Soil erosion in future scenario using CMIP5 models and earth observation datasets

Swati Maurya¹, Prashant K Srivastava*,¹,², Aradhana Yaduvanshi³, Akash Anand¹, George P. Petropoulos⁴, Lu Zhuo⁵, R K Mall²

¹Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi-221005, India
²DST-Mahamana Centre of Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi-221005, India
³Department of Civil Engineering, Indian Institute of Technology Bombay, Powai-400076, India
⁴Department of Geography, Harokopio University, of Athens, Athens, Greece
⁵Department of Civil and Structural Engineering, University of Sheffield, Sheffield, UK

*Correspondence: prashant.just@gmail.com; Tel.: +917571927744

Abstract:

Rainfall and land use/land cover changes are significant factors that impact the soil erosion processes. Therefore, the present study aims to investigate the impact of rainfall and land use/land cover changes in the current and future scenarios to deduce the soil erosion losses using the state-of-the-art Revised Universal Soil Loss Equation (RUSLE). In this study, we evaluated the long-term changes (period 1981-2040) in the land use/land cover and rainfall through the statistical measures and used subsequently in the soil erosion loss
The future land use/land cover changes are produced using the Cellular Automata Markov Chain model (CA-Markov) simulation using multi-temporal Landsat datasets, while long term rainfall data was obtained from the Coupled Model Intercomparison Project v5 (CMIP5) and Indian Meteorological Department. In total seven CMIP5 model projections viz Ensemble mean, MRI-CGCM3, INMCM4, canESM2, MPI-ESM-LR, GFDL-ESM2M and GFDL-CM3 of rainfall were used. The future projections (2011-2040) of soil erosion losses were then made after calibrating the soil erosion model on the historic datasets. The applicability of the proposed method has been tested over the Mahi River Basin (MRB), a region of key environmental significance in India. The finding represents that rainfall-runoff erosivity gradually decreases from 475.18 MJ mm/h/y (1981-1990) to 425.72 MJ mm/h/y (1991-2000). A value of 428.53 MJ mm/h/y was obtained in 2001-2010, while a significantly high values 661.47 MJ mm/h/y is reported for the 2011-2040 in the ensemble model mean output of CMIP5. The combined results of rainfall and land use/land cover changes reveal that the soil erosion loss occurred during 1981-1990 was 55.23 t/ha/y (1981-1990), which is gradually increased to 56.78 t/ha/y in 1991-2000 and 57.35 t/ha/y in 2000-2010. The projected results showed that it would increase to 71.46 t/h/y in 2011-2040. The outcome of this study can be used to provide reasonable assistance in identifying suitable conservation practices in the MRB.

**Keywords:** Soil erosion; CMIP5 model; CA-Markov; Mahi River Basin; GIS; remote sensing
1. Introduction

Climate and land use changes are inter-related with each other. Direct effect of climate change in terms of rainfall intensity, duration, magnitude (Renschler et al., 1999; Pandey et al., 2007; Jain and Kumar, 2012; Rajeevan and Nayak, 2017) and indirect effect of land use change in term of urban sprawl, deforestation and other human activity caused an increase in the soil erosion losses. Therefore, the consequences of these climate and land use changes are essential to quantify the soil erosion rate for sustainable agricultural and environmental development. In India, almost 167 Mha of the area is found vulnerable to water and wind erosion (Das, 2014). Food and Agriculture Organization (FAO) reported that 25 to 40 billion tons of topsoil are degraded every year and it eventually impact the crop yield and soil properties (Montanarella et al., 2015). In general, soil erosion is a natural geological process that results in the removal of soil particles by water or wind and it is transported with the stream (Ganasri and Ramesh, 2016). Soil erosion is a major issue worldwide, which causes losses of soil nutrients, increasing sedimentation in rivers, degradation of agricultural land, high runoff and so forth. Therefore, it is imperative that natural resources should be managed on a sustainable basis to ensure long-term productivity and food security (Renschler et al., 1999; Pandey et al., 2007; Gajbhiye et al., 2014). Earth Observation (EO) provides detailed information about land, topography, watersheds characteristics, including soil types, land use and land cover and geomorphology. This information can also be easily integrated with Geographical Information Systems (GIS) to provide a quantitative measure of soil erosion.
Various models developed in the past for soil losses assessment such as Water Erosion Prediction Project (WEPP), Soil and Water Assessment Tool (SWAT), Universal Soil Loss Equation (USLE), Revised Universal Soil Loss Equation (RUSLE) and others. Among all updated version of USLE i.e. RUSLE model is widely used and worldwide accepted due to its ability to provide an accurate estimation of soil erosion both quantitatively and spatially (Renard et al., 1991; Kouli et al., 2009; Bonilla et al., 2010; Nagaraju et al., 2011a; Prasannakumar et al., 2012; Tirkey et al., 2013; Karamesouti et al., 2016). A lot of studies conducted over the Indian region such as Thomas et al. (2018) reported a severe rate of soil loss in the tropical mountain river basin of Western Ghats, India using RUSLE with the transport limited sediment delivery (TLSD) function (Thomas et al., 2018). Kumar et al. (2014) suggested that soil erosion in the Himalayan watershed is a very sensitive factor as high slope and depleting forest covers are major causes of erosion (Kumar et al., 2014). In the last few decades, with the advancements in satellite observations and data quality, there is a substantial increase in the research studies on the impact of land use and rainfall on soil erosion. (Markose and Jayappa, 2016) used the RUSLE model in a tropical humid climatic zone that is experiencing a severe loss in soil due to natural factors, whereas, (Wang et al., 2018) compared the effects of rainfall and land use land cover patterns on soil erosion for different watersheds which is likely to play a crucial role in modelling and management of multi-scale watersheds. Another study by (Wei et al., 2007) considered the influence of different rainfall patterns to estimate the impact of land use on the soil erosion, and concluded that the concentration as well as high intensity with short duration rainfall events influences the soil erosion processes.
Additionally, Global Climate Models (GCMs) have been successfully used in the scientific community for future climate projections. In general, their resolution is not enough to produce the regional climatic condition. Therefore in this study the NEX-GDDP (NASA Earth Exchange Global Daily Downscaled Projections) based Coupled Model Intercomparison Project Phase (CMIP5) data at fine resolution 0.25° x 0.25° (Bao and Wen, 2017) is employed. In the purview of the above, the focus of this study is to assess the impact of both climate and land use/land cover changes on soil erosion using the RUSLE model. In order to achieve the objectives, we investigated the NEX-GDDP-CMIP5 model performance over the study area for rainfall and estimated the land use/land cover changes using the multidate Landsat satellite images. Future projections of landscape changes are also estimated through CA-Markov and by using the classified multidate satellite images of the historical time period. Afterwards, soil erosion losses were provided for the baseline and future scenarios.

