Redesigning the Morgan-Morgan-Finney Soil Erosion Model for Global High-Resolution Application.

By using remote sensing data as the main source of input and coupling runoff calculation to the IHE hydrological model, WaterPix.

C.E.M Luger[12]; Supervised by: **Claire Michailovsky[1], Wim Bastiaanssen[12]** and **Thom Bogaard[2]**

[1]Water Accounting, IHE Delft Institute for Water Education, Westvest 7, 2611 AX Delft, The Netherlands

[2] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

Submitted for review: 21 June 2018

1. Abstract

The Morgan-Morgan-Finney (MMF) model is identified as the most appropriate hillslope soil erosion model for high resolution global application with the objective of identifying relative erosion risk areas and modelling the effect of land cover changes on erosion at a catchment scale. This is justified by its simple model structure, low input requirements, semi-empirical basis and distributed application. However, transferability to global application is hindered by its reliance on empirical data for input requirements. This paper proposes methodologies of generating input data based on remotely sensed products which improves the spatial and temporal accuracy of the input rasters. The MMF model is further redesigned by coupling it with the IHE Delft Institute for Water Education inhouse hydrological model, WaterPix, to replace runoff calculations in order to improve its treatment of infiltration and by applying the model in monthly timesteps to analyse erosional differences within the seasonal crop calendar. The redesigned MMF model algorithm is coded in python and applied over a headwater in the Ganga basin located in the Madhya Pradesh state of India. The model produces realistic erosion rates and distributions and provides additional information on the spatial and temporal variation of erosion in the study area. However, validation of the redesigned model with field data needs to be prioritized before it is utilized.

2. Introduction

Soil erosion models are used to simulate the movement and transfer of sediment forced by hydrological, hydraulic, meteorological and geological parameters in order to quantify sediment yield and identify soil erosion sources and sinks. The importance of such tools in land use management and in obtaining system insight into erosional processes has inspired an extensive inventory of models. A review by Karydas, Panagos and Gitas (2014) classified an unexhaustive list of 82 soil erosion models which are commonly categorized as empirical and process models based on the underlying principles of their development. This variety is owing to the multifaceted and complex character of environmental processes and conditions the simulations aim to replicate.

Empirical models run on equations derived from statistically evaluated field observations and experiments. Considered largely as a black-box approach, it does not consider individual process of erosion but rather lumps them into empirical parameters overarchingly describing the interrelating and interacting processes involved. These models are widely used owing to their simplicity, low input and compute requirements but has been criticized in its exploratory applications under different conditions of their development.

Process based models are created according to universal natural rules of conservation of mass, energy and continuity, and attempts a descriptive representation of the physical system. This independence from empirical relationships is the common justification for their transferrable application. The degree of physics represented in these models assures many earth-surface process researchers of their rigorousness. However, they rarely simulate the physical ideal and approximations and parameterizations are nearly always present (Beven, 1989; Dunin, 1975 qtd in Merritt, Letcher and Jakeman, 2003; Murray, 2013) Attempts at detailed representation also requires numerous parameters related to each involved erosional process which at large scales become extremely compute intensive, limiting application to often plot and hillslope scales.

Lack of available data to satisfy the intensive input requirements often becomes one of the main draw-backs of these models (Aiello, Adamo and Canora, 2015).

Within the overarching empirical and process-based categorization, soil erosion models are further classified according to the scales they are designed to imitate and the resolutions they are applied. Models are developed to a certain spatial scale (i.e. plot, hillslope or catchment) where different erosional processes dominate. Sheet and rill erosion dominate at plot and hillslope scale, whereas gully erosion attributes up to 44% of sediment yield at the catchment scale (De Vente and Poesen, 2005; Brazier, 2013; Mitasova *et al.*, 2013). Furthermore, an inverse relationship in erosion and modelled area has been observed because sediment yield is controlled by both erosion and storage (Brazier, 2013). Therefore, scaling a model developed for plot scale to catchment scale would thus simplify erosional processes that cannot be resolved by an application of a simple factor.

Soil erosion models also differ in their spatial apportionment of the landscape (i.e. spatially discretized and spatially distributed models). Discretized models spatially average parameters over the entire study area or over designed hydrologic units which are expected to behave similarly under static conditions. Distributed models represent input variables and modelled values as continuous fields usually also discretized as regular grids. Physics inclined researchers often argue that distributed models are more fundamental as they simulate processes at smaller scales, however in the emergent phenomena viewpoint, discretized models can simulate variables and interactions that emerge at the large scale (Murray, 2013). Furthermore, the viability of lumping up physical relationships often developed in small scale to interconnecting cells in a large scale grid is questionable (Beven, 1989). Distributed models still have an advantage because it can model erosion at any point depending on the

resolution of the simulation (Mitasova *et al.*, 2013).

Temporal scale and resolution is another important characterization of a model. Many simple erosion models assume steady erosion over time and provide long-term estimates of erosion (most empirical models treat temporal variability in this way) (Mitasova *et al.*, 2013). This has implications on how the mass of the overall long-term record is distributed across event magnitudes causing an overestimation of low magnitude events and underestimation of high magnitude events (Furl, Sharif and Jeong, 2015). Process based models are proposed to perform better at apportioning mass with respect to extreme events. However, due to the complex, multi-scale interactions and limited temporal resolution of the input data, most physical based models resort to performing a steady state case using peak flows of individual erosion events (Mitasova *et al.*, 2013; Furl, Sharif and Jeong, 2015). Regardless of the empirical or process basis of the models, they are limited in fully representing the random component of erosional process, thus several studies referenced in Nearing (1998) have reported the bias in erosion predictions during extreme events.

Soil erosion model comparisons are often inappropriate as models were developed under unique stipulations. An underperforming model can outperform another in a different context (De Vente and Poesen, 2005; Murray, 2013) There can thus be no consensus on the best performing model across all scales and conditions. There is an agreement in literature however that validation of many soil erosion models are limited (Brazier, 2013). Future resources in soil erosion model research needs to focus on validation however difficult as appropriate measurement are rarely available.

3. Research Objective and Approach

The Water Accounting+ Framework supported by IHE provides a global standard for the impact assessment of climate change and land-use changes. This study focused on modelling hillslope erosion, as its mitigation is an important ecosystem services that varies in response to land-use changes.

The objective of this research is to identify and adapt a suitable hillslope soil erosion model that can recognize erosion risk areas and model the effect of land use changes on erosion at high-resolution (spatial) catchment scale which can be applied globally. The approach of the study to fulfill the objective consists of the following sequential steps:

i. Identifying an appropriate model through a literature review;
ii. Adapting model design to better suit the objective;
iii. Improving input parameter definition to better suit the objective;
iv. Analysing model results; and
v. Comparing model results to a simplistic heuristic map to determine whether a more resource intensive soil erosion model application is justified.

4. Materials and Methods

4.1. Study Area

The study area is a headwater in the Ganga basin located in the Madhya Pradesh state of India (Figure 1). Elevation is highest to the south with a gradual slope towards the north. The area is subtropical with the highest rainfall occurring in the monsoon season from June to September (Mondal, Khare and Kundu, 2018). The area crosses the district boundaries of Dewas, Rajgarh and Shajapur where the main crops grown are soya bean, wheat and gram which are planted over two seasons (Table 1).

The study area was chosen as it had available WaterPix output data from IHE and was at a reasonable size for preliminary model application and analysis.

Table 1. Crop Calendar of the main crops grown in the Dewas, Rajgarh and Shajapur Districts.

Crop	Area * [ha]	Season	Growth Period											
			J	F	M	A	M	J	J	A	S	O	N	D
Soyabean	935284	Kharif						P				H		
Wheat	512706	Rabi				H						P		
Gram	261934	Rabi				H						P		

***Total crop area over the whole of Dewas, Rajgarh and Shajapur**
***P=Plant; H=Harvest**
***Dark Green=Always Planted; Light Green = Early Plant of late Harvest**
Source: ***Central Bank of India and Government of India (2014)***

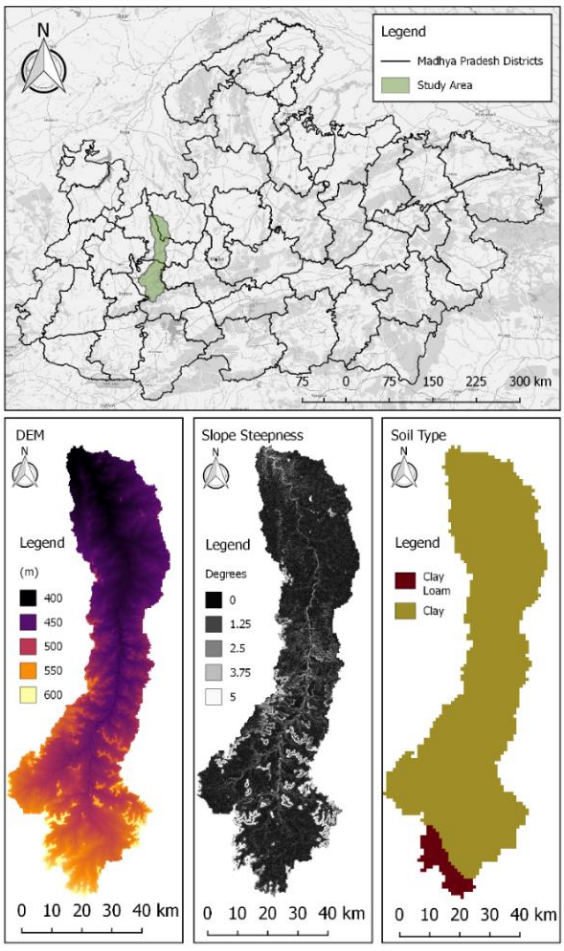


Figure 1. Study area.

4.2. Model Selection Criteria

There is an extensive inventory of soil erosion models available for selection in this research, developed under different degrees of physical basis, scales and contexts. This is owing to the multifaceted and complex character of environmental processes and conditions the simulations aim to replicate.

The initial approach for the model selection was to identify a single best performing model through a literature review of comparison studies and model descriptions. The approach proved to be difficult because of the number of models available and the largely contradicting conclusions of the comparison studies. Soil erosion models do not simulate the ideal and consequently, performance depends on the context they are applied to. Thus, a selection criterion for the model was created based on the objective of the research. This is discussed in the following paragraphs.

The intended use for the soil erosion model in this context is to identify erosion risk areas and model the effect of land use changes on erosion. To identify erosion risk areas, the chosen model has to run in a distributed manner in order to ensure an appropriate spatial detail to generate meaningful conclusions. Absolute values of sediment yield are not necessary as relative values will give sufficient information on higher and lower risk areas. To model the effect of land use changes, the model needs to be compatible with GIS inputs and raster data. For this use, event and daily time resolutions are not necessary because land use changes at larger temporal resolutions.

The model must be capable of generating high spatial resolution outputs at a catchment scale. This limits the selection to distributed models developed for catchment scale applications to ensure dominant processes at this scale are represented. In addition, simple model structures are preferred to ensure reasonable compute requirements when modelling at high resolution over a large area.

Finally, the model is intended for global application in gauged and ungauged basins

which fundamentally adds uncertainty to data availability. To satisfy the criteria, the model must have low input requirements and/or easily gathered input parameters. Additionally, the model needs to be robust in exploratory applications.

