Estimation of the soil erodibility factor of a peri-urban catchment by group method of data handling and enhanced particle swarm optimization method

Chander Kant [1] *, Mrinmoy Majumder [2], Dharmendra Kumar Tyagi [3], Ashish Prabhat Singh [3]

[1] PG student NIT Agartala Tripura
[2] Assistant Professor NIT Agartala
[3] PG Student NIT Agartala
Corresponding author E-mail: tchanderkant@gmail.com

Abstract

Soil erosion has become a major deterrent in any watershed management program. The erodibility of the soil from the river banks has degraded watersheds all over the world. That is why in any watershed development programmers' erodibility of soil becomes a significant design parameter. However, there is lack of efficient simulation model for estimation of soil erosion. The existing models are location sensitive and mostly empirical nature in the present investigation, the authors tried to estimate the soil erodibility factor of the USLE method with the help of Enhanced PSO. The data for development of model is generated by Normalized Design of Experiment method which assumes that maximum and minimum value can be represented by I and O respectively. The same model was developed with the help of GMDH also. As per the model matrices of GMDH model shows better reliability. The selected model was applied to predict soil erodibility factor for 21 no's of location in west Tripura region. From the prediction and comparison with the actual data it was found that the selected models have an accuracy of 99.8% in predicted model and 89.8% in case study.

Keywords: GMDH (Group Method of Data Handling); PSO (Particle Swarm Optimization); Soil Erodibility Factor; USLE (Universal Soil Loss Equation).

1. Introduction

In this paper we have estimated the soil erodibility factor K from Universal soil loss equation by using GMDH and PSO software so that value of this K can be used as a general value and case study is taken as west Tripura and annual rainfall data from Indian water portal and annual soil loss in ton per acre is taken from meteriologic department .Erosion is a natural geological process resulting from removal of uppermost soil by natural agencies like wind, water transporting them while some human intervention can significantly increase rate of erosion . It is one of major agricultural problem and also the major global environmental issues in present days. Erosion is generated by a combination of factors such as slope gradient, climate (e.g. long dry periods followed by substantial rainfall), unsuitable land use, land cover patterns (vegetation) and ecological catastrophes (like forest fires).The Universal Soil Loss Equation, developed by A.R.S scientists W. Wischmeier and D. Smith, has been the most widely accepted and utilized soil loss equation for over 30 years[1]Planned as a method to predict average annual soil loss triggered by sheet and rill erosion, the USLE is often criticized for its lack of applications[2]. While it can estimate long - term annual soil loss and monitor ecologists on proper cropping, administration, and safeguarding practices, it can only be applied for a specific location due to its empirical nature.
The USLE for calculate approximately average annual soil erosion is:

$$A = R \times K \times LS \times CP \qquad (1)$$

Where
A= average annual soil loss in (tons per acre per year).
R=Rainfall erosivity index (MJ mm/ha/hr./yr.).
K= erodibility factor.
LS = topographic factor L is for slope length& Sis for slope in meter.
C=cropping factor.
P = conservation practice factor.
The USLE has another concept of experimental significance, the unit plot concept. The unit plot is demarcated as the standard plot condition to determine the soil's erodibility [3]. These conditions are when the LS factor = 1 (slope 9% and length = 72.6 feet) where the area is crop-free and tillage is up and down slope and no conservation practices are functional (CP=1). So we can mark this as

$$K=A/R \qquad (2)$$

A simpler method to estimate K was presented by Wischmeier [1] which take account of the particle size of the soil, organic matter content, and soil structure and outline permeability. The soil erodibility factor "K" can be estimated from a nomograph if this information is known. The LS factors can easily be determined from a slope effect chart by knowing the length and gradient of the slope. The cropping management factor (C) and conservation practices factor (P) are more challenging to obtain and must be determined empirically from plot data. They are defined in soil loss ratios (C or P with / C or P without).As we can see that value of K is varying from place to place and in this paper we have tried

to generalize the value of K by using optimization software like GMDH and PSO.

