International Journal of Scientific Modern Research and Technology
(Volume: 24, Issue: 03, Number: 05) Paper ID: IJSMRT - 26090305
ISSN: 2582-8150

A Review Paper on Assessing Soil Erosion and Developing Conservation Strategies using Machine Learning in the Part of Upper Narmada Basin

Richa Shrivas
Research Scholar,, Department of Soil and Water Engineering, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India, 482004

Dr. R. N. Shrivastav
Professor, Department of Soil and Water Engineering, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India, 482004

Dr. S. K. Pyasi
Professor and Dean, Department of Soil and Water Engineering, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India, 482004

Dr. S. K. Nema
Scientist, National Institute of Hydrology (NIH), North Western Regional Centre, Jodhpur, India, 342003

Dr. Y. K. Tiwari
Professor, Department of Soil and Water Engineering, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India, 482004

Dr. Namrata Jain
Professor, Department of Agronomy, Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India, 482004

Abstract. Accelerated water erosion is a principal driver of land degradation in the Upper Narmada Basin (UNB) of central India, where steep Maikal and Satpura uplands, erodible black and red soils, an intense four-month monsoon, deforestation, and expanding cultivation combine to remove fertile topsoil and silt up major reservoirs. Climate change is now intensifying rainfall variability and extreme events, yet erosion assessments in the basin remain almost entirely static: they rely on empirical models such as the Revised Universal Soil Loss Equation (RUSLE) that cannot capture non-linear interactions among terrain, soil, land use, and rainfall and that have never been coupled with future climate projections from the Coupled Model Intercomparison Project Phase 6 (CMIP6) under Shared Socioeconomic Pathway (SSP) scenarios. This review synthesizes the literature on conventional, machine learning (ML), and climate-scenario approaches to erosion assessment and frames it around three research objectives for the UNB up to Jabalpur: (i) developing and validating an ML-based baseline erosion map, (ii) predicting spatio-temporal changes in erosion under SSP2-4.5 and SSP5-8.5 for 2041–2060 and 2061–2080, and (iii) formulating a climate-resilient conservation management plan. Tree-based ensembles, notably random forest and gradient boosting, are the most consistently accurate learners for erosion susceptibility mapping, while studies of future erosivity show that projected rainfall changes can substantially raise erosion risk. We identify six gaps: scarce erosion inventories, optimistic validation, limited explainability, weak coupling of maps with conservation decisions, the absence of CMIP6-SSP scenarios in basin-scale erosion modelling, and a lack of downscaled, bias-corrected rainfall erosivity. An integrated methodological framework—data acquisition, factor-map generation, random forest modelling, CMIP6-driven future erosion mapping, hotspot identification, and conservation planning—is proposed as a roadmap for climate-resilient soil conservation in the basin.

Keywords: Soil erosion · Machine learning · random forest · RUSLE · CMIP6 · Shared Socioeconomic Pathways · Erosion hotspots · Climate-resilient conservation · Narmada Basin.

*How to cite this article: Richa Shrivas, Dr. R. N. Shrivastav, Dr. S. K. Pyasi, Dr. S. K. Nema, Dr. Y. K. Tiwari, Dr. Namrata Jain (2026). A Review Paper on Assessing Soil Erosion and Developing Conservation Strategies using Machine Learning in the Part of Upper Narmada Basin, International Journal of Scientific Modern Research and Technology (IJSMRT), ISSN: 2582-8150, Volume-24, Issue-03, Number 05, Sep-2026, pp. **33-44**, URL: https://www.ijsmrt.com/wp-content/uploads/2026/10/IJSMRT-26090305.pdf*

Copyright © 2026 by author (s) and International Journal of Scientific Modern Research and Technology Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0)

IJSMRT-26090305

I. INTRODUCTION

Soil is a finite and effectively non-renewable resource; once lost, its natural replacement takes centuries. Soil erosion by water is among the most widespread forms of land degradation worldwide, and global modelling indicates that land-use change in the early twenty-first century has measurably increased erosion rates, with agricultural expansion in the tropics a major contributor (Borrelli et al., 2017). In India, erosion is the dominant soil-degradation process, reducing crop productivity, depleting nutrients and organic carbon, and degrading downstream water bodies (Bhattacharyya et al., 2015). The country has been estimated to lose about 5,334 million tonnes of soil annually (Narayana & Babu, 1983), and national sediment budgets show that a substantial share of this material is transported into rivers and stored in reservoirs, shortening their design life and reducing irrigation and hydropower capacity (Sharda & Ojasvi, 2016).

The Narmada, the largest west-flowing river of peninsular India, drains much of Madhya Pradesh. Its upper reach, the Upper Narmada Basin (UNB), extends from the source at Amarkantak through the Dindori and Mandla districts to the Jabalpur gauging station, and is a critical agricultural and hydrological zone that supports millions of farmers and feeds the Bargi reservoir. The basin combines forested and increasingly cultivated uplands with erodible vertisols and red–yellow soils and receives most of its annual rainfall in a few monsoon months. Studies of its watersheds have repeatedly highlighted erosion, land degradation and sediment delivery to reservoirs as management concerns (Meshram et al., 2021, 2022).

Climate change adds to further, time-varying pressure. Rainfall erosivity scales strongly with rainfall intensity, so projected increases in heavy-rainfall events translate into disproportionate increases in erosive energy (Nearing, 2001; Panagos et al., 2022). Global assessments driven by Shared Socioeconomic Pathway (SSP) scenarios project that climate change could increase water erosion substantially by 2070 (Borrelli et al., 2020), and hydrological studies of the upper Narmada project higher monsoon flows under future climate (Pandey et al., 2021; Rickards et al., 2020). Yet the erosion literature for the basin remains almost entirely static.

