Sediment yield modelling using SDR and MUSLE with high resolution satellite precipitation dataset in an ungauged basin

Vinodh Kumar Sampath (✉ vinothkumar1995@gmail.com )
National Institute of Technology Tiruchirappalli

Nisha Radhakrishnan
National Institute of Technology Tiruchirappalli

Research Article

Keywords: SY, MUSLE, SDR, LULC, CHIRPS

Posted Date: March 1st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2633330/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations:
No competing interests reported.

Tables are available in Supplementary Files section.
Sediment yield modelling using SDR and MUSLE with high resolution satellite precipitation dataset in an ungauged basin

Vinoth Kumar Sampath¹, Nisha Radhakrishnan¹

¹ Department of Civil Engineering, National Institute of Technology Tiruchirappalli, Tamil Nadu, India

* E-mail: vinothkumar1995@gmail.com

Abstract

Erosion has become one of the extreme ecological dangers in up to date. Decrease of minerals in the upper layer of soil leads to failure in sustainable crop practices. Many researchers have developed prediction models of Sediment Yield (SY) in gauged basins. But modelling for an ungauged basin is very difficult due to the difficulty in validating the predicted model. The primary goal of the study was to identify the sedimentation in the study area using multiple(SY) methodologies, including Sediment Delivery Ratio (SDR) and Modified Universal Soil Loss Equation (MUSLE), in an ungauged basin with a high-resolution satellite precipitation dataset. Different attributes such as LULC (land use / land cover), soil texture, precipitation, topography, etc. was incorporated to estimate the SY in Ponnaniyar river basin. The generated SY map from SDR and MUSLE was evaluated by the receiver operating characteristic curve (ROC). The SDR model was found to be an efficient method for determining the SY for Ponnaniyar river basin, and also satisfied the criteria of AUC value of 0.752. The severely affected sub-watershed of Ponnaniyar river basin was identified with the help of erosion and yield spatial map. The obtained results will help prioritize the sub-watershed for locating water harvesting structures in further studies. This study suggests placing the gauging station in Ponnaniyar river basin to monitor the daily observation of discharge and SY estimation to prevent soil loss during flash flooding.

Keywords: SY, MUSLE, SDR, LULC, CHIRPS

1. Introduction

Erosion is a naturally occurring phenomenon that substantially degrades the environment and crop productivity and has a negative effect on the socioeconomic and habitat progress of the area under consideration. Nevertheless, compared to natural causes, human activities such as overgrazing, deforestation, and other land conversions have significantly altered the rate of soil erosion (SE) (Pradhan 2010). An enormous danger to a region’s economy comes from SE caused by the agricultural industry (Sartori et al. 2019). The removal of top soil layer leads to nutrition loss, making the soil unsuitable for agriculture and thereby affecting crop production stability. Intensive agriculture practice would also trigger SE and lead to soil fertility loss (Rahmati et al. 2016, and Abdo and Salloum 2017). To solve the aforementioned issues, conducting best practices that can address identifying soil erosion-prone areas at the sub-watershed scale was important for its evaluation (Alexakis et al. 2013; Patel et al. 2022).

While the word "soil erosion" often refers to erosion that takes place inside a watershed, "sediment yield" refers to the amount of soil that the watershed has left over. Sediment Yields (SY) from catchment restricts the sustainable utilization of natural resources. Predicting the SY is crucial because it influences the factors like reservoir storage, watershed characteristics, and channel shape, as well as watershed health and efficacy of watershed management practices (Sadeghi et al. 2007, Noor et al. 2013). The hydrometric station (gauging station) is provided at the outlet of the watershed to monitor the daily SY and discharge of water. Only a few gauging stations, managed by the Central Water Commission (CWC), India, are available in Tamil Nadu State for significant watersheds. Researchers have created various modelling techniques such as USLE (Universal Soil Loss Equatio
Equation), MUSLE (Modified Universal Soil Loss Equation), and RUSLE (Revised Universal Soil Loss Equation) to determine the erosion and SY in catchment scale due to the lack of actual sediment loss and yield data in ungaged basins or inaccessible regions. Initially, the loss of soil rate was estimated using Universal Soil Loss Equation (USLE) developed by Wischmeier and Smith 1978. There are some limitations in the USLE method's calculations of SE, which don't include tillage, direct runoff, stream channels, or gullies. Only slopes with low gradients are eligible for this USLE. (Ketema and Dwarakish 2021). In order to overcome the drawbacks of USLE, Renard et al. 1997 developed the Revised Universal Soil Loss Equation (RUSLE). The RUSLE model has altered the USLE model parameters such as R Factor (rain erosivity), K Factor (soil erodibility), C Factor (cover management), and P Factors (agriculture practicing). In USLE model, Length (L) and Steepness (S) factors are calculated separately, but in RULSE model, both factors are combinedly calculated. Geographical information systems (GIS) and remote sensing (RS) have developed into very promising tools that can be used to analyse different hydrological modelling applications, such as watershed delineation (Jones et al. 1990), groundwater potential assessment (Abijith et al. 2020), hydrologic extreme events, such as flood (Lyu et al. 2018) and drought (Tsesmelis et al. 2022), sediment yield (Ahmad et al. 2022). To calculate the amount of sediment discharged through the catchment's outflow, the RUSLE-based SDR (Sediment Delivery Ratio) approach is used, which also reduces the overestimation of sediment output. The SDR is an amount of highland excess erosion that is carried from a certain place. This SDR is determined by various equations such as Vanoni equation (Vanoni 1975), USDA (USDA 1979), Renfro equation (Renfro 1975), Maner equation (Maner 1958).

