Physics and Chemistry of the Earth 34 (2009) 741–748

Contents lists available at ScienceDirect

Physics and Chemistry of the Earth

journal homepage: www.elsevier.com/locate/pce

Estimating gully erosion contribution to large catchment sediment yield rate in Tanzania

Preksedis Marco Ndomba [a,*], Felix Mtalo [a], Aanund Killingtveit [b]

[a] University of Dar es Salaam, Department of Water Resources Engineering, P.O. Box 35131, Dar es Salaam, Tanzania

[b] University of Science and Technology, Department of Hydraulic and Environmental Engineering, Faculty of Engineering, N-7491 Trondheim, Norway

ARTICLE INFO

Article history:
Received 21 January 2009
Received in revised form 16 June 2009
Accepted 30 June 2009
Available online 9 July 2009

Keywords:
Cubist
Data mining tool
Gully erosion
Sediment delivery ratio
Sediment yield

ABSTRACT

The objective of this paper is to report on the issues and proposed approaches in estimating the contribution of gully erosion to sediment yield at large catchment. The case study is the upstream of Pangani River Basin (PRB) located in the North Eastern part of Tanzania. Little has been done by other researchers to study and/or extrapolate gully erosion results from plot or field scale to large catchment. In this study multi-temporal aerial photos at selected sampling sites were used to estimate gully size and morphology changes over time. The laboratory aerial photo interpretation results were groundtruthed. A data mining tool, Cubist, was used to develop predictive gully density stepwise regression models using aerial photos and environment variables. The delivery ratio was applied to estimate the sediment yield rate. The spatial variations of gully density were mapped under Arc View GIS Environment. Gully erosion sediment yield contribution was estimated as a ratio between gully erosion sediment yield and total sediment yield at the catchment outlet. The general observation is that gullies are localized features and not continuous spatially and mostly located on some mountains' foot slopes. The estimated sediment yield rate from gullies erosion is 6800 t/year, which is about 1.6% of the long-term total catchment sediment yield rate. The result is comparable to other study findings in the same catchment. In order to improve the result larger scale aerial photos and high resolution spatial data on soil-textural class and saturated hydraulic conductivity – are recommended.

1. Introduction

Gully erosion features can be categorised into two main types, namely classical gullies and ephemeral gullies. Ephemeral gullies will reform in the same location in a field where flow from land slope region concentrates (Chmelova and Sarapatka, 2002). Sometimes ephemeral gullies are called gullying sheet erosion, when the channels are between 10 and 20 cm deep but as much as several metres wide. Classical gullies are an advanced stage of channel erosion. They are formed when channel development has progressed to the point where the gully is too wide and too deep to be tilled across (Chmelova and Sarapatka, 2002). The channels are at least 50 cm deep and, more specifically, they can no longer be eliminated through cropping techniques. These channels carry large amount of water after rains and deposit eroded material at the foot of the gully. Gullies form in various shapes including V-shaped, with an even slope down to the lowest point; U-shaped, with vertical sides; and still others develop through tunnelling and subsidence.

Most of the available study methods (field measurements or modelling) of gully erosion work well at small basins and preferably at plot scales (Ndomba, 2007). They have the disadvantage that their results cannot easily be extrapolated to larger basins. For instance, field measurements as direct methods have been reported by other workers to give un-reliable estimates (Peart and Walling, 1988). Other workers such as Wasson (2002) have noted the transferability problem of plot or micro scale studies results to larger catchment.

Erosion models that have been developed for modelling gully erosion range from simple empirical models such as SedNet (Wilkinson et al., 2004) and physics based models such as SHETRAN (Bathurst, 2002) and Limburg Soil Erosion Model (LISEM) Gullies, European Soil Erosion Model (EUROWISE) (Jetten, 2002). In a physically deterministic way for instance, incision can be modelled by using shear stress of the flow along the sides and bottom of the gully and compare it to the shear strength of the profile (Jetten, 2002). This would require a complete knowledge of the soil cohesion and variation of cohesion with depth, and the capability of simulating the flow conditions in detail. Such kind of information is not readily available in ungauged (i.e. poorly gauged) and at a catchment scale. Therefore, in most cases the only option is to use simple tools and readily available data (Ndomba, 2007).

* Corresponding author. Tel.: +255 754 635 272.
E-mail address: pmndomba2002@yahoo.co.uk (P.M. Ndomba).

1474-7065/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.pce.2009.06.009

Empirically, The USDA (1977) determined that at least six factors significantly influenced the rate of headcut advance: tributary watershed area, storm precipitation, soil erodibility, slope of approach channel above the headcut, groundwater level, and changes in runoff due to land use changes above the gully. The tributary watershed area and precipitation were found to best explain the rate of gully headcut advance. It should be noted that regression equations so far developed are meant for individual gullies within specific regions (Ndomba, 2007).