There is a certain dichotomy in the aforementioned model requirements. Only process-based models can be exploratory, however typically these have high input requirements and complex model structures. Empirical models, while simple and easy to apply, characteristically cannot be transferrable. Nevertheless, empirical models, such as the Universal Soil Loss Equation (USLE) and its revisions, are arguably the most utilized soil erosion models even in exploratory applications.

The USLE was developed for uniform plot scale applications on 22.1 m hill slope lengths (Wischmeier and Smith, 1978). When applied on catchment scale and areas with complex topography, the model and its variations was found to produce subpar results (Romero-Díaz, Alonso-Sarriá and Martínez-Lloris, 2007; Oliveira *et al.*, 2013; Furl, Sharif and Jeong, 2015; Arkowitz, 2017). Some argue that by applying the USLE in a distributed manner over a catchment, it inherently reverts to a plot scale application. However, erosional processes dominant in plot scale, with which USLE was developed for, may not be dominant in the catchment scale. Such a lumped application of a plot scale model may simplify the erosional processes at the catchment scale.

Other frequently recurring models during the literature review are the process-based WEPP and SWAT models. WEPP is input and compute intensive with a complex model structure which is not applicable to global studies and at high resolution over a catchment scale (Arkowitz, 2017). SWAT is a holistic model which computes hydrologic and erosional responses separately. For the latter the model uses USLE-based equations and inherits the same drawbacks (Arnold *et al.*, 1998). For these reasons, widely used models did not satisfy the requirements for this research. Rather, the lesser known semi-empirical Morgan-Morgan-Finney (MMF) model was chosen. Specific justification for the selection and a description of the MMF model will be discussed in the following subsection 4.2.1.

4.2.1. Morgan-Morgan-Finney (MMF) Model Selection Justification and Description

The Morgan–Morgan–Finney (MMF) model is semi-empirical in that it applies concepts by Meyer and Wischmeier (1969) to provide a stronger physical basis than the USLE while retaining the characteristic flexibility and simplicity of empirical models (Morgan, 2001). The model needs less data than most of the other processed-based models (Shrestha, 1997). Furthermore, the MMF model is compatible with GIS input data and grid distributed application at any resolution. These characteristics make the model ideal for global and high-resolution application.

Output rigorousness is sacrificed for the simplicity of the model. However it is suitable as a rapid first-approximation of erosion risk areas which is the intended use of the model for this research (Morgan, 2001).

There are three versions of the Morgan-Morgan-Finney (MMF) model. It was first developed by Morgan, Morgan and Finney in 1984 for plot scale application. A revision of the model published by Morgan in 2001, included routing for application at the catchment scale and changed the way soil particle detachment by raindrop and flow is simulated. The revised version showed an improvement in model performance when compared to the original model (Morgan, 2001). Another modification was made by Morgan and Duzant in 2008, in which the effect of vegetation was elaborated and depositional processes were considered. The Morgan and Duzant modification does add several input parameters to the model with seemingly low performance improvements (Morgan, 2001; Morgan and Duzant, 2008). Therefore, to maintain the simplicity of the model, the revised MMF model from 2001 is applied in this study.

Finally, comparison studies of the earlier versions of the MMF model against common erosion models such as the USLE and WEPP show that the 1984 version of the MMF model generally underperforms and a clear improvement in relative performance is seen when comparisons are made against the 2001 version of the MMF model (Table 2). Comparison studies of the 2008 modification of the MMF model could not be found during the literature review. Although comparison studies cannot by itself function as a good justification of a model selection, it gives an indication of an improvement in relative performance which further justifies the model selection.

4.3. MMF model design and application

The model calculates soil erosion in two parts, the water phase and erosion phase and considers splash and run-off detachment as defined on Figure 2. A conceptual diagram of the model is shown on Figure 3 with the operating functions listed on Table 3. A description of the input parameters used in the equations of Table 3 and their references are listed on Table 4.

The water phase determines the kinetic energy from direct throughfall, KE(DT), and leaf drainage, KE(LD), available to detach soil particles. In this phase, effective rainfall (ER) is first calculated based on an interception factor (A) which represents the proportion of rainfall permanently intercepted by the cover. The proportion of the effective rainfall reaching the ground as leaf drainage (LD) or direct throughfall (DT) is further divided according to the percentage canopy cover (CC). The kinetic energy of leaf drainage is then calculated based on the plant height (PH) according to an equation by Brandt (1990). The kinetic energy of direct throughfall is dependent on the characteristic intensity of erosive rain (I) for the climate of the application area. Typical energy-intensity relationships used as the operating function for different climate types are listed in Morgan (2001). The operating function listed on Table 4 is specific for the subtropical climate of the study site.

The kinetic energy values derived in the water phase are used in the soil phase of the model to calculate soil particle detachment by rainfall (F) and run-off (H) using a simplification of soil erosion equations described by Meyer and Wischmeier (1969). The former is dependent on the detachability of soil (K) whereas the latter is dependent on the soil cohesion (COH). Soil detachment due to runoff and the transport capacity is further dependent on the volume of runoff (Q), cover factors (GC and C) and the slope (S).

The sum of the erosion caused by rainfall and runoff are compared to the transport capacity (TC). The lower of the two values is taken as the final erosion rate.

Table 2. Preferred model from Comparison studies of the 1984 and 2001 version of the MMF against other soil erosion models.

Comparison Study	MMF Version	Soil Erosion Models	Preferred Model
Vigiak and Sterk (2001)	1984	USLE	MMF
Svorin (2003)	1984	USLE, RULSE, SLEMSA	USLE
Pandey *et al.* (2009)	1984	USLE	USLE
Bayramov, Buchroithner and Mcgurty (2013)	1984	USLE	USLE
Mondal, Khare and Kundu (2018)	1984	USLE, RUSLE	RUSLE
Yazidhi (2003)	2001	RUSLE	RULSE
Fernández, Vega and Vieira (2010)	2001	RUSLE	MMF
Jha and Paudel (2010)	2001	RUSLE	MMF
Tesfahunegn, Tamene and Vlek (2014)	2001	SWAT	MMF, SWAT
Vieira *et al.* (2014)	2001	RUSLE, WEPP	MMF
Li *et al.* (2017)	2001	RULE, MUSLE, Zheng's Model, WEPP, SWAT, DYRIM, WATEM/SEDEM, Tian's Model, Si's Model, Yang's Model	MMF, WEPP, Si's Model

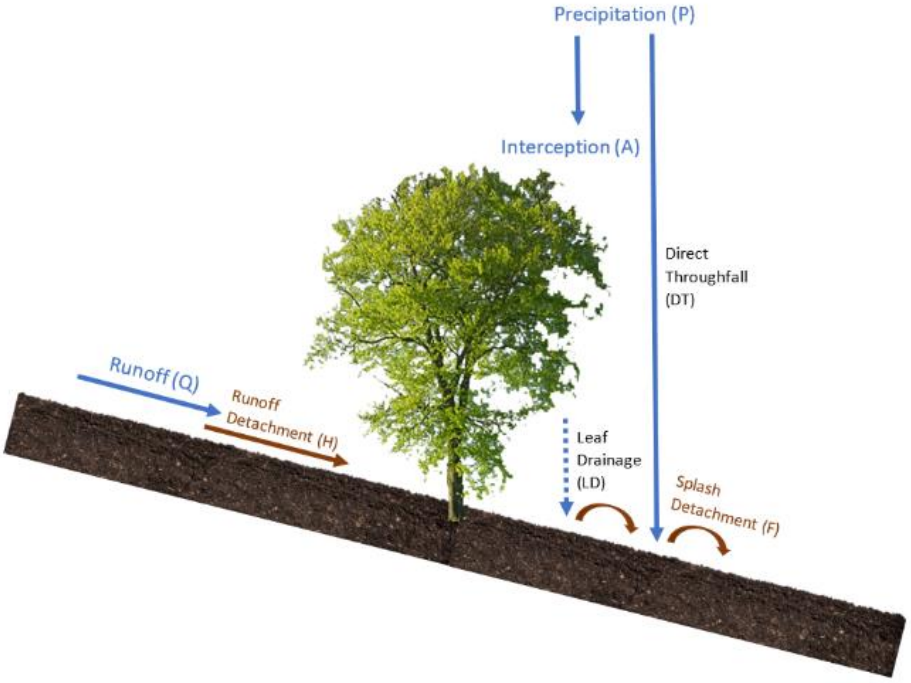


Figure 2. Soil particle detachment by raindrop and runoff considered in the MMF model.

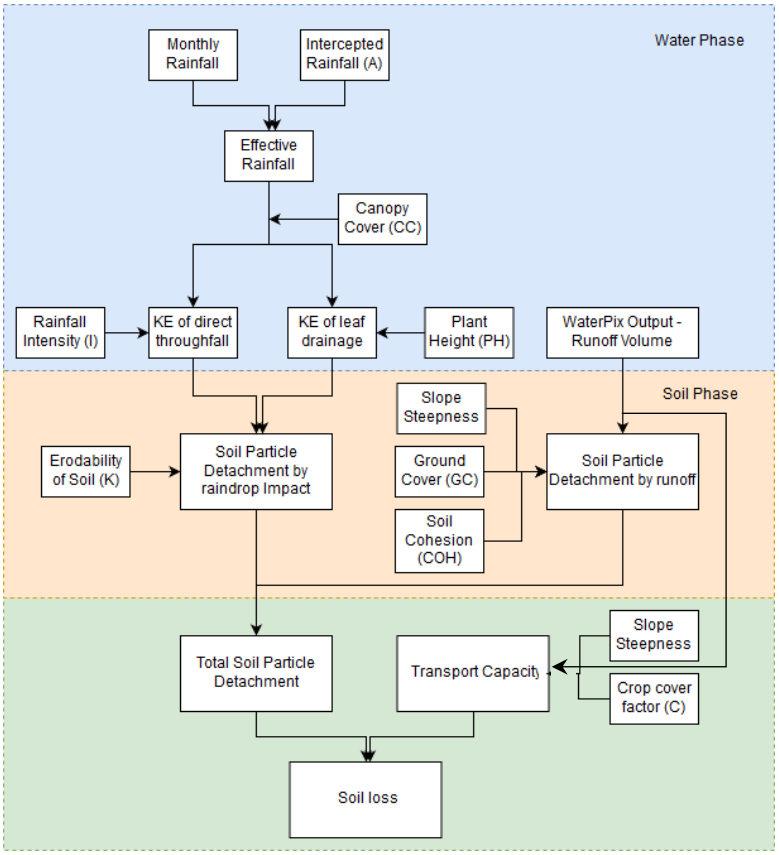


Figure 3. Conceptual diagram of redesigned MMF model.