MUSLE innovated by Williams 1975 determines the SY based on Runoff Volume, Peak Discharge, Land use / Land Cover (LULC), topography. In MUSLE method, the runoff and discharge factor was used instead of the kinetic energy of rainfall value. The impact of rainfall, runoff removes the top soil layer and it transfers the particles from one place to another and finally discharged at outlet of station (Jain and Das 2010). The SCS-CN method was applied for the determination of runoff volume and peak discharge in the MUSLE model to estimate the SY of catchment USDA (1979).

From various literature, the prediction of SY under different landforms and the climatic condition is done with MUSLE and SDR models and validated with the hydrometric station (gauged station). But determining the SY in an ungaged basin is very complicated due to the validation of prediction model outcomes. Also, identifying severely affected zones in the study area is quite difficult and complex. In order to resolve this fore mentioned conflict, the river basin is further divided into sub-watersheds (SW) used to evaluate the soil erosion (SE) and sediment yield (SY) in detail. This sub watershed is a subunit of the whole catchment and is also used for better understanding the factors that influence the erosion and yield process (Surya and Mudgal 2012). This methodology reduces erosion and sedimentation by constructing water harvesting structures.

The primary goal of this study was (i) To develop the factors of SDR and MUSLE models in the Ponnaniyar river basin (study area) with satellite precipitation dataset; (ii) To develop the SY map using SDR and MUSLE in spatial aspect; (iii) To validate generated SY map by using ROC / AUC (iv) To identify the severely affected SWs in Ponnaniyar river basin.

Due to data scarcity in the Ponnaniyar river basin, the satellite precipitation dataset is utilized for SY modelling. Satellite precipitation dataset is a gridded form of rainfall dataset available from TRMM, CHIRPS, IMD, CPC, PERSIANN_CDR, MSWEP with resolution of 0.25° x 0.25°, 0.05° x 0.05°, 0.25° x 0.25°, 0.5° x 0.5°, 0.25° x 0.25°, 0.1° x 0.1° (Reddy and Saravanan 2022). Among the gridded dataset, the high-resolution CHIRPS is selected, which is almost equal to the average rainfall of the study area. There are only a few applications of the gridded dataset in SY modelling (Singh and Saravanan 2022; Ijaz et al 2022; Magginoni and Massari 2018). Also, the precipitation dataset is studied in gauged station region. The study's
innovative endeavour was to use the high resolution satellite precipitation dataset (CHIRPS) to estimate the SY in an ungauged basin using MUSLE and SDR technique and to identify the Sub Watersheds (SW) in the study region that were most adversely impacted.

2. Description of Study Area

The Ponnaniyar river basin is one part of Cavery river basin situated in Trichy district, Tamil Nadu, India. This study area covers 1803.53 sq.km geographical area. The basin extends from latitude of 10°10'00"N to 10°50"00"N and departure of 78°10'00"E to 78°50'00"E. The relief of the river basin lies from -42 m to 674 m. The average lowest and highest temperature lies between 23°C and 38.56°C (Abijith et al. 2020). The major soil texture found in the study area is clay, sandy loam, loamy sand and sandy clay loam. The major crop practices in the study area is paddy, banana, sugarcane, millets and pulses. Many studies have been done in the Ponnaniyar river basin, such as soil erosion, site suitability analysis of maize and millets, flood hazard delineation, and impacts of climate change on paddy crops (Sampath and Radhakrishnan 2022, Sabareshwari 2018, Sabareshwari and Basker 2018, Natarajan and Radhakrishnan 2021, Sivaraj 2016). But, studies on SY in the Ponnaniyar river basin have not yet been conducted. For a detailed analysis of erosion and yield at the micro and macro level, the basin was delineated into 13 sub-watersheds (SWs) i.e. SW 1, SW 2, SW 3,……, SW13 by using terrain analysis in ArcGIS platform with DEM data (Natarajan and Radhakrishnan 2021).

Figure 1 Location of Ponnaniyar river basin

3. Datasets and Methodology

3.1 Datasets

The datasets for SY modelling includes Boundary map of basin (iamwarm.gov.in) for extracting the boundary of watershed and features; satellite data such as Landsat 8 OLI/TIRS (multispectral satellite image) of date 24.12.2021 with 30 m resolution obtained from USGS earth explorer and with path and row being 143 and 53 respectively which would be used for Land Use / Cover preparation, Cartosat -1, a high spatial resolution image (Digital Elevation Model) acquired from NRSC India applied for analysis of fill, flow direction and accumulation; Soil Map of 1 : 50000 scale from FAO data bank (Food and Agriculture Organization) to classify the Hydrological Soil Group (A,B,C and D); Rainfall data obtained from CHIRPS (Climate Hazards Group Infrared Precipitation with Station) with resolution of 0.05° x 0.05° i.e. 5km x 5km used to calculate erosivity factor, runoff depth, volume and peak discharge for SY modelling. The Sub Watersheds (SW) of Ponnaniyar river basin has been created by using Terrain Processing Methods with DEM data in the ArcGIS platform.

3.2 Methodology

The methodology of SY modelling of Ponnaniyar river basin is represented in the flow chart shown in Figure 2.

