Optimizing sediment sampling programme in a catchment. The case of Pangani River Basin

Preksedis Marco NDOMBA[1]

[1]University of Dar es Salaam, Department of Water Resources Engineering, P.O. Box 35131, Dar es Salaam, Tanzania; Email: pmndomba2002@yahoo.co.uk

Abstract

In previous sedimentation studies in Pangani River Basin located in the Northeastern part of Tanzania, it has been assumed that data set available for rating curve construction or estimating sediment yield/load represents a near-optimum data collection scheme and that estimates of the errors involved are minimum estimates. Other researchers elsewhere have done little as well, in quantifying errors due to varying number of data points. Besides, the performance of the sediment sampling programmes have been compromised due to lack of information on informed hydro-climatic variables triggering much upland soil loss and delivery of sediment loads to catchment outlets. Therefore, this study quantitatively estimated the number of data points required for optimal sediment sampling through developing an "efficient" rating curve with informed hydro-climatic condition. This was achieved by optimization techniques using readily available 291 sub-daily data points on sediment concentration and stream flows. A total of 274 randomly generated samples of different sizes were used to fit the rating curves by Ordinary Least Square (OLS) approach. However, the analysis was constrained to using more than 16 data points as minimum number for a two-parameter model such as rating curve parameterization/regression. The developed rating curves were used to estimate the annual sediment loads during the sediment sampling programme period. The Relative error in percent (RE) as index of performance between estimated load and "actual load" was computed. RE less than 20% was considered as satisfactory for engineering practices. Findings showed that a sediment concentration sample of less than 30 data points; *i.e*., 17, 26, and 27; covering critical hydrological conditions is adequate. An independent analysis indicates that these samples are from selected runoff events in months of March, April, May, and June. Events of April mostly dominated the sediment delivery to the catchment outlet. As this result is based on one year sediment sampling programme, thus there is a need to validate the results using recent sediment flow data in the same catchment. The sampling programme proposed is site specific, so it may not be extended easily to other catchments. Further studies should also analyze the reliability of the proposed sample size for optimal sampling programme in Pangani River Basin and other catchments.

***Keywords**:* Optimization, Pangani River Basin, Sediment rating curve, Sediment sampling programme, Sedimentation.

1. INTRODUCTION

The objective of suspended-sediment sampling is to determine the mean discharge-weighted sediment concentration and grain size in the cross section, despite the variation in sediment concentration and the velocity distribution across the stream and within each sampling vertical. Having sampled the discharge-weighted mean concentration, suspended-sediment load is computed as the product of mean concentration and discharge (Morris and Fan, 1998).

Experience in both large and small watersheds has indicate that majority of suspended sediments are transported by high-flow and/or storm events (Horowitz, 2010; Ndomba, 2007; Sadeghi, 2010). In smaller watersheds, the percentage may be even higher (Horowitz, 2010). That means sampling efforts need to be concentrated on storm events to better delimit the annual fluxes of suspended sediments and to establish error limits for unsampled events when estimates are made using a variety of techniques such as rating curves. However, the logistics of sampling storm events in small watersheds can be extremely difficult due to time constraints and logistical issues (Horowitz, 2010; Ndomba et al. 2008).

Measuring and estimating suspended sediment yields in rivers has long been subject to confusion and uncertainty (Thomas, 1985; Sadeghi, 2010). Many methods have been developed for collecting data and estimating yields, a fact that suggests lack of a compelling measurement methodology (Thomas, 1985).

Early studies on sediment load determination have focused on accuracy of field sampling techniques, particularly in relation to the accuracy of various sediment samplers. Further studies were concerned with accuracy of laboratory techniques used for sediment concentration determination (Loughran, 1976). More recently, considerable research has been carried out assessing the procedures used for load calculation, particularly in relation to sediment rating curves (Walling and Webb, 1981, 1988; Ferguson, 1986; Ndomba, 2007; Ndomba et al. 2008). The most common way of combining intermittent concentration data with continuous discharge data uses a rating curve to predict unmeasured concentrations/loads from the discharge (Ferguson, 1986; Walling, 1977).

The use of sediment rating curves has been practiced for many years (Walling, 1977). A suspended sediment rating curve or transport curve is usually presented in one of two basic forms; either as a suspended sediment concentration/streamflow (Equation 1), or a suspended sediment discharge/streamflow (Equation 2) relationship (Walling,1977).

$$C_i = \alpha Q^{\beta} \qquad [1]$$

$$Q_s = aQ^{b} \qquad [2]$$

Where, C_i is sediment concentration in mg/l; Q_s is sediment load in t/day; Q is streamflow discharge in m^3/s; and α and a, and β and b are coefficients and exponents of the rating relationships. In both cases a logarithmic plot is commonly used, with least squares regression employed to fit a straight line through the scatter of points (Walling, 1977). In most cases, rating curves are constructed from instantaneous observations of discharge and either sediment concentration or load, but several specific variants have been proposed (Walling, 1977). Colby (1956) has classified rating relationships, according to temporal resolution of the data, into instantaneous, daily, monthly, annual and flood period curves and, according to particle size criteria, into clay-silt ratings and sand-sized ratings. Other workers have subdivided instantaneous data according to stage and season, constructing separate rating relationships for rising and falling stages (Loughran, 1976).

The relation between discharge and suspended sediment concentration typically is site specific and rarely can be applied to another location (Vongvixay et al., 2010). Even at a single location, the relation can vary depending on the season of the year (Walling, 1977). The procedure used to combine the rating relationship and the associated streamflow data could give rise to underestimation of loads by as much as 50 percent and the inherent inaccuracy of using a rating curve to predict sediment concentration or loads could give rise to errors of as much as +50 percent (Walling,1977). Critical evaluation of sediment load data is necessary if this is to be used in further analysis and it is suggested that rating curve procedures should only be used if the potential errors are recognized and judged acceptable for the purposes involved (Walling,1977).