Table 3. Operating Functions

Water Phase:	
Effective Rainfall	$ER\ [mm] = R \times (1 - A)$
Leaf Drainage	$LD\ [mm] = ER \times CC$
Direct Throughfall	$DT\ [mm] = ER - LD$
Kinetic Energy of Direct Throughfall	$KE(DT)\ \left[\frac{J}{m^2}\right] = DT \times (11.8 + 8.73\ \times logI)$
Kinetic Energy of Leaf Drainage	$KE(LD)\ \left[\frac{J}{m^2}\right] = LD \times (15.8 \times PH^{0.5}) - 5.87$
Soil Phase:	
Soil Particle Detachment by Raindrop Impact	$F\ \left[\frac{kg}{m^2}\right] = K \times (KE(DT) + KE(LD)) \times 10^{-3}$
Soil Particle Detachment by Runoff	$H\ \left[\frac{kg}{m^2}\right] = \frac{(\sum Q_m)^{1.5}}{\sum({Q_m}^{1.5})} \cdot \frac{[{Q_m}^{1.5} \times \sin S\ \times (1 - GC) \times 10^{-3}]}{0.5\ \times COH}$
Transport Capacity	$TC\left[\frac{kg}{m^2}\right] = \frac{(\sum Q_m)^2}{\sum({Q_m}^2)} \cdot [{Q_m}^2 \times \sin S\ \times (C) \times 10^{-3}]$

The application of the model in this study deviates from the proposed application in four main points. The first deviation comes from the coupling of the MMF model to the IHE in house hydrological model, WaterPix, which is used in place of the Kirkby (1976) equations, as quoted in Morgan (2001), to generate the volume of runoff as input to the soil phase of the MMF model. This deviation has the advantage of accounting for infiltration as runoff is routed through the catchment. The ignorance of the process in the MMF model was identified by Feng et al. (2014) as a significant source of error in the original application.

The second deviation is the application of runoff routing for the calculation of the soil particle detachment by runoff in the soil phase of the model. The proposed application of routing was limited to transport capacity calculations in the original publication.

The third deviation is in the preparation of the input data. Recommended empirically based parameters are replaced by values calculated from remote sensing products in order to improve on the temporal/spatial accuracy and transferability of the inputs. By increasing reliance on remotely sensed products, the ease for global application of the model is improved.

This is particularly important to accommodate the final deviation of study which is the reduction of the modelling timestep from yearly to monthly. This deviation preserves the resolution of the WaterPix overland flow output and gives further insight into erosion risk throughout the life and agricultural cycle of the crops as well as the effects of seasonal weather differences. In order to accommodate a reduction in timestep, the operating function for transport capacity (TC) and soil particle detachment from runoff (H) on Table 3 are normalized against their yearly value as recommended by Davison et al. (2008).

The model was applied over the study area using inputs from 2010. The semi-empirical nature of the model allows application in a distributed manner so it can provide useful information on the source areas of sediment. For application in this study, the interest area was distributed into 2.4x2.4 km^2 grid cells to mimic the output resolution of WaterPix. The model was developed for catchment scale application and this will be maintained in this study.

The MMF model algorithm is written in python language. Instructions to run the coded model are on Appendix 1 while a copy of the scripts are on Appendix 2 and 3.

Table 4. Input Parameters to MMF

Factor	Parameter	Definition	Type	Source	Reference
Overland Flow	O	Overland Flow (mm)	Raster	Overland flow WaterPix Output	-
Climate	R	Monthly rainfall (mm)	Raster	CHIRPS	Funk *et al.* (2015)
	I	Typical value for intensity of erosive rain (mm/h)	Value or Type	Literature Review	Mondal, Khare and Kundu (2018)*
	Type	Climate Type	Type	Literature Review	Mondal, Khare and Kundu (2018)*
Soil	K	Soil detachability index (g/J) defined as the weight of soil detached from the soil mass per unit of rainfall energy	Raster	Empirical value Soil Map	Morgan (2001) Land and Development Division, FAO, ROME (2003)
	COH	Cohesion of the surface soil (kPa) as measured with a torvane under saturated conditions	Raster	Empirical value Soil Map	Morgan (2001) Land and Development Division, FAO, ROME (2003)
Landform	S	Slope steepness (°)	Raster	DEM	Lehner, Verdin and Jarvis (2008)(HydroShed Database) – accessed from WA toolbox
Land Cover	A	Proportion (between 0 and 1) of the rainfall contributing to permanent interception and stemflow.	Raster	LAI Number of Rainy days	NASA LAP DAAC (2010) – accessed from WA toolbox Funk *et al.* (2015)
	C	Crop cover management factor	Raster	NDVI	NASA LAP DAAC (2010) – accessed from WA toolbox
	CC	Canopy cover expressed as a proportion (between 0 and 1) of the soil surface protected by the vegetation or crop canopy	Raster	LAI	NASA LAP DAAC (2010) – accessed from WA toolbox
	GC	Ground cover expressed as a proportion (between 0 and 1) of the soil surface protected by vegetation or crop cover on the ground	Raster	LAI	NASA LAP DAAC (2010) – accessed from WA toolbox
	PH	Plant height (m)	Raster	Ecocrop database Land Use Map	FAO (2013) Created by IHE based on information from Goethe Universit (n.d.), ESA (n.d.) and NIRCA (n.d.)

*Reference is specific to study site

5. Input Data and Pre-processing

The required inputs listed on Table 4 to the MMF model are pre-processed to the same resolution as the overland flow output of WaterPix. The inputs are derived from a combination of remotely sensed products and empirically derived values to create raster maps covering the study area. The processing of several input parameters diverge from the recommended methods stated in Morgan (2001) to improve on their spatial and temporal accuracy as suggested methods often rely on empirically derived values that would likely not account for the monthly land use and land cover changes. By increasing reliance on remotely sensed products for input data, the model is more capable of being applied globally where lack of data is often the biggest hindrance. These divergences are elaborated in the sections below discussing the processing of input parameters where required. Any generalized values proposed here should be replaced when actual field data is available.

5.1. Overland Flow

The overland flow in this study diverges from the original methodology in two ways: (i) it uses the output of the IHE inhouse hydrological parameterization, WaterPix, in order to account

for infiltration; and (ii) a single-pathway routing algorithm (GRASS – r.watershed module) is applied to the overland flow raster to generate a distributed overland flow accumulation map in order to account for the routing of flow (De Jong *et al.*, 1999; Shrestha and Jetten, 2018).

5.2. Climate

The monthly rainfall input to the model is the same as the monthly rainfall used to run the hydrological parameterization, WaterPix, for consistency. The climate type of the study area is subtropical which informs the selection of the intensity of erosive rain (25 mm/h) and the operating function for the kinetic energy of direct through fall on Table 3. Other typical values and relationships for specific climate types are listed by (Morgan, 2001).

5.3. Soil

The soil type informs values for the detachability of soil (K) and soil cohesion (COH). A soil map was first generated using the "The Digital Soil Map of the World" maintained by The Land and Development Division of FAO maintains. This is a standardized dataset of percentage compositions of clay, sand and silt for 106 soil units, derived from a statistical analysis of 4553 soil profiles held in the WISE (World Inventory of Soil Emission) database (Batjes, 1997). Based on a soil texture diagram, the soil type is identified for every cell generating the soil map shown on Figure 1. The soil type is assigned a numerical value according to Table 5 and is input to the model where the values of K and COH are inherently applied.

5.4. Landform

The slope steepness is calculated from the DEM using the r.slope.aspect module from GRASS where the degree format can be specified.

Table 5. Soil Type numerical value and K&COH values.

Soil Type	Value	K (Morgan 2001)	COH (Morgan 2001)
Sand	1	1.2	2
Loamy sand	2	0.3	2
Sandy loam	3	0.7	2
Loam	4	0.8	3
Silt	5	1.0	3*
Silt loam	6	0.9	3
Sandy clay loam	7	0.1	3
Clay loam	8	0.7	10
Silty clay loam	9	0.8	9
Sandy clay	10	0.3	9*
Silty clay	11	0.5	10
Clay	12	0.05	12

*Values for these soil types were not available. The values on this table are assumed based on similar soil types

5.5. Landcover

Input rasters under the landcover category are conventionally derived by assigning empirical constants to specific crop types or land use classifications on a land use map. The validity of these values is certain only when they are applied under the same environmental conditions as the empirical tests which are often not carried out over the life and agricultural cycle of the crop and may not reflect the appropriate temporal variation for outputs at a monthly timescale. The spatial accuracy of input rasters generated from this methodology is further depended on the quality of the land use map.

Inputs derived from remotely sensed data improves on the conventional approach by (i) removing dependence on the accuracy of the land-use map; (ii) removing inaccuracies inherent in empirical values; (iii) retaining accuracy when up sampling (i.e. whereas the conventional approach can only be up sampled using the mode statistic, remotely sensed data can be averaged therefore retaining information of lumped crop types within an up sampled cell) and (iv) improves the efficiency and ease of

application as it does not require an extensive literature review to identify the appropriate empirical values, which may be time consuming as an encompassing database is non-existent. In some cases, an appropriate value may not be found in literature.

The methods used for deriving the input rasters based on remotely sensed data are discussed in detail in the following sections which includes comparison to empirically derived values.

Interception Factor

The interception factor (A) has values between 0 and 1 and is described as the proportion of rainfall contributing to permanent interception. The input raster was generated in this study as a function of interception (i), precipitation (P), number of rainy days (n) and the LAI product which describes the percentage of plant cover over the percentage of ground cover. For cells where n = 0 while P > 0, the n values are assigned to 1.

$$i = n * LAI(1 - \left(1 + \frac{P(1 - e^{-0.5*LAI})}{n * LAI}\right)^{-1})$$

$$A = \frac{i}{P}$$

A timeseries of the LAI-based A-values for several sampling locations on different land use classes are on Figure 4(a). For comparison, empirical values for the study site gathered from a literature review are listed on Table 6 and plotted on Figure 4(a).

Values generated from remote sensing products are significantly smaller than their empirical counterparts for all dominant land use types. The temporal variability of the LAI-based values is more realistic as they follow the crop calendar patterns on Table 1. However, they do not have the clear distinction that empirical values have between land use types.

The magnitude difference between the two methods when calculating the effective rainfall increases with the magnitude of precipitation, as the empirical based values assume a linear relationship and the LAI-based values are computed as a function of precipitation.

Crop Cover Management Factor

The crop cover management factor (C-factor) is the product of the Cropping Management Factor (C) and the Support Practice Index (P) of the USLE. The former is the ratio of soil loss from land cropped under specified conditions (combination of cover, crop sequence, and management practices) to the corresponding loss from clean-tilled, continuous fallow. The latter is the ratio of soil loss with a specific support practice(i.e. contouring, strop cropping and terrace contour farming) to the corresponding loss with up-and-down-slope culture (Wischmeier and Smith, 1978). Both factors range from 0 to 1. A value of 1 indicating no cover effects and no erosion control.

The C-factor is calculated in this study using the NDVI product according to the function below (Van Der Knijff, Jones and Montanarella 1999). The recommended values for alpha and beta are 2 and 1 respectively. Artefacts/noise likely due to clouds were observed in the raw NDVI data. Cells affected were removed and interpolated.

$$C = e^{-\alpha \frac{NDVI}{(\beta - NDVI)}}$$

$$C - factor = C * P$$

The support practice index (P) is based on direct observations and thus difficult estimate. Pham, Degener and Kappas (2018) lists studies that have attempted to relate it to slope inclination and land-use maps. However due to the low performance of the proposed methodologies, this study will maintain the empirical value on Table 6.

Figure 4(b) shows a time-series comparison of the empirical C-factor values and the NDVI-based values for sampling locations over the dominant land use classes. The latter values have strong temporal variation significantly exceeding the empirical range but follows the crop calendar pattern on Table 1. High variability proves the importance of a temporally varying C-factor value for calculating monthly erosion estimates. The NDVI-based values have no apparent distinction between crop types. The year

average values are 0.34, 0.31 and 0.31 for wheat, soybean and gram respectively thus aligning only to the empirical value for Soybean.

