitting climate simulations, incorporating high-temporal-resolution precipitation products where available, coupling emulator outputs with physically based soil erosion and sediment transport models, exploring advanced ensemble-learning and deep-learning architectures, and evaluating the transferability of the proposed framework across contrasting climatic and physiographic regions.

Overall, this study demonstrates that climate-compatible machine-learning emulators provide an effective bridge between standardized climate projections and rainfall erosivity assessment. Together with the parsimonious regression framework developed by Ince et al. (2026), the results indicate that transferable climate-compatible models can provide consistent large-scale assessments of future rainfall erosivity despite differences in modelling philosophy. This convergence strengthens confidence in the projected regional patterns across Türkiye and provides a robust scientific basis for climate-adaptive soil conservation, watershed management, and evidence-based land management planning under future climate change.

Declarations

Competing Interests

The authors declare that they have no relevant financial or non-financial interests to disclose.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Author Contribution

K.İ. and G.E. wrote the main manuscript text and prepared figures. All authors reviewed the manuscript.

Acknowledgements

The authors express their gratitude to the staff of the General Directorate of Combating Desertification and Erosion and the General Directorate of Meteorology of the Ministry of Environment, Urbanization, and Climate Change of the Republic of Türkiye for providing data and valuable support throughout this research.

Data Availability:

Data will be made available upon request.

References

1. Almagro A, Oliveira PTS, Nearing MA, Hagemann S (2017) Projected climate change impacts in rainfall erosivity over Brazil. *Sci Rep* 7:8130
2. Bagarello V, Dallan E, Ferro V, Pampalone V, Borga M (2025) Evaluation and projected changes in rainfall erosivity: Topography dependence revealed by Convection-Permitting climate projections for the Mediterranean island of Sicily. *CATENA*, 254 [10.1016/j.catena.2025.108975]
3. Borrelli P, Robinson DA, Panagos P, Lugato E, Yang JE, Alewell C, Wuepper D, Montanarella L, Ballabio C (2020) Land use and climate change impacts on global soil erosion by water (2015–2070). *Proc. Natl. Acad. Sci.* 117 (36), 21994–22001 sep
4. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
5. Brown LC, Foster GR (1987) Storm erosivity using idealized intensity distributions. *Trans ASAE* 30(2):379–386
6. Chen Y, Wei T, Li J, Xin Y, Ding M (2024) Future changes in global rainfall erosivity: Insights from the precipitation changes. *J Hydrol* 638. <https://doi.org/10.1016/j.jhydrol.2024.131435>
7. Dalfes HN, Karaca M, Şen ÖL (2007) Climate change scenarios for Turkey. İçinde: Ç. Güven (Editör) *Climate Change and Turkey Impacts, Sectoral Analyses, Socio-Economic Dimensions*. United Nations Development Programme (UNDP) Turkey Office, Ankara, pp 11–18
8. Das S, Jain MK (2023) Unravelling the future changes in rainfall erosivity over India under shared socio-economic pathways, vol 232. *CATENA*. <https://doi.org/10.1016/j.catena.2023.107417>
9. Das S, Jain MK, Gupta V (2024) *Sci Data* 11:926. <https://doi.org/10.1038/s41597-024-03756-5>. GloRESatE: A dataset for global rainfall erosivity derived from multi-source data
10. Demir İ, Kılıç G, Coşkun M (2008) Climate predictions for Turkey using PRECIS Regional Climate Model: Scenario HaDAMP3 SRES A2. *Proceedings of the International Fourth Symposium on Atmospheric Sciences*, 25–28 Mart 2008, İstanbul, Turkey
11. Erpul G, İnce K, Demirhan A, Küçümen A, Akdağ MA, Demirtaş İ, Sarıhan B, Çetin E, Şahin S (2020) *Provincial Statistics of Water Erosion – Soil Erosion Control Strategies (Sustainable Land/Soil*