Figure 2 Methodology of Study

3.3 Land Use / Land Cover and Accuracy Assessment

One of the factors considered in SY modelling is land use. For the study area, LULC classes such as forest, settlement, vegetation, Barren land, agriculture, and water bodies were classified using Landsat – 8 data from USGS (30m resolution) by random trees classifier (RTC). A type of decision tree based on machine learning (ML) technique that can conduct both regression and classification in various types of applications was chosen to handle problems (i.e., ensemble learning) utilising Random Trees.
The classification of LULC in Ponnaniyar river basin was done by Level 1 classification under Indian ecosystem defined by National Remote Sensing Agency (NRSA 2006). The accuracy assessment was made to validate the classified satellite image with real-time earth features extracted by Google Earth Pro software. In this accuracy assessment, various statistical methods involved, such as overall accuracy, and kappa coefficient was applied to validate the classified LULC imagery.

### 3.4 Revised Universal Soil Loss Equation (RUSLE)

The annual soil loss of Ponnaniyar river basin in 2021 was determined using the RUSLE model. Equation 1 is applied to determine annual SE

\[
A = R \times K \times L \times S \times C \times P
\]  

Where, \(A\) expresses the yearly soil erosion of given area in tons ha\(^{-1}\) yr\(^{-1}\), \(R\) Factor expresses the rainfall erosivity in mm/ha/h/yr, \(K\) Factor expresses the erodibility of soil in tons ha/h/MJ/ha mm, \(LS\) Factor expresses the geographical of the study area, \(C\) Factor expresses the land management in catchment and \(P\) Factor expresses the type of agriculture practicing. In \(C\), \(P\) and \(LS\) factors are unitless.

#### 3.2.1 Rainfall Erosivity Factor (R Factor)

Erosion is triggered by the precipitation’s abrupt effect on the soil. Estimating erosivity, erosion, and sedimentation has become a complicated task as a result of the limited number of rain gauges with discontinuity of information. Hence, a satellite-based precipitation dataset was used for the analysis. High resolution (5km x 5km) satellite precipitation data (CHIRPS) was used for R factor calculation. Open-source satellite precipitation data from CHIRPS can be compared to local rain gauge readings in the research region and is almost as accurate. It is available daily, monthly, and yearly in gridded dataset format. In R factor analysis, the monthly and yearly dataset available for 2021 was used for the analysis. Initially, the gridded dataset was transformed into point data, and its attributes i.e., precipitation data were used to determine the R factor. Finally, the R value was calculated with Arnoldus erosivity formula (Arnoldus 1980) then it was interpolated using Inverse Distance Weightage (IDW), an interpolation technique used to represent spatially. Equation 2 was used for the erosivity calculation.

\[
R = 1.735 \times 10^{(1.5\log_{10}\sum_{i=12}^{n} p_i^2 - 0.8188)}
\]  

Where, \(P_i\) is monthly rainfall in mm, \(P\) is annual rainfall in mm

#### 3.2.2 Soil Erodibility Factor (K Factor)

Soil Erodibility Factor is a major factor in determining SE. The variation in SE rate depends on the percentage of sand, silt, clay, organic matter, and type of soil texture. In clay soil, the greater the plasticity, the lower the rate of erodibility. In comparison to silt, sandy soil has less erodibility rate. All types of soil have some degree of erodibility, although those with more silt have a higher degree of erodibility. This erodibility factor is estimated using Equation (3) and (4) acquired from Wischmeier and Smith 1978. The structure and permeability code developed by Tran et al. (2011) was adopted to calculate the K factor in the SE model. The percentage of silt, sand, clay, and organic matter was derived from Food and Agriculture Organization (FAO) data hub. The soil texture categorization was made in Ponnaniyar river basin based on United States Department of Agriculture (USDA 1979) guidelines and applied for the erodibility analysis.
\[ K = \frac{1}{759.4} \left(2.1 \times 10^{-4} (12 - OM) + M^{1.14} + 3.25 (s - 2) + 2.5 (p - 3)\right) \]  
(3)

Where, \( M \) is represented as the particle size fractions in m

\[ M = (\% \text{ of } silt + \% \text{ of } very \text{ fine } sand) \times 100 - \% \text{ of } clay \]  
(4)

Where, \( s \) and \( p \) represents the structure code and permeability code of the soil, OM represents the organic matter of soil respectively; \( s \) and \( p \) are dimensionless.

### 3.2.3 Length and Steepness Factor (LS Factor)

This LS Factor relates the topography of study area which influence the erosion and transport of soil by the combined effect of slope and length. Topography with higher slope value tends to have a high erosion rate and vice versa (Srinivasan et al. 2020).

To generate LS Factor map, the main parameters utilized were the slope of the basin, flow accumulation, and flow direction, which was determined using Cartosat-1 Digital Elevation model (DEM) created in the ArcGIS platform. Equation 5 was used for LS factor analysis

\[ LS = \left(\frac{\text{Slope Length}}{12.5}\right)^{0.4} \times \left(\frac{0.0174 + \sin \theta}{0.0896}\right)^{1.4} \times 1.4 \]  
(5)

Where, Length of gradient is calculated using pixel size of elevation data, and the gradient represent in degree.

### 3.2.4 Cover Management Factor (C Factor)

Land Management Factor (C Factor) represents the interconnection between the SE and land use/cover. The value of the C factor lies between zero and one, where the value ‘1’ expresses a high-risk zone and the value ‘0’ expresses a risk-free zone of erosion.