Percentage Canopy Cover/ Ground Cover

Morgan (2001) does not explicitly state a methodology for calculating the percentage canopy cover (CC) and ground cover (GC) although empirical values were added in a lookup table in the publication of the Morgan and Duzant (2008) modification.

To determine the percentage canopy cover, an LAI-based approach was applied in this study following the same methodology as in WaterPix. Canopy cover (CC) is calculated as a function of LAI according to the equation below.

$$CC(\%) = 1 - e^{-0.5*LAI}$$

The percentage ground cover may be inferred from the crop cover management (C) values of the previous section. Although GC only accounts for vegetation cover in the ground while C-values account for vegetation cover as well as foliage, trash and management practices, these would disperse overland flow energy in a similar way. In fact, the earlier version of the MMF model used the C-factor in place on GC (Morgan, Morgan and Finney, 1984).

$$GC(\%) = 1 - C$$

Plant Height

Plant Height is difficult to estimate with remotely sensed products. This study assigned averaged values from the Ecocrop database of FAO according to the land-use map (Table 6).

Table 6. Empirical values for MMF input

LULC Classification	Dominant crop	Cover (%)	A (%)	Reference	C Value	Reference	P Value	Reference	PH (m)	Reference
Herbaceous cover	-	0.43	32.5	Morgan (1995)	0.04	Jain and Kothyari (2000)	1	Mondal, Khare and Kundu (2018)	0.3	Morgan and Duzant (2008)
Rainfed crops - cereals	Wheat	7.91	43	Morgan (1995)	0.2	Morgan (1995)			1.2	EcoCrop database by FAO (2013)
Rainfed crops - oilseed	Soybean	34.31	15	Dunne and Leopold (1978)	0.35	Morgan (1995)			1	EcoCrop database by FAO (2013)
Irrigated crops - leguminous	Chickpea	57.06	25*	-	0.53	Vigiak and Sterk (2001)			0.6	EcoCrop database by FAO (2013)
Irrigated crops - fruit and nuts	-	0.14	25*	-	0.29	Jain and Kothyari (2000)			1*	-
Urban paved Surface	-	0.14	0	Morgan (1995)	1	-			0	-

*No information found. Value is assumed.

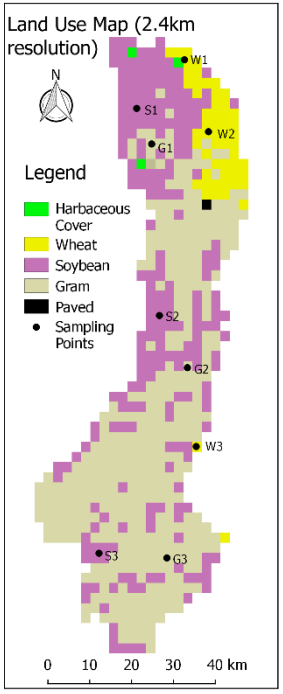


Figure 5. Sampling locations for the dominant land use classes, Wheat(W1,W2,W3), Soybean(S1,S2,S3) and Gram(G1,G2,G3).

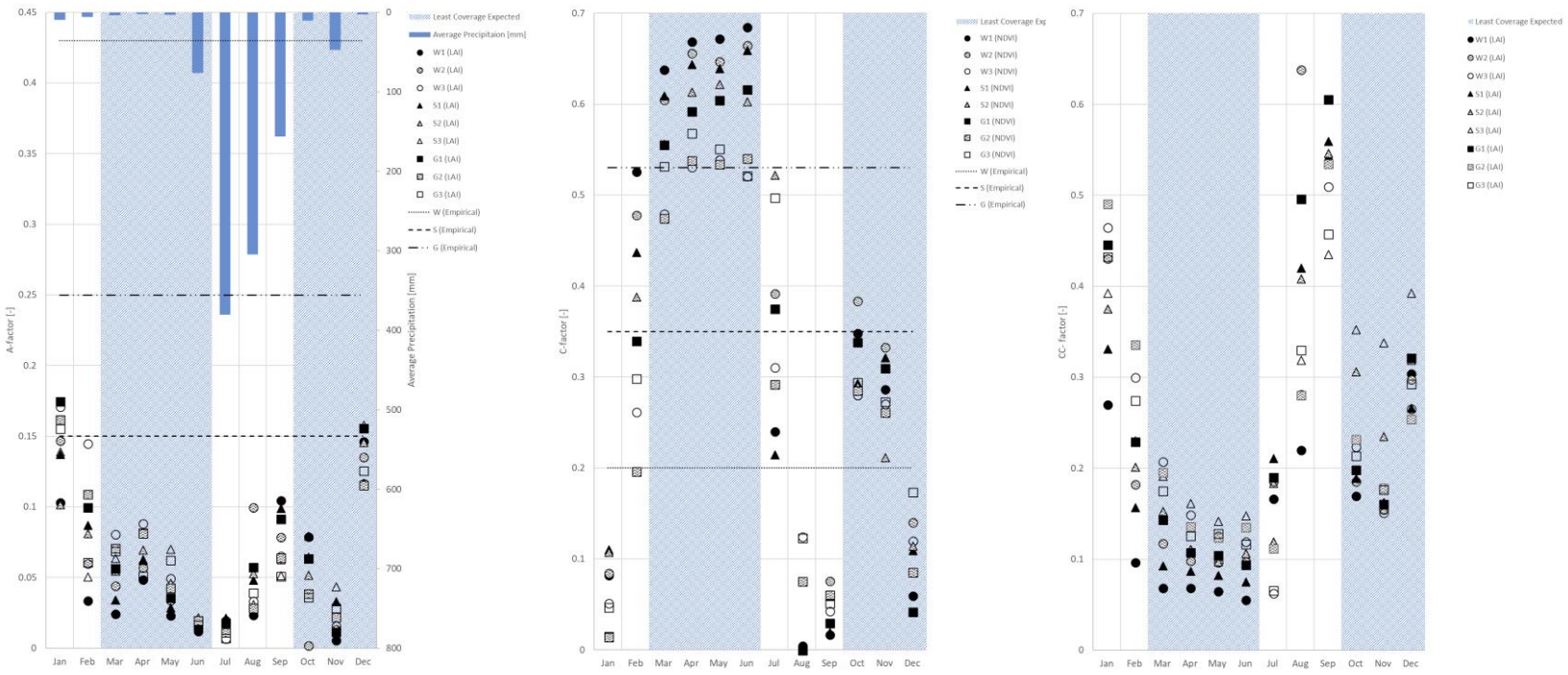


Figure 4. Left(a): A-factor empirical and LAI-based values for sampling locations on Figure 5. Middle (b): C-factor empirical and NDVI-based values for sampling locations on Figure 5. Right (c): CC values calculated from LAI for sampling locations on Figure 5.

6. Limitations and Uncertainties of Input Data and Model

The semi-empirical nature of the MMF model allows some physical basis while maintaining the simplicity and minimal input requirement of an empirical model. This allows grid application globally but does not exempt it fully from the transferability uncertainties often attached to empirical models. Furthermore, it was initially developed for plot scale application which was then adapted to catchment scale by applying routing of flow and sediment so scalability issues may be inherited.

The model is further limited by its lumped treatment of events. It essentially smooths out high and low magnitude events as it deals with long term totals of runoff. This adds to the model's uncertainty as not all precipitation events result in erosion. A month with frequent low magnitude events can represent itself as a relatively high monthly runoff total but may in reality have little to no erosion (Andresen, 2017). Soil type can exacerbate the uncertainty. Clayey soils, for example, require a significant amount of runoff to erode during a particular event. The model is likely to overpredict erosion in these areas during months of frequent low intensity events. The model can also underestimate erosion because it smooths out upper end events where the highest erosion rates occur (Furl, Sharif and Jeong, 2015). In fact, when tested on plots in the Loess region of China, the 1984 and 2001 version of the MMF model could not predict the very high erosion rates (Morgan, 2001). These inaccuracies however can also be attributed to the soil erosion models, in general, not being able to capture the random component of erosion within a deterministic model (Nearing, 1998).

The rigorousness of the operating functions as well as the quality of the input data are other characteristics that may limit the performance of the model. The outstanding inaccuracies in the operating functions of the original 2001 revision of the MMF model are: (i) ignorance to deposition (ii) neglection of infiltration in the water phase (iii) the assumption of erosion only occurring on bare ground; and (iv) neglecting convergence and divergence of overland flow for calculating soil detachment by runoff.

The first point is resolved in the 2008 modification of the model. However, this version increased input requirements without assuring an improvement in performance. Therefore, it was not selected for application in this research. Deposition can be applied in future development of this version by defining the deposition rate in cells to be the difference between detachment capacity and transport capacity when the latter is exceeded.

The second point is resolved in this study by using the overland flow output from WaterPix which considers infiltration. This implies, however, that the validity of the model is essentially linked to the validity of the hydrological parameterization. Validation of the hydrological parameterization would inform the level of uncertainty its utilization would incur.

The final point is resolved by passing the runoff output of WaterPix through a single pathway routing algorithm to generate the distributed overland flow accumulation map. This flow direction algorithm assumes a concentrative flow where water moves only to one cell, although dispersive flow frequently occurs in reality (Oliveira *et al.*, 2013). Future works could apply a dispersive multi-direction flow algorithm instead.

Originally, the majority of input parameters to the model are based on empirical values. This study aimed to reduce reliance on empirical data by proposing methods of input derivation based on remotely sensed data. The remaining empirical values used in this redesign are the intensity of erosive rain(I), plant height (PH) and soil detachability (k) and cohesiveness (COH). These values are characteristically only valid when applied in similar environments they were derived in. Often, they do not reflect the characteristic changes during the life cycle and agricultural cycle of the site. For example,

plant height values would vary temporally according to the growth and harvesting stages of the crop. Inaccuracies due to static plant height inputs in this research are partly alleviated because plant height is scaled indirectly by crop cover calculated from remotely sensed data.

The accuracy of the spatial distribution of empirically based raster inputs also depend on the accuracy and resolution of the GIS maps (i.e. land-use map and soil map) they are attached to. The soil map used in this study was generated from a statistical analysis of sampling locations inferring some degree of interpolation. Hence, it shows very little variation in the study site. As another example, the land-use map used general classifications forcing assumption on the specific crop type. These GIS-based maps are often static and rarely reflect monthly variations which may be particularly essential for land use changes when double cropping is involved or multi-season harvesting.

Certainty of input rasters generated from remotely sensed data are also not assured without costly verification in field. The accuracy of the data is further restricted to available resolutions and can be affected by artefacts. For example, the uncertainty in a DEM increases in areas with high relief and a low resolution may fail to capture the correct route of water accumulation (Andresen, 2017). When modelling sediment erosion specifically, Chaplot (2005) found the best results at resolutions <50m.

Due to the limitations and uncertainties stated above, it is recommended perform a field verification regardless of the nature of the input data presented in this study. Ultimately, the results from this study should not be taken as absolute values of soil erosion but rather as a good rapid first approximation of erosion rates to identify source areas of sediments in the catchment (Morgan, 2001). But the biggest limitation of this research, much like many soil erosion modelling publications, is the lack of validation against field data.