- Management Practices and Approaches) (in Turkish). Publications of the General Directorate of Combating Desertification and Erosion, ISBN: 978-605-7599-36-0. Ankara, Türkiye
12. Erpul G, Şahin S, Akgöz R, İnce K, Güden A, Çetin E (2016) Rainfall Properties of Turkey and R Factor of the Revised Universal Soil Loss Equation (In Turkish), Publications of the General Directorate of Combatting Desertification and Erosion, ISBN: 978-605-4610-93-8, Ankara
 13. IPCC (2023) Summary for Policymakers. In: Climate Change. Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]
 14. İnce K (2025) The Impact Of Climate Change On Precipitation Intensity And Energy Flow, And The Evaluation Of Soil Erosion At The National Scale In Türkiye. (Doctoral dissertation, Ankara University, Ankara, Türkiye)
 15. İnce K, Atasoy T, Erpul G (2026) Assessment of climate change impacts on rainfall erosivity (R) across Türkiye using panel-data regression, Soil and Tillage Research, December 2026, Article: 107371, 264. 10.1016/j.still.2026.107371
 16. Kaya P, Aydın S, Yilman FE, Ozsayar M, Erpul G, Bayramin I (2006) Kinetic energy calculations for RUSLE technology using long-term recorded data in Turkey. 18th International Soil Meeting (ISM) on Soil Sustaining Life on Earth, Managing Soil and Technology, May 22–26, 2006 Şanlıurfa – Turkey, Proceedings, Volume 2: 555–560
 17. Kilic OM, Gunal H (2021) Spatial-temporal changes in rainfall erosivity in Turkey using CMIP5 global climate change scenario. Arabian Journal of Geosciences
 18. Kundzewicz ZW, Ulbrich U, Graczyk D, Krüger A, Leckebusch GC, Menzel L, Pinskiwar I, Radziejewski M, Szwed M (2005) Summer floods in Central Europe-Climate change track. Nat Hazards 36(1–2):165–189
 19. Nearing GS, Kratzert F, Sampson AK, Pelissier CS, Klotz D, Frame JM, Gupta HV (2021) What role does hydrological science play in the age of machine learning? Water Resour Res, 57(3)
 20. Nearing MA et al (2004) Expected climate change impacts on soil erosion rates: A review. Journal of Soil and Water Conservation, Vol: 59, No:1
 21. Nearing MA, Garbrecht JD, Steiner JL (2015) Can conservation trump impacts of climate change on soil erosion? An assessment from winter wheat cropland in the Southern Great Plains of the United States. Weather Clim Extremes 10:36–43
 22. Önoğlu B, Semazzi FM (2009) Regionalization of Climate Change Simulations over the Eastern Mediterranean. J Clim 22:1944–1961
 23. Panagos P, Borrelli P, Matthews F, Liakos L, Bezak N, Diodato N, Ballabio C (2022) Global rainfall erosivity projections for 2050 and 2070. J Hydrol 610:127865
 24. Papalexiou E, Montanari A (2019) Global and regional increase of precipitation extremes under global warming. Water Resour Res 55(6):4901–4914
 25. Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Aroita G, Dormann CF (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure.

26. Senanayake S, Pradhan B, Wedathanthirige H, Alamri A, Park H-J (2024) Monitoring soil erosion in support of achieving SDGs: A special focus on rainfall variation and farming systems vulnerability, CATENA. 234. <https://doi.org/10.1016/j.catena.2023.107537>
27. Slater LJ, Arnal L, Boucher MA, Chang AYY, Moulds S, Murphy C, Zappa M (2023) Hybrid forecasting: blending climate predictions with AI models. *Hydrol Earth Syst Sci* 27(9):1865–1889
28. Sun Y et al (2007) How often will it rain? Changes in extreme precipitation frequency and intensity under global warming. *Geophys Res Lett* 34:L17706
29. Takhellambam BS, Srivastava P, Lamba J, McGehee RP, Kumar H, Tian D (2023) Projected mid-century rainfall erosivity under climate change over the southeastern United States. *Sci Total Environ* 865:161119. [10.1016/j.scitotenv.2022.161119](https://doi.org/10.1016/j.scitotenv.2022.161119)
30. Trenberth KE (2011) Changes in precipitation with climate change. *Clim Res* 47:123
31. Türkeş M (1999) Vulnerability of Turkey to desertification with respect to precipitation and aridity conditions. *Turk J Eng Environ Sci* 23:363–380
32. Türkeş M (2003) Global climate change and our future climate. *Proceedings of the Panel on Our Future Climate, World Meteorological Day, March 23, 12–37*. General Directorate of Meteorology, Ankara, March 23, 2003
33. Yang D, Kanae S, Oki T, Koike T, Musiake K (2003) Global potential soil erosion with reference to land use and climate changes. *Hydrol Process* 17(14):2913–2928
34. Zhao Y, Zhu D, Wu Z, Cao Z (2024) Extreme rainfall erosivity: Research advances and future perspectives. *Science of the Total Environment*