According to the LULC classification, the vegetation classification was given a C value of 0.015 (FAO 1977), the agriculture classification was given a C value of 0.6 (Ganasri and Ramesh 2016), the forest classification was given a C value of 0.05 (Wang et al. 2016), and the barren land and water bodies were given C values of 1 and 0, respectively. (Erdogan et al. 2007). The study adopts highly accurate LULC map generated by random trees classifier, a machine learning-based technique for developing the C factor map.

### 3.2.5 Support Practice Factor (P Factor)

P factor represents the effect of agricultural practice triggering erosion. The P factor value was generated from the LULC and gradient of the catchment. Generally, P factor value lies from 0.1 to 1. The highest value of P factor 1 was assigned to non-agricultural land such as barren land, forest, settlement, vegetation, water bodies, and agriculture land with a slope greater than 100%. The rest of P value in agriculture land was given for different slope lies from 0 to 5, 5 to 10, 10 to 20, 20 to 30, 30 to 50, 50 to 100 % as 0.1, 0.12, 0.14, 0.19, 0.25 and 0.33 as adopted from Wishchmeier and Smith (1978).

### 3.3 Sediment Delivery Ratio (SDR)

In RUSLE model, it is not possible to quantify directly the total amount of SY reaching the outlet of the basin due to a few quantities of sediment deposited in the basin itself (Wu et al. 2013, Kumar et al. 2014, Da Cunha et al. 2017, Gupta and Kumar 2017, Vatandas and Yavuz 2017). The process of SDR is transporting the total sediment to the outlet of the watershed by the action of runoff water. The SDR lies between 0 and 1, always less than the erosion rate (Fernandez et al 2003). The SDR is
identified using an equation 6 proposed by Vanoni 1975. In various past studies (Ben Cheikha et al. (2021); Banasik (2021); Bhattacharya et al. (2020); Woznicki and Nejadhashemi (2013); Fernandez et al. (2003)), the Vanoni equation is almost equal to SY from gauging stations.

\[ SDR = 0.42 A^{-0.125} \]  

(6)

Where, SDR denotes Sediment Delivery Ratio (no units), and A is the area of the catchment (sq.km).

### 3.5 Sediment Yield (SY) Estimation

The SY is interconnected to SE but differs from SE. The term erosion is defined as the process of detachment and shifting the particles from one location to another location (Fenta et al 2016; Vatandaslar and Yavuz 2017). The soil fragments that break off from flat fields with little or no surface runoff move only a small distance and are not carried to the outlet of the watershed (Kumar et al. 2020). The yearly SY was identified by the yearly rate of SE and sediment delivery ratio. Equation 7 (Fistikoglu and Harmancioglu 2002) was used to identify the yearly SY of the catchment.

\[ SY = \sum_{i=1}^{n} SDR \times SE \]  

(7)

Where, SY denotes the sediment yield of the basin (tons/ha/yr), SE denotes annual soil loss (tons/ha/yr)

### 3.4 Modified Universal Soil Loss Equation (MUSLE)

In order to empirically relate the rainfall period, the soil yields to highland erosion rate evidenced by USLE (Wischmeier and Smith 1978) erosion elements and the transit effectiveness of runoff water revealed by the mechanism of total rainfall runoff volume and peak runoff rate, the MUSLE model was developed (Williams 1972; Williams 1975). For small and semi dry rangeland watersheds, the return time related to peak discharge is investigated with respect to watershed runoff parameters, and the significance of major rainfall-runoff events on long duration SY is explored (Jain and Das 2010). Only in a large river basin can the yearly sediment production by USLE be calculated by multiplying the projected annual soil loss along with SDR. The following MUSLE Equation (Williams and Berndt 1974) was used to model the SY of the study area.

\[ SY = 11.8 \left( Q_v \times Q_p \right)^{0.56} \times K \times LS \times C \times P \]  

(8)

Where, SY denotes Sediment Yield (tons/yr), Qv denotes the total volume of runoff (cu. m), Qp denotes peak discharge of the watershed (cumec), and the rest of the variables is similar to the RUSLE model.

#### 3.4.1. Calculation of Runoff Volume and Peak Discharge

The SCS – CN (Soil Conservation Service Curve Number) under AMC (Antecedent Moisture Condition) based on water balance Equations 9, 10, and 11 was used to determine the runoff depth. These equations developed by USDA (United states Department of Agriculture) estimate the runoff volume and peak discharge using the parameters such as daily rainfall, LULC, HSG (Hydrologic Soil Group), and runoff depth.

Water balance equation:

\[ P = I_a + F_c + Q \]  

(9)
Proportional equality hypothesis:
\[ \frac{Q}{P-I_a} = \frac{F_c}{S} \]  

(10)

Hypothesis:
\[ I_a = \lambda S \]  

(11)

Where, P denotes the total precipitation (mm), Ia denotes Initial abstraction, Q denotes the direct runoff (m³), S represents the maximum runoff potential (mm), Fc denotes the cumulative infiltration, and it excludes Ia. Runoff depth was calculated using the S and Ia, while Ia = 0.2S i.e., \( \lambda = 0.2 \) from Equation 11 under Indian circumstances. Using the SCS-CN (now known as National Resource Service Curve Number (NRCS)), the runoff depth is calculated using Equations 12 and 13.

\[ Q = \frac{(P-0.2S)^2}{(P+0.85)} \text{ if } P \geq I_a \]  

(12)

\[ Q = 0 \text{ if } P < I_a \]  

(13)

The maximum surface runoff potential was calculated using Equation 14 and 15
\[ S = \frac{2540}{CN_w} - 254 \]  

(14)

\[ CN_w \text{ or } CN_{II} = \frac{\sum (CN_i * A_i)}{\sum A} \]  

(15)

Where CN_w / CN_{II} represents the Weighted Curve Number (CN) of sub watershed and CNi represents Curve Number of hydrological soil group and LULC type on overlayed map, A_i represents the area with CN_i and \( \sum A \) represents the total area of sub watershed (sq.km).