2. Methods used

For this we have used the latest technique of optimization like GMDH and PSO and west Tripura is taken as case study for this paper it is an managerial district in the state of Tripura in India. The district headquarters are situated at Agartala. As of 2012 it is the most populous district of Tripura. While Tripura lies around between the north latitude of 22 degree 56' and 24 degree 32' and between longitude 91 degree 0' and 92 degree20' east. the West Tripura district lies about latitude 23 degrees 16' to 24 degrees 14' north and longitude 91 degrees09' east to 91 degree 47' east. The West Tripura District is surrounded by Bangladesh in the north and west by Khowai district in the east and by Sepahijala region in the south. The district headquarters is situated at Agartala, which is also the capital of the State.

2.1. Methodology in detail

GMDH shell
GMDH is a family of inductive algorithm for computer, which is based on mathematic modeling of multi-parametric database which makes it fully automatic. This method of Optimization was developed by Prof. Alexey G. Ivakhnenko at the institute of Cybemetics in Kyiv (Ukraine). This is a Polynomial neural network, which is 'Self Organizing' in nature, which means that in this method the connections between neurons are not fixed but rather are selected during training to optimize the network. It also select the layers automatically in the network, to produce the maximum accuracy, without any over fitting. And at the time of Training of data it also select the neurons from the pool of candidates and add it with hidden layers.

It is a data mining, discovery, system modeling, optimization and pattern recognition tool, which perform better than the classical forecasting algorithms as Single Exponential Smooth, Double Exponential Smooth, Back propagation neural Network[]

$$Y = A_0 + \sum_{i=1}^{m} \sum_{j=1}^{m} A_i X_i X_j + \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} A_i X_j X_k.$$

Particle Swarm Optimization (PSO)
Particle Swarm Optimization (PSO) is an evolutionary computation technique developed by Kennedy and Eberhart in 1995[4, 5].
In a PSO system, it starts with the random initialization of a population (Swarm) of individuals (particles) in the search space and works on the social behavior in the swarm. The position and the velocity of the ith particle in the d-dimensional search space can be represented as

$$X_i = [X_1, X_2, X_3 \dots\dots\dots\dots X_d] \tag{3}$$

$$\text{And } V_i = [V_1, V_2, V_3 \dots\dots\dots V_d] \tag{4}$$

Each particle has its own best position (P best)

$$P_i = [P_1, P_2, P_3 \dots\dots P_d] \tag{5}$$

Corresponding to the personal best objective value obtained so far at time t. The global best particle (g best) is denoted by P_g, which represents the best particle found so far at time t in the entire swarm. The new velocity of each particle is calculated as follows;

$$V_{i,j}(t + 1) = wv_{i,j}(t) + C_1 \text{rand}_1 \left(P_{i,j} - X_{i,j}(t)\right) + C_2 \text{rand}_2 \left(P_{g,j} - X_{i,j}(t)\right) \tag{6}$$
Where: j = 1, 2, 3 …………d

Where c_1 and c_2 are acceleration coefficients, w is inertia factor, rand_1 and rand_2 are two independent random numbers uniformly distributed in the range of [0, 1].
Thus, the position of each particle is updated in each generation according to the following equation:

$$x_{i,j}(t + 1) = x_{i,j}(t) + v_{i,j}(t + 1), \tag{7}$$

$$j = 1,2,3, \dots\dots\dots\dots\dots. d.$$

Generally, the value of each component in V_i can be clamped to the range $[-v_{max}, v_{max}]$ to control excessive roaming of particles outside the search space. Then, the particle flies toward a new position according to (8). This process is repeated until a user-defined stopping criterion is reached.
The procedure of standard PSO is summarized as follows.
1) Initialize a population of particles with random positions and velocities, where each particle contains d variables (i.e., d = n).
2) Evaluate the objective values of all particles; let P best of each particle and its objective value equal to its current position and objective value; and let g best and its objective value equal to the position and objective value of the best initial particle.
3) Update the velocity and position of each particle according to (i) and (ii).
4) Evaluate the objective values of all particles.
5) For each particle, compare its current objective value with the objective value of its p best. If current value is better, then update P best and its objective value with the current position and objective value.
6) Determine the best particle of the current swarm with the best objective value. If the objective value is better than the objective value of g best, then update g best and its objective value with the position and objective value of the current best particle.
7) If a stopping criterion is met, then output g best and its objective value; otherwise go back to (3).
Enhanced PSO
The enhanced PSO variant is based on the work of Fourier and Groenwold, as follows;
The max. Velocity of individual, travel by individual in one time step, to a certain space is given by:

$$v_0^{max} = \gamma(X_{UB} - X_{LB}) \tag{8}$$

If the best solution found does not exist in the whole swarm than it is assumed that the velocities are large and the algorithm cannot locate better solution due to Overshooting, for this reason both the inertia factor and maximum velocity are decreased as follows:

$$\left(P_g\right)\big|_k = OF(P_g)\big|_{k+1} \Rightarrow W_{k+1} = a\, W_k, V_{k+1}^{max} = \beta\, V_k^{max} \tag{9}$$