7. Results and Discussion

The model returns spatially distributed monthly erosion rates over the study area. Figure 6 depicts an image compilation of the model results for 2010. Table 8 lists additional monthly information, namely: (i) the percentage area composition under different soil erosion classes based on the definition used by Mondal, Khare and Kund (2018); (ii) Total erosion over the entire study site; (iii) Percentage of area composition under the detachment and transport limiting cases; (iv) Percentage area composition dominated by raindrop impact or runoff detachment; (v) Total monthly erosion for different slope classes; and (vi) Total monthly erosion for different land use classes.

The cell-by-cell calculated erosion rates range from a minimum of 2.69×10^{-13} kg/m^2 in April to a maximum of 62.7 kg/m^2 in July. However, magnitudes greater than 10 kg/m^2 are only found along the stream network where huge quantities of flow converge maintaining here a detachment limited state. Estimates over the stream network are likely inaccurate as the model does not consider hydraulic factors.

It is evident that the study area is most sensitive to erosion in the month of July when according to the crop calendar on Table 1, agricultural fields of gram and wheat have been harvested and soybean fields are young. Input parameters that represent coverage such as the C-factor and CC% does show that this month has lower coverage (Figure 4). Though not the lowest, erosion is exacerbated as this month experiences the highest precipitation amounts in the year Figure 4(a).

The breakdown of unscaled erosion in the detachment limited case due to raindrop impact and overland flow for this month verifies that the former dominates detachment in the majority area (Figure 7). This may be an artefact of the analysis as runoff dominates in-stream where flow converges which makes up a significantly smaller area percentage than the on-site coverage where raindrop erosion dominates. Also evident in this figure is the soil

map footprint which implies the importance of the factor in the calculations of raindrop detachment. Raindrop detachment is the dominant cause of erosion for the majority area for most of the year as seen on Table 8, therefore it is important to improve on the quality of the soil map in future runs. The relative dominance of raindrop detachment may be attributed to the clayey soils in the study area.

Another outstanding observation from Figure 6, is the contrast of erosion magnitudes at the southern and northern section of the study site during August 2010. To a lesser extent a similar pattern is observed in September. A significant section of the area in these months become transport limited as seen on Table 6, due to an increase in coverage denoted by the C-factor in the northern part of the study area as seen on Figure 4(b). This suggests distinct agricultural practices between the two areas. Based on the seasonal pattern of erosion and coverage, it may be inferred that the north section grows crops during the Kharif season (i.e. soybean) and the southern section grows during the Rabi season (i.e. Gram or wheat).

Slight soil loss (up to 0.05 kg/m^2) is observed in the majority area at the beginning and later part of the year coinciding with the Rabi crop season. Areas classified under severe loss during these months are over stream networks where flow converge. The coincidence of high precipitation and low coverage in July increases the severe soil loss area percentage (greater than 0.6 kg/m^2) to 43.7% which gradually decreases towards the end of the year as coverage increases and precipitation decreases.

Areas with 2-5-degree slopes experience the greatest erosion as they are found often along the stream network. The lowest slope class has a greater average soil loss than the steepest class for most months of the year. Lower slopes are often found downstream and thus have more upstream contributing flow. These numbers may also be misleading as lower slopes are found along the stream network with high flow convergence. August is an exception to this trend because of significantly lower coverage in the "steeper" southern section of the site causing more erosion. There is no indication in these results that slope is big determiner for erosion. Rather, erosion magnitudes are more dependent on flow convergence.

The highest soil losses in the land use classes are found under Gram. These results may be misleading because the majority area is classified under this class, thus streams with high values may skew the result. Furthermore, the results may be compromised by inaccuracies in the land use map. It became clear from the erosion and C-factor spatial distribution in the month of august that the northern and southern section of the study area have very distinct land use practices. Whereas the land use map does not show such an evident distinction indicating some inaccuracies.

The majority area during months with low runoff relative to the precipitation (i.e. May, June, September and November) are transport limited. Whereas the area during months with high runoff relative to the precipitation (i.e. May and October) is mostly detachment limited. This trend emerges because raindrop detachment is the dominant erosional process for the majority area while the transport capacity is calculated based on runoff. This indication that raindrop impact is the dominant cause of erosion can inform the appropriate practices for mitigation.

From the results it may be inferred that for the transport limited case, coverage and flow are the determining factors while for the detachment limited case, rainfall and soil type become dominant parameters.

Mondal, Khare and Kundu (2018) published results of a comparative study of three erosion models, MMF, USLE and RUSLE, south of the study site on a part of the Narmada river basin in Madhya Pradesh. Though not directly comparable as this study was done on a different location and time period, it can serve as a sanity check for the results. The soil loss classes and their percentage distribution have a similar distribution as the year average results of this study. The three erosion models computed that 50.63-61.87% of the area have slight erosion (<5 t/h/yr) and severe erosion

areas were estimated at 0.89-1.95% of the study site. High estimates in this study for severe erosion areas are attributed to the consideration of erosion along the stream network which were ignored in the 2018 publication.

When looking at the yearly averaged results of this study and the annual results of Mondal, Khare and Kundu (2018), one can reach a conclusion that erosion mitigation measures are not required as the majority area experiences only slight soil loss. However, by simulating monthly erosion patterns, it becomes evident that there is significant monthly variation and the majority area shifts to very severe loss in July when land cover is low and precipitation is high. There is a significant amount of information achieved by redesigning the model to run on monthly timesteps.

7.1. Heuristic Map Comparison

Soil erosion requires a significant amount of resources to model while the confidence on model accuracy for most of the scientific community remains small owing to the difficulty of validation. A simple heuristic map can be created based on commonly known rules of hillslope erosion from easily accessible data that require no processing. This section explores the added value of modelling over the information provided by a heuristic map.

The heuristic map is created following a similar logic to the MMF model, erosion is divided into splash and runoff erosion and are affected positively and negatively by factors stated in Table 7. The specific values of these factors are normalised to 1 with 0 being the lowest occurring value and 1 being the highest occurring value. The normalised values are then summed for each cell and their relative values reflect their relative erosion severity. Factors that affect both splash and runoff erosion are only added once.

Table 7. Factors used to create a Heuristic map of hillslope erosion.

Splash Erosion	**Runoff Erosion**
Precipitation (+)	Precipitation (+)
Ground Cover* (-)	Ground Cover* (-)
Soil Type – Detachability (K) (+)	Soil Type – Cohesiveness (COH) (-)
	Upstream contributing area (+)
	Slope (+)

*The simplest raw representation of ground cover is LAI

The difficulty in this method is assigning the correct weights to the factors. Investigation into the appropriate weights is beyond the scope of this study so for simplicity, the factors are assumed to have equal weights. The cells are then binned into three classes based on their relative values: (i) The lowest 20th percentile the values constitute the low erosion class, (ii) Values between the 20th and 80th percentile are classified as moderate erosion and; (iii) the highest 20th percentile constitute the high erosion class. The generated heuristic maps for the year, 2010, are shown on Figure 9. MMF model results classified under the same definition are shown on Figure 8 for comparison.

There is a general trend in the MMF model results of the northern section having lower erosion than the southern section. On the contrary, the heuristic map predicts sudden shifts in the areas of high and low erosion because it becomes very sensitive to shifts in the inputs without regard of the actual magnitudes as these are normalized.

The soil map foot print is consistently evident in the heuristic map and the erosion along the stream network is not continuous. Convergence of flow along the network should generate a continuity of erosion as evident in the MMF results. These results suggest that without the appropriate weights, a heuristic map may not give very meaningful results and the identification of these weights may require similar resources to applying a model.

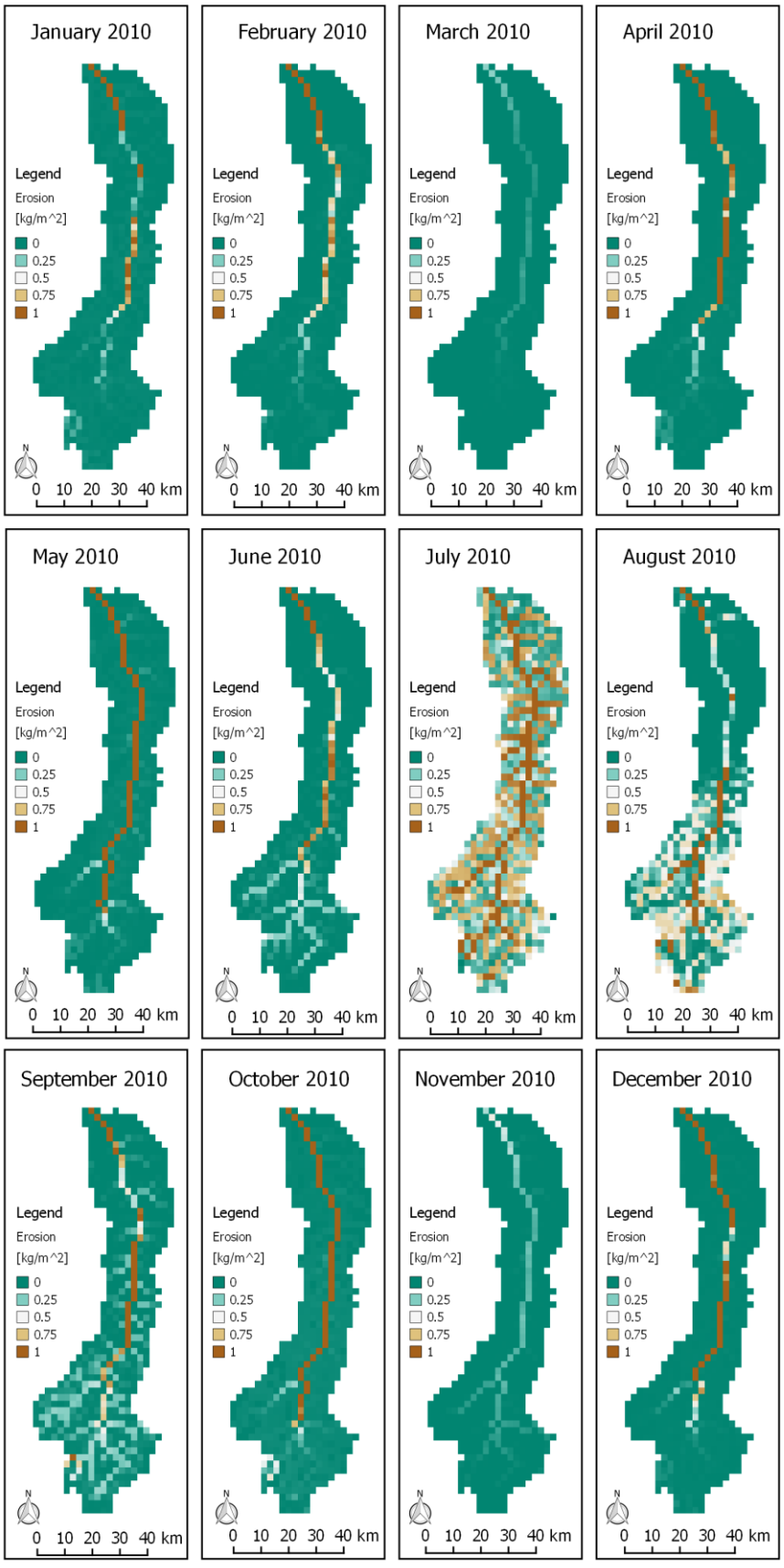
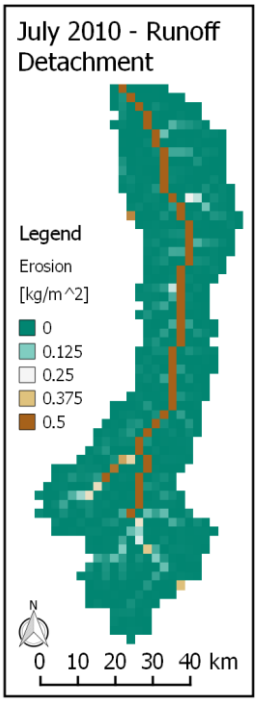