Erosion assessment in India has long relied on the empirical Universal Soil Loss Equation (USLE) (Wischmeier & Smith, 1978) and its revised form, RUSLE (Renard et al., 1996), implemented in a geographic information system (GIS) with remote-sensing inputs (Ganasri & Ramesh, 2015). These models are transparent and data-efficient, but they predict long-term average sheet and rill erosion only, rely on factor values transferred from other environments, cannot capture non-linear spatial patterns, and do not integrate future climate signals. Over the past decade, machine learning (ML) has emerged as a data-driven, spatially explicit complement: learners such as random forest (RF), support vector machines (SVM) and boosted trees map erosion susceptibility with high accuracy from readily available geospatial predictors (Arabameri et al., 2019; Gayen et al., 2019). Combining such learners with CMIP6 climate projections offers a route to forward-looking, climate-resilient conservation planning, but this combination has not yet been attempted for the UNB. This review examines the evidence needed to build it.

II. PROBLEM STATEMENT

Soil erosion in the Upper Narmada Basin is driven by three compounding forces: accelerating rainfall variability under climate change, widespread deforestation, and poor land-use management. Several studies have assessed erosion in the region using GIS, morphometric analysis and empirical models (RUSLE/USLE) (Meshram et al., 2021, 2022; Sharma et al., 2014), but none has integrated future climate projections from CMIP6 under SSP scenarios into erosion-risk modelling. Existing studies also rely on static, empirical equations that cannot capture the complex non-linear interactions among terrain, soil, land use and rainfall, and their outputs are rarely translated into prioritized, scenario-tested conservation plans.

The research gap is therefore the absence of a climate-resilient conservation management plan for the UNB that is founded on ML-based erosion modelling and CMIP6-SSP projections. Closing this gap requires (a) a validated, data-driven baseline of present-day erosion risk, (b) projections of how that risk will shift in space and time under plausible emission pathways, and (c) a transparent procedure for converting persistent high-risk areas into prioritized conservation interventions.

III. OBJECTIVES

Within this problem setting, the proposed research pursues three objectives, and the review is organized to provide the methodological basis for each:

- To develop, validate and simulate baseline soil erosion in the upper part Upper Narmada Basin
- To predict spatio-temporal changes in soil erosion under different climate scenarios.
- To formulate a climate resilient management plan for conserving soil erosion.

Accordingly, this review aims to (i) summarize the conventional and ML-based approaches used for erosion assessment and evaluate their strengths and limitations in data, validation and interpretability (Objective 1); (ii) review how CMIP6 and SSP scenarios have been used to project rainfall erosivity and future erosion (Objective 2); (iii) review erosion-related and conservation-prioritization studies in the Narmada basin (Objective 3); and (iv) propose an integrated methodological framework linking the three objectives.

IV. REVIEW METHODOLOGY

A structured narrative review was conducted. Literature was identified from Scopus, Web of Science, Google Scholar and the publishers' portals of Springer, Elsevier and MDPI using combinations of the terms "soil erosion," "gully erosion," "sediment yield," "susceptibility," "machine learning," "random forest," "RUSLE," "rainfall erosivity," "climate change," "CMIP6," "SSP," "Narmada," and "Madhya Pradesh." Peer-reviewed journal articles, conference papers and authoritative technical handbooks in English were considered. Studies were retained if they (a) applied ML, deep learning or a hybrid data-driven model to soil erosion, gully erosion or sediment yield; (b) projected rainfall erosivity, erosion or basin hydrology under climate-change scenarios; or (c) assessed erosion, land degradation or watershed priority within the Narmada basin by any method. Foundational sources on algorithms, climate-model ensembles and datasets were also included. Evidence was synthesized thematically and mapped to the three research objectives.

V. STUDY AREA: THE UPPER NARMADA BASIN UP TO JABALPUR

The study area is the headwater portion of the Narmada system in eastern Madhya Pradesh, from the source on the Maikal plateau at Amarkantak to the confluence of the Narmada with the Burhner at Manot, which forms the basin's outlet (Fig. 1). This headwater catchment covers principally the Dindori and Mandla districts; Jabalpur lies further downstream, outside the delineated basin, and is shown in Fig. 1 only as a nearby reference town. The river flows westward through a structurally controlled valley between the ranges to the south. Principal tributaries in this reach include the Burhner, Banjar and Halon. Downstream of the study area outlet, the Narmada continues through the Bargi reservoir before reaching Jabalpur. Geology is dominated by Deccan Trap basalts with Gondwana and Vindhyan sedimentary formations, giving rise to deep black vertisols in the valley and shallower red and lateritic soils on the uplands. The climate is sub-humid, with the Vindhyan scarps to the north and the Satpura ranges to the south, and the southwest monsoon (June–September) supplying most of the annual rainfall, often in high-intensity events that generate erosive runoff.