As Ferguson (1986) suggests in his research that most estimates of river load by rating curve method would have been too low. Some researchers have proposed a statistic bias correction factor to remove the degree of underestimation by the rating curve method (Ferguson, 1986). Walling and Webb (1988) research findings have indicated that statistical bias-correction procedures do not provide accurate estimates in their study rivers and that other sources of error associated with rating curves contribute more in producing inaccurate estimates. However, other researchers used long term sampling programme field data to derive a correction factor (Thodsen et al., 2004). It should be noted that some researchers have indicated that excellent rating curve can even be based on single flood events (Summer et al. 1992; Sadeghi, 2010). The relation between suspended sediment concentration and water discharge during floods is very variable (Sadeghi, 2010). Some clarifications on how to develop a rating curve from one hydrological year sediment sampling programme have been recently discussed in Ndomba et al. (2008).

However, in previous sedimentation studies in the Pangani River Basin, it has been assumed that data set available for constructing rating curves or estimating sediment yield/load represents a near-optimum data collection scheme and that estimates of the errors involved are minimum estimates (Ndomba, 2007; Ndomba et al. 2008). Other researchers elsewhere have done little as well, in quantifying errors due to varying number of data points (Ferguson, 1986; Walling, 1977). Besides, the performance of the sediment sampling programmes have been compromised due to lack of an informed hydro-climatic variables information triggering much upland soil loss and delivery of sediment loads to outlet of the catchments.

Therefore, this study aims at quantitatively estimating the number of data points required to develop an "efficient" rating curve with corresponding hydro-climatic conditions. The findings of this study will be used as inputs into optimizing sediment sampling programme in the Pangani River Basin.

2. MATERIALS AND METHODS

2.1 Description of the Study Area

The study catchment which is located in the upstream of Pangani River Basin (PRB) covers an area of 7280 km^2 (Fig. 1). The main outlet of the catchment is situated at 1DD1 flow gauging station. The catchment covers mainly the Arusha region and Moshi district in Kilimanjaro region. Population densities of more than 600 persons/km^2 are found on the slopes of Mt. Kilimanjaro. More than 50% of the basin, mainly the lowland plains are arid or semi-arid with an annual precipitation of 500−600 mm/year. High levels precipitation can be found in the southern slopes of the mountain areas with an annual precipitation of between 1000 − 2000 mm/year.

The rainfall pattern is bimodal with two distinct rainy seasons, long rains from March to June and Short rains from November to December. Rivers and streams draining the 1DD1 run generally in the North-South and South East directions. This includes the flow from Mt. Kilimanjaro, West of Moshi, flows from the Kikuletwa, Kware Springs and streams from the Southern slopes of Mount Meru. 1DD1 subcatchment contains spring discharges, which include Chemka spring having a yield of 10 m^3/s located 10 km East of Kilimanjaro International Airport (KIA) and is part of Rundugai springs, Shiri spring (0.2m^3/s) and Nsere (0.16 m^3/s). The study area forms a headwater of the main PRB. The mountain slopes of Mt. Kilimanjaro can be divided into five ecological zones, the lower slopes (900−1800 masl), the forest (1800−2800 masl), the heath and moorland, the highland desert (4000−5000 masl) and the summit (above 5000 masl).

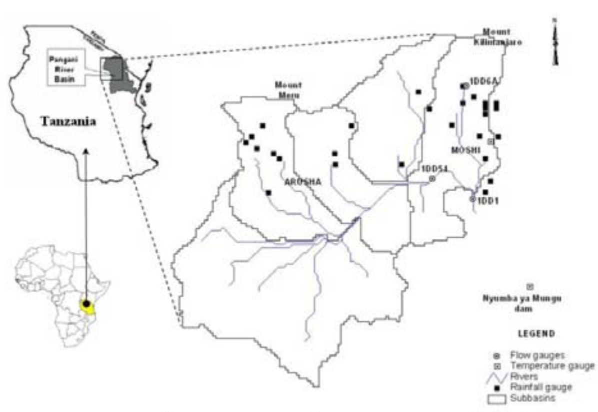

Figure 1:Location map of the study area, upstream of Pangani River Basin

The surrounding area on the plains below can be classified as tropical savannah. The human settlements and agricultural activities are located on the plains and on the lower slopes between 900 and 1800 masl. The area above this level forms the Kilimanjaro Forest Reserve and Kilimanjaro National Park. In this area, there are, in principle, no settlement, or human activities related to agriculture and land use. Wild fires, illegal timber collecting, land pressure and over usage of natural resources are among Mt. Kilimanjaro's biggest problems. Based on the Soil Atlas of Tanzania (Hathout, 1983) and analysis by Ndomba (2007) the main soil type in the upper PRB is clay with good drainage.

The sediment-sampling station at 1DD1 is located upstream of a hydro-electric Nyumba Ya Mungu dam (NYM). 1DD1 sub-catchment is one of the main runoff-sediment contributors to NYM dam (Ndomba, 2007). Generally, the selected site is ideal based on the fact that the hydrological information can be obtained or extracted from the neighbouring gauging stations, as pre-requisite information in sediment sampling program planning (Ndomba, 2007), (Fig. 2)