Figures

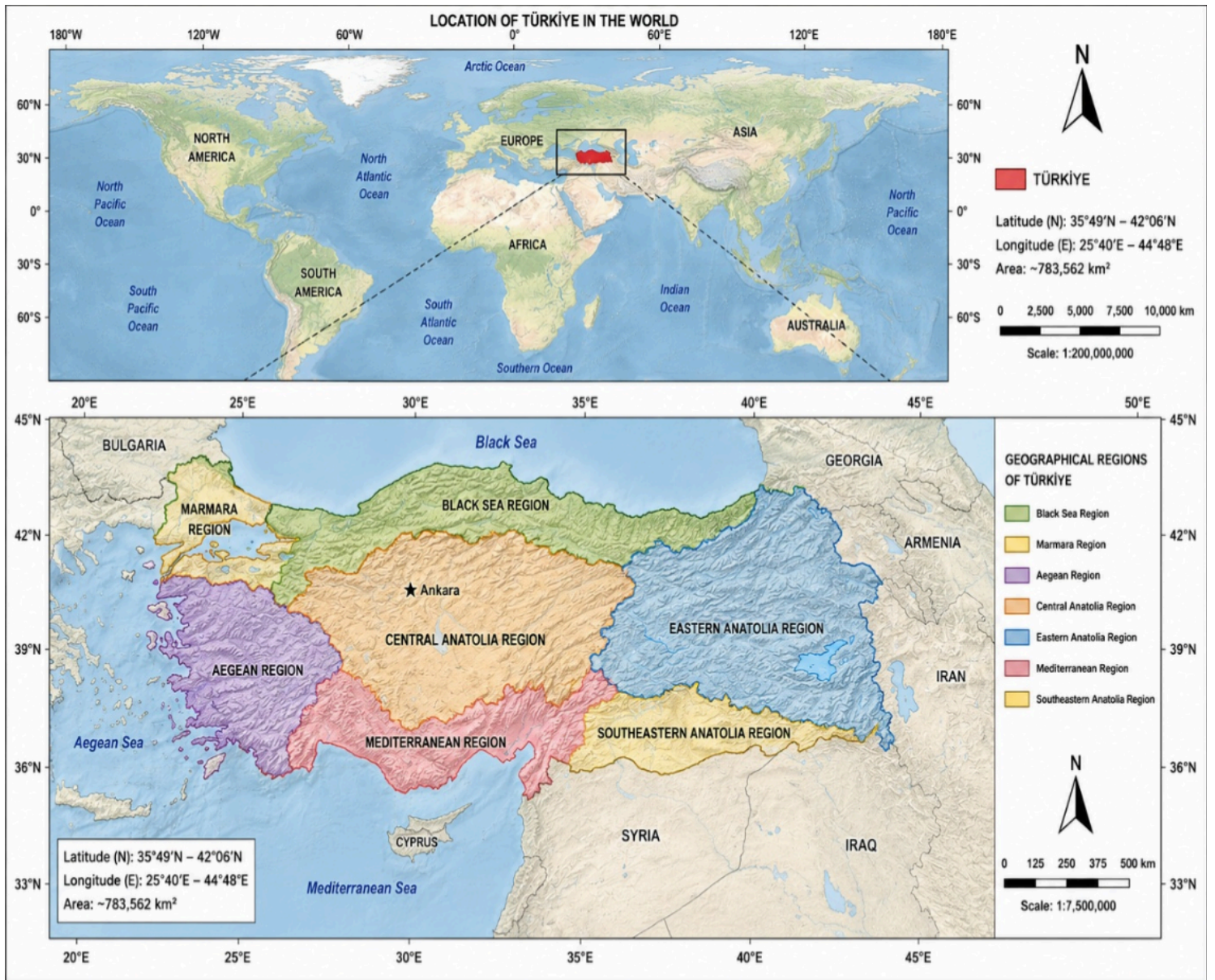


Figure 1

Study Area: Location of Türkiye in the World and Its Geographical Regions