 ISSN: 2582-8150

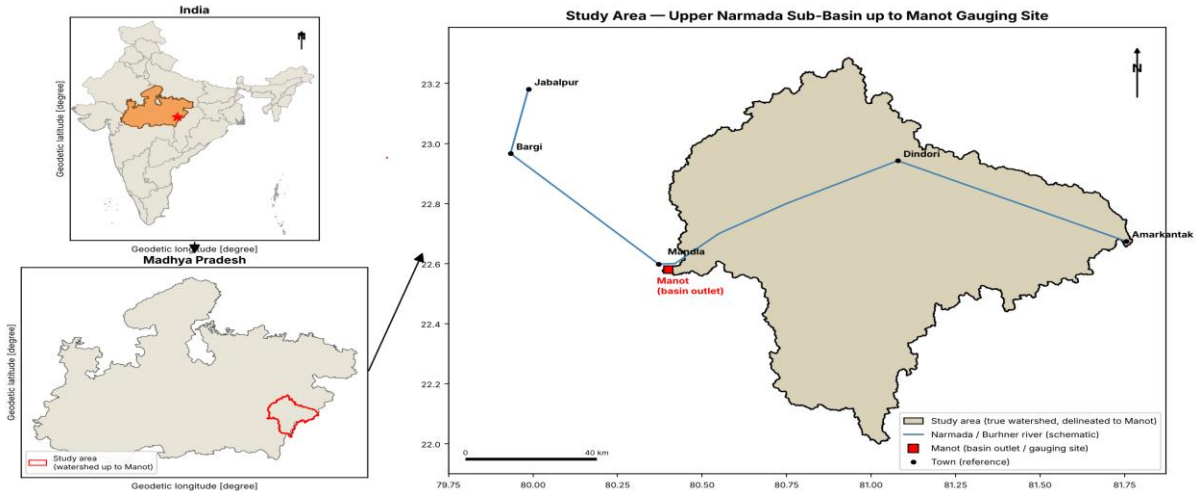

Fig. 1. Location of the study area: the Upper Narmada Basin up to Jabalpur, Madhya Pradesh, India.

Three characteristics make the UNB a demanding setting for erosion modelling. First, steep forested headwaters adjoin intensively cropped plains, so erosion drivers vary sharply over short distances. Second, rapid land-use changes, including deforestation and forest-to-agriculture conversion on slopes, alters cover and management conditions over time. Third, field observations, gully inventories, plot measurements and suspended-sediment records are sparse relative to the basin's size. Remote sensing and freely available datasets (Table 1) therefore form the backbone of most assessments; for the proposed research these are assembled for 1995–2025 to define the baseline period.

Table 1. Principal datasets for baseline and future erosion modelling in the UNB.

Factor group	Dataset	Use in erosion models
Rainfall	IMD 0.25° gridded daily rainfall (Pai et al., 2014); global erosivity maps (Panagos et al., 2017)	Rainfall erosivity (R factor), event intensity
Topography	ALOS World 3D 30 m DEM (Tadono et al., 2014); SRTM 30 m (Farr et al., 2007)	Slope, LS factor, curvature, drainage density
Soil	NBSS&LUP soil maps; SoilGrids 250 m (Hengl et al., 2017)	Texture, organic carbon, erodibility (K factor)
Land cover	Landsat, Sentinel-2 via Google Earth Engine (Gorelick et al., 2017)	LULC, NDVI, cover–management (C factor)
Conservation	Field and agency records of practices	Support-practice (P factor)
Observations	CWC drainage-area and sediment data; geotagged erosion points	Training labels and validation for ML models
Future climate	CMIP6 GCM ensemble (Eyring et al., 2016); NEX-GDDP-CMIP6 (Thrasher et al., 2022)	Future rainfall and R factor under SSP2-4.5 and SSP5-8.5

VI. CONVENTIONAL EROSION MODELS AND THEIR LIMITS

RUSLE estimates mean annual soil loss A (t ha^{-1} yr^{-1}) as the product $A = R{\cdot}K{\cdot}LS{\cdot}C{\cdot}P$, where R is rainfall erosivity, K soil erodibility, LS the slope length–steepness factor, C cover–management and P support practice (Renard et al., 1996). Its GIS implementation is the standard method in Indian watershed studies; a representative application to the Nethravathi basin derived all factors from remote sensing and GIS layers to locate erosion-prone zones (Ganasri & Ramesh, 2015), and USLE–GIS analysis has been used to identify critical erosion-prone areas of small agricultural watersheds for priority treatment (Pandey et al., 2006). Where only monthly rainfall is available,

R is commonly estimated from the Modified Fournier Index (Arnoldus, 1980).

In the Narmada system, Sharma et al. (Sharma et al., 2014) combined satellite-derived slope, drainage, soil, land-use and morphometric layers by GIS overlay to map erosion-prone zones of the Barchha Nala watershed, finding the highest risk on steep, sparsely vegetated slopes, but did not consider temporal change. Morphometric analysis and sediment-yield indices have also been widely used to rank watersheds; Meshram et al. [5] developed morphometric deterministic models from SRTM-derived parameters of 49 UNB watersheds and obtained coefficients of determination above 0.70 for the sediment yield index.

Process-based models such as the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998) simulate runoff and sediment transport continuously, but their calibration demands long hydrometric records that are available at only a few UNB gauging stations. Empirical and process-based models share a structural limitation: they encode predefined functional relationships. They cannot exploit additional predictors, struggle to capture threshold behaviour, offer no direct means of learning from observed erosion features, and—as applied in the UNB—have been run for present-day conditions only. These limitations motivate the data-driven approaches reviewed next (Objective 1) and the climate-scenario approaches reviewed in Sect. 8 (Objective 2).

VII. MACHINE LEARNING APPROACHES

ML applications to erosion fall into four families (Fig. 2): susceptibility classification of erosion presence or absence, regression of soil loss or sediment yield, deep learning on large tabular or image data, and hybrid systems that combine learners with optimization, empirical models or decision analysis. Table 2 summarizes the algorithms, and Table 3 lists representative studies.