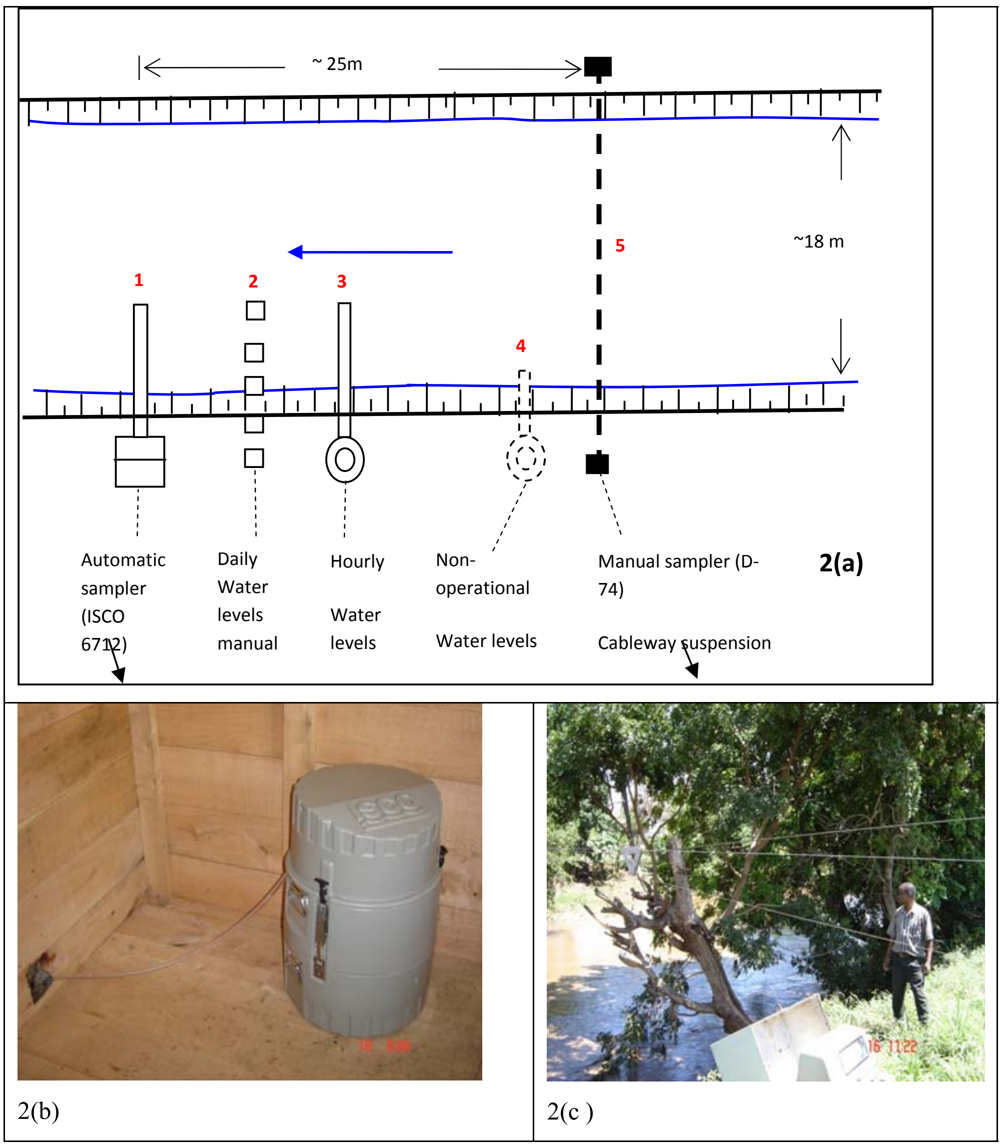

Fig. 2(a-c) Suspended sediment sampling equipment and layout design at the outlet of the test catchment 1DD1-Kikuletwa station.

At 1DD1 sampling station (Fig. 2), the gauge height at zero flow and historical lowest water gauge height are 0.51m and 0.88m, respectively. An Automatic sediment-pumping sampler (ISCO 6712) collects sub-daily suspended sediment samples at the outlet of the larger tributary called 1DD1-Kikuletwa. A sampling tube head at the latter station is located 0.08m below the lowest observed water stage and 0.29 m above the riverbed. A daily observer's samples are also taken at the same station using a Depth integrating sampler, D-74, at midway location number 5. The daily samples are taken at a single vertical located midway of the sampling cross-section. Intermittent cross section samples are taken along the channel width using D-74 trailed over cableway suspension facilities (Fig. 2c). The site is about 25 m long and with an average water surface top width of 18 m (Figure 2a). The set of equipment used in this study is known to sample sediment concentrations iso-kinetically, by many other workers (Gurnel et al., 1992; and Yuzyk et al., 1992). Besides, it is recommended that the ISCO sampling tube head be installed in a turbulent river reach (Bogen, 1992).

The major erosion processes in the study area is sheet erosion from agricultural fields in the headwater regions as sediment sources. These are zones of maximum biological activity - the topsoil (i.e. Ahorizon) or plow layer in slopes of Mounts Kilimanjaro and Meru slopes (Ndomba et al. 2007)

2.2 Data Types and Collection Approach

The primary data on sediment flow was collected. These are continuous subdaily sediment concentrations data. A sampling frequency of up to 2-12 samples per day was achieved during the higher flows. Digital recording data logger measures hourly flow water levels (Fig. 2a). It was envisaged that the sampling tube head would sample suspended sediments most of the time. The daily observer's samples are taken at a single vertical located midway (about 8m out into the flow) of the sampling cross-section.

2.3 Quality Control and Assurance

The single point and vertical measurements as sampled by ISCO and D-74 samplers were correlated to check if they register consistent results (Table 1). This approach was adopted in this study so as to calibrate the concentrations of ISCO machine. As shown in Fig. 2 above the two sampling locations for the equipments are close enough, about 25 m apart. The banks of the sampling reach are so high that no lateral inflow of sediment from intervening catchment is expected.

Table 1: t-test Paired Two samples for means of suspended sediment concentrations determined from samples taken concurrently by ISCO and D-74

Equipment used	ISCO 6712	D-74
Mean suspended sediment concentration [mg/l]	121.21	151.96
Variance	941677	212642
Observations	60	60
Pearson Correlation, r	0.97	
Hypothesized Mean Difference	0	
Degree of freedom, Df	59	
t stat	-1.308	
t critical two tail at 5% level of significance	2.001	

Based on the t statistic test in Table 1 above, where | t stat | less than t critical two tail, the two concentration values from ISCO 6712 and D-74 samples are NOT significantly different at a 5% level of significance. The paired samples are concurrent samples taken at the same hour mostly before noon. The samples are temporally distributed all-year round so they reflect different flow conditions. However, laboratory sample analysis suggests that the observer samples overestimate the ISCO machine samples during extremely higher floods. Probably, the observer fails to estimate both the sampling transit rate and depth of sampling such that the manual sampler scoops the bed. Other researchers such as Gurnell et al. (1992) have reported that although differences in the estimates from ISCO in comparison with DH-48 sampler may exists, yet ISCO sampler provides unbiased estimates of suspended sediment concentration in comparison with the DH-48. Besides, intermittent cross sectional sediment samples are taken. The sampling time in a day is estimated such that corresponding samples from ISCO and/or observer are drawn somewhere between the period. Two intermittent sample analyses results for wet and dry seasons are presented in details in Tables 2 and 3 below.

Table 2: Suspended sediment concentration distribution across 1DD1 cross section on 18/4/2006 between 15:30 and 16:30 hours during rainy season at 1DD1 station.