Figure 6. Monthly erosion [kg/m^2] MMF Result



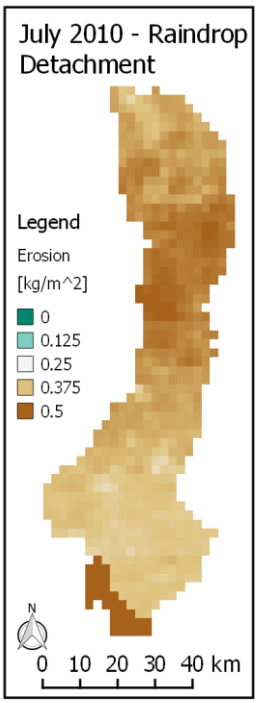


Figure 7. Runoff and raindrop detachment for July 2010

Table 8. Monthly MMF Model Erosion Results

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
Area in Soil loss Class (%)														**Average**
Class	**Soil loss (kg/m^2)**													
slight loss	<=0.05	91.9%	92.6%	94.1%	91.5%	86.7%	82.8%	2.2%	49.6%	69.8%	87.7%	90.0%	91.1%	77.5%
moderate	0.05-0.1	1.0%	0.3%	4.0%	1.6%	4.6%	3.5%	10.0%	6.8%	5.9%	2.7%	2.3%	1.6%	3.7%
high loss	0.1-0.15	1.3%	0.7%	1.0%	0.0%	0.4%	0.9%	10.5%	3.6%	2.7%	0.7%	1.6%	0.1%	2.0%
very high loss	0.15-0.3	1.7%	0.7%	0.9%	0.7%	0.6%	5.3%	19.5%	7.4%	11.7%	1.0%	5.1%	0.6%	4.6%
severe loss	0.3-0.6	0.4%	1.0%	0.0%	0.7%	0.6%	2.2%	14.1%	18.8%	4.0%	0.9%	1.0%	0.9%	3.7%
very severe loss	>0.6	3.6%	4.6%	0.0%	5.5%	7.1%	5.3%	43.7%	13.9%	5.8%	6.9%	0.0%	5.8%	8.5%
Total Soil loss (kg/m^2)		121.70	61.61	4.85	57.56	530.41	57.60	1484.64	281.19	87.04	275.06	13.00	113.14	
Area with Limiting Case (%)														
	Detachment	43.9%	30.3%	10.0%	36.8%	69.0%	11.7%	38.5%	25.0%	18.0%	69.0%	9.5%	58.3%	
	Transport	56.1%	69.7%	90.0%	63.2%	31.0%	88.3%	61.5%	75.0%	82.0%	31.0%	90.5%	41.7%	
Area with Dominant Erosion Process (%)														
	In-stream (Runoff)	8.7%	9.1%	6.5%	10.4%	25.1%	6.6%	7.5%	4.6%	6.9%	12.3%	2.7%	12.0%	
	On-Site (Raindrop)	91.3%	90.9%	93.5%	89.6%	74.9%	93.4%	92.5%	95.4%	93.1%	87.7%	97.3%	88.0%	
Average Soil loss in Slope Class(kg/m^2)														
	<=2	0.1466	0.0838	0.0070	0.0830	0.7463	0.0856	2.2139	0.3918	0.1353	0.3914	0.0195	0.1504	0.3712
	2-5	0.3601	0.1244	0.0073	0.0852	0.9203	0.0716	1.8068	0.4806	0.0723	0.4499	0.0148	0.2492	0.3869
	>5	0.0211	0.0042	0.0001	0.0469	0.0155	0.0120	0.1916	0.5976	0.0086	0.0159	0.0024	0.0171	0.0777
Average Soil loss in Land Use Class(kg/m^2)														
	Wheat	0.0043	0.0014	0.0001	0.0004	0.0060	0.0005	0.3570	0.0325	0.0073	0.0077	0.0006	0.0015	0.0349
	Soybean	0.3168	0.1198	0.0069	0.0840	0.8494	0.0612	1.8782	0.3887	0.0729	0.4201	0.0141	0.1983	0.3675
	Gram	0.0698	0.0688	0.0074	0.0859	0.7251	0.1031	2.4129	0.4521	0.1716	0.3935	0.0229	0.1324	0.3871

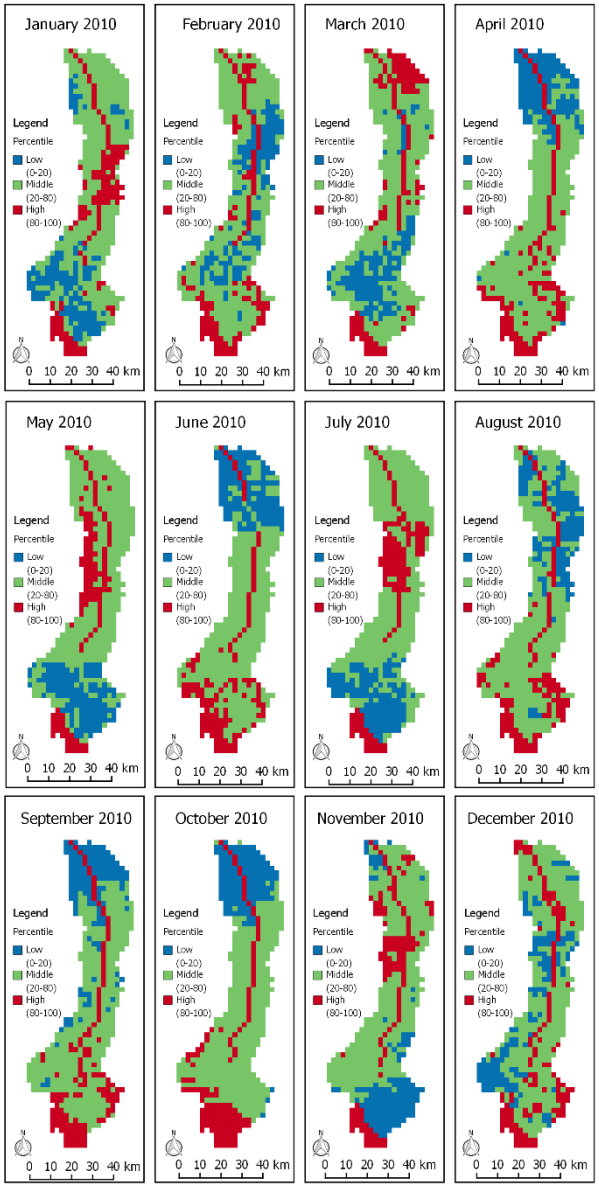


Figure 9. Spatial distribution of Low, Medium and High erosion classes based on a Heuristic map.

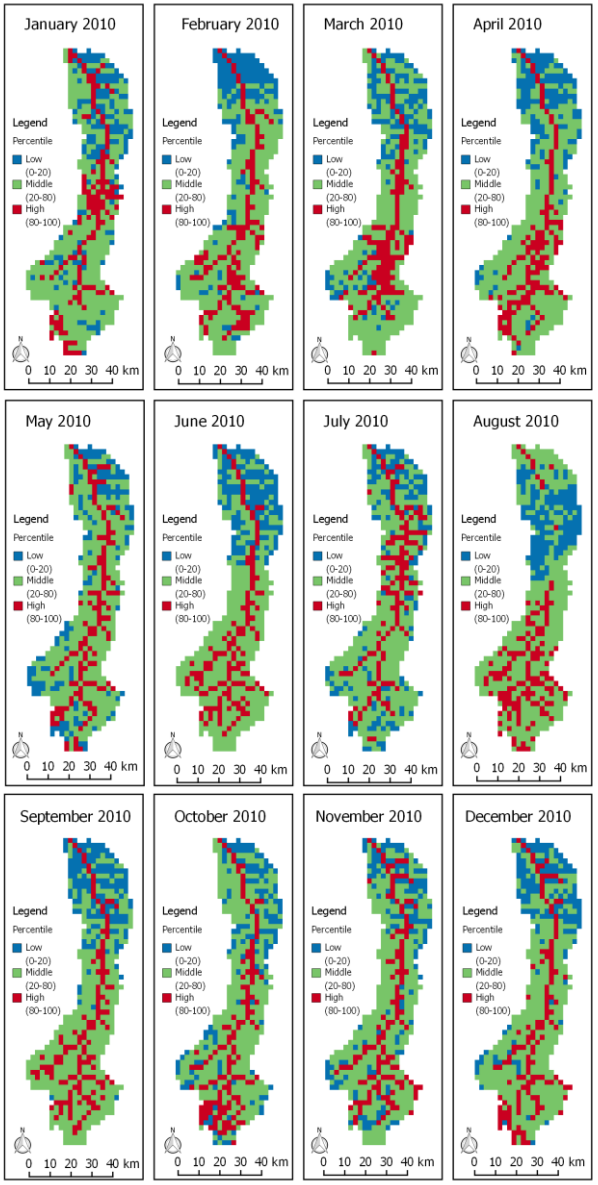


Figure 8. Spatial distribution of Low, Medium and High erosion classes based on the MMF results.

8. Conclusion

Selecting a soil erosion model appropriate for this research proved challenging because of the extensive library of models available and the contradicting conclusions of comparison studies. Model performance ostensibly relies on the context it is applied to. The model selection was thus executed by defining specific criterion according to the objectives of this research which resulted in the selection of the Morgan-Morgan-Finney (MMF) Model.

The MMF model was then redesigned to better fit the objectives of the study and alleviate some of its limitations. The modelled erosion result trends across the year are consistent with the seasonal crop calendar in the area. Furthermore, a sanity check with a comparison study validates to small degree, the soil erosion magnitudes and spatial distribution generated by the MMF model. However, the performance and reliability of the model has not been validated due to a lack of field data at the study area.

The redesigned MMF model improves on the original by providing more information on the distribution of erosion risk both temporally and spatially to help prioritize and identify high risk areas and inform mitigation measures. Additionally, by reducing its reliance on empirical information and by generating input parameters from remotely sensed data, exploratory application of the model globally is justifiable.

9. Recommendations

It is recommended that validation of the model with field data is prioritised in future works. After which, the model configuration can be improved by:

- Performing a comparison analysis of model results using empirical inputs and inputs derived from remotely sensed products;
- Applying a multi-flow direction routing algorithm, such as D∞, to account for divergent flow;
- Accounting for deposition in the same way as the USPED model which assumes that if erosion exceeds transport capacity deposition occurs at this cell;
- Testing application of the model with higher resolution data (i.e. 250x250m);
- Excluding streams/lakes from the model run as erosion prediction over these are likely inaccurate because hydraulics is not considered; and
- Replace empirical P-factor values with values derived from remote sensing products.

The model code can further be improved by:

- Allowing single month runs; and
- Incorporating input pre-processing.