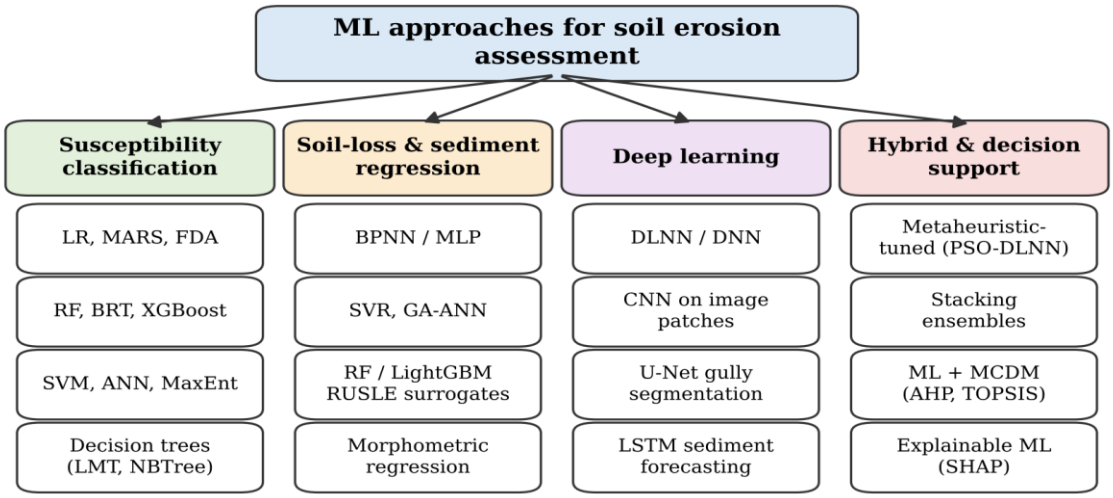

Fig. 2. Taxonomy of machine learning approaches applied to soil erosion assessment.

Tree-Based Ensembles

Random forest (Breiman, 2001) aggregates many decorrelated decision trees and provides out-of-bag error estimates and variable importance with minimal tuning. It is the most frequently reported best performer in erosion susceptibility mapping, which is why it is adopted as the core learner for the baseline model (Objective 1). In the Pathro River basin of eastern India, Gayen et al. [15] compared multivariate adaptive regression splines (MARS), flexible discriminant analysis (FDA), RF and SVM using 174 field-surveyed gullies and found RF most accurate in terms of area under the receiver operating characteristic curve (AUC). Saha et al. (Saha et al., 2020) obtained comparable results for the Hinglo River basin, where tree-based learners identified roughly 7% of the basin as highly susceptible. Gradient boosting, particularly XGBoost (Chen & Guestrin, 2016), adds regularization and handles missing values efficiently. Decision-tree variants can, however, produce divergent maps despite similar accuracy statistics: Arabameri et al. [16]

showed that only the logistic model tree yielded geomorphologically plausible patterns among three tree models with similar validation scores, underlining the need to inspect maps rather than metrics alone.

Kernel Methods and Neural Networks

Support vector machines (Cortes & Vapnik, 1995) construct maximum-margin decision boundaries in a transformed feature space and perform well with small, high-dimensional samples, a common situation for erosion inventories. Artificial neural networks (ANNs), typically multilayer perceptrons trained by back-propagation, have long been applied to suspended-sediment and sediment-yield prediction. In susceptibility studies, ANNs and RF were trained with four-fold cross-validation after multicollinearity screening by Ghorbanzadeh et al. (Ghorbanzadeh et al., 2020), demonstrating that careful factor selection matters as much as algorithm choice. Neural models are sensitive to architecture and hyperparameters, and their results are harder to interpret than those of tree ensembles.

Deep Learning

Deep learning extends neural networks to many layers and can learn features directly from raw inputs. Band et al. (Band et al., 2020) coupled a deep learning neural network (DLNN) with particle swarm optimization and reported improved gully susceptibility prediction over conventional ANN and SVM models. Encoder–decoder architectures such as U-Net (Ronneberger et al., 2015) enable semantic segmentation of gullies and ravines from high-resolution imagery—an attractive route to building the erosion inventories the UNB currently lacks—and long short-term memory (LSTM) networks (Hochreiter & Schmidhuber, 1997) can forecast sediment loads from rainfall–runoff sequences. Deep models, however, require large labelled datasets and substantial computing, and they risk overfitting in data-scarce basins.

Hybrid and Ensemble Frameworks

Hybrid frameworks aim to combine complementary strengths. Arabameri et al. (Arabameri et al., 2020) evaluated multi-criteria decision-making, statistical and ML models and their ensembles for gully susceptibility mapping and concluded that combined models generally improved performance. Three hybrid patterns are particularly relevant to the UNB: (i) metaheuristic tuning of learner hyperparameters; (ii) stacking, in which predictions of several base learners feed a meta-learner; and (iii) RUSLE–ML coupling, in which RUSLE factor maps serve as predictors (or RUSLE soil-loss maps as targets) so that the ML model can incorporate additional predictors and be re-run rapidly with modified factor layers. The third pattern is central to Objective 2, because it allows a trained model to be driven with future *R*-factor surfaces derived from climate projections.

A recent demonstration of this third pattern comes from the Chota Nagpur Plateau of eastern India, where Tiwari et al. (Tiwari et al., 2026) used RUSLE-derived soil-loss maps as the machine-learning target across seven predictor scenarios evaluated with multiple linear regression, artificial neural networks (ANN) and random forest; LASSO feature selection combined with variance-inflation-factor screening reduced the predictor set to six layers (LS, SOC, C, P, NDVI and K), and the ANN model on this reduced set achieved the highest accuracy ($R^2 = 0.984 \pm 0.006$, RMSE = 2.35 ± 0.43 t ha^{-1} yr^{-1}). This result illustrates that a parsimonious RUSLE–ML framework can support reliable soil erosion prediction and conservation prioritization in data-scarce, monsoon-dominated Indian watersheds broadly comparable in setting to the UNB.

Table 2. Comparison of ML algorithm families for soil erosion assessment.

Family	Strengths	Limitations
LR, MARS, FDA	Simple, interpretable coefficients; low data demand	Limited to near-linear or additive effects
RF, BRT, XGBoost	High accuracy; robust to noise and mixed data types; built-in importance	Importance can be biased toward correlated factors; extrapolates poorly
SVM	Effective with small samples and many predictors	Kernel and parameter tuning; probability output is indirect
ANN / DLNN	Captures complex non-linearity; flexible inputs	Needs larger data and tuning; weak interpretability

CNN, U-Net, LSTM	Learns spatial or temporal features; can map gullies from imagery	Large labelled datasets and GPUs required
Hybrid / stacked	Often best accuracy; can embed physical (RUSLE) factors	Complexity, risk of overfitting, reproducibility issues

Table 3. Representative methods relevant to ML-based erosion assessment.