Sample number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Location from left bank [m]	1.2	3.6	6.0	8.4	10.8	13.2	15.6	18.0	20.4	22.8
Sample volume [ml]	86	102	70	100	100	100	100	218	100	15
Weight of dry sample [mg]	23.3	25.4	14.8	60.9	22.6	22.8	20.8	48.2	17.7	2.1
Suspended sediment concentrations at verticals [mg/l]	271	249	211	609	226	228	208	221	177	140
Mean suspended sediment concentration of the composite sample [mg/l]	260.0mg/l									

The cross section mean sediment concentration derived from a composite sample on the same date and time is 260 mg/l. The corresponding ISCO sampler sample pumped at 16:00 hours is analyzed to have suspended sediment concentration of 264mg/l. The computed cross section coefficient is 260/264=0.986. In Table 3, the cross section samples results during a low flow (dry season) are presented.

Table 3: Suspended sediment concentration distribution across 1DD1 cross section on 24/11/2006 between 15:30 and 16:22 hours during low flow condition.

Sample number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Location from left bank [m]	2	4	6	8	10	12	14	16	18
Sample volume [ml]	150	250	250	150	100	150	100	30	30
Weight of dry sample [mg]	7.0	4.2	6.0	3.1	2.2	5.2	2.5	1.7	1.1
Suspended sediment concentrations at verticals [mg/l]	46.7	16.8	24.0	20.7	22.0	34.7	25.0	56.7	36.7
Mean suspended sediment concentration of a composite sample [mg/l]	27.3 mg/l								

The cross section mean sediment concentration derived from a composite sample on the same date and time is 27.3mg/l. A concurrent ISCO sampler grab sample pumped at 16:00 hours was analyzed to have suspended sediment concentration of 28.0mg/l. The computed cross section coefficient is 27.3/28.0=0.975. With cross section coefficients values as computed above usually no correction of point or vertical sample concentration is applied. A value less that ± 5% from unit signifies that no need of correction (Guy and Norman,1970; MoW,1979). Although 20 sampling verticals are considered sufficient to give mean concentration (Guy and Norman,1970; MoW,1979), technical and practical issues have hindered its applicability in this study.

For instance, it took us about an hour to take sample from 10 verticals. Temporal variability within a day of transported sediment is high with coefficient of variation, Cv, of up to 166% during high discharges. With highly variable sediment transport characteristics within a day as observed for 1DD1 gauging station, it is obvious that more verticals samples imply more sampling time is required and hence the derived mean concentration would not be comparable to single observer's sample especially in a wet season. On the other hand the temporal variability during the low flows is insignificant with Cv down to 2%. But during this time of the year flowing water is clear from sediments, therefore more verticals samples are not justified from technical point of view. However, the accuracy of estimating mean concentration does not merely depend on number of sampling verticals but rather depends on the lateral variation of sediment concentration (MoW,1979). From analyses above one would conclude that a satisfactory uniform distribution of sediment across a cross section exists.

Besides, the degree of (random) spatial variability across the river cross section was illustrated by expressing the standard deviation as a proportion of the mean (coefficient of variation, C_v) for the 1DD1 sampling station cross-section for two sampling occasions (Table 4).

Table 4: The degree of (random) spatial variability of sediment concentrations across the river cross sections at 1DD1 sampling station

Sampling station	Cross sectional spatial variability of sediment concentrations, C_v, as per wet and dry seasons		
	Wet season	Dry season	Mean
1DD1	0.499	0.489	0.494

The values of this ratio as presented in Table 4 above are 0.499 and 0.489 for wet and dry seasons, respectively, with a mean of 0.494 as derived from two sampling occasions at 1DD1 station. These results imply that the sediment is well mixed across the flow and that although suspended sediment concentration may vary from point to point, sometimes with higher concentrations near midway (Table 2, sample no.4) and sometimes near banks (Table 3, sample no.8), is essentially random and therefore, sampling at any point one location should give an unbiased estimate of section suspended sediment concentration. Other researchers working elsewhere such as Gurnell, et al. (1992) and Yuzyk, et al. (1992) reported a similar observation.

A quantitative assessment as discussed above suggests that the two measurements methods give comparable results and the concentrations from these samples need not be corrected. Therefore, mean concentration of sediments for the cross section was assumed equal to fixed-point sample concentrations from ISCO sampler. As discussed above, the majority of sediments transported in this river are fine-grained sediments particles. Other researchers have shown that sediment concentration is nearly uniform across the width and along the vertical when majority of transported sediments are fine-grained sediments (Morris and Fan, 1998; Yuzyk, et al. 1992).

As Table 5 below stipulates, several statistics were analyzed to assess the data quality. They include standard deviation, coefficient of variation and standard error of the mean.

Table 5: A summary of sediment flow data for the Nyumba Ya Mungu Reservoir catchment in the upstream part of the Pangani River catchment as sampled between March 18 and November 10, 2005 by an ISCO 6712 machine at 1DD1 gauging station

Statistic	Subdaily suspended sediment concentration [mg/l]	Gauge Height [m]	Streamflow discharge [m^3/s]
Number of data points	291	291	291
Maximum	9110.0	4.44	256.53
Minimum	16.0	0.89	12.19
Mean	282.5	1.32	34.79
Standard Deviation, STD	801.7	0.49	30.02
Coefficient of Variation, Cv (%)	283.8	36.69	86.27
Standard Error of the Mean, SEM	47.0	0.03	1.76

2.4 Determination of Suspended Sediment Actual load

The sediment transport, G_s, was calculated for hourly intervals from Equation 3.

$$G_s = \int_{t1}^{t2} Q(t)C(t)dt \qquad [3]$$

Water discharge, Q(t), was calculated from the continuous record of water stage. Hourly sediment concentrations, C(t), were obtained by linear interpolation between the known concentrations of collected samples. And t signifies time intervals in hours. Different interpolation algorithms were applied to some sampled flood events and linear interpolation was found the best. Bogen (1992) also computed sediment transport from a sampling programme involving two to four samples a day with a linear interpolation between samples. Missing concentrations data during ISCO sampler down time (i.e. maintenance) were filled with daily observers' samples. A maximum of two days machine downtime has been experienced. Literally, no data-missing situation has happened in 1DD1 sampling site because paired concurrent samples at 9:00 hours both from sub-daily ISCO machine samples and daily samples by observer were deliberately considered in sampling design.