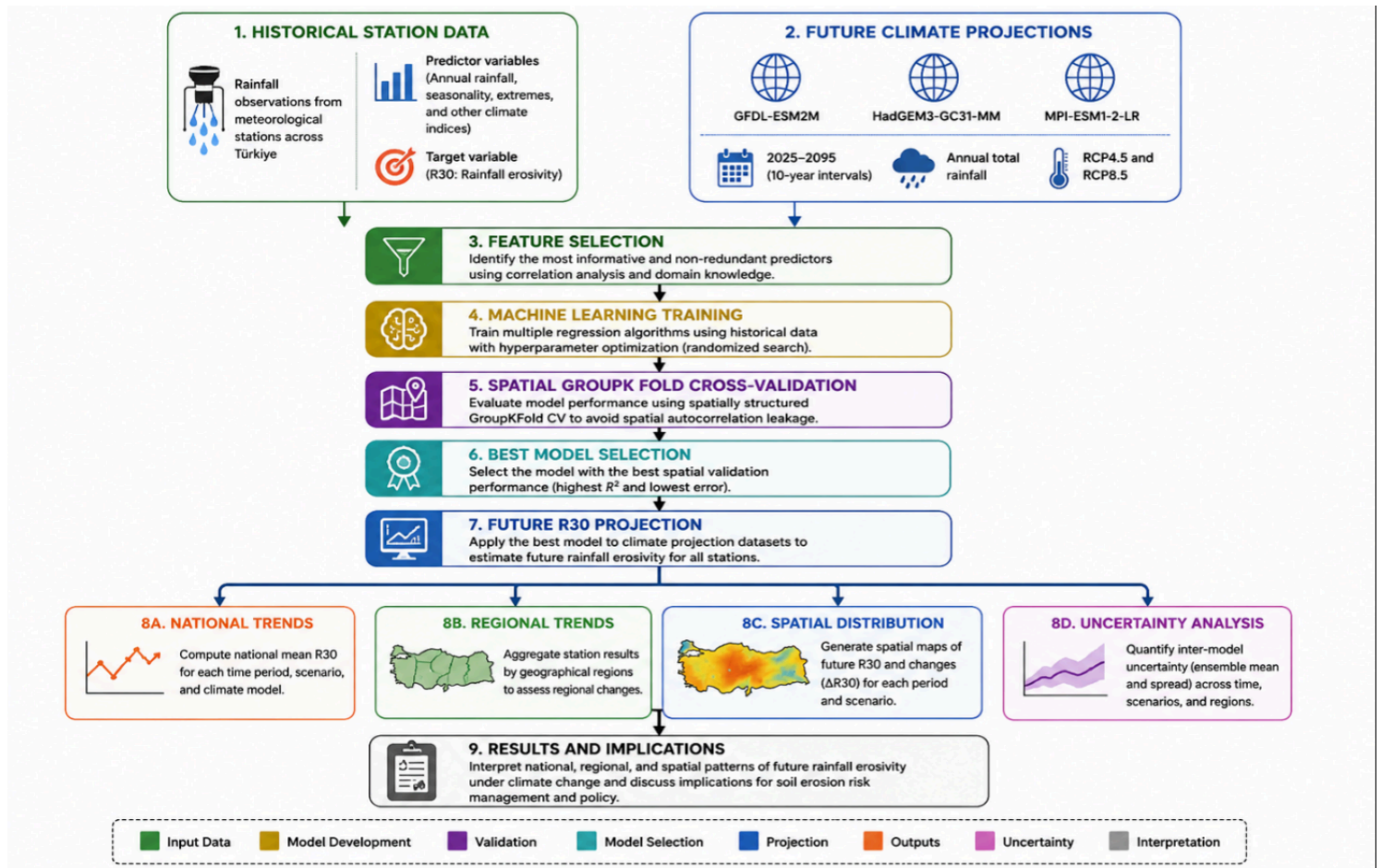


Figure 2

Overall methodological framework for projecting future rainfall erosivity under climate change.