2. Study area

Mahi River is one of the largest rivers in India passing through the three geographically larger states Madhya Pradesh, Rajasthan and Gujarat and terminated at the Gulf of Khambhat as shown in Figure 1. The MRB covers an area of 34,842 km². The basin can be divided into three parts-lower, middle and upper part. The upper part of the basin is having mostly hills and forests with some plain area in Madhya Pradesh. The middle part is having developed lands and mostly found in Gujarat. The Gujarat region is also encompassing most of the lower basin, which is very fertile with alluvial soil. In MRB, the area that can be used for agriculture
is around 2.21 Mha. The other soil types which are found in the basin are red and black soils. Hydro-geologically the basin is dominated by basaltic rocks with trappean. The average rainfall in MRB is approx. 785 mm. Apart from agriculture, it is one of the important sources for irrigation, drinking water and industrial water demand.

![Fig.1. Location map of Mahi River Basin, India.](image)

3. Materials and Methods

In this study, the NASA-NEX-GDDP-CMIP5 model output, IMD (observed) datasets, Land use/land cover from Landsat were used. Along with the assessment, the future land cover expansion and climate change scenarios are also considered for their potential impacts on
soil erosion in MRB. To achieve this objective, an integrated approach of an erosion model, climate model and land use/land cover datasets has been used. The methodology of the present study has been summarized in Figure 2. The detailed description of datasets and methodology are provided in sub sections.

3.1 Digital Elevation Model (DEM)

The Shuttle Radar Topography Mission (SRTM) launched in collaboration between NASA and the National Geospatial Intelligence Agency (NGA). It provides void filled elevation data globally (http://www.cgiarcsi.org). In the present study, a 30 m DEM (v.3) is used for the extraction of slope of the study area using the spatial analyst tool of Arc GIS 10.1 software (in Figure 3(a)). Slope expressed the inclination of landform associated with the physical feature. Higher slope value leads to rapid runoff with potential soil erosion (Stefanidis and Stathis, 2018).
3.2 IMD Rainfall datasets

The Indian Meteorological Department (IMD) provided the gridded daily rainfall data at 0.25° X 0.25°. The daily rainfall recorded from 6955 rain gauge stations of National Data Centre, IMD, Pune, India (Pai et al., 2014). IMD uses the Inverse Distance Weighted interpolation technique along with the radial distance to convert the point-based gauge data into grid data. 30 years (1981-2010) of annual average rainfall data have been used, obtained for the meteorological stations Dhariawad, Mataji, Rangeli, Chakaliya, Paderibadi, Khanpur in the study area (Figure 3(b)).

Fig.2. Workflow of the methodology developed in this study

Fig.3. (a) Slope map (b) Annual average rainfall (1981-2010)
3.3 Soil map

Soil map data is obtained from the FAO, United Nations, at 1:5000,000 scale and the dataset can be obtained at no cost from FAO(http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/faounesco-soil-map-of-the-world). It provide information related to soil properties at the depth 0 – 30 cm (topsoil) and 30 – 100 cm (subsoil) with various parameters as Organic Carbon, pH(H2O), Calcium carbonate, Sand fraction, Silt fraction, Clay fraction, Bulk Density and so on. The data showed that the study region is mainly covered by eight soil classes as shown in Figure 4).