References

Aiello, A., Adamo, M. and Canora, F. (2015) 'Remote sensing and GIS to assess soil erosion with RUSLE3D and USPED at river basin scale in southern Italy', CATENA. Elsevier, 131, pp. 174–185. doi: 10.1016/J.CATENA.2015.04.003.

Andresen, D. A. (2017) 'Modelling Sediment Yield in the Sink Creek and Purgatory Creek Watersheds Near San Marcos, Texas'. Available at: https://digital.library.txstate.edu/handle/10877/6929 (Accessed: 18 January 2018).

Arkowitz, A. (2017) 'Using GIS to compare leading process and empirically based soil erosion models within headwater watersheds'. Available at: http://openknowledge.nau.edu/4921/ (Accessed: 22 January 2018).

Arnold, J. G. et al. (1998) 'Large area hydrologic modelling and assessment part I: model development', JAWRA Journal of the American Water Resources Association. Wiley Online Library, 34(1), pp. 73–89.

Batjes, N. H. (1997) 'A world dataset of derived soil properties by FAO-UNESCO soil unit for global modelling', SoilUseandManagement, 13. doi: 10.1111/j.1475-2743.1997.tb00550.x.

Bayramov, E., Buchroithner, M. F. and Mcgurty, E. (2013) 'Differences of MMF and USLE Models for Soil Loss Prediction along BTC and SCP Pipelines'. doi: 10.1061/(ASCE)PS.1949-1204.0000117.

Beven, K. (1989) 'Changing ideas in hydrology—the case of physically-based models', Journal of hydrology. Elsevier, 105(1–2), pp. 157–172.

Brandt, C. J. (1990) ‘Simulation of the size distribution and erosivity of raindrops and throughfall drops’, Earth Surface Processes and Landforms. Wiley Online Library, 15(8), pp. 687–698.

Brazier, R. (2013) ‘Hillslope Soil Erosion Modeling’, in Treatise on Geomorphology, pp. 135–146. doi: 10.1016/B978-0-12-374739-6.00029-4.

Central Bank of India (no date) Crop Season crops in M.P. Available at: http://www.slbcmadhyapradesh.in/docs/CropSeason in M.P..pdf (Accessed: 18 April 2018).

Chaplot, V. (2005) ‘Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO 3 –N loads predictions’. doi: 10.1016/j.jhydrol.2005.02.017.

Davison, P. S. et al. (2008) ‘PSYCHIC – A process-based model of phosphorus and sediment mobilisation and delivery within agricultural catchments. Part 1: Model description and parameterisation’, Journal of Hydrology, 350, pp. 290–302. doi: 10.1016/j.jhydrol.2007.10.036.

Dunne, T. and Leopold, L. (1978) Water In Environmental Planning, XF2006179213.

ESA (no date) Map of Forest and protected areas. Available at: https://www.esa-landcover-cci.org/.

FAO (2013) Ecocrop - Plant Search Form. Available at: http://ecocrop.fao.org/ecocrop/srv/en/cropFindForm (Accessed: 3 May 2018).

Feng, T. et al. (2014) ‘Modeling soil erosion using a spatially distributed model in a karst catchment of northwest Guangxi, China’, Earth Surface Processes and Landforms, 39(15), pp. 2121–2130. doi: 10.1002/esp.3625.

Fernández, C., Vega, J. A. and Vieira, D. C. S. (2010) ‘Assessing soil erosion after fire and rehabilitation treatments in NW Spain: Performance of rusle and revised Morgan-Morgan-Finney models’, Land Degradation and Development. John Wiley & Sons, Ltd., 21(1), pp. 58–67. doi: 10.1002/ldr.965.

Funk, C. et al. (2015) ‘The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes’, Scientific data, (2), p. 150066. Available at: chg.geog.ucsb.edu/data/chirps/.

Furl, C., Sharif, H. and Jeong, J. (2015) ‘Analysis and simulation of large erosion events at central Texas unit source watersheds’, Journal of hydrology, 527, pp. 494–504. doi: 10.1016/j.jhydrol.2015.05.014.

Goethe University (no date) Global data set of monthly irrigated and rainfed crop areas around the year 2000 (MIRCA2000). Available at: https://www.uni-frankfurt.de/45218023/MIRCA.

Jain, M. K. and Kothyari, U. C. (2000) ‘Estimation of soil erosion and sediment yield using GIS’, Hydrological Sciences Journal. Taylor & Francis Group, 45(5), pp. 771–786. doi: 10.1080/02626660009492376.

Jha, M. K. and Paudel, R. C. (2010) ‘Erosion Predictions by Empirical Models in a Mountainous Watershed in Nepal’, Journal of Spatial Hydrology, 10(1). Available at: http://spatialhydrology.net/index.php/JOSH/article/viewFile/91/90 (Accessed: 26 February 2018).

De Jong, S. M. et al. (1999) ‘Regional assessment of soil erosion using the distributed model SEMMED and remotely sensed data’, Catena, 37, pp. 291–308. Available at: www.elsevier.comrlocatercatena (Accessed: 15 March 2018).

Karydas, C. G., Panagos, P. and Gitas, I. Z. (2014) ‘A classification of water erosion models according to their geospatial characteristics’, International Journal of Digital Earth, 7(310), pp. 229–250. Available at: http://www.tandfonline.com/doi/full/10.1080/17538947.2012.671380#.Uvj95fmSwTM (Accessed: 26 February 2018).

Van Der Knijff, J. M., Jones, R. J. A. and Montanarella, L. (1999) Soil Erosion Risk Assessment in Italy - European Soil Bureau. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.397.2309&rep=rep1&type=pdf (Accessed: 20 March 2018).

Land and Development Division (FAO:ROME) (2003) ‘The Digital Soil Map of the World’, in Food and Agriculture Organization of the United Nations. Version 3.

Lehner, B., Verdin, K. and Jarvis, A. (2008) ‘New global hydrography derived from spaceborne elevation data’, Eos, Transactions, AGU, 89(10), pp. 93–94.

Li, P. et al. (2017) ‘Comparison of soil erosion models used to study the Chinese Loess Plateau’. doi: 10.1016/j.earscirev.2017.05.005.

Merritt, W. S., Letcher, R. A. and Jakeman, A. J. (2003) ‘A review of erosion and sediment transport models’, Environmental Modelling & Software, 18, pp. 761–799. doi: 10.1016/S1364-8152(03)00078-1.

Meyer, L. D. and Wischmeier, W. H. (1969) ‘Mathematical simulation of the process of soil erosion by water’, Transactions of the ASAE. American Society of Agricultural and Biological Engineers, 12(6), pp. 754–758.

Mitasova, H. et al. (2013) ‘GIS-Based Soil Erosion Modeling’, in Treatise on Geomorphology. Elsevier, pp. 228–258. doi: 10.1016/B978-0-12-374739-6.00052-X.

Mondal, A., Khare, D. and Kundu, S. (2018) ‘A comparative study of soil erosion modelling by MMF, USLE and RUSLE’, Geocarto International. Taylor & Francis, 33(1), pp. 89–103. doi: 10.1080/10106049.2016.1232313.

Morgan, R. P. C. (1995) Soil Erosion & Conservation. 2nd edn. Longman Group Limited.

Morgan, R. P. C. (2001) ‘A simple approach to soil loss prediction: a revised Morgan–Morgan–Finney model’, Catena, 44, pp. 305–322. Available at: www.elsevier.comrlocatercatena (Accessed: 14 February 2018).

Morgan, R. P. C. and Duzant, J. H. (2008) ‘Modified MMF (Morgan–Morgan–Finney) model for evaluating effects of crops and vegetation cover on soil erosion’, Earth Surface Processes and Landforms. John Wiley & Sons, Ltd., 33(1), pp. 90–106. doi: 10.1002/esp.1530.

Morgan, R. P. C., Morgan, D. D. V. and Finney, H. J. (1984) ‘A predictive model for the assessment of soil erosion risk’, Journal of Agricultural Engineering Research. Academic Press, 30, pp. 245–253. doi: 10.1016/S0021-8634(84)80025-6.

Murray, A. B. (2013b) ‘Which Models Are Good (Enough), and When?’, in Treatise on Geomorphology, pp. 50–58. doi: 10.1016/B978-0-12-374739-6.00027-0.

NASA LAP DAAC (2010) ‘NDVI and LAI data. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota

(https://lpdaac.usgs.gov)'. Available at: https://e4ftl01.cr.usgs.gov/MOLT/ (Accessed: 15 March 2018).

Nearing, M. A. (1998) 'Why soil erosion models over-predict small soil losses and under-predict large soil losses', Catena, 32, pp. 15–22. Available at: https://ac.els-cdn.com/S0341816297000520/1-s2.0-S0341816297000520-main.pdf?_tid=fdae6d2e-080c-11e8-87db-00000aab0f27&acdnat=1517571432_b1418aa976e601a74373c09d3cb3a47f (Accessed: 2 February 2018).

NIRCA (no date) Data on major crops/district. Available at: http://www.nicra-icar.in/nicrarevised/index.php/home1 (Accessed: 14 May 2018).

Oliveira, A. H. et al. (2013) 'Development of Topographic Factor Modeling for Application in Soil Erosion Models'. doi: 10.5772/54439.

Pandey, A. et al. (2009) 'Soil erosion modeling of a Himalayan watershed using RS and GIS', Environmental Earth Sciences. Springer-Verlag, 59(2), pp. 399–410. doi: 10.1007/s12665-009-0038-0.

Pham, T. G., Degener, J. and Kappas, M. (2018) 'Author's Accepted Manuscript - Intergrated Universal Soil Loss Equation (USLE) and Geographical Inforamation System (GIS) for Soil Erosion Estimation in a SAP Basin; Central Vietname'. doi: 10.1016/j.iswcr.2018.01.001.

Romero-Díaz, A., Alonso-Sarriá, F. and Martínez-Lloris, M. (2007) 'Erosion rates obtained from check-dam sedimentation (SE Spain). A multi-method comparison', Catena. Elsevier, 71(1), pp. 172–178.

Shrestha, D. P. (1997) 'Assessment of Soil Erosion in the Napalese Himalaya, a Case Study in Likhu Khola Valley, Middle Mountain Region'. Oxford & IBH Publishing Co. Pvt. Ltd, 2(1), pp. 59–80. Available at: http://www.thinkgeography.org.uk/Year 10 Geog/Year 10/Soils/erosion.pdf (Accessed: 15 March 2018).

Shrestha, D. P. and Jetten, V. G. (2018) 'Modelling erosion on a daily basis, an adaptation of the MMF approach', International Journal of Applied Earth Observation and Geoinformation. Elsevier, 64, pp. 117–131. doi: 10.1016/J.JAG.2017.09.003.

Svorin, J. (2003) 'A test of three soil erosion models incorporated into a geographical information system', Hydrological Processes, 17(5), pp. 967–977. doi: 10.1002/hyp.1174.

Tesfahunegn, G. B., Tamene, L. and Vlek, P. L. G. (2014) 'Soil erosion prediction using Morgan-Morgan-Finney model in a GIS environment in northern Ethiopia catchment', Applied and Environmental Soil Science, 2014. doi: 10.1155/2014/468751.

De Vente, J. and Poesen, J. (2005) 'Predicting soil erosion and sediment yield at the basin scale : Scale issues and semi - quantitative models', Earth - Science Reviews, 71, pp. 95–125. doi: 10.1016/j.earscirev.2005.02.002.