As Rieger et al., (1988) suggest where soil types in the basin are dominated by fines as for PRB, the use of suspended sediment concentration can be rationalized and time based values of concentrations can be obtained using automatic water sampler or turbidity monitors. Many other researchers adopted the same approach (Walling, 1977; Walling and Webber, 1981).

2.5 Generating Random Sample Size

A random sample in this context is defined as a sample in which any one individual measurement in the population is as likely to be included as any other (Alder and Roessler, 1972). A variety of sampling scheme exist, including convenience sampling, purposive sampling and random sampling. Sampling is random, if it involves a randomization device, such as a table of random numbers or the use of random number generator from a computer. The most basic type of random sampling is simple random sampling. In this study, the Sampling analysis tool in Microsoft Excel 2007 spreadsheet was explored to generate/create a random sediment sample from a population by treating the input range as a population.

This study uses a population of 291 data points of sediment concentrations. It is considered that a random sample will help in making important decisions from a smaller selection of data. The following steps were executed to create the samples of varying sizes (Tables 6 and 7 below). Given data as presented in Table 6 below, "=RAND()" (no quotes) formula was typed in the "random" column to generate random numbers. All of the data in the spreadsheet along with the corresponding random numbers were selected. For this purpose titles/headings were not selected. The "Sort" button under "Data" pull-down menu was clicked. The "Random Column" was chosen from the "Sort by Column" drop-down list and "Smallest to Largest" from the "Sort by Order" drop-down list. The first nth top numbers of rows were chosen to make up a random sample (Tables 6 and 7). These steps are repeated any time a sample of size n is generated.

Table 6: The first 17th sediment flow data records with unranked random numbers - illustration

Serial Number	Date (DD/MM/YY)	Time (HHMM)	Suspended Sediment concentration, SS (mg/L)	Stage, H (m)	Flow, Q (m^3/s)	Sediment load, Q_s (t/day)	Random Numbers
1	18/03/2005	09:00	18.6	0.89	12.192	19.593	0.180130794
2	18/03/2005	21:00	29.3	0.90	12.614	31.933	0.827238984
3	19/03/2005	09:00	32.0	0.90	12.614	34.875	0.785888048
4	19/03/2005	21:00	30.1	0.94	14.116	36.711	0.767730099
5	20/03/2005	09:00	368.5	0.94	14.334	456.372	0.619354896
6	20/03/2005	21:00	615.0	0.95	14.937	793.692	0.671990068
7	21/03/2005	09:00	48.4	0.95	14.772	61.773	0.786546996
8	21/03/2005	21:00	120.7	0.92	13.575	141.567	0.739006756
9	22/03/2005	09:00	65.6	0.92	13.468	76.334	0.370318365
10	22/03/2005	21:00	99.5	1.10	21.856	187.892	0.664520505
11	23/03/2005	09:00	4632.0	1.13	23.126	9255.136	0.369607792
12	23/03/2005	15:49	5080.0	1.24	28.654	12576.584	0.425025597
13	23/03/2005	21:49	1603.0	0.96	15.157	2099.232	0.98023105
14	24/03/2005	03:49	1054.0	0.94	14.347	1306.518	0.473440996
15	24/03/2005	09:49	613.0	0.94	14.334	759.175	0.182822995
16	24/03/2005	15:49	410.0	0.94	14.334	507.768	0.343399808
17	24/03/2005	21:49	266.0	0.94	14.334	329.430	0.522532793

Table 7: The first 17th sediment flow data records with ranked random number from smallest to largest

Serial Number	Date (DDMMYY)	Time (HHMM)	Suspended Sediment concentration, SS (mg/L)	Stage, H (m)	Flow, Q (m^3/s)	Sediment load, Q_s (t/day)	Random Numbers
74	16/04/2005	16:06	2690.0	2.13	80.899	18802.222	0.00139806
33	31/03/2005	12:00	46.4	0.95	14.772	59.220	0.00202847
254	18/07/2005	09:00	34.0	1.08	20.673	60.729	0.002858462
213	15/06/2005	09:00	30.0	1.22	27.451	71.153	0.006387988
184	01/06/2005	18:00	109.6	1.79	59.584	564.227	0.01740348
135	12/05/2005	10:00	71.2	1.33	33.098	203.608	0.026143873
59	08/04/2005	21:00	115.0	1.20	26.592	264.218	0.029977548
222	21/06/2005	09:00	27.0	1.19	26.090	60.863	0.03099512
261	16/08/2005	09:00	45.0	0.99	16.552	64.354	0.033374949
126	07/05/2005	22:00	32.4	1.12	22.592	63.243	0.037752779
129	09/05/2005	10:00	27.2	1.12	22.543	52.978	0.046477555
238	29/06/2005	09:00	34.0	1.17	25.192	74.004	0.048909198
265	20/08/2005	09:00	46.0	0.97	15.790	62.756	0.050456405
286	05/11/2005	09:00	112.0	0.97	15.523	150.213	0.064695746

90	19/04/2005	08:06	149.0	1.52	43.712	562.731	0.065302589
204	10/06/2005	21:00	52.5	1.34	33.894	153.743	0.067789588
82	18/04/2005	00:06	225.0	1.45	39.804	773.790	0.085505827

2.6 Development of Sediment Rating Curve Using Ordinary Least-Square Approach

This study uses instantaneous sub-daily sediment samples from ISCO sampler to derive a rating curve. The rating relationships have been determined using Ordinary Least-Square (OLS) regression of the logarithmic transformed data. The coefficients and exponents of equations 1 and 2 are estimated. This follows the approach adopted by many previous workers and can be justified on statistical grounds in terms of data normality, linearity of the relationship and considerations of homoscedasticity (*i.e.*, variance is fixed throughout a distribution with a scatter plot of data looks oval), (Walling,1977). The estimated load based on sediment rating curve, Estimated, and actual loads, Actual, as computed above are used to compute Relative Error in percent, RE (Equation 4).

$$RE(\%) = \left|\left(\frac{Actual - Estimated}{Actual}\right)\right| 100 \qquad [4]$$

Relative error was used as performance index and in this context good performance is confirmed when RE in percent is less than 20%. This threshold is considered satisfactory for most of engineering practices.