If the random velocity assign by the operator is moving away from the swarm then the operator is activated with a probability P_{cr} as follows:
The craziness operator assigns a random velocity vector to an individual resulting in its moving away from

$$\text{if } r < P_{cr} \Rightarrow \text{randomly assign } V_{k+1} \text{ with } 0 < V_{k+1} \leq V_{k+1}^{max} \ \forall \text{ particle d} \tag{10}$$

Where, r denotes a random variable, whose distribution is uniform in the interval [0, 1].
Both the elite particle and an elite velocity are employed by the algorithm. And in this algorithm the worst performance is moved towards the best position of swarm due to which there is a gradual shift towards the region where good solution reside:
$$X^{pe} = P_g$$

In addition of this, if velocity vector V_k^d resulted in an improvement of P_g then:

$$X_{k+1}^d = P_g + C_3 r_3 \times V^{pe} \tag{11}$$

Where, $V^{pe} = V_k^d$, C_3 is a parameter of the algorithm and C_3 represents a vector of random variables, which is distributed uniformly in the interval [0, 1].

2.2. Case study

West Tripura was taken as a case study for this paper rainfall data of year 2002 was taken from India Water Portal official site and value of average annual soil loss is taken from meteorological department and further this site is thoroughly studied, the value of soil erodibility factor was found to be as shown in following table Variation of erodibility factor in West Tripura

Table 1: Table Showing Values of Soil Erodibility Factor of Different Location of West Tripura

| LOCATION	A (TON PER HAC PER YEAR)	R(MM)	K
JOYSINDHUBARI	48.03	2598	0.018
MAICHARI	35.74	2598	0.013
TRIMURBARI	48.24	2598	0.018
DURAI CHHARA	29.53	2598	0.011
NAPUI CHHARA	146.42	2598	0.056
LOTMA CHHARA	156.56	2598	0.060
KATLUTM CHHARA	164.91	2598	0.063
MAHARANI CHHARA	132.46	2598	0.050
DATU CHHARA	46.68	2598	0.017
KULAI CHHARA	104.54	2598	0.040
TAPHIAUNGMA CHHARA	34.36	2598	0.013
LAMPHA CHHARA	26.64	2598	0.010
BAHWRA CHHARA	223.74	2598	0.086
CHANDRAL CHHARA	188.4	2598	0.072
CHLINGMA CHHARA	186.56	2598	0.071
JAMTHUM CHHARA	65.64	2598	0.025
SURMA CHHARA	67.01	2598	0.025
SONAROY CHHARA	72.26	2598	0.027
HALAM CHHARA	198.98	2598	0.076
ABHANGA NALA CHHARA	246.88	2598	0.095
BALRAM CHHARA	63.48	2598	0.024
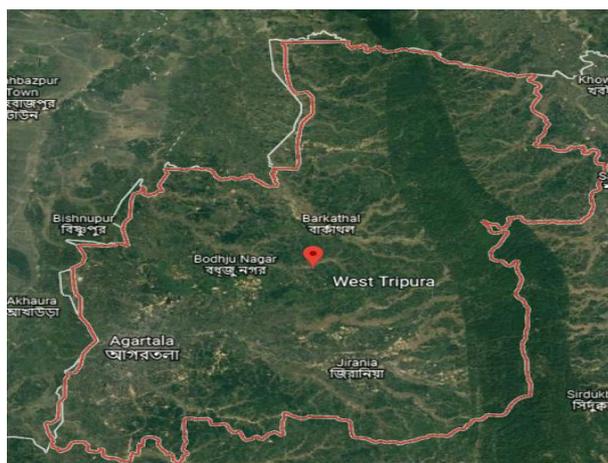
Fig. 1: Showing the Location of Related Study Area.
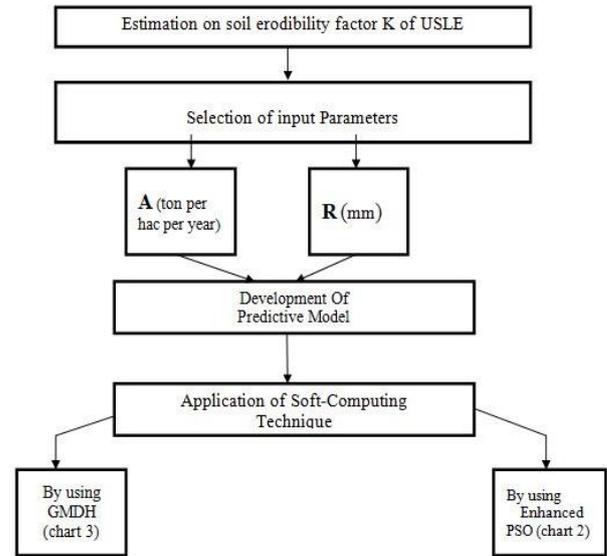
3. Model flowchart diagram