Study	Region	Method / data	Key finding
Gayen et al. [15]	Pathro basin, India	MARS, FDA, RF, SVM; 174 gullies, 12 factors	RF most accurate; all models performed well
Saha et al. (Saha et al., 2020)	Hinglo basin, India	RF, GBRT, NBTree, tree ensemble; 120 gullies	≈7% of basin highly susceptible
Arabameri et al. [16]	Bastam, Iran	ADTree, NBTree, LMT with index of entropy	Similar metrics but only LMT map plausible
Ghorbanzadeh et al. (Ghorbanzadeh et al., 2020)	Lamerd, Iran	ANN, RF; VIF screening; 4-fold CV; 232 gullies	Factor selection improves robustness
Band et al. (Band et al., 2020)	Shirahan, Iran	DLNN, PSO-DLNN vs ANN, SVM; 132 gullies	Optimized deep model outperformed baselines
Meshram et al. [5]	Upper Narmada, India	Morphometric models, PCA; 49 watersheds	Sediment yield index predicted with $R^2 > 0.70$
Meshram et al. [6]	Narmada, India	Game-theory prioritization (Condorcet, fallback bargaining)	Mohgaon watershed ranked most vulnerable
Tiwari et al. (Tiwari et al., 2026)	Chota Nagpur Plateau, India	RUSLE + MLR/ANN/RF; LASSO + VIF feature selection; 7 predictor scenarios	Reduced 6-predictor ANN model best overall (R^2 = 0.984)

Explainability

A recurring criticism of ML is opacity. RF variable importance offers a first, global view of which factors drive predictions, and model-agnostic methods, especially S Hapley Additive ex Planations (SHAP) (Lundberg & Lee, 2017), attribute each prediction to its input factors in a consistent, additive way. In erosion studies these tools reveal which factors—slope, rainfall erosivity, vegetation cover, land use—drive susceptibility at basin and pixel scales, and whether their effects show thresholds. For the UNB, explainability is not merely academic: it is the step that links hotspots to their dominant drivers (Objective 3), so that conservation agencies know *why* a sub-watershed is ranked critical and can select appropriate measures.

VIII. CLIMATE CHANGE, CMIP6 AND SSP SCENARIOS IN EROSION MODELLING

The Coupled Model Intercomparison Project Phase 6 (CMIP6), coordinated by the World Climate Research Programme, provides standardized simulations from multiple global climate models (GCMs), with improved representation of extreme rainfall relative to earlier phases (Eyring et al., 2016). Its ScenarioMIP experiments are driven by Shared Socioeconomic Pathways (SSPs), which describe how society, economy and emissions may evolve to 2100 and underpin the IPCC Sixth Assessment Report (O'Neill et al., 2016; Riahi et al., 2016). Three pathways are most commonly used: SSP1-2.6 (low emissions, sustainable development), SSP2-4.5 (intermediate) and SSP5-8.5 (high emissions). Because raw GCM output is too coarse and biased for basin studies, statistically downscaled and bias-corrected products such as NASA NEX-GDDP-CMIP6 at 0.25° are typically used (Thrasher et al., 2022).

Rainfall erosivity is the channel through which climate change enters erosion models. Early scenario studies showed that warming-driven increases in rainfall and intensity raise erosivity and hence soil loss [7]. At the global scale, Panagos et al. [8] projected rainfall erosivity for 2050 and 2070 from CMIP-based climate

 ISSN: 2582-8150

data and found increases over most regions, and Borrelli et al. (Borrelli et al., 2020) combined SSP-based climate and land-use scenarios with RUSLE-type modelling to show that climate change is likely to increase water erosion substantially by 2070. Using CMIP6 bias-corrected projections, Shrestha et al. (Shrestha et al., 2021) reported increases in extreme precipitation and runoff in Asian river basins under high-emission SSPs, while noting that basin-scale downscaling remains necessary.

For Narmada, climate-scenario work has so far concentrated on hydrology. Rickards et al. [10] projected long-term increases in monsoon-season flow in the upper Narmada for 2031–2060 from a CMIP5 ensemble, and Pandey et al. [11] showed that land-use dynamics and climate change jointly alter the basin's hydrological response, calling for adaptive watershed management. No study has yet propagated CMIP6-SSP rainfall into erosion-risk maps for the UNB (Table 4).

Table 4. Climate-scenario and Narmada-basin studies relevant to the research objectives.

Study	Region / method	Key finding	Limitation
Nearing [7]	USA; erosivity under climate scenarios	Warming increases erosivity and soil loss	Region-specific; earlier-generation models
Borrelli et al. (Borrelli et al., 2020)	Global; RUSLE-based model with SSP scenarios	Climate change likely to increase water erosion by 2070	Coarse resolution; sheet and rill erosion only
Panagos et al. [8]	Global; projected R factor for 2050, 2070	Erosivity increases over most regions	Global scale; needs local validation
Shrestha et al. (Shrestha et al., 2021)	Asian basins; CMIP6 bias-corrected + hydrological model	More extreme precipitation and runoff under high-emission SSPs	Basin-scale downscaling needed
Rickards et al. [10]	Upper Narmada; GWAVA with CMIP5 ensemble	Higher monsoon flows for 2031–2060	Hydrology only; no erosion
Pandey et al. [11]	Upper Narmada; hydrological modelling with climate and LULC scenarios	Climate and land-use change significantly affect hydrology	Focus on hydrology rather than conservation planning
Sharma et al. (Sharma et al., 2014)	Barchha Nala, Narmada; RS–GIS overlay	High risk on steep, low-vegetation slopes	Temporal change not included