![Soil Map](image)

**Fig. 4. Soil map of the area**

3.4 Land use/land cover estimation and prediction
Landsat satellite data is used for land use/land cover estimation. Landsat is a collaborative effort of the US Geological Survey and the National Aeronautics and Space Administration (NASA). In this study, Landsat 1-5 having MSS (Multispectral scanner) and TM (Thematic Mapper) sensors data are used to prepare land use/land cover maps for the years 1981, 1991, 2001, 2011. Before the classification of the images, they are geo-referenced and projected to WGS 1984 UTM Zone 43N coordinate system. In ENVI software, Support Vector Machine (SVM) algorithm based supervised classification system is applied to classify the images. SVM is found to be the best algorithm for land use/land cover classification by many researchers (Srivastava et al., 2012; Singh et al., 2014; Nandi et al., 2017; Fragou et al., 2020). The study area is classified into five classes namely, Waterbody, Cropland, Grassland, Barren, Urban and Forest land respectively. Table 1 is showing the overall classification accuracy and the Kappa performance statistics, which is 78.3%, 82.7%, 80.8%, 88.4% and 0.76, 0.79, 0.77, 0.85 respectively for the classified images of the year 1981, 1991, 2001 and 2011. Further, the state of the arts CA-Markov has been used for the prediction of land use/land cover classes of 2040 as shown in Figure 5. CA-Markov model is one of the most commonly used and consistent model for simulating land use/land cover changes, it combines cellular automata and Markov chain to predict the changes through space and time (Weng, 2002). CA-Markov is widely used in several studies such as in ecological modelling (Ghosh et al., 2017), watershed management (Yulianto et al., 2018), urban growth (Aburas et al., 2017) and land use policy designing (Liu et al., 2017). Mathematical expression for the CA-Markov model can be understood through Eq. 1 and 2

\[
S(t, t + 1) = P_{ij} * S(t)
\]  

(1)
\[\|P_{ij}\| = \begin{bmatrix}
P_{1,1} & P_{1,2} & \cdots & P_{1,n} \\
P_{2,1} & P_{2,2} & \cdots & P_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
P_{n,1} & P_{n,2} & \cdots & P_{n,n}
\end{bmatrix}\] (2)

Where \(S(t)\) is the image at time \(t\), \(S(t+1)\) is the image at time \(t+1\) and \(P_{ij}\) is the transition probability matrix in which \(i\) is the current state and \(j\) is the future state. The value of \(P_{ij}\) varies from 0 to 1 in which the low transition probability will be near to 0 and high transition probability will be near to 1.

Table 1: Accuracy assessment of land use/land cover classification

<table>
<thead>
<tr>
<th>Land Use/Land Cover Classes</th>
<th>1981 PA(%)</th>
<th>1981 UA(%)</th>
<th>1991 PA(%)</th>
<th>1991 UA(%)</th>
<th>2001 PA(%)</th>
<th>2001 UA(%)</th>
<th>2011 PA(%)</th>
<th>2011 UA(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterbody</td>
<td>100</td>
<td>98.8</td>
<td>96.6</td>
<td>100</td>
<td>100</td>
<td>96.2</td>
<td>100</td>
<td>98.5</td>
</tr>
<tr>
<td>Forest</td>
<td>82.4</td>
<td>78.6</td>
<td>92.0</td>
<td>88.2</td>
<td>87.5</td>
<td>93.6</td>
<td>94.2</td>
<td>91.7</td>
</tr>
<tr>
<td>Grassland</td>
<td>84.5</td>
<td>78.3</td>
<td>88.0</td>
<td>84.5</td>
<td>76.2</td>
<td>72.5</td>
<td>87.2</td>
<td>88.5</td>
</tr>
<tr>
<td>Cropland</td>
<td>70.3</td>
<td>78.5</td>
<td>83.3</td>
<td>80.0</td>
<td>77.2</td>
<td>81.0</td>
<td>82.5</td>
<td>84.1</td>
</tr>
<tr>
<td>Barren</td>
<td>88.0</td>
<td>84.5</td>
<td>78.2</td>
<td>75.5</td>
<td>85.6</td>
<td>88.2</td>
<td>79.6</td>
<td>75.5</td>
</tr>
<tr>
<td>Urban</td>
<td>67.2</td>
<td>69.2</td>
<td>75.5</td>
<td>79.4</td>
<td>68.4</td>
<td>71.2</td>
<td>84.2</td>
<td>87.1</td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>78.3</td>
<td>82.7</td>
<td>80.8</td>
<td>88.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa Accuracy</td>
<td>0.76</td>
<td>0.79</td>
<td>0.77</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Producer Accuracy (PA), User Accuracy (UA)*

Fig. 5 Spatial distribution of land use/land cover (a) 1981, (b) 1991, (c) 2001,
3.5 Global Climate Model data

The NEX-GDDP datasets are downscaled climate scenarios derived from the General Circulation Model (GCM) simulations of the Coupled Model Intercomparison Project Phase 5 (CMIP5). The four major greenhouse gas emissions scenarios are considered as Representative Concentration Pathways (RCPs) based on IPCC AR5 (Intergovernmental Panel on Climate Change–Fifth report). The NEX-GDDP dataset uses statistical downscaling approach namely-Bias-Corrected Spatial Disaggregation (BCSD) method to downscale the projections for RCP 4.5 and RCP 8.5 from the 21 CMIP5 models (Wood et al., 2004; Maurer and Hidalgo, 2008). Detail document is available at https://cds.nccs.nasa.gov. Daily scale data for maximum temperature, minimum temperature and precipitation at fine resolution 0.25° (~25km×25km) are available at https://cds.nccs.nasa.gov/nex-gddp/. In this study, seven GCMs of CMIP5 were selected, which work well over the Indian region and have been validated by (Bokhari et al., 2018; Jain et al., 2019; Sahany et al., 2019). The institution, country and spatial resolution of the seven models are shown in Table 2. The long term rainfall datasets from (1981-2040) were obtained for all the seven models using the NEX-GDDP-CMIP5.