The weighted CN wereis calibrated for dry (I) and wet condition (III) under AMC to reduce the uncertainty in the runoff calculation. The conversion method for CN_I and CN_{III} was done using the Sobani formula (Sobani 1976) and Hawkins formula (Hawkins et al. 1985) as shown in Equations 16 and 17, respectively.

\[ CN_I = \frac{CN_{III}}{2.334-0.01334 CN_{III}} \]  

(16)

\[ CN_{III} = \frac{CN_{II}}{0.427+0.00573 CN_{II}} \]  

(17)

\[ Q_p = \frac{0.208 * A * Q^{0.5}}{D + 0.6} \]  

(18)

Where, Qp represents peak discharge (m³/s), A denotes as area of watershed (km²), Q represents direct runoff (mm), D represents duration of rainfall (hrs), and it is assumed to be 24 hrs due to uneven rainfall in monsoon day in this study area (Kamath et al 2011).

The time of concentration was used to calculate the peak discharge of the sub watershed and was determined by using the lag method that is shown in Equation 19 (Folmar and Miller 2008).

\[ t_c = \frac{L^{0.8} (S+1)^{0.7}}{1140 + Y^{0.5}} \]  

(19)
Where, \( tc \) is the time of concentration (hrs), \( L \) is the flow length (ft), \( S \) is the surface potential (inch), and \( Y \) is the average slope (\%) of sub watersheds.

4. Validation

The Receiver Operating Curve (ROC) / Area Under Curve (AUC) provides a mathematical characterization of how well the stochastic deterministic identification and prediction system performs. Based on Google Earth Pro and field observation in erosion susceptible zones of the study area, randomly 25 erosive points, i.e., True positive point (TP), and 25 non-erosive points, i.e., True negative point (TN) were selected, and created in ArcMap software to plot the ROC / AUC curve with the help of ArcSDM (Statistical Data Modular), an extension tool of ArcMap software. This ROC / AUC plot was used to identify the level of accuracy on a developed spatial map of SY by MUSLE and SDR model. Generally, the value of AUC differs from 0 to 1; the value 0 or nearest to 0 represents low level accuracy, while the value 1 or nearest to 1 represents high level accuracy. From Mandrekar 2010, the level of accuracy in ROC / AUC was identified in five levels of agreement as excellent with lies between 0.9 and 1.0, very good with lies between 0.8 and 0.9, good with lies between 0.7 and 0.8, satisfactory with lies between 0.6 and 0.7, unsatisfactory with lies between 0.5 and 0.6 respectively. Equations 20 and 21 determine the TPR and FPR.

\[
\text{True Positive Rate (TPR)} = \frac{TP}{(TP+FN)} \tag{20}
\]

\[
\text{False Positive Rate (FPR)} = \frac{TN}{(TN+FP)} \tag{21}
\]

Where, TP indicates True Positive, TN indicates True Negative, FP indicates False Positive, and FN indicates False Negative.

5. Results and Discussion

5.1 Analysis of Land Use / Land Cover

LULC is classified using random trees classifier, a supervised machine learning classifier. In the Ponnaniyar river basin, the overall classified LULC shows that the major domination in this region is covered by vegetation category, followed by agriculture and barren land with 48.09%, 18.94%, and 10.99% respectively. The least land use/land cover was found to be settlement followed by forest and waterbodies represented as 6.41%, 7.7% and 7.87%. The classified image is shown in Figure 4. From Table 1, in barren land classification, major coverage was found in SW 2 followed by SW 11 and SW 9 with 35.95%, 24.21%, and 11.04%, while the lowest was found in SW 7 followed by SW 6 and SW 13 with 0.10%, 0.17%, and 0.42%. In forest classification, the highest coverage was found in SW 5 and SW 4, with 51.78% and 27.58%. In SW 7, there is no forest cover found in the sub watershed. The least coverage was found in SW 6 with 0.23%. In vegetation classification, major dominance was found in SW 10 with 19.77%, followed by SW 4 with 13.80% and SW 5 with 12.13%. A minor dominance was found in SW 7 with 0.69%. In settlement classification, SW 13, SW 2, and SW 10, with 18.41%, 14.00%, and 12.27% had the highest coverage, while SW 7 with 1.35% was the least coverage. Moderate coverage of settlement was found in SW 4 and SW 5 with 8.21% and 8.55%. In forest bodies classification, SW 10 followed by SW 2 and SW 11 with 20.20%, 12.55%, and 10.84%, was found to be of highest coverage, while the lowest coverage was found in SW 7 with 1.29%. Moderate coverage of water bodies was found to be in SW 8 and SW 9 with 7.94% and 8.28%, respectively. In agriculture, SW 2 and SW 5, with 27.13% and 12.80% were found to have major dominance. Moderate and lowest coverage was found in SW 4 and SW 7 with 5.08% and 0.62%. The barren land and agriculture category has a major contribution in triggering the erosion and yield in the region. SW with high coverage of barren land and agriculture category denotes high possibility of erosion and sedimentation, while the SW
with high coverage of settlement and forest category denotes the low possibility of erosion and sedimentation. In water bodies
classification, there is no occurrence of SE. In vegetation classification, it prevents direct impact of raindrop impacts and reduces
the kinematic action of water. In accuracy assessment, the overall accuracy and kappa coefficient were estimated as 87% and
0.84, which has a good agreement in classification. The higher accuracy of LULC improves the accuracy of SE and SY.