Vieira, D. C. S. et al. (2014) 'Modelling runoff and erosion, and their mitigation, in burned Portuguese forest using the revised Morganâ€"Morganâ€"Finney model'. doi: 10.1016/j.foreco.2013.12.006.

Vigiak, 0 and Sterk, G. (2001) 'Empirical water erosion modelling for soil and water conservation planning at catchment-scale'. Available at: https://www.witpress.com/Secure/elibrary/papers/ECO01/ECO01022FU.pdf (Accessed: 26 February 2018).

Wischmeier, W. H. and Smith, D. D. (1978) 'Predicting rainfall erosion losses', Agriculture handbook no. 537, (537), pp. 285–291. doi: 10.1029/TR039i002p00285.

Yazidhi, B. (2003) 'A comparative study of soil erosion modelling in Lom Kao-Phetchabun, Thailand'. Available at: https://www.itc.nl/library/Papers_2003/msc/ereg/bamutaze_yazidi.pdf (Accessed: 15 March 2018).

Appendices

Appendix 1: Morgan-Morgan-Finney Model Running Instructions

The MMF model generated for this research is simulated using two python scripts: RUN.py and MMF.py. These scripts can be found in Appendix 2 and Appendix 3 respectively. RUN.py is the executable file which imports the inputs and executes the MMF.py script. The user is required to enter file handles for the input files and the preferred location for the output files **(Note: monthly input file names must have 2-integer numbering from 01 to 12 referring their associated months).** The model accepts exactly 12 files for monthly data as it has been programmed to run for 1 year.

MMF.py contains the operating functions organized in classes associated with the main phases of the model, *WaterPhase, SoilPhase and TransportCapacity*. The *output* class applies the scaling factor, determines the final erosion rates and generates the output of the model. The model calculates the distributed monthly erosion at each month in the same resolution and extent of the input data. The files are written in geotiff format and can be viewed in GIS software.

Appendix 2: RUN.py

```
import numpy as np
import wa
import becgis
import matplotlib.pyplot as plt
import glob
import netCDF4 as nc
from collections import Counter
import MMF

## Input parameters/filehandles are specified here

outputfh = #file handle where output tif's are saved

Itype= #0=Temperate,1=Tropical,2=Strongly Seasonal
KEtype= #0=North America east of the Rocky Mountains,1=North-western Europe, 2=Mediterreanean-type,
        #3=Western Mediterranean, 4=Tropical, 5=Eastern Asia,
        # 6=Temperate southern hemisphere climates (i.e. Australia)

## Single files

PH = becgis.OpenAsArray('.../Plant_Height.tif')
SoilMap = becgis.OpenAsArray('.../Soil_map.tif')
Steepness= becgis.OpenAsArray('.../Steepnes.tif')
Pfactor = becgis.OpenAsArray('.../Pfactor.tif')

## Multiple files

Runofflist = glob.glob('.../Monthly_Runoff_*.tif')
LAIlist = glob.glob('.../Monthly_LAI_*.tif')
NDVIlist = glob.glob('.../Monthly_NDVI_*.tif')

##Files input to WaterPix

NC = nc.Dataset('.../Forcing_input_from_WaterPix.nc')
Year = 2010

Pall = NC.variables['Precipitation_M']
Nall = NC.variables['RainyDays_M']

list = np.array([t/100 for t in NC.variables['time_yyyymm'][:]])
P = Pall[np.where(list==Year)]
N = Nall[np.where(list==Year)]

geoinfo = becgis.GetGeoInfo(fh=Runofflist[0])

MMF = MMF.output(LAIlist,P,N,Itype,KEtype,Runofflist,NDVIlist,geoinfo, PH, SoilMap, Steepness,Pfactor,outputfh)
```

Appendix 3: MMF.py

```
import numpy as np
import wa
import becgis
import matplotlib.pyplot as plt
import glob
import netCDF4 as nc
from collections import Counter

class WaterPhase():
    def __init__(self,LAIM,PM,NM,Itype,KEtype,PH,Ilib,KElib):

        m,n = np.shape(PM)

        for i in range (m):
            for j in range (n):
                if PM[i,j]>0 and NM[i,j]<1:
                    NM[i,j]=1

        self.Cov = self.Cover(LAIM,m,n)
        self.Int = self.Interception(LAIM,PM,NM,m,n)
        self.ke = self.KE(PM,Itype,KEtype,Ilib,KElib, PH, m,n)

    def Interception(self,LAIM,PM,NM,m,n):

        A_m = np.zeros([m,n])

        for i in range(m):
            for j in range(n):
                try:
                    I=LAIM[i,j] * (1 - (1 + (PM[i,j]/NM[i,j]) * (1 - np.exp(-0.5 * LAIM[i,j])) * (1/LAIM[i,j]))**-1)  * NM[i,j]
                    A_m[i,j] = (I/PM[i,j])
                except:
                    A_m[i,j] = np.nan
        return A_m

    def Cover(self,LAIM,m,n):

        CC= np.zeros([m,n])

        for i in range(m):
            for j in range(n):
                try:
                    CC[i,j] = 1-np.exp(-0.5 * LAIM[i,j])
                except:
                    CC[i,j] = np.nan
        return CC

    def KE(self,PM,Itype,KEtype, Ilib, KElib, PH, m,n):

        KE = np.zeros([m,n])
        KE_DT = np.zeros([m,n])
        KE_LD = np.zeros([m,n])

        for i in range(m):
            for j in range(n):
                try:
                    ER= (1-self.Int[i,j])*PM[i,j]
                    LD = ER*self.Cov[i,j]
                    DT = ER - LD

                    I = Ilib[Itype]

                    KE_DT[i,j] = (eval(KElib[KEtype]))*DT
                    KE_LD[i,j] = LD*((15.8*PH[i,j]**0.5)-5.87)

                    if KE_LD[i,j] < 0:
                        KE[i,j] = KE_DT[i,j]
                    else:
```

```
                    KE[i,j] = KE_DT[i,j]+KE_LD[i,j]

                except:
                    KE[i,j] = np.nan

        return KE,KE_DT,KE_LD

class SoilPhase():
    def __init__(self,
                 Soillib,
                 SoilMap,
                 Steepness,
                 RunoffM,
                 KE,
                 GC):

        m,n = np.shape(Steepness)

        self.k = self.Soil(Soillib,SoilMap,m,n)[0]
        self.COH = self.Soil(Soillib,SoilMap,m,n)[1]

        #Calculate detachment by rainfall

        self.F = self.k*KE*10**-3

        #Calculate detachment by runoff

        self.H = (1/(0.5*self.COH))*(RunoffM**1.5)*np.sin(np.deg2rad(Steepness))*(1-GC)*10**-3

        self.Total = self.F+self.H

    def Soil(self,Soillib,SoilMap,m,n):

        k = np.zeros([m,n])
        COH = np.zeros([m,n])

        for i in range(m):
            for j in range(n):
                try:
                    k[i,j]= Soillib[int(SoilMap[i,j]-1)][0]
                    COH[i,j]= Soillib[int(SoilMap[i,j]-1)][1]
                except:
                    k[i,j] = np.nan
                    COH[i,j] = np.nan
        return k,COH

class TransportCapacity():
    def __init__(self,
                 NDVIM,
                 Steepness,
                 RunoffM,
                 Pfactor):

        m,n = np.shape(NDVIM)

        self.Cfactor = self.CFactor(NDVIM,Pfactor,m,n)

        self.Cap = self.Cfactor[1]*(RunoffM**2)*np.sin(np.deg2rad(Steepness))*(10**-3)

    def CFactor (self,NDVIM, Pfactor, m,n):

        C = np.zeros([m,n])
        CF = np.zeros([m,n])

        for i in range(m):
            for j in range(n):
                try:
                    C[i,j]= np.exp(-2*NDVIM[i,j]/(1-NDVIM[i,j]))
                    CF[i,j]=C[i,j]*Pfactor[i,j]
                except:
```

```
                C[i,j] = np.nan
                CF[i,j] = np.nan
        return C,CF

class output():
    def __init__(self,LAIlist,P,N,Itype,KEtype,Runofflist,NDVIlist,geoinfo, PH, SoilMap,Steepness, Pfactor,outputfh):

        l,m,n = np.shape(P)

        Ilib = [10,25,30] #[Temperate,Tropical,Strongly Seasonal]

        KElib = ['11.8+8.73*(np.log10(I))', #North America east of the Rocky Mountains
                 '8.95+8.44*(np.log10(I))', #North-western Europe
                 '9.81+11.25*(np.log10(I))', #Mediterreanean-type
                 '35.9*(1-0.56*np.exp(-0.034*I))', #Western Mediterranean
                 '29.8-(127.5/I)', #Tropical
                 '9.81+10.60*(np.log10(I))', #Eastern Asia
                 '29*(1-0.6*np.exp(-0.04*I))'] #Temperate southern hemisphere climates (i.e. Australia)

        Soillib = ([1.2,2],[0.3,2],[0.7,2],[0.8,3],[1,3],[0.9,3],[0.1,3],[0.7,10],[0.8,9],[0.3,9],[0.5,10],[0.05,12])

        self.Actualtemp = np.zeros([m,n,12])
        self.Actual = np.zeros([m,n,12])
        self.Capacitytemp = np.zeros([m,n,12])
        self.Capacity = np.zeros([m,n,12])


        for i in range (12):
            RunoffM = np.abs(becgis.OpenAsArray(Runofflist[i]))
            #Negative numbers in the runoff raster indicate that those cells possibly have surface runoff from
            #outside of the current geographic region. As this model is developed for catchmetn application
            #it is assumed that surface runoff comes from within the geographic region and thus the absolute of the
            #raster is taken.
            LAIM = becgis.OpenAsArray(LAIlist[i])
            NDVIM = becgis.OpenAsArray(NDVIlist[i])
            PM = P[i,:,:]
            NM = N[i,:,:]

            WP = WaterPhase(LAIM,PM,NM,Itype,KEtype,PH, Ilib, KElib)
            ke = WP.ke[0]
            Atemp = WP.Int
            CC = WP.Cov

            TC = TransportCapacity(NDVIM,Steepness,RunoffM, Pfactor)
            self.Capacitytemp[:,:,i]=TC.Cap
            GC = 1-TC.Cfactor[0]

            SP = SoilPhase(Soillib,SoilMap,Steepness,RunoffM,ke,GC)
            self.Actualtemp[:,:,i] = SP.Total

        ScaleA = ((np.sum(self.Actualtemp,axis=2))**1.5)/(np.sum((self.Actualtemp**1.5),axis=2))

        self.Actual = self.Actualtemp*(np.repeat(ScaleA[:, :, np.newaxis], 12, axis=2))

        ScaleC = ((np.sum(self.Capacitytemp,axis=2))**2)/(np.sum((self.Capacitytemp**2),axis=2))
        self.Capacity = self.Capacitytemp*(np.repeat(ScaleC[:, :, np.newaxis], 12, axis=2))

        self.Erosion = np.minimum(self.Actual,self.Capacity)

        for i in range(12):

            becgis.CreateGeoTiff(outputfh+'_'+str(i+1)+'.tif',
                                 self.Erosion[:,:,i],
                                 geoinfo[0],
                                 geoinfo[1],
                                 geoinfo[2],
                                 geoinfo[3],
                                 geoinfo[4],
                                 geoinfo[5])
```