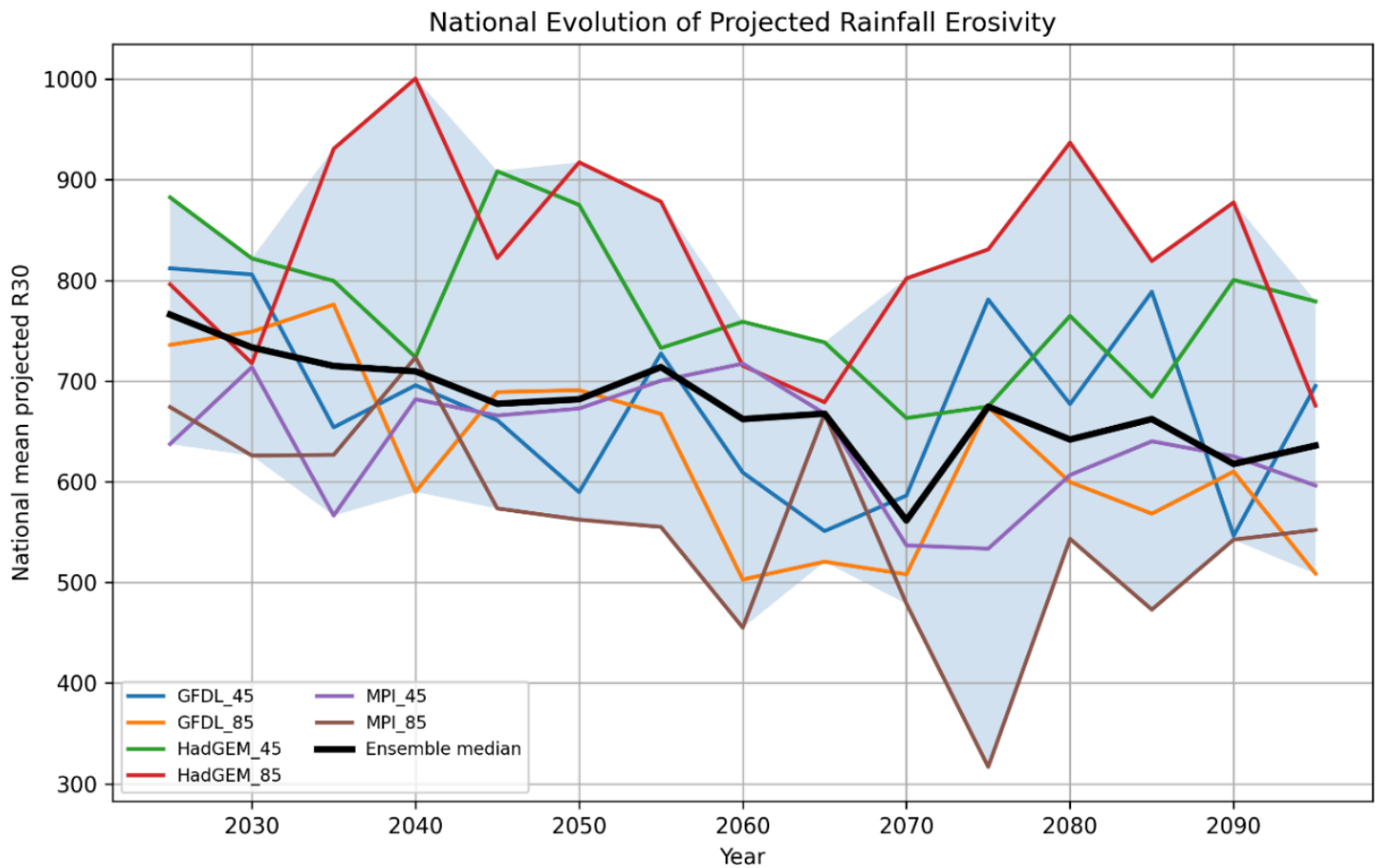


Figure 3

National evolution of projected rainfall erosivity under multiple climate model–scenario combinations (2025–2095)