Table 1 Area and its Percentage of LULC in Each Sub Watersheds

5.2 Parameter of RUSLE

In the erosivity factor (R Factor), the rainfall data in the form of a gridded dataset with a resolution of 5km x 5km (CHIRPS) in
monthly and yearly data is transformed into a shape file, i.e., point data and the erosivity was calculated using the Arnoldus
equation. The IDW technique was applied in the ArcGIS software to represent the erosivity value spatially. The highest and
lowest erosivity in the basin scale was 1242.35 mm/ha/h/yr and 895.93 mm/ha/h/yr. The overall mean erosivity of the watershed
was found to be 1044.45 mm/ha/h/yr. From Table 2, the highest mean erosivity was found in SW 5 followed by SW 4 and SW
11 with 1156.139 mm/ha/h/yr, 1115.45 mm/ha/h/yr, while the lowest mean erosivity was found in SW 13 followed by SW 12
and SW 11 with 946.745 mm/ha/h/yr, 955.876 mm/ha/h/yr and 958.130 mm/ha/h/yr. The moderate erosivity was found in SW
6, SW 3, and SW 10 with 1061.861 mm/ha/h/yr, 1066.662 mm/ha/h/yr and 1079.565 mm/ha/h/yr respectively (Figure 4a). In
this erodibility factor (K Factor), the soil map was download from Food and Agriculture Organization (FAO). Based on the
percentage of clay, silt, sand, and organic matter, the soil texture was classified as sandy loam, loamy sand, clay, and sandy clay
loam under the USDA texture classification. The erodibility of soil lies from 0.05288 to 0.09438 tons ha h/ha/MJ/mm (Figure
4b). The mean erodibility of the Ponnaniyar river basin was found to be 0.0798 tons ha h/ha/MJ/mm. From Table 2, the highest
mean erodibility was found in SW 2 with 0.0916 tons ha h/ha/MJ/mm, while the lowest mean erodibility was found in SW 3
followed by SW 8 with 0.0592 tons ha h/ha/MJ/mm and 0.0645 tons ha h/ha/MJ/mm. SW 13, SW 4, and SW 5 with 0.0713 tons
ha h/ha/MJ/mm, 0.0729 tons ha h/ha/MJ/mm and 0.0744 tons ha h/ha/MJ/mm were found to have moderate erodibility. In SW
2, high erodibility was observed due to the presence of Loamy sand soil type, while the least erodibility in SW 3 was found to
be clay soil. The Length and Steepness/Topography Factor (LS Factor), one of the major parameters in the RUSLE model triggers
SE. This factor was calculated using Equation 5 with flow accumulation and percentage of slope data. The LS factor value lies
from 0 to 47.73. The mean LS value of the Ponnaniyar river basin was found to be 0.281. From Table 2, SW 6 with 0.536,
followed by SW 7 with 0.471, was found to have the highest LS value, while the least LS value was found in SW 8 at 0.273.
Moderate LS values were found in SW 1, 13, and SW 4 with 0.374, 0.341, and 0.304. More the steepness more will be erosion
rate. The spatial map of the LS factor is represented in Figure 4c. In C Factor, the C value depends upon the LULC of the study
area. The LULC is classified as water bodies, barren land, settlement, vegetation, agriculture, and forest. Generally, this C factor
value lies from zero to one. The C factor value 1 indicates higher erosion, i.e., barren land, and 0 indicates no erosion, i.e., water
bodies. The overall mean C value of this study area was found to be 0.233. From Table 2, the lowest C value was found in SW
4 and SW 13 with 0.104, while the C value was found in SW 11 with 0.512, followed by SW 2 with 0.442. Moderate value was
found in SW 5 and SW 7 as 0.165 and 0.166 respectively. The Northeast side of the study area, i.e. SW 2, SW 9, and SW 11
majorly, were found to be a barren land region (Figure 4d). In P Factor, the P value is designated based on information on LULC
and gradient with the assistance of RS and GIS tools. In the Ponnaniyar river basin, the P value lies between 0.1 and 1,
respectively. The mean value of the P factor of the basin was found to be 0.832. From Table 2, the highest average value of P
factor was found in SW 4 followed by SW 3 and SW 12 with 0.925, 0.904 and 0.885, while the moderate mean value of P factor was found in SW 5 and SW 6 with 0.848 and 0.879. The lowest mean was found in SW 2 with 0.715, followed by SW 11 with 0.737 and SW 9 with 0.795. The gradient and LULC play a major role in generating the P factor in the study area. The generated spatial map of P factor is shown in Figure 4e.