3.6 Evaluation of the CMIP5 Model output

The performances of seven models of NEX-GDDP-CMIP5 (six model output and one ensemble) were assessed by both statistical measures and spatial patterns of mean annual
precipitation. Taylor diagram (Taylor, 2001) is a suitable tool for the assessment of the model performance through the statistical measures in terms of spatial correlation coefficient, centred pattern Root Mean Square (RMS), and the ratio of spatial standard deviations. Taylor diagram is user-friendly because of three metrics at a single platform. The circle centred at the observed point represents the RMS and the circle centred at the origin point represents the standard deviation and the correlation coefficient. For the best performance in terms of the spatial correlation and standard deviation, the value should be close to 1 and for RMS the value should be close to 0.

Table 2 Features of the six CMIP5 global climate models.

<table>
<thead>
<tr>
<th>CMIP5 Models</th>
<th>Institution, Country</th>
<th>Atmospheric Resolution</th>
<th>NEX-GDDP resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Geophysical Fluid Dynamics Laboratory Climate Model, version3 (GFDL-CM3)</td>
<td>National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, U.S.A</td>
<td>2.5° X 2°</td>
<td>0.25°X 0.25°</td>
</tr>
<tr>
<td>2-Institute of Numerical Mathematics</td>
<td>Institute of Numerical Mathematics, Russia</td>
<td>2°X1.5°</td>
<td>0.25°X 0.25°</td>
</tr>
<tr>
<td>Coupled Model, version 4.0 (INMCM-4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>3-Max Plank Institute Earth System Model, low resolution (MPI-ESM-LR)</td>
<td>Max Plank Institute for Meteorology, Germany</td>
<td>1.875°X1.8653°</td>
<td>0.25°X 0.25°</td>
</tr>
<tr>
<td>4-Meteorological Research Institute Coupled Atmosphere–Ocean General Circulation Model, version 3 (MRI-CGCM3)</td>
<td>Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, Japan</td>
<td>1.125°X1.1215°</td>
<td>0.25°X 0.25°</td>
</tr>
<tr>
<td>5-The second–generation Canadian Earth System model (CanESM2)</td>
<td>Canadian Centre for Climate Modelling and Analysis, Canada</td>
<td>2.8125°X2.7906°</td>
<td>0.25°X 0.25°</td>
</tr>
</tbody>
</table>
4. Revised Universal Soil Loss Equation (RUSLE) model

RUSLE was invented by the USDA-Agricultural Research Service for the conservation planning and management. Originally USLE (Wischmeier and Smith, 1978) was developed to predict soil loss by unit plot condition in tropics region based on rainfall, soil type, topography, crop pattern and management practices. The revised version i.e. RUSLE was later proposed with some modifications in the algorithm of USLE factors (Moore and Wilson, 1992; Renard et al., 1997). RUSLE is a spatially distributed model and does not required too much data for the computation as well as it provide valuable results verified by various research articles. (Fernandez et al., 2003; Yue-Qing et al., 2008; Demirci and Karaburun, 2012; Naqvi et al., 2013; Pan and Wen, 2014; Pradeep et al., 2015). It provide the annual average soil loss in (t/ha/y) by the following equation (Renard, 1997):

\[ A = R \times K \times LS \times C \times P \]  

Where \( A = \) Average Soil Loss Per Unit Area (t/ha/y); \( R = \) Rainfall-Runoff Erosivity Factor (MJ mm ha\(^{-1}\)h\(^{-1}\)year\(^{-1}\)); \( K = \) Soil Erodibility Factor (metric tons ha\(^{-1}\)MJ\(^{-1}\)mm\(^{-1}\)); \( LS = \)
Topographic Factor (dimensionless); C = Cover Management Factor (dimensionless); and P = Conservation Practice Factor (dimensionless). Detailed descriptions of each of the RUSLE component are covered in the following subsections.

### 4.1 Soil Erodibility Factor (K)

The K factor represents the susceptibility of soil detachment, or transportation of soil particles due to rainfall. K factor significantly affected by soil structure, texture, organic content, and hydraulic properties of soil. The K values (tons/ha/MJ) can be calculated by the following equation (Sharpley and Williams, 1990).

\[ K = A \times B \times C \times D \times 0.1317 \]  

(4)

where:

\[ A = [0.2 + 0.3 \exp(-0.0256 \cdot SAN(1 - SIL/100))] \]  

(5)

\[ B = \left[ \frac{SIL}{CLA + SIL} \right]^{0.3} \]  

(6)

\[ C = \left[ 1.0 - \frac{0.25C}{C + \exp(3.72 - 2.95)} \right] \]  

(7)

\[ D = 1.0 - \frac{0.70SN1}{SN1 + \exp([-5.41 + 22.9SN1])} \]  

(8)

Where; SAN, SIL and CLA represents the percentage of sand, silt and clay, respectively; C = organic carbon content; SN1 = sand content subtracted from 1, divided by 100.