3. RESULTS AND DISCUSSIONS

This section discusses a number of key issues, mainly, trying to answer the following: what is the minimum (cost effective) number of data points required for optimal sampling programme in the upstream of Nyumba Ya Mungu dam, when should these be sampled, How do the results compare with other similar studies, and limitations of the study. Figs. 3 through 5 illustrate/present the main results of this study.

One would note from Fig. 3 below that Relative Error, RE in percent converges to 45% with increasing sample size. There is a significant scatter with decreasing sample size from sample size of 120. All data points were used in the past studies in the study area to develop a field-data based corrected rating curve. A rating curve correction factor of 2.78 was used in the previous study to estimate long term catchment sediment load (Ndomba et al. 2008). It is evident that the RE for sediment sample size of less than 60 data points (i.e., sample size 17, 26, 27, 34, 44, 55 and 57) have good perform with RE less than 20%.as inscribed by a dotted red box (Fig. 4). They represent hydrological conditions of months of March, April, May, June, July, August, October and November with monthly total sampling frequency in percent of 11, 31, 19, 21, 4, 6, 3 and 5, respectively (Fig. 5). As this study was interested in small sample size less than 30, cost effective size, one would note that sample size of 17, 26, 27 data points qualify as cost effective sample size (Fig. 4). As these samples were randomly generated, the astonishing results is their persistent prevalence

Such results compare with that of Walling and Webb (1981, 1988) whereby they asserted that sediment loads are always underestimated/overestimated by ±50%. Ferguson (1986) attributed it to inbuilt assumption of Ordinary Linear Square (OLS) approach of constructing a sediment rating curve. As evidenced from Fig. 3 the greater the sample size the smaller the uncertainty/spread, this suggests that collection of small sample data size is subject to high uncertainty. So caution should be exercised and particularly in representing/capturing various hydrological regime information/data.

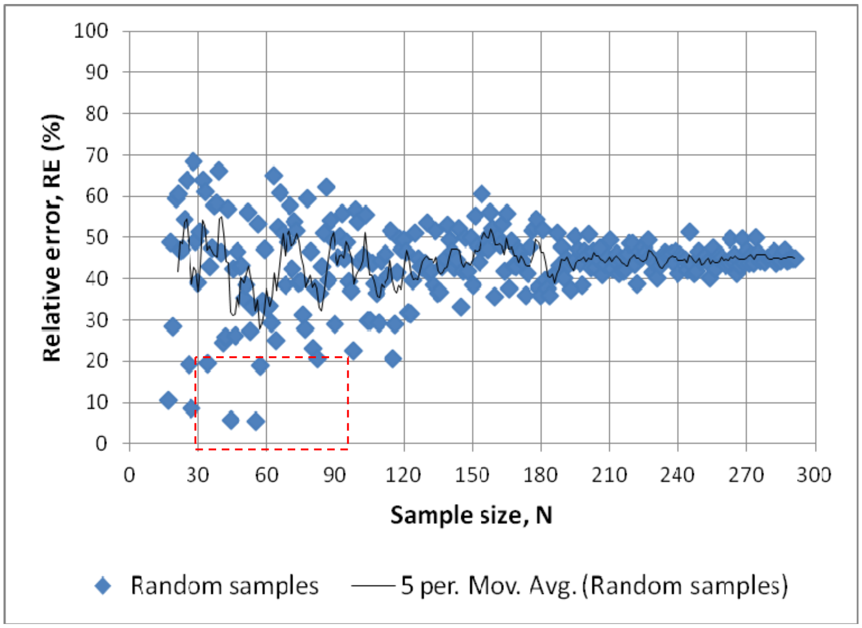

Figure: 3 Plot of Relative error versus number of data points/sample size, for all the samples Note: Inscribed in a box are samples with RE less than 20%

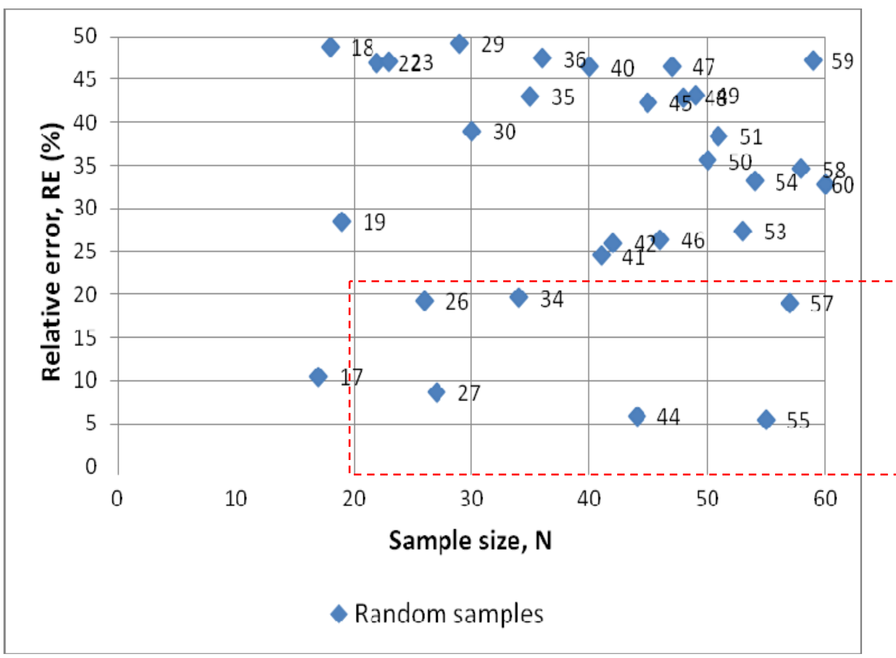

Figure 4: Plot of Relative error versus sample size with good performance inscribed in a dotted box