### Table 2 Values of RUSLE Parameters in Each Sub watershed

#### 5.3. Soil Erosion (SE)

In this river basin, the erosion lies between 0 and 2690.71 to tons ha\(^{-1}\) yr\(^{-1}\). The average SE of the Ponnaniyar river basin was found to be 1.95 tons/ha/yr. From Figure 4f, Soil Loss which is less than 5 tons ha\(^{-1}\) yr\(^{-1}\) denotes very low category, loss lies from 5 to 10 tons ha\(^{-1}\) yr\(^{-1}\) denotes low category, loss lies from 10 to 15 tons ha\(^{-1}\) yr\(^{-1}\) denotes moderate category, loss lies from 15 to 20 tons ha\(^{-1}\) yr\(^{-1}\) denotes high category while the loss lies beyond 20 tons ha\(^{-1}\) yr\(^{-1}\) denotes very high category. The highest mean soil loss was found in SW 11, followed by SW 2, SW 9 represented 5.65 tons/ha/yr, 5.06 tons ha\(^{-1}\) yr\(^{-1}\), and 3.98 tons ha\(^{-1}\) yr\(^{-1}\), while the lowest mean soil loss was found in SW 8 followed by SW 3 and SW 12 represented with 0.46 tons ha\(^{-1}\) yr\(^{-1}\), 0.68 tons ha\(^{-1}\) yr\(^{-1}\) and 0.72 tons ha\(^{-1}\) yr\(^{-1}\) respectively. Moderate mean soil loss was found in SW 1, SW 5, and SW 10 represented by 1.47 tons ha\(^{-1}\) yr\(^{-1}\), 1.46 tons/ha/yr and 1.41 tons ha\(^{-1}\) yr\(^{-1}\). SW 11, SW 2, and SW 9 were severely affected due to the highest cover of barren land in the sub watersheds. Also, these SWs having moderate erosivity value tend to have high SE and SY. Results of SW 2, SW 11 and SW 9 strongly show the high soil erodibility due to poor cover management. Also, the lowest P value was found in aforementioned SWs. The spatial map of SE of the basin is represented in Figure 4f.

#### 5.4 Sediment Yield (SY) by SDR

The sediment delivery ratio (SDR) of the Ponnaniyar river basin was determined from the Vanoni equation. The calculated SDR was used to predict the SY of the study area. The average SY of 13 sub watersheds ranged between 0.11 and 1.30 tons ha\(^{-1}\) yr\(^{-1}\). The mean SY of the Ponnaniyar river basin was found to be 0.43 tons ha\(^{-1}\) yr\(^{-1}\). From Table 6, the highest mean SY was found in SW 11 followed by SW 2 and SW 9 with 1.30 tons ha\(^{-1}\) yr\(^{-1}\), 1.06 tons ha\(^{-1}\) yr\(^{-1}\), and 0.96 tons ha\(^{-1}\) yr\(^{-1}\), while the lowest mean SY was found in SW 8 with 0.11 tons/ha/yr followed by SW 3 and SW 12 with 0.16 tons/ha/yr and SW 12 with 0.17 tons ha\(^{-1}\) yr\(^{-1}\). Moderate mean SY was found in SW 6, 7 and 10 with 0.30 tons/ha/yr respectively. From Figure 6a, the SY from SDR and MUSLE is always less than SE. SW 2, SW 9, and SW 11 have the highest SY and strong interference in erosion in these sub watersheds. The LULC has high interference in SY but has no direct causes.

#### 5.5 Sediment Yield by MUSLE

5.5.1 Curve Number (CN), Runoff Depth (Q) and Peak Discharge (Q\(_p\))

The Curve Number (CN) of the Ponnaniyar river basin was calculated based on the SCS – CN method (Currently NRCS i.e., Natural Resource Conservation Service) using the Hydrologic Soil Group (HSG), i.e., soil types B, C, and D and LULC categories such as Agriculture, Barren Land, Forest, Settlement, Vegetation, and Water bodies. Based on overlay analysis of HSG and LULC, the weighted CN is generated for each sub watersheds. The determined weighted CN is calibrated for dry and wet weather under antecedent moisture conditions (AMC). Generally, the CN value lies from 0 to 100. In this study area, the maximum CN value was 87 in SW 3, SW 6, SW 8, while minimum CN value was 62 in SW 2. The intermediate CN values were found in SW 13, SW 11, SW 09, and SW 04 with 75, 81, 82, and 83, respectively. The lower CN value represents the lower runoff potential
and vice versa. The generated spatial map of the weighted CN value is shown in Figure 5a. Runoff volume is a main criteria in
the MUSLE model for determining the yield of sediment. The runoff depth was based on CN values, i.e., CNI, CNII, and CNIII,
maximum surface potential (S), area of sub watershed, and daily precipitation data (P) by SCS CN method estimated for each
sub watersheds. After calculating the runoff depth of each SW, it is multiplied by the respective extent of the SW and the runoff
volume. The runoff depth lies from 38.69 mm to 130.98mm in this study area. The mean runoff depth of the basin was found to
be 74.01 mm. From Table 4, the highest mean runoff depth was found in SW 5 at 102.46 mm, followed by SW 4 at 83.47 mm
and SW 10 at 78.60 mm, while the lowest mean was found in SW 1 at 47.7 4mm. The moderate runoff was found in SW 2, SW
6, and SW 9, with 73.18 mm, 70.27 mm, and 72.11mm, respectively. The annual runoff depth of the Ponnaniyar river basin is
shown in Figure 5b. Peak discharge (Q_p) of each SW is calculated by the runoff volume (Q_v) and time of concentration (tc).
Different methods were used to determine time of concentration, such as kripich, lag time, william, synder, etc. Many of the past
studies followed the lag time method to determine the time of concentration (tc) in both gauged and ungauged basin. This study
determines the time of concentration with the lag time method. The peak discharge of this river basin lies between 0.33 m³/s and
1.35 m³/s. The mean peak discharge of the entire basin was found to be 0.73m³/s. From Table 4, the mean highest peak discharge
was found in SW 5, followed by SW 4 and SW 8 with 1.15 m³/s, 0.90 m³/s, and 0.89 m³/s, while the lowest mean was found in
SW 1, SW 12, and SW 2 with 0.45 m³/s, 0.46 m³/s, and 0.53 m³/s. Moderate mean discharge was found in SW 3, SW 7, SW 13
with 0.74 m³/s, 0.78 m³/s, and 0.77 m³/s, respectively. The peak discharge of the basin is shown in Figure 5c.