Soil maps are the basic layer for the estimation of the K factor. Firstly, the vector layer of the soil map is converted into raster format by ArcGIS 10.1 software. After which, k values are assigned to the map by using reclassify tool of the ArcGIS 10.1.
4.2 Rainfall-runoff Erosivity (R) factor

R represents how the rainfall frequency, intensity, duration of rainfall and rate of runoff affects the soil erosion. Originally, R factor estimated by the long term average of rainfall kinetic energy and the maximum 30 min intensity during the storm event (Arnoldous, 1980). Due to the scarcity of the data, here we used the equation based on the annual average rainfall datasets (Wischmeier and Smith, 1978).

\[ R = 38.5 + 0.35r \]  
\( (9) \)

Where; \( R \) = Rainfall Erosivity Factor (MJ mm ha/h/year); \( r \) = Annual Average Rainfall (mm).

4.3 Conservation Practice Factor (P)

The P factor represent the support practices that are applied in the field to reduce the rate of runoff, to control the flow and velocity of runoff, to change the pattern of runoff and so forth. P is the ratio of soil loss with a specific support practice to the corresponding slope tillage (Wischmeier and Smith, 1978; Renard et al., 1997). P factor values varies from 0 – 1 (Renard et al., 1997). P of 1 assign to those areas where have poor conservation practices (i.e., scrub land, wasteland, Urban) while 0 or 0.3 value assigned to those areas where have good conservation practices.

4.4 Topographic Factor (LS)
Slope length (L) and slope steepness (S) are jointly expressed as LS. L is defined as the distance of flow path from the origin of overland flow to the point where deposition begins or runoff water enters in a flow channel, and S is the steepness of slope (Pradhan et al., 2012). LS can be evaluated by field measurement or using DEM via the following equation:

$$LS = (\text{flow accumulation} \times \text{cell size/22.13})^{0.4} \times \sin(\text{Slope/0.896})^{1.3}$$  \hspace{1cm} (10)

Where flow accumulation represents the number of grid cells that shows the flow downward; cell size is the grid cell size (30m is used in this study); sin Slope is the slope degree in sin.

### 4.5 Crop Management Factor (C)

C-factor is the most important factor after the topography. It shows the cropping pattern, management practices and the erosion control measure of soil loss (Mati et al., 2000). The C-factor is decided based on land use/land cover classes as shown in Table 3.

#### Table 3 C-Factor of the Mahi River Basin taken from the different studies

<table>
<thead>
<tr>
<th>Land Use/Land Cover</th>
<th>C-factor</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed forest</td>
<td>0.003</td>
<td>(Ganasri and Ramesh, 2016)</td>
</tr>
<tr>
<td>Shrubland</td>
<td>0.18</td>
<td>(Rao, 1981b)</td>
</tr>
<tr>
<td>Grassland</td>
<td>0.05</td>
<td>(Rao, 1981b)</td>
</tr>
<tr>
<td>Cropland</td>
<td>0.28</td>
<td>(Rao, 1981b)</td>
</tr>
<tr>
<td>Urban</td>
<td>1.0</td>
<td>(Tirkey et al., 2013)</td>
</tr>
<tr>
<td>Barren or Sparsely vegetated</td>
<td>0.33</td>
<td>(Rao, 1981b)</td>
</tr>
</tbody>
</table>
6. Results and Discussion

6.1 Performance assessment NEX-GDDP-CMIP5 outputs

Taylor diagram presents a comparison of IMD data (i.e., the station observations) with the NEX-GDDP-CMIP5’s six models output data and ensemble for the period 1981-2010. Figure 6. Taylor diagram shows that all individual model and ensemble mean cluster lies in between a correlation coefficient of 0.5 to 0.85. However, standard deviation value of MRI-CGCM3, INMCM4 and Ensemble mean is close to 0.75 mm/day with an RMS value approx. 0.075 mm/day. The INMCM4 and MRI-CGCM3 showed a slightly higher RMS (0.18 and 0.13mm/day) than Ensemble model. Moreover, ensemble value reduces the uncertainty (i.e., parametric, structural and response) of individual model and showed a good performance (Giorgi and Mearns, 2002; Hagedorn et al., 2005; Palmer et al., 2005; Chaturvedi et al., 2012). The monthly mean rainfall of the individual models and ensemble mean climatology over the MRB is shown in Figure 7. These plots illustrate that the MRI-CGCM3 and INMCM4 along with the ensemble mean are all underestimated but show similar pattern to the IMD, while the other models (i.e., canESM2, MPI-ESM-LR, GFDL-ESM2M and GFDL-CM3) indicated a large inter-model difference.
Fig. 6 Performances of NEX-GDDP-CMIP5 model outputs during the monsoon months (1981-2010)

Fig. 7 Annual mean rainfall of the IMD, NEX-GDDP–CMIP5 models and the Ensemble mean during the period 1981-2010
Furthermore, the spatial variabilities of the annual mean rainfall for the IMD and the NEX-GDDP-CMIP5 models are shown in figure 8 (a-h). IMD has the highest rainfall gradient occurred in the north-east and the north-west parts, with moderate to low rainfall that is occurred in the north-west part of the MRB. A similar spatial distribution observed in the best performing models i.e., MRI-CGCM3, INMCM4 and ensemble mean in comparison to other models.
Fig. 8 Spatial distribution of the annual mean rainfall during the time period 1981-2010:

(a) IMD, (b) Ensemble mean, (c) MRI-CGCM3, (d) INMCM4, (e) GFDL-CM3, (f) GFDL-ESM, (g) MPI-ESM-LR, (h) CanESM2.