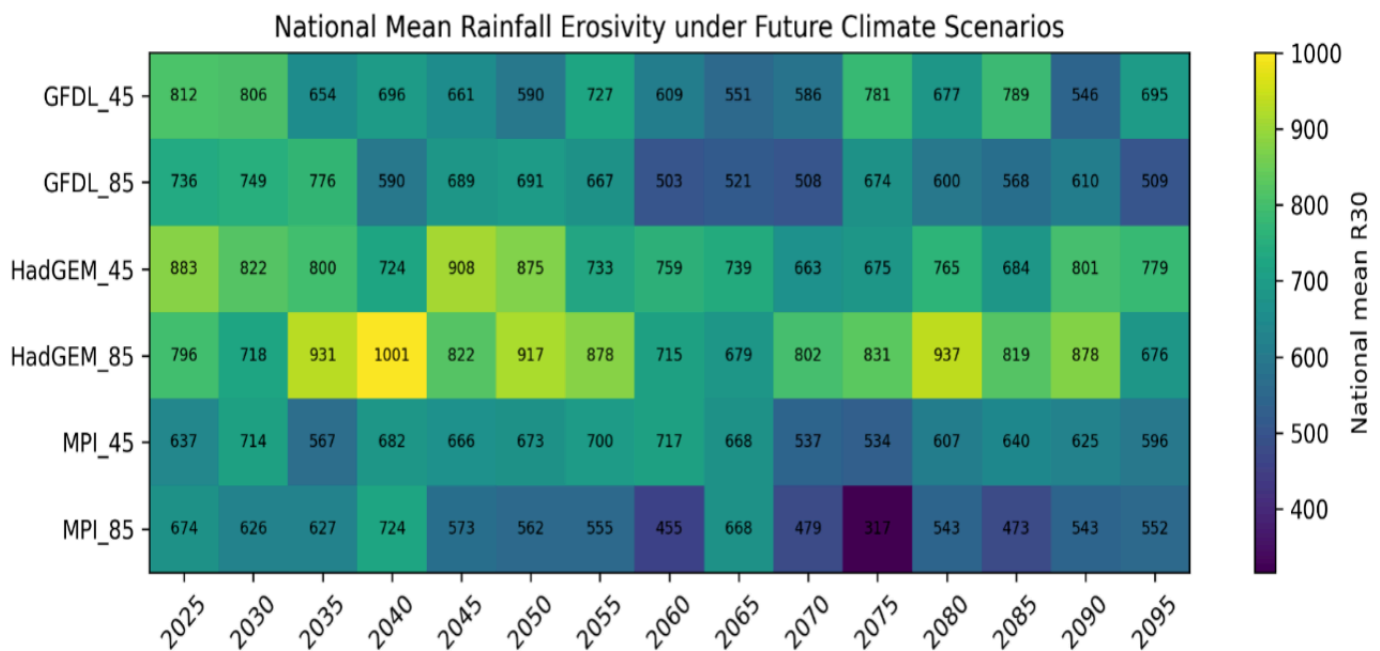


Figure 4

Heatmap comparison of national mean projected rainfall erosivity across climate models and projection years