IX. VALIDATION PRACTICE

Most susceptibility studies split erosion and non-erosion points randomly, commonly 70:30, and report AUC, accuracy, confusion-matrix statistics or Cohen's kappa; regression studies report RMSE, R^2 and Nash–Sutcliffe efficiency. Random splits are convenient and comparable across studies, but neighboring samples are spatially autocorrelated, so test points are not fully independent of training points and accuracy can be overestimated. Spatially blocked cross-validation, which holds out contiguous regions, gives a more honest estimate of transferability (Roberts et al., 2016). Inventories also frequently contain sampling bias toward accessible sites near roads and settlements. For the UNB, we therefore recommend the conventional 70:30 split with confusion matrix and ROC–AUC as the primary evaluation, supplemented by sub-watershed-wise (spatial) cross-validation, comparison of predicted high-risk zones with field GPS points and historical satellite imagery, and, where possible, checks against CWC sediment records. For future scenarios, which cannot be validated directly, uncertainty should be expressed through the spread across GCMs and SSPs.

X. PROPOSED METHODOLOGICAL FRAMEWORK

Drawing on the evidence reviewed, Fig. 3 presents an integrated framework that operates the three research objectives. Outputs of each stage feed the next: the trained baseline model is reused to generate future maps, the baseline map is the reference for change

detection, and persistent high-risk areas become the targets of the management plan.

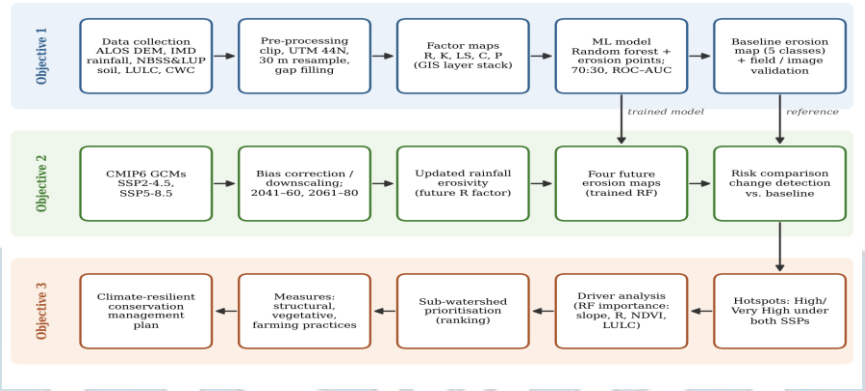

Fig. 3. Flow diagram of the proposed methodology for baseline and future soil-erosion assessment and climate-resilient conservation planning in the Upper Narmada Basin.

Objective 1: Baseline Erosion Modeling

Data for 1995–2024 are assembled as in Table 1: ALOS DEM for slope and *LS*, IMD rainfall for *R*, NBSS&LUP soil data for *K*, LULC maps for *C*, conservation-practice data for *P*, and CWC drainage-area and sediment data. Pre-processing clips all layers to the basin boundary up to the Jabalpur gauging station, reprojects them to UTM Zone 44N (WGS 1984), resamples them to 30 m × 30 m (bilinear interpolation for continuous layers), aggregates monthly rainfall to an annual *R* factor using the Modified Fournier Index or the Wischmeier equation [12, 27], and fills gaps by nearest-neighbor or temporal interpolation. The resulting factor maps (*R*, *K*, *LS*, *C*, *P*) are stacked with geotagged erosion and non-erosion points; the dataset is split 70:30 for training and testing, and an RF model is trained and evaluated with the confusion matrix and ROC–AUC. Applying the trained model to the full raster stack with present-day rainfall and current LULC yields the baseline erosion map, classified into five zones (Very Low, Low, Moderate, High, Very High) with areal statistics for each class and validated against field GPS points and historical imagery.

Objective 2: Future Erosion under CMIP6-SSP Scenarios

Bias-corrected, downscaled CMIP6 rainfall for SSP2-4.5 and SSP5-8.5 is used to compute future *R*-factor surfaces for two periods, 2041–2060 and 2061–2080, while the other factors are held at baseline values so that the climate signal is isolated. Running the trained RF model with each future *R* surface produces four future erosion maps (Table 5). Comparison then quantifies (i) the percentage of basin area shifting from lower to higher risk classes relative to the baseline, (ii) areas that are consistently high-risk across all four scenarios, and (iii) a ranking of sub-watersheds by projected increase in erosion. Using several GCMs and reporting their spread is recommended to communicate projection uncertainty.

Table 5. Future scenario matrix for erosion mapping.

Emission pathway	Period	Output
SSP2-4.5 (intermediate)	2040–2060	Future erosion map 1
SSP2-4.5 (intermediate)	2060–2080	Future erosion map 2
SSP5-8.5 (high)	2040–2060	Future erosion map 3
SSP5-8.5 (high)	2060–2080	Future erosion map 4

Objective 3: Hotspots and Climate-Resilient Management Plan

Erosion risk is classified by predicted soil loss into severity classes commonly used in Indian studies: Very

Low (< 5 t ha^{-1} yr^{-1}), Low (5–10), Moderate (10–20), High (20–40) and Very High (> 40). Sub-watersheds that fall in the High or Very High class under both SSP2-4.5 and SSP5-8.5 are designated *climate-sensitive hotspots*. RF importance (and SHAP, where applied) identifies the dominant controlling factors in each hotspot—typically steep slopes (> 15°), high rainfall erosivity, low vegetation cover (NDVI < 0.3) and degraded land use—and sub-watersheds are ranked for intervention, building on existing prioritization approaches [6]. Hotspot maps are overlaid on administrative boundaries for planning. Measures are then matched to drivers: structural measures (contour and graded bunds, terracing, check dams, gully plugs) where slope and drainage dominate; vegetative measures (afforestation, grassed waterways, vegetative barriers) where cover is degraded; and sustainable farming practices (contour cultivation, cover crops, conservation tillage, residue retention) on cultivated slopes. Because the hotspots are defined under future climate, the resulting plan is climate-resilient by design.