The box-whisker plots of the annual mean rainfall datasets for the period 1981-2010 and the 2011-2040 are shown in Figure 9 (a-b). In the plot, boxes are having the upper quartile,
median line (center) and the lower quartile. The whiskers are represented as the dotted line at each end of the box, and outliers are shown incircle. The annual mean rainfall of models has median in the center which represents a uniform distribution of the rainfall.

![Box-Whisker plot of the annual mean rainfall datasets during the time periods](image)

**Fig.9** Box-Whisker plot of the annual mean rainfall datasets during the time periods (a) 1981-2010 and (b) 2011-2040.

### 6.3 Input parameters of RUSLE

The five major factors of RUSLE (R, K, LS, P and C) were estimated through the rainfall data, soil datasets, land use/land cover, DEM and satellite images as discussed in the following sections:
6.3.1 Soil Erodibility Factor (K) and Topographic Factor (LS)

The K factor varies from 0.034-0.052. The smaller value of K factor indicates lower permeability, low antecedent moisture content of soil and vice versa (Ganasri and Ramesh, 2016). The results indicated that the north part of the MRB showed the highest erodibility (0.052), and the central part and the north-east part show moderate to low erodibility (0.04-0.034) of the MRB as shown in Figure 10(a).

The 0 value of LS is obtained in the south-west region of the MRB with the lowest elevation (1.79° - 4.42°), while a value of 0.324 can be seen in the north-west part having the steepest slope (15.74°-50.59°) Figure 10(b). The overall results suggested that the LS factor varies significantly between the north-west and the central part of the watershed.

Fig.10 (a) Soil erodibility factor, (b) and Topographic factor of the study area

6.3.2 Crop Management Factor (C)

The value of C factor is assigned for particular land use class according to the literature survey (Rao, 1981a; Alexakis et al., 2013). In general, the minimum value of C implies that
the crop management practices are good and vice versa (Benkobi et al., 1994; Biesemans et al., 2000; Kouli et al., 2009). The C factor of the base period 1981, 1991, 2001, 2011 and future 2040 land use/land cover are shown in Figure 11, while Table 4 illustrated the percentages of the area occupied. On comparison with the baseline time period, finding indicates that the C-factor of Urban, Barren, Cropland and Grassland area are increasing, while for Water and Forest areas, a decreasing value is observed in 2040.

Table 4. Percent land area for each C value calculated using the classified images of different years.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterbody</td>
<td>6.50%</td>
<td>4.80%</td>
<td>4.50%</td>
<td>4.94%</td>
<td>4.06%</td>
</tr>
<tr>
<td>Forest</td>
<td>23.37%</td>
<td>45.00%</td>
<td>40.36%</td>
<td>25.66%</td>
<td>22.43%</td>
</tr>
<tr>
<td>Grassland</td>
<td>19.04%</td>
<td>7.40%</td>
<td>9.67%</td>
<td>6.11%</td>
<td>44.76%</td>
</tr>
<tr>
<td>Cropland</td>
<td>25.72%</td>
<td>28.17%</td>
<td>24.74%</td>
<td>48.00%</td>
<td>17.36%</td>
</tr>
<tr>
<td>Barren</td>
<td>22.34%</td>
<td>11.27%</td>
<td>15.67%</td>
<td>7.33%</td>
<td>11.65%</td>
</tr>
<tr>
<td>Urban</td>
<td>2.40%</td>
<td>3.00%</td>
<td>5.01%</td>
<td>3.19%</td>
<td>5.71%</td>
</tr>
</tbody>
</table>
Fig. 11 C-factor of the study area in the year (a) 1981 (b) 1991 (c) 2001 (d) 2011 and (e) 2040

6.3.3 Conservation Practice Factor (P)
In this study due to the absence of the field observation, the value of P factor is assigned on the basis of earlier studies (Mati et al., 2000; Ganasri and Ramesh, 2016). The P-factor of the base period 1981, 1991, 2001, 2011 and future 2040 land use/land cover classes are shown in Figure (12) and Table 5, which illustrated the P-Factor percentage area occupied by different classes. On comparison with the base time period, the forest, grassland and cropland were found increasing while barren and water areas were decreased due to poor conservation practices.

Table 5. P-Factor calculated using the classified images of different years.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Water and Barren</td>
<td>29.84%</td>
<td>16.67%</td>
<td>20.20%</td>
<td>16.27%</td>
<td>23.71%</td>
</tr>
<tr>
<td>Cropland, Forest and Grassland</td>
<td>70.15%</td>
<td>80.91%</td>
<td>74.77%</td>
<td>80.78%</td>
<td>80.57%</td>
</tr>
<tr>
<td>Urban</td>
<td>2.40%</td>
<td>3.00%</td>
<td>5.01%</td>
<td>3.19%</td>
<td>5.71%</td>
</tr>
</tbody>
</table>
Fig. 12 P-factor of the study area in the year (a) 1981, (b) 1991, (c) 2001, (d) 2011, (e) and 2040