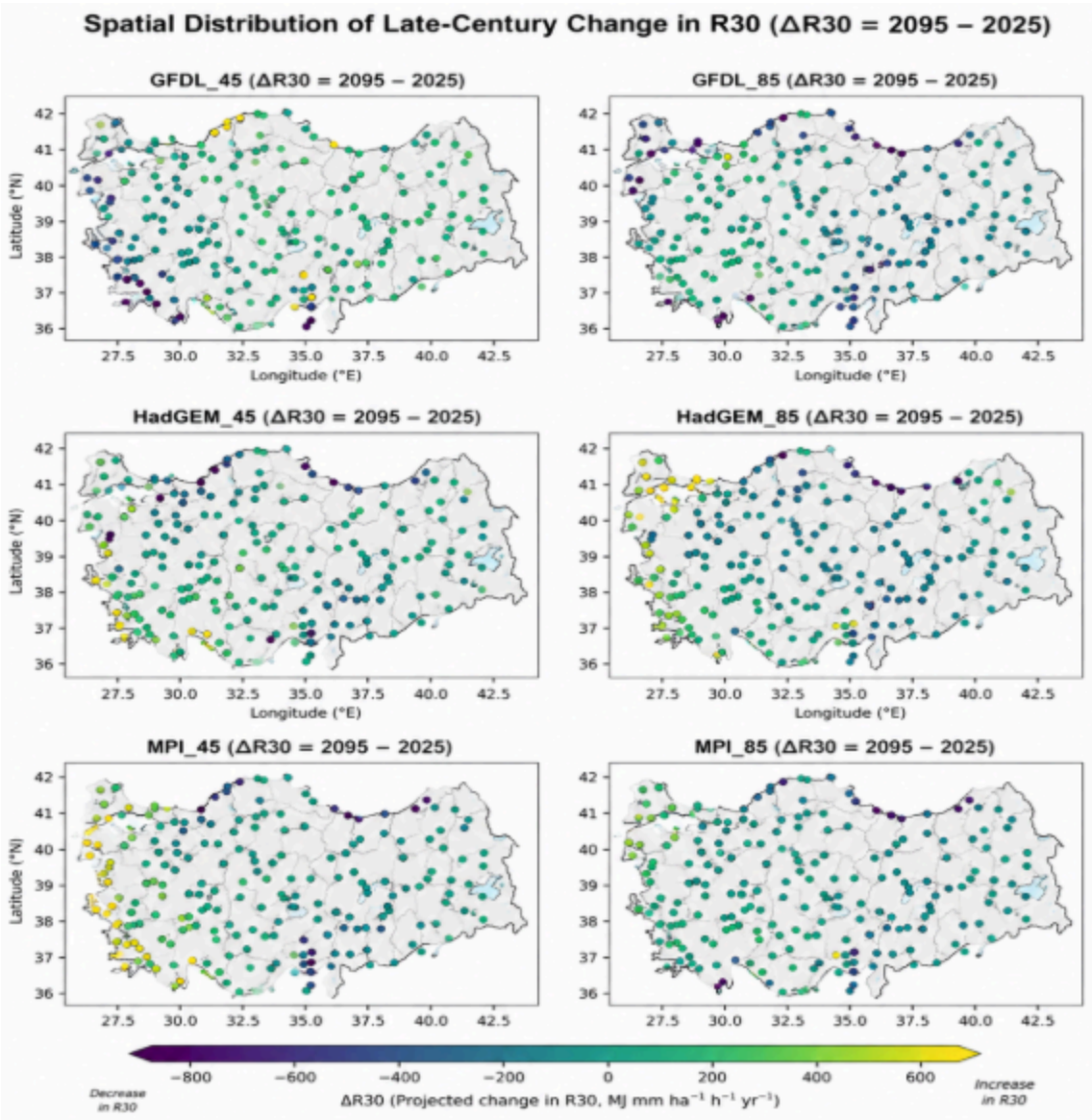


Figure 5

Spatial distribution of projected late-century changes in rainfall erosivity ($\Delta R_{30} = 2095 - 2025$) across Türkiye under six climate model–scenario combinations.