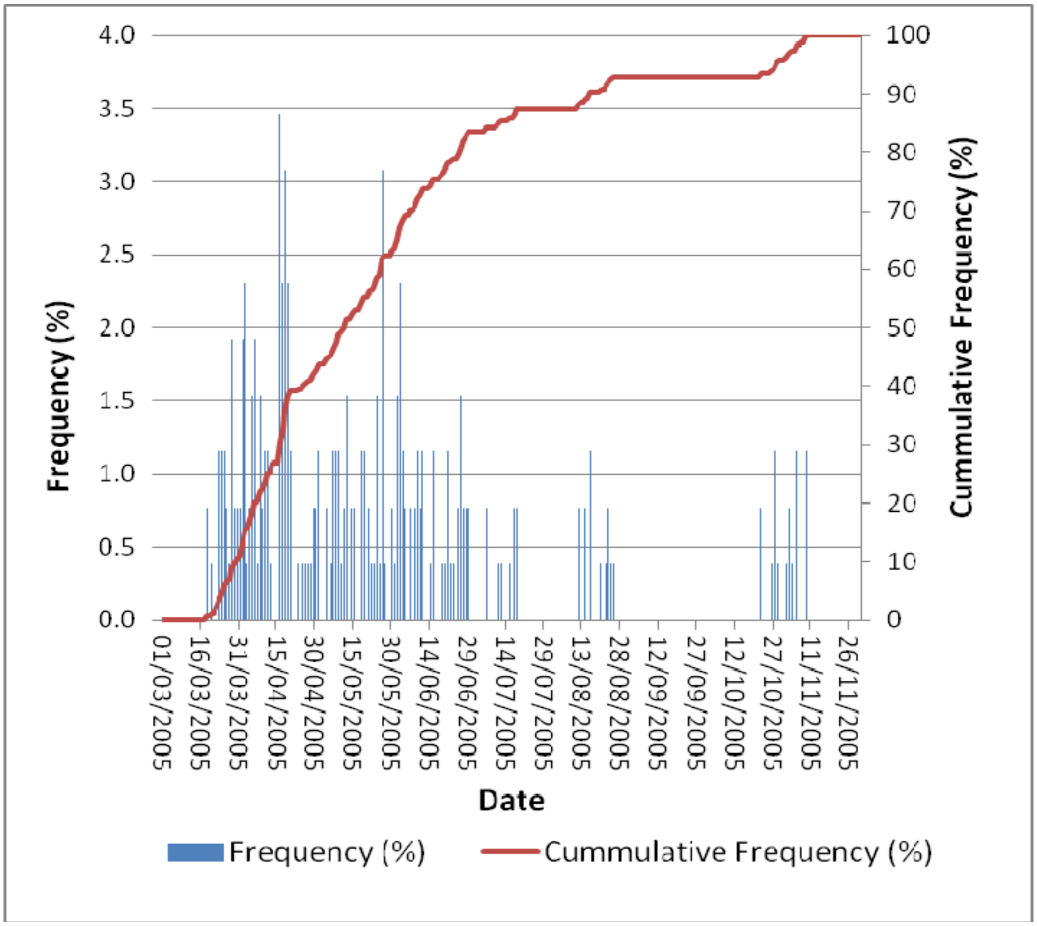

Figure 5: Proposed programme of sampling frequency as derived from frequencies of sample size with RE less than 20%

Therefore, sampling programme should be planned for the period between March and November with little attention to months of September and October according to suggested frequencies (Figure 5). Therefore, sampling programme should focus on months of March, April, May, and June (Fig. 5).

It should be noted however that these findings are based on one year hydrological sediment sampling programme. So these findings lack validation independent sediment flow data set. This study used a sedigraph developed from ISCO-6712 pumping sampler-subdaily sediment flow data to estimate actual load. The temporal variability of these is normally high as logistically it is impractical to sample the entire runoff event (Ndomba, 2007). In order to minimize the temporal variability, higher frequent samples from turbidity sensors would be required in future research to minimize the uncertainty of actual load estimation as well as the relative error computations. Besides, this study has not analyzed the reliability of proposed sample sizes for optimal sampling programme. Notwithstanding, the study has unfolded the unknown especially on the optimal sampling programme in terms of cost effective sample size for estimating actual sediment load through rating curve development. This study findings suggest that even without correcting rating curve, as proposed by others (Ferguson, 1986; Walling and Webb, 1981, 1988) relative error of much less than 20% could be achieved through a well-scheduled and timely sampling programme as suggested/proposed. Further analysis indicates that the proposed sampling programme will be able to capture sediment loads from temporal varying sediment sources in the upstream of Pangani River Basin as reported in Ndomba et al. (2007).

4. CONCLUSION AND RECOMMENDATIONS

The preliminary findings show that a cost effective sediment concentration sample of less than 30 data points; i.e., 17, 26, and 27; covering critical hydrological condition is adequate. An independent analysis indicates that these samples are from selected runoff events of months of April, May and June. Events of month of April dominated mostly the sediment delivery/transport to the catchment outlet. Notwithstanding, the optimal sample size should represent all hydrological seasons in the catchment with varying degree/frequency.

The results of this study need to be validated using recent sediment flow data in the same catchment. The sampling programme proposed is site specific, thus may not be extended easily to other catchments. Further studies should analyze the reliability of the proposed sample size for optimal sampling programme in the Pangani River Basin as little was done in this work.

5. ACKNOWLEDGEMENTS

The author would like to extend special gratitude to The Norwegian Programme for Development, Research and Education (NUFU), the Postgraduate Programme for Water Management Project at University of Dar es Salaam, Pangani River Basin Water Office in Moshi District, UNESCO-IHE Partnership Research Fund (UPaRF) ACCION(Adaptation to climate change impacts on the Nile River Basin research project), UPaRF MARAFLOWS (Environmental flows for people and ecosystems in the Mara River Basin); Project number 60043191) for providing financial and/or logistical support to carry out this study. Besides, the author wishes to express his gratitude to Dr. Kahimba, F. who helped to improve this paper through his thorough review.

6. AUTHOR BIOGRAPHY

Dr. Preksedis Marco Ndomba, is a Senior Lecturer in the Department of Water Resources Engineering, College of Engineering and Technology at University of Dar es Salaam. He is also a coordinator of regional research projects in the Nile Basin Capacity Building Network (NBCBN) and Nile FRIEND initiatives.