6.3.4 Rainfall-Runoff Erosivity Factor (R)
Many studies have suggested that the soil loss of a catchment is primarily affected by rainfall (Pandey et al., 2007; Nagaraju et al., 2011b; Chatterjee et al., 2014; Samanta and Bhunia, 2016). The mean annual rainfall-runoff erosivity of the base scenario (1981-1990), (1991-2000), (2001-2010) and future scenario (2011-2040) are shown in Figure (13). From Figure 13 (a)-(c) the spatial distribution represents the highest erosivity in the north and the northwest parts 290-450 MJ mm ha/h/y (1981-1990), 300-420 MJ mm ha/h/y (1991-2000), 345.45-426.53 MJ mm ha/h/y (2001-2010), the moderate value has been found in the central part, and the lowest value observed in the east-south part 160-260 MJ mm ha/h/year (1981-1990), 170-260 MJ mm ha/h/y (1991-2000), 238.55-318.9 MJ mm ha/h/y (2001-2010) in the MRB.

However, during 2011-2040, the rainfall-runoff erosivity are estimated to be 675.16, 661.45 and 625.56 MJ mm ha/h/year for the MRI-CGCM3, the ensemble mean and the INMCM4 respectively, as shown in Figure 13(e)-(f). By comparing with the base time period, it can be seen that the rainfall-runoff erosivity increases gradually in the future scenario (2011-2040) to approx. 36.88%, 35.57% and 31.88% in the MRI-CGCM3, the ensemble means and the INMCM4 respectively.
Fig. 13 Rainfall-runoff erosivity during the time period (a) 1981-1990, (b) 1991-2000, (c) 2001-2010 of IMD, and (d-f) for the Ensemble mean, the MRI-CGCM3 and the INMCM4 respectively, during the period 2011-2040.

6.4 The soil erosion assessment of the base scenario and validation

Slope and terrain properties play a major role in shaping rate of soil erosion. Steep slopes are prone to more soil erosion as compared to the less steep slope. In the findings, the north-west, the east and the central region of MRB are highly affected by the soil erosion problem due to the steep slope and poor conservation practices along with intense rainfall. However, the annual average soil loss was reported as 55.23 t/ha/y (1981-1990), 56.78 t/ha/y (1991-2000), 57.35 t/ha/y (2000-2010) and categorized into five zones; very slight, slight, moderate, moderate severe, and severe (see Figure 14 (a)-(c)).

South west portion of the MRB has coverage of very slight soil loss class zone. With each passing decade the soil loss has increased by 1.55 t/ha/y and 0.57 t/ha/y. Increase in soil loss could potentially occur due to the heavy rains and change in land use/land cover pattern. We further explored the impact of land use and rainfall change impact on the soil erosion rate in current and future scenarios. The National Bureau of Soil Survey and Land Use Planning (NBBS & LUP)’s point based soil loss datasets (http://www.bhoomigeoportal-nbsslup.in/) are also in line with the obtained results. The datasets are categorized into very slight (<5 t/ha/y), slight (5-10 t/ha/y), moderate (10-15 t/ha/y), moderate severe (15-20 t/ha/y), severe (20-40 t/ha/y), very severe classes (40-80 t/ha/y), and extremely severe classes (>80 t/ha/y) are available from the site http://www.bhoomigeoportal-nbsslup.in/. The datasets showed a
similar soil loss values as obtained from the RUSLE model and the overall accuracy is found as 85%. The category wise accuracy can be varied from very slight, slight, moderate to severely eroded. Therefore, the result suggested that the RUSLE is a promising approach for this type of the study as well as cost-effective in the identification of vulnerable area for soil erosion risk.

6.5 Soil erosion for the base and future scenarios

Based on rainfall-runoff erosivity and land use change, soil erosion is predicted while other factors influenced by the soil type and topography are kept constant while performing the future projection. The changes in C-factor and P-factor along with R-factor increases significantly in the future time series (2011-2040) in comparison to the present time series (1981-2010). Similarly, the rate of the annual average soil erosion increases to 71.56, 66.34, and 60.56 t/h/year in the MRI-CGCM3, the ensemble means and the INMCM4 model respectively in future time series (2011-2040) Figure 14 (d)-(f). As compared to the base scenario, the annual average soil erosion increases to 29.56%, 20.11% and 11.21% in the MRI-CGCM3, the INMCM4 and the ensemble mean model, respectively. As compared to the soil erosion based on land use /land cover area, we find significant results, as the highest soil erosion rate is recorded in forest class which is 217.13 to 327.45 t/ha/y and cropland 239.43 to 312.87 t/ha/y as shown in Table 6. The forest and cropland land cover area decrease by 42.23% and 33.13% in the future scenario (2040), it may be the result of the expansion in grassland and urban areas. Similarly, moderate soil erosion rates were found in the grassland that is 110.63 to 128.96 t/ha/y along with a significant increase in land area of approximately 47.34% due to the transition of forest and cropland areas and barren areas has
shown a soil erosion rates of 178.21 to 146.59 t/ha/y with an overall decrease in the land area of -1.23% due to the expansion of urban areas. While in urban area, the soil erosion rate was found to be the lowest 21.25 to 58.4 t/ha/y but the land area increased significantly to 72.32% from base to predicted future scenario. Projected increase in barren land and settlement area might affect the local rainfall mechanism in the basin but at the same time intense rainfall could exacerbates the rate and magnitude of land degradation by increased soil loss. With decrease in crop land and forest area in future scenario pose threat to natural ecosystem and biodiversity. Projected increase in a water body area is a good sign as far as future water demand and supply is concern in the MRB.