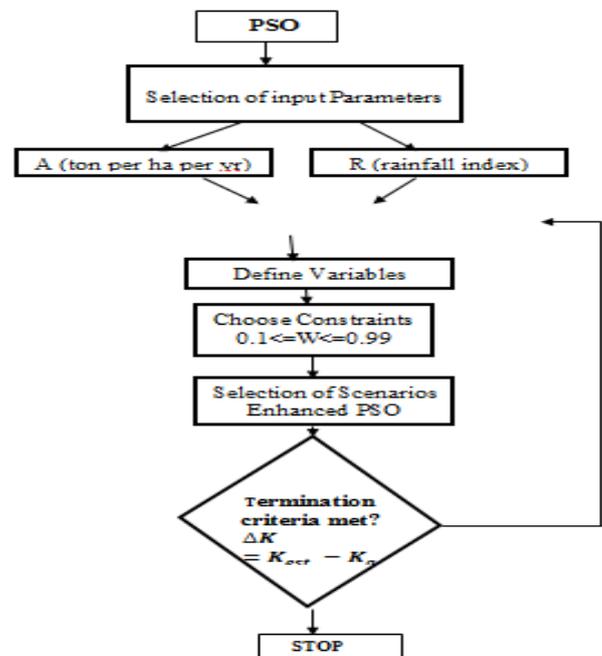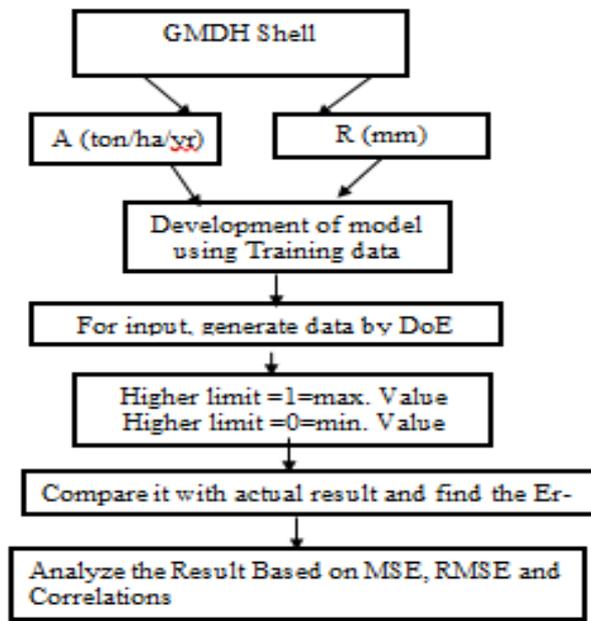**Chart. 1:** Flowchart Showing Soft Technique Used for Modeling.

3.1. Model for PSO

Chart. 2: Flowchart Showing PSO.

3.2. Model for GMDH

Flowchart 3: Chart Showing Stepwise Working of GMDH.