Expected Outputs

- A validated ML–GIS framework for spatial erosion susceptibility mapping in the UNB.
- A baseline erosion risk map (1995–2025) with sub-watershed-level statistics.
- Four future erosion risk maps under SSP2-4.5 and SSP5-8.5 for 2040–2060 and 2060–2080.
- Identification of climate-sensitive erosion hotspots and priority conservation areas.
- A climate-resilient conservation strategy for watershed planners.

XI. RESEARCH GAPS AND FUTURE DIRECTIONS

The review identifies the following gaps for the UNB and analogous central Indian basins, each of which the proposed framework addresses:

- No climate-scenario erosion modelling: CMIP6-SSP projections have not been integrated into erosion-risk modelling for the UNB; coupling trained ML models with future *R* factors fills this gap (Objective 2).
- Static, empirical assessments: Existing studies rely on fixed equations and present-day conditions; ML models capture non-linear factor interactions and can be re-run for scenarios (Objective 1).
- Downscaling and erosivity uncertainty Basin-scale, bias-corrected rainfall and erosivity projections are scarce; multi-GCM ensembles and explicit uncertainty reporting are needed.
- Inventory scarcity and optimistic validation: No public basin-wide erosion inventory exists; field GPS surveys, image-based mapping, and spatial cross-validation should accompany the standard 70:30 evaluation.
- Limited interpretability: Few studies explain predictions; RF importance and SHAP should accompany every model intended for planning.
- Weak decision coupling: Susceptibility maps are rarely translated into ranked, climate-tested conservation plans; hotspot persistence across scenarios offers a transparent basis for prioritization (Objective 3).

Future work should also incorporate projected land-use change alongside climate change, share code and data openly, and foster collaboration among universities, the soil survey and state watershed agencies.

XII. CONCLUSION

Machine learning has matured from an experimental alternative into a credible complement to empirical erosion models, and tree-based ensembles, particularly random forests, deliver consistently strong susceptibility predictions with built-in measures of factor importance. Separately, CMIP6 and SSP-based studies show that future rainfall erosivity, and therefore erosion, is likely to rise in many regions. In the Upper Narmada Basin, however, erosion assessment remains static and largely empirical, and no study has linked ML-based erosion modelling with CMIP6-SSP projections or used the result to plan conservation. By framing the literature around three objectives—an ML-based baseline, scenario-based future erosion, and a hotspot-driven management plan—this review provides the methodological basis for a climate-resilient conservation strategy that can protect both the basin's agricultural soils and its reservoirs.

REFERENCES

[1] Borrelli, P., Robinson, D.A., Fleischer, L.R., Lugato, E., Ballabio, C., Alewell, C., et al.: An assessment of the global impact of 21st century land use change on soil erosion. Nat. Commun. 8, 2013 (2017)

[2] Bhattacharyya, R., Ghosh, B.N., Mishra, P.K., Mandal, B., Rao, C.S., Sarkar, D., et al.: Soil degradation in India: challenges and potential solutions. Sustainability 7(4), 3528–3570 (2015)

[3] Narayana, D.V.V., Babu, R.: Estimation of soil erosion in India. J. Irrig. Drain. Eng. 109(4), 419–434 (1983)

[4] Sharda, V.N., Ojasvi, P.R.: A revised soil erosion budget for India: role of reservoir sedimentation and land-use protection measures. Earth Surf. Process. Landf. 41(14), 2007–2023 (2016)

[5] Meshram, S.G., Meshram, C., Hasan, M.A., Khan, M.A., Islam, S.: Morphometric deterministic model for prediction of sediment yield index for selected watersheds in upper Narmada Basin. Appl. Water Sci. 12(7), 1–11 (2022). https://doi.org/10.1007/s13201-022-01644-0

[6] Meshram, S.G., Adhami, M., Kisi, O., Meshram, C., Duc, P.A., Khedher, K.M.: Identification of critical watershed for soil conservation using game theory-based approaches. Water Resour. Manag. 35 (2021). https://doi.org/10.1007/s11269-021-02856-w

[7] Nearing, M.A.: Potential changes in rainfall erosivity in the U.S. with climate change during the 21st century. J. Soil Water Conserv. 56(3), 229–232 (2001)

[8] Panagos, P., Borrelli, P., Matthews, F., Liakos, L., Bezak, N., Diodato, N., Ballabio, C.: Global rainfall erosivity projections for 2050 and 2070. J. Hydrol. 610, 127865 (2022)

[9] Borrelli, P., Robinson, D.A., Panagos, P., Lugato, E., Yang, J.E., Alewell, C., Wuepper, D., Montanarella, L., Ballabio, C.: Land use and climate change impacts on global soil erosion by water (2015–2070). Proc. Natl. Acad. Sci. USA 117(36), 21994–22001 (2020)

[10] Rickards, N., Thomas, T., Kaelin, A., Houghton-Carr, H., Jain, S.K., Mishra, P.K., et al.: Understanding future water challenges in a highly regulated Indian river basin—modelling the impact of climate change on the hydrology of the upper Narmada. Water 12(6), 1762 (2020)

[11] Pandey, B.K., Khare, D., Kawasaki, A., Meshesha, T.W.: Integrated approach to simulate hydrological responses to land use dynamics and climate change scenarios employing scoring method in upper Narmada basin, India. J. Hydrol. 598, 126429 (2021)