7. REFERENCES

1. Alder, H. L., & Roessler, E. B., 1972, *“Introduction to Probabilities and Statistics”*, Fifth Edition, W.H. Freeman and Company.
2. Bogen, J., 1992,*“Monitoring grain size of suspended sediments in rivers”,* In: Erosion and Sediment 3.
3. Transport Monitoring Programmes in River Basin (Proc. Oslo Symposium, August 1992), 183190. IAHS Publ. 210. IAHS Press Wallingford, UK.
4. Colby, B.R., 1956, “*Relationship of sediment discharge to streamflow*”, *US Geol. Survey Open File Report*.
5. Ferguson, R.I., 1986, “*River Loads Underestimated by Rating Curves*”, *Water Resour. Res.* 22(1), 74−76.
6. Gurnell, A.M., Clark, M.J., Hill, C.T., & Greenhalgh, J., 1992, “*Reliability and Representativeness of a suspended concentration monitoring programme for a remote alpine proglacial river*”, In: Erosion and Sediment Transport Monitoring Programmes in River Basins, (Proc. Oslo, Norway, August 1992), 191−200 IAHS Publ. 210, IAHS Press Wallingford, UK.
7. Guy, H.P., & Norman, V.W., 1970, “*Field methods for measurements of fluvial sediment*”, Techniques of Water Resource Investigation of the US Geological Survey, C2 (3).
8. Hathout, S.A., 1983, “*Soil Atlas of Tanzania”,* Tanzania Publishing House, Dar es Salaam.

9. Horowitz, A.J., 2010, "*The Use of instrumentally collected-composite samples to estimate the annual fluxes of suspended sediment and sediment-associated chemical constituents*", Sediment Dynamics for a Changing Future, *Proceedings of the ICCE symposium held at Warsaw University of Life Sciences – SGGW, Poland, 14-18 June 2010, IAHS Publ. 337. Pp 273-281.*
10. Loughran, R.J., 1976, "*The calculation of suspended-sediment transport from concentration vs discharge curves: Chandler River*", NSW. *Catena* 3, 45-61.
11. Morris, G.L., and Fan, J.,(eds), 2008, "*Reservoir Sedimentation handbook, Design and management of dams, reservoirs, and watersheds for sustainable use",* Copyright © 1998 by The McGraw-Hill Companies.
12. MoW, Ministry of Water Energy and Minerals, 1979, "*Sediment transport in streams sampling, analysis and computation",* In: Manual on Procedures in operational hydrology. Vol.5.
13. Ndomba, P.M. 2007, "*Modelling of Erosion Processes and Reservoir Sedimentation Upstream of Nyumba Ya Mungu Reservoir in the Pangani River Basin*", A PhD Thesis (Water Resources Engineering) of University of Dar es Salaam.
14. Ndomba, P.M., Mtalo, F. & Killingtveit, A., 2007, "*A Proposed Approach Of Sediment Sources And Erosion Processes Identification At Large Catchments*", Journal of Urban and Environmental Engineering, V.1, n.2 (2007) 79–86. ISSN 1982-3932.
15. Ndomba, P.M., Mtalo, F. & Killingtveit, A., 2008, "*Developing an Excellent Sediment Rating Curve From One Hydrological Year Sampling Programme Data: Approach*", Journal of Urban and Environmental Engineering, V.2,n.1,p.21-27. ISSN 1982-3932, doi:10.4090/juee.2008.v2n1.021027.
16. Rieger, W.A., Olive, L.J. and Gippel, C.J., 1988, "*Channel sediment behavior as a basis for modelling delivery processes. Sediment budgets",* Proceedings of the Porto Alegre Symposium, December, 1988. *IAHS Publ.no. 174, pp541-548.*
17. Sadeghi, S.H., 2010, "*How Important is sediment graph development in Iran? Sediment Dynamics for a Changing Future*", Proceedings of the ICCE symposium held at Warsaw University of Life Sciences – SGGW, Poland, 14-18 June 2010, IAHS Publ. 337. Pp361-365.
18. Summer, W., Klaghofer, E., Abizeid, I., & Villeneuve, J.P., 1992, "*Critical reflections on long term sediment monitoring programmes demonstrated on the Austrian Danube*", In*:* Erosion and Sediment Transport Monitoring Programmes in River Basins *(*Proceedings of the Oslo Symposium, August 1992), 255–262. IAHS Publ. 210. IAHS Press Wallingford, UK.
19. Thodsen, H., Hasholt, B., & Pejrup, M., 2004, "*Transport of phosphorous, wash load and suspended sediment in the River Varde Å in Southwest Jutland, Denmark*", In: Sediment Transfer through the Fluvial System (Proc. Moscow Symposium, August 2004), 466473. IAHS Pub.288, IAHS Press Wallingford, UK.
20. Thomas, R.B., 1985, "*Estimating Total Suspended Sediment Yield With Probability Sampling*", Water Resources Research, Vol. 21, No. 9, Pages 1381–1388, September 1985.
21. Vongvixay, A., Grimaldi, C., Gascuel-Odoux, C., Laguionie, P., Faucheux, M., Gilliet, N. & Mayet, M., 2010, "*Analysis of Suspended Sediment Concentration and Discharge Relations to Identify Particle Origins in Small Agricultural Watersheds*", Sediment Dynamics for a Changing Future, *Proceedings of the ICCE symposium held at Warsaw University of Life Sciences – SGGW, Poland, 14-18 June 2010, IAHS Publ. 337. Pp 76-83.*
22. Walling, D.E., 1977, "*Limitation of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers*", In: *Erosion and Solid Matter Transport in Inland Waters* (Proc. Paris Symposium, July 1977), 34–48. IAHS Pub. 122. IAHS Press Wallingford, UK.
23. Walling, D.E., & Webb, B.W.,, 1981, "*The reliability of suspended sediment load data",* Symposium on Erosion and Sediment Transport Measurements in Florence. IAHS Publ., vol.133, pp 177-194.
24. Walling, D.E., & Webb, B.W., 1988, "*The reliability of rating curve estimates of suspended sediment yield: some further comments*", Symposium on Sediment Budgets in Porto Alegre. IAHS Publ., Vol.174, pp 337-350.
25. Yuzyk, T.R., Gummer, WM.D., & Churchland, L.M., 1992, "*A comparison of methods used to measure suspended sediment in Canada's federal monitoring programs"*, In: *Erosion and Sediment Transport Monitoring Programmes in River Basins (*Proceedings of the Oslo Symposium, August 1992), 289297. IAHS Publ. 210. IAHS Press Wallingford, UK.