4. Result and discussion

Model that we have developed using data determination is now compared with each other and results are obtained.On plotting the actual predicted data with GMDH and PSO data following plots are obtained

Following parameters are used for the error estimation

- Absolute error (AE),
- Relative error (RE) (%) and
- Mean-square error (MSE)

These are defined as:

Absolute Error = |actual value − predicted value|

$$\text{Relative Error} = \frac{|\text{actual value} - \text{predicted value}|}{\text{Actual value}}$$

$$\text{Mean Square Error Error} = \frac{\sum_{i=1}^{N}(\text{actual value} - \text{predicted value})^2}{N}$$

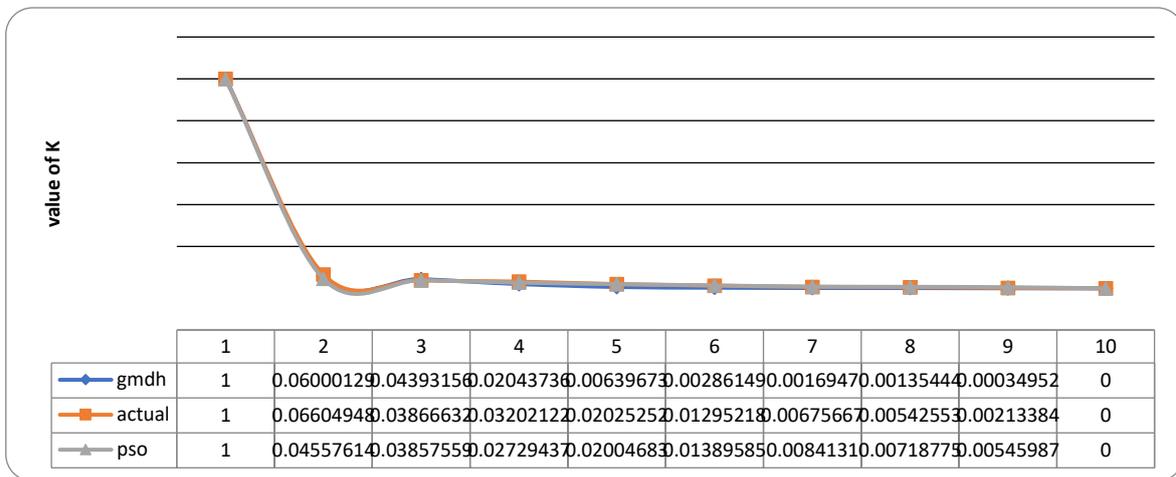Where, the term i denotes the ith term of data and N denotes total no

	1	2	3	4	5	6	7	8	9	10
gmdh	1	0.0600012	0.0439315	0.0204373	0.0063967	0.00286149	0.0016947	0.0013544	0.00034952	0
actual	1	0.0660494	0.0386663	0.0320212	0.0202525	0.01295218	0.0067566	0.0054255	0.00213384	0
pso	1	0.0455761	0.0385755	0.0272943	0.0200468	0.01389585	0.0084131	0.0071877	0.00545987	0

Fig. 2: Figure Showing the Value of K As Estimated by USLE, GMDH, PSO Method.
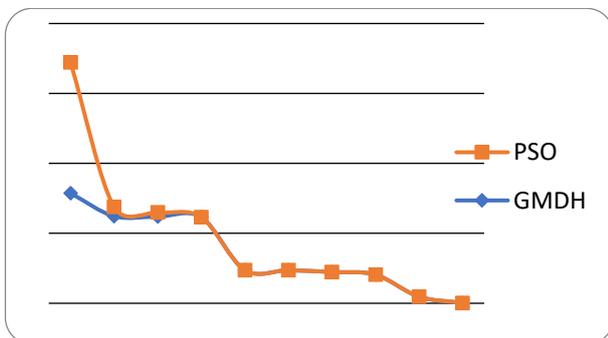
4.1. Error in predicted models

Fig. 3: Figure Showing the Absolute Error of GMDH and PSO Model.