[12] Wischmeier, W.H., Smith, D.D.: Predicting Rainfall Erosion Losses: A Guide to Conservation Planning. Agriculture Handbook No. 537, USDA, Washington, DC (1978)

[13] Renard, K.G., Foster, G.R., Weesies, G.A., McCool, D.K., Yoder, D.C.: Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE). Agriculture Handbook No. 703, USDA, Washington, DC (1997)

[14] Ganasri, B.P., Ramesh, H.: Assessment of soil erosion by RUSLE model using remote sensing and GIS – a case study of Nethravathi Basin. Geosci. Front. 7(6), 953–961 (2016)

[15] Gayen, A., Pourghasemi, H.R., Saha, S., Keesstra, S., Bai, S.: Gully erosion susceptibility assessment and management of hazard-prone areas in India using different machine learning algorithms. Sci. Total Environ. 668, 124–138 (2019)

[16] Arabameri, A., Chen, W., Loche, M., Zhao, X., Li, Y., Lombardo, L., Cerda, A., Pradhan, B., Bui, D.T.: Comparison of machine learning models for gully erosion susceptibility mapping. Geosci. Front. 11(5), 1609–1620 (2020)

[17] Sharma, S.K., Gajbhiye, S., Nema, R.K., Tignath, S.: Assessing vulnerability to soil erosion of a watershed of Narmada Basin using remote sensing and GIS. Int. J. Environ. Res. Dev. 4(2), 153–160 (2014)

[18] Pai, D.S., Sridhar, L., Rajeevan, M., Sreejith, O.P., Satbhai, N.S., Mukhopadhyay, B.: Development of a new high spatial resolution (0.25° × 0.25°) long period (1901–2010) daily gridded rainfall data set over India. Mausam 65(1), 1–18 (2014)

[19] Panagos, P., Borrelli, P., Meusburger, K., Yu, B., Klik, A., Lim, K.J., et al.: Global rainfall erosivity assessment based on high-temporal resolution rainfall records. Sci. Rep. 7, 4175 (2017)

[20] Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., Iwamoto, H.: Precise global DEM generation by ALOS PRISM. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. II-4, 71–76 (2014)

[21] Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al.: The Shuttle Radar Topography Mission. Rev. Geophys. 45(2), RG2004 (2007)

[22] Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al.: SoilGrids250m: global gridded soil information based on machine learning. PLoS ONE 12(2), e0169748 (2017)

[23] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R.: Google

Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27 (2017)

[24] Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9(5), 1937–1958 (2016)

[25] Thrasher, B., Wang, W., Michaelis, A., Melton, F., Lee, T., Nemani, R.: NASA Global Daily Downscaled Projections, CMIP6. Sci. Data 9, 262 (2022)

[26] Pandey, A., Chowdary, V.M., Mal, B.C.: Identification of critical erosion prone areas in the small agricultural watershed using USLE, GIS and remote sensing. Water Resour. Manag. 21(4), 729–746 (2007)

[27] Arnoldus, H.M.J.: An approximation of the rainfall factor in the Universal Soil Loss Equation. In: De Boodt, M., Gabriels, D. (eds.) Assessment of Erosion, pp. 127–132. Wiley, Chichester (1980)

[28] Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R.: Large area hydrologic modeling and assessment. Part I: model development. J. Am. Water Resour. Assoc. 34(1), 73–89 (1998)

[29] Breiman, L.: Random forests. Mach. Learn. 45(1), 5–32 (2001)

[30] Saha, S., Roy, J., Arabameri, A., Blaschke, T., Tien Bui, D.: Machine learning-based gully erosion susceptibility mapping: a case study of Eastern India. Sensors 20(5), 1313 (2020)

[31] Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, pp. 785–794. ACM (2016)

[32] Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. 20(3), 273–297 (1995)

[33] Ghorbanzadeh, O., Shahabi, H., Mirchooli, F., Valizadeh Kamran, K., Lim, S., Aryal, J., Jarihani, B., Blaschke, T.: Gully erosion susceptibility mapping (GESM) using machine learning methods optimized by the multi-collinearity analysis and K-fold cross-validation. Geomat. Nat. Hazards Risk 11(1), 1653–1678 (2020)

[34] Band, S.S., Janizadeh, S., Chandra Pal, S., Saha, A., Chakrabortty, R., Shokri, M., Mosavi, A.: Novel ensemble approach of deep learning neural network (DLNN) model and particle swarm optimization (PSO) algorithm for prediction of gully erosion susceptibility. Sensors 20(19), 5609 (2020)

[35] Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., et al. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015)

[36] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. 9(8), 1735–1780 (1997)

[37] Arabameri, A., Blaschke, T., Pradhan, B., Pourghasemi, H.R., Tiefenbacher, J.P., Bui, D.T.: Evaluation of recent advanced soft computing techniques for gully erosion susceptibility mapping: a comparative study. Sensors 20(2), 335 (2020)

[38] Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, vol. 30, pp. 4765–4774 (2017)

[39] O'Neill, B.C., Tebaldi, C., van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., et al.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci. Model Dev. 9(9), 3461–3482 (2016)

[40] Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., et al.: The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. Glob. Environ. Change 42, 153–168 (2017)

[41] Shrestha, S., Bae, D.-H., Hok, P., Ghimire, S., Pokhrel, Y.: Future hydrology and hydrological extremes under climate change in Asian river basins. Sci. Rep. 11, 17089 (2021)

[42] Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., et al.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40(8), 913–929 (2017)

[43] Tiwari, M.K., Guru, P.K., Kumar, S., Rajwade, Y.A., Chandel, N.S.: Integrated assessment of soil erosion drivers using RUSLE, remote sensing, and scenario-based machine learning in a data-scarce watershed of the Chota Nagpur Plateau, eastern India. J. Indian Soc. Remote Sens. (2026). https://doi.org/10.1007/s12524-026-02576-x