5.5.2 Annual Sediment Yield

The SY of the study area was identified using the MUSLE model along with influencing parameters such as runoff volume, peak
discharge, erodibility factor, topographic factor, land management, and agriculture practice factor. In this river basin, the SY
from MUSLE lies between 0 and 358.24 tons ha⁻¹ yr⁻¹. The average SY of the Ponnaniyar river basin, determined using the
MUSLE model, was 0.34 tons ha⁻¹ yr⁻¹. From the MUSLE model in Table 4, the highest mean SY was found in SW 6 with 0.73
tons ha⁻¹ yr⁻¹, followed by SW 4 and SW 5 with 0.54 tons ha⁻¹ yr⁻¹ and 0.53 tons ha⁻¹ yr⁻¹ while the lowest found in SW 13 with
0.18 followed by SW 12 with 0.20 tons ha⁻¹ yr⁻¹ respectively. Moderate mean SY was found in SW 7, SW 9, and SW 10 with
0.33 tons ha⁻¹ yr⁻¹ and 0.35 tons ha⁻¹ yr⁻¹. SW 5 has a high mean slope, which tends to high erosion, discharge, and sediment.
Figure 8 shows that the SE of 13 sub watershed is more than the SY estimated by MUSLE and SDR. The spatial map of SY map
by the MUSLE model is shown in Figure 6b.

6. Validation

The curve of ROC is determined to evaluate the accuracy level of the predicted SY map. From the ROC curve, the generated
yield map based on MUSLE and SDR was found to be good predictors. The results of ROC / AUC show that the accuracy of the
SY map of MUSLE and SDR is 0.645 and 0.768. The specific modelling techniques and validation of ungauged basin were still
not found yet. Many studies modelled the SY in gauged, but only a few modelled in the ungauged basin but validated with
minimum one-gauge station data. Another conceptual model was incorporated to evaluate the SY in the ungauged basin, which
was done by the regionalization concept. This concept defines the process by which transfer of information from gauged station
to an ungauged station, if the catchment is comparable i.e., similar hydrological properties (Razavi and Coulibaly 2013). Due to
dissimilarities of hydrological properties with the gauged station, this concept could not be incorporated into this study area. In
this study area, there is no gauge station found and the rain gauge station data have a discontinuous and restricted number of gauges. This study attempted different aspect view to validate the SY map based on soil susceptible locations to locate the sedimentation. The SDR model was found to be in good agreement with SY prediction in this region and has a good accuracy level in the ROC curve compared with MUSLE model. The main limitation of MUSLE model is that prediction is limited only for to SY and is not capable of estimating soil loss in the long term. Also, it is an event or storm-based model to predict the SY of the basin. RUSLE – SDR based model can apply to any region with different gradients, LULC, and climatic condition. The basic statistics of erosion and yield are shown in Table 3.

Table 3 Basic Statistics of Erosion and Yield

7. Conclusions

The Ponnaniyar river basin's study area was susceptible to high SE and SY. The primary motive of this research was to estimate the SY in the ungauged basin using various methods such as MUSLE and SDR models along with high resolution satellite precipitation dataset. The SY was estimated, validated, and represented in a spatial aspect. The use of satellite precipitation products estimate the yield in inaccessible locations, i.e., remote areas. SW 11, SW 2, and SW 9 were found to be severely affected by both SE and SY and also strongly interrelated with a high time of concentration, runoff depth, and CN value. From the ROC curve, the SDR with 0.768 has a good agreement than the MUSLE model with 0.645. In order to avoid the conjunction in priority of sub watershed between MUSLE and SDR model, it is found that RUSLE based SDR model is effective in the identification of SY in the ungauged basin. This study will be useful in constructing the water harvesting structures and provides the support system for policy makers. Also, this study helps locate the gauging station to observe the daily variation of discharge and SY to reduce the sudden impact of flooding and erosion.

Declaration

Author’s Contributions

Vinoth Kumar Sampath: Conception, Methodology Formulation, Investigation, Data Preservation, Original Manuscript Writing;

Nisha Radhakrishnan: Guidance, Monitoring, Evaluating, Formatting, approved the final Manuscript.

Funding

Stating that no funds is being allocated for this work

Availability of data and materials

Depending on valid requests, data, and information will be provided

Competing Interest

The authors would like to state that the findings and information from this investigative research project are independent of any conflicting interests
References


Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). Geoscientific Model Development Discussions, 7(1), 1525-1534.


Figure 1

Location of Ponnaniyar river basin
Figure 2

Methodology of Study
Figure 3

Land Use/Land Cover Map
Figure 4

(a) R Factor (b) K Factor (c) LS Factor (d) C Factor (e) P Factor (f) Annual Soil Loss
Figure 5

Parameters of MUSLE (a) Weighted CN (b) Runoff Depth (c) Peak Discharge

Figure 6
Sediment Yield (a) SDR (b) MUSLE

Figure 7

ROC Curve
Figure 8

Comparison of Soil Erosion and Sediment Yield

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.docx