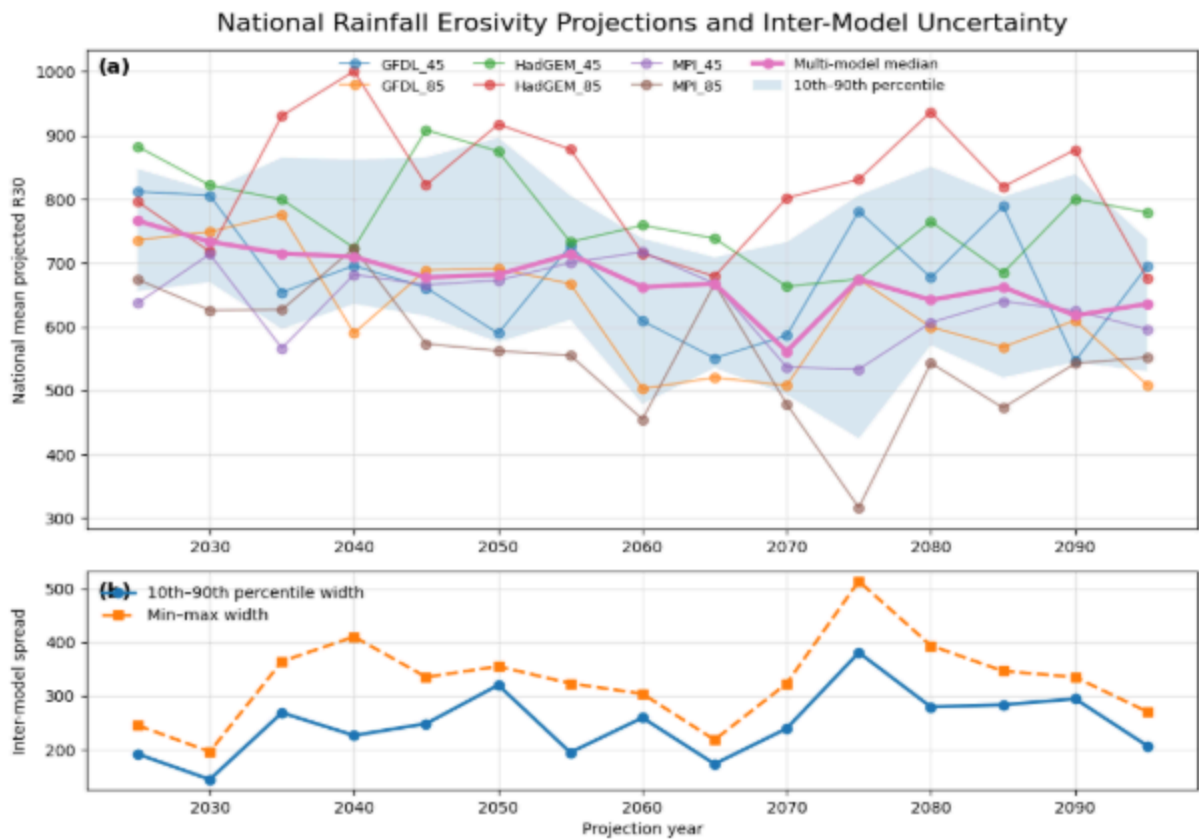


Figure 6

Temporal evolution of national rainfall erosivity projections and associated ensemble uncertainty during 2025–2095

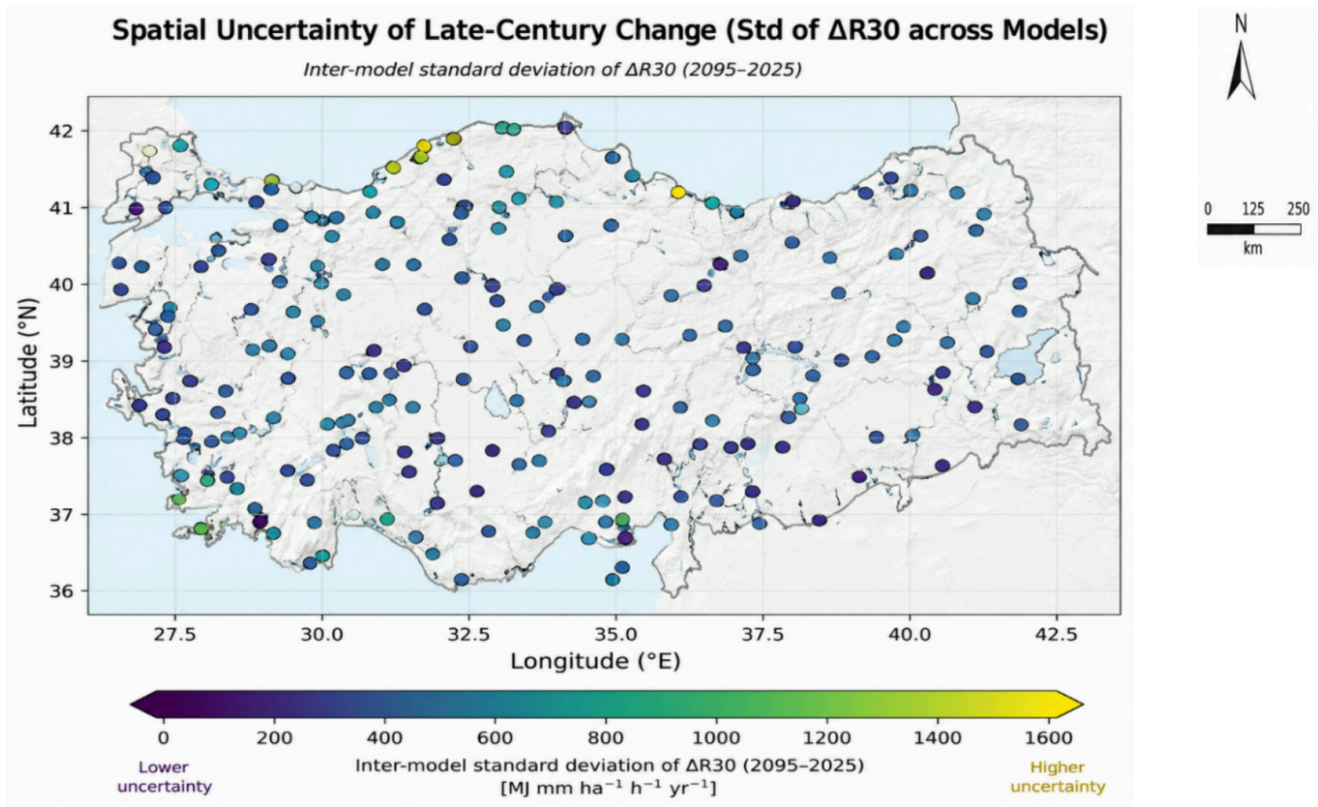


Figure 7

Spatial distribution of inter-model uncertainty in projected late-century rainfall erosivity change across Türkiye