Gully erosion processes are not well understood in literature (Ndomba, 2007). Therefore, in this study gully erosion was modelled by data mining tools. This method usually involves the fitting of very complex "generic" models that are not related to any reasoning or theoretical understanding of underlying causal processes; instead, these techniques can be shown to generate accurate predictions in validation samples (Statsoft, 2006). Regression analysis was conducted between environmental variables (i.e. spatial data on landuse, mean annual rainfall, relief, soil characteristics such as soil texture) and gully density as dependent variable. It should be noted that the role of some of these environmental variables as degradation factors in the spatial and temporal domain of soil degradation intensity have been well documented by a number of workers (Hudson, 1971; and Yanda, 1995). However, Yanda (1995) cautioned that the interplay of these factors is site specific; in some places, the slope may be most important factor, while in others the human factors may be the most important. Rainfall is an important soil erosion factor as it determines vegetation cover and acts as a detaching as well as transporting agent (Yanda, 1995). Relief factors such as slope and aspect influence erosion. For instance, steep slopes render much of the area susceptible to erosion. This is because steep slopes increase the volume and velocity of surface runoff, which causes soil detachment and transportation. On steeper slopes the process can be intense enough to form gullies. Also, bare rock surfaces found at many hill crests permit instantaneous generation of surface runoff (Christiansson, 1972). Long slopes are more affected and particularly those, which drain a large subcatchment area. In Mwisanga catchment in Tanzania the total soil stripping on such steeply sloping parts of the catchment has been facilitated by the original relative shallowness of the soil and weathering profiles in such areas (Yanda, 1995). Landuse factors play an important role in land degradation. This is mainly through vegetation clearing, which exposes the soil surface to water impact.

Data mining is an analytic process designed to explore data usually large amounts of data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data (Statsoft, 2006). Data mining is also known as knowledge discovery, machine learning or computational learning theory and the ultimate goal of data mining is prediction (Statsoft, 2006). The data mining tool used in this study is Cubist (Rulequest Research, 2004). Cubist is a tool for generating rule based predictive model from data (Rulequest Research, 2004).

Gully erosion in Mount Meru slopes in the Pangani River Basin (PRB) has been reported by Semu et al. (1992) as significant land degradation processes and sources of sediment to PRB. They reported presence of very big gullies up to 10 or more metres wide, with a depth of up to 10 m. Some of them are several 100 m long or more (Semu et al., 1992). However, little was known on the factors that influence the formation of gullies and contribution of gully erosion to the sediment yield in the study area.

Ideally, the data requirement to map and/or simulate gully size and morphological changes is huge. Some of data set used by others (Martinez-Casasnovas et al., 2003; Hughes and Prosser, 2003; and Yanda, 1995) includes: gully depth and width, thickness of soil horizon on the gully walls, documentation of gully morphologies and gullying processes based on photos, multi-temporal of black and white aerial photos taken during the dry season for the selected sites in the catchment with active growing gullies, ground truthing report, landform units, soil geological classification, soil characteristics such as soil thickness, A-horizon texture and B-horizon texture, climate indices such as temperature seasonality, minimum temperature-coldest period, temperature-annual change, mean annual precipitation, lowest period moisture index, moisture index seasonality, landuse map, Relief such as slope and hill slope length derived from Digital Elevation Model (DEM), High resolution DEM (i.e. 1 m by 1 m), annual average ground cover derived from Normalized Differential Vegetation Index (NDVI) based on Advanced Very High Resolution Radiometer (AVHRR) remote sensing. It should be noted that it was not possible to obtain all these data at a catchment scale for our study case. Thus, only 10 environmental variables as presented in Table 1 below were used.

Therefore, the objective of this paper is to report on the issues and proposed approaches in estimating the contribution of gully erosion to sediment yield at large catchment.

2. Materials and methods

2.1. Description of the study area

The study area is located in the upstream of Pangani River Basin (PRB), in the northeastern part of Tanzania (Fig. 1). The catchment outlet is located at Nyumba Ya Mungu reservoir (NYM) dam. The NYM reservoir catchment regulates about 12,000 km^2 area of upper PRB drained by rivers Kikuletwa (1DD1) and Ruvu (1DC1) (Ndomba, 2007). 1DD1 and 1DC1 subcatchments cover 7280 km^2 and 2590 km^2, respectively. The 1DD1 and 1DC1 gauging stations are located about 15 km and 5 km, respectively upstream of the reservoir (Ndomba, 2007). Gauging station at 1D8C near NYM dam is a most downstream boundary (Fig. 1). The NYM reservoir catchment comprises of complex geological formations such as North Pare Mountains, Mount Kilimanjaro, Mount Meru, Lakes

Table 1
Statistics of environmental variables data (i.e. 12179 data points) for Nyumba Ya Mungu (NYM) reservoir catchment as extracted from digital spatial data.

Statistic	Environmental variables									
	Elevation (masl)	Aspect (°)	SOL_WHC (mm)	SOL_texture (–)	SOL_k (mm/h)	SOL_Z (mm)	Slope (°)	Sll (m)	MAR (mm)	CN2 (–)
Minimum	737	0	10	2	0.1	10	0	61	372	57
Maximum	5825	360	154	3	0.1	185	26.1	122	2110	98
Mean	1243	162	86.2	2	0.1	96	2.3	95	699	77
Standard deviation	481	88	26.2	0	0	38	3.2	25	333	5
Coefficient of variation Cv (%)	38.7	54.3	30.4	0.0	0.0	39.6	139.1	26	47.6	6.5

Note: Elevation = ground elevation of the upland catchment, aspect = steepest down-slope direction with reference to north, SOL_WHC = soil water holding capacity, SOL_texture = textural class of the soil, SOL_k = saturated hydraulic conductivity of the soil, SOL_Z = soil depth, slope = upland ground slope, Sll = slope length, MAR = long-term mean annual rainfall, and CN2 = curve number for moderate moisture condition.