These results indicate that the change in soil erosion rate follows the rainfall and land use changes, which has been validated by various research articles, as Sharma et al., suggested that mean soil erosion potential of the watershed was increased slightly due to the transition of LULC categories to cropland (Sharma et al., 2011). Zare et al., results indicate that mean soil erosion increases by 45% from the base period to future period, because of the most significant transition observed in the forest area to settlement (Zare et al., 2017). Mondal and Gupta et al., studies have reported that the increasing trend of precipitation and land use changes could increase the future rate of soil erosion over the Himalayan and Narmada River basin (Mondal et al., 2016; Gupta and Kumar, 2017).

**Table.6 Average annual soil loss (t/ha/y) of different land use land covers classes.**
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>217.13</td>
<td>318.89</td>
<td>322.34</td>
<td>315.21</td>
</tr>
<tr>
<td>Grassland</td>
<td>110.63</td>
<td>117.32</td>
<td>125.25</td>
<td>131.89</td>
</tr>
<tr>
<td>Cropland</td>
<td>239.43</td>
<td>246.15</td>
<td>320.21</td>
<td>205.38</td>
</tr>
<tr>
<td>Barren</td>
<td>178.21</td>
<td>162.35</td>
<td>199.90</td>
<td>235.21</td>
</tr>
<tr>
<td>Urban</td>
<td>21.25</td>
<td>28.54</td>
<td>20.12</td>
<td>42.26</td>
</tr>
</tbody>
</table>
Legend - soil loss (t/ha/y)

- Very slight
- Slight
- Moderate
- Moderate severe
- Severe
- No soil loss
Fig.14. Soil erosion rate during the time period (a) 1981-1990 (b) 1991-2000 (c) 2001-2010 of IMD, and (d-f) for the Ensemble mean, the MRI-CGCM3 and the INMCM4 respectively, during the time periods (2011-2040).

7. Conclusion

The study demonstrated the potential impact of long-term rainfall and land use/land cover changes on soil erosion using the state-of-the-art RUSLE and NEX-GDDP-CMIP5 models. The results indicate that the RUSLE has potential to capture catchment characteristics including climatic variables such as rainfall distribution, soil properties (texture, organic carbon), topography (slope, flow accumulation), land use (crop pattern, management and practices), and hence can help in the quantification of the soil erosion losses. The MRI-CGCM3, INMCM4 and ensemble mean are the most suitable models to capture the spatial variability of the precipitation with high spatial correlation (0.65-0.83) and low error rate (0.52 mm/day) with respect to the observed (IMD) datasets, during the time period 1981-2010. The finding of land use changes during the time period 2040 reported that urban, barren, cropland and grassland area with poor crop management practices are increasing while water and forest area are decreasing. Furthermore, it is concluded that in near future the rainfall erosivity factor may increase which can lead to high soil erosion rate. The outcome of this study would be of important help in evaluating the landform and their processes, agricultural productivity, hazardous mitigation and so forth within the study area and for deducing the changes in the future. In addition, the results obtained from this study can be utilized by various government agencies, developers and policymaker for a better soil
and water conservation in the MRB. Furthermore, the implementation of the proposed technique is robust as it is based on satellite imagery and ancillary datasets provided globally at no cost. The method is straight-forward, and requires low computational facility and hence can be easily reapplied in other parts of the world to cover a broad spectrum of catchments.

Acknowledgements

The first author is highly thankful to the Department of Science and Technology, Government of India for funding the research work. The authors would like to thank the Banaras Hindu University for supporting this PhD research work and for providing necessary funds.

Conflicts of Interest: The authors declare no conflict of interest.

References


Zare, M., Panagopoulos, T., Loures, L., 2017. Simulating the impacts of future land use change on soil erosion in the Kasilian watershed, Iran. Land Use Policy 67, 558-572.
Highlights:
Comprehensive evaluation of long term precipitation and satellite derived landscape changes in the Mahi River Basin
Assessment of CMIP5 rainfall projections from seven different models using ground dataset
Evaluation of landscape past, current and future prediction using SVM, CA-Markov Chain model and Earth observation dataset
Integrated framework for prediction of soil erosion rates using projected rainfall from CMIP5 and SVM-CA-Markov Chain derived landscape dataset
Credit Author Statement:

Conceptualization, Prashant K Srivastava; Data curation, Swati Maurya, Aradhana Yaduvanshi; Formal analysis, Swati Maurya, Akash Anand, Aradhana Yaduvanshi; Funding acquisition, Prashant K Srivastava and R K Mall; Investigation, Prashant K Srivastava and Swati Maurya; Methodology, Prashant K Srivastava and Swati Maurya; Project administration, Prashant K Srivastava; Resources, Prashant K Srivastava, George P. Petropoulos and R K Mall; Software, Prashant K Srivastava and Swati Maurya; Supervision, Prashant K Srivastava; Validation, Swati Maurya, Akash Anand and Aradhana Yaduvanshi; Visualization, Lu Zhuo and Swati Maurya; Writing – original draft, Prashant K Srivastava, Aradhana Yaduvanshi and Swati Maurya; Writing – review & editing, Prashant K Srivastava, George P. Petropoulos, R K Mall, and Lu Zhuo.