Error in predicted model:

Table 2: Showing the Errors and Correlation Factor of Predicted Model

	MSE	RMSE	CORRELATION
GMDH	3.52	4.44	0.998
PSO	1.03	2.96	0.997

4.2. Case study

In case study values from different location is calculated and validation of data is done by using GMDH model that was developed previously in model development phase. After running these values following graph are obtained
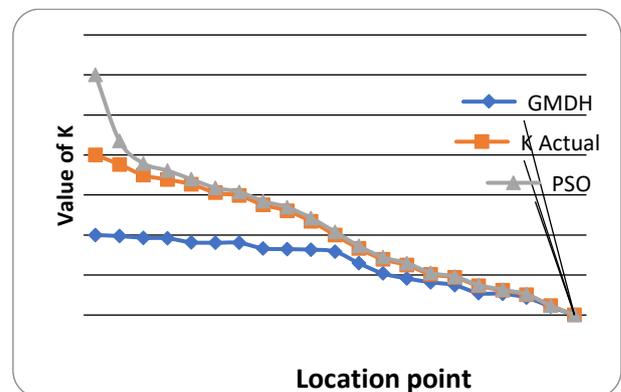

Fig. 4: Figure Showing the Soil Erodibility Factor (K) with Respect to Related Location as Estimated by USLE, GMDH and PSO.

4.3. Error graph for case study: error graph on comparison with actual data of case study the following graph is obtained
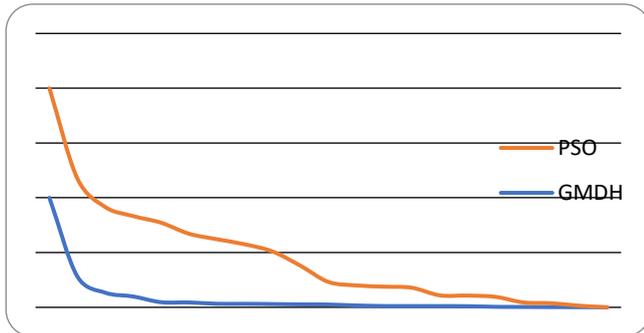

Fig. 5: Figure Showing the Difference between K Predicted by PSO and GMDH and K Obtained by USLE.

Error in case study:

Table 3: Showing the Max Absolute Error, Root Mean Square Error and Correlation Factor

	MAE	RMSE	CORRELATION
GMDH	0.374	0.489	0.898
PSO	0.086	0.230	0.636

5. Conclusions

Although USLE is having its own limitations that it can only applied to sheet erosion since source of energy is rain only and applicable only for hilly region having slope of 1-20%. This model is only applied for average value of previous data and cannot apply for particular storm. Moreover it neglects the certain interaction between factors in order to distinguish more easily the individual effect of each factor. Thus through this paper we can add that PSO gives us a least error as comparison to GMDH as correlation factor value of PSO graph gives the min value as comparison to GMDH for estimation of soil erodibility factor K.

Acknowledgement

References

[1] WischmeierW.H. C.B. Johnson and B.V. Cross. 1971. "A soil erodibilitynomograph for farmland and construction sites." Journal of soil and Water Conservation 26:189-193 ISSN 1941-3300.

[2] Lane, L.J., E.D. Shirley, andV.P.Singh. 1988."Modeling erosion on hill slopes."P.287-308. In: M.G.A nderson (ed)"Modeling Geomorphological system." John Wiley, Pubi., NY.

[3] ID Moore, JP Wilson - Journal of soil and water conservation, 1992 "Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation.

[4] Wischmeier W.H., andD.D.Smith, 1960."A universal soil loss equation to guide conservation farm planning"Trans. Int. Congr. Soil SCI., 7th, p. 418-425.

[5] Wischmeier.W.H.andD.D.Smith.1978."Predicting Rainfall Erosion losses: A Guide to Conservation Planning." Agriculture handbook No, 507, USDA/Science and Education Administration, US Govt.printing office, Washington, DC.58pp.

[6] National resources Conservation Service, U.S.Department of Agriculture.Washinton, DC.61.FR27998 "Technical assis-tance."1996-06-04.

[7] "Assessment of soil loss of Dhalai River Basin Tripura INDIA using USLE". Kapil Gosh, Sunil Kumar De, Shreya Bandyopadh-yay, S ushmitaSaha. Department of Geography and Disaster Management, Tripura University, Suryaminagar, India.

[8] S.Nanna,"A Geo-Information Theoretical Approach to Inductive Erosion Modeling Based on Terrain Mapping Units,"Ph.D. Thesis, Wageningen Agriculture University,Wageningen, 1996.

[9] Y.S.Elirehema,"Soil Water Erosion Modeling in selected Watershed in Southern Spain," IFA, ITC, Enschede, 2001.