Malaysian Journal of Mathematical Sciences 4(1): 53-75 (2010)

Adaptive Cluster Sampling for a Temporal-Scale Population

[1,2]Arabkhedri, M., [1]Lai, F.S., [3]Noor Akma, I., [1]Mohamad-Roslan, M.K.

[1]Faculty of Forestry, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

[2]Soil Conservation and Watershed Management Research Institute, P.O.Box: 13445-1136, Tehran, Iran

[3]Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

E-mail: mahmood.arabkhedri@gmail.com, arabkhedri@scwmri.ac.ir

ABSTRACT

Adaptive cluster sampling (ACS) is appropriate for rare clustered populations with localization tendencies. Up to now, it has been used exclusively for investigating spatial-scale problems rather than temporal-scale such as this study is dealing with, i.e. sediment transport in rivers. Suspended sediment load is carried mostly during relatively short periods coincide with high flows otherwise negligible. In ACS, more samples from critical river stages can be taken with respect to the aggregation tendencies of sediment loads during transport; thus increasing the level of representativeness of samples. Adoption of ACS to this new area needs further verification and adaptation such as definition of the sampling unit, population frame, neighborhood relation, and threshold. In this study, several scenarios were defined for the purpose of evaluating the ACS in sediment estimation. Numerous sample sets were taken from intensive discharge-load records of Sg. Pangsun River, Malaysia. These sample sets are different with respect to initial sample size, neighborhood relation, and discharge threshold. Total suspended sediment loads were then estimated using modified Horvitz-Thompson method. The comparison made between the symmetric neighborhood relation and the forward method suggested in this study showed that the latter could be used instead of the former in sediment studies without losing the accuracy. The findings also suggested the flow duration curve is a useful tool for ranking initial samples in order to determine an optimum discharge threshold.

Keywords: Adaptive Cluster Sampling, Sediment load estimation, Sg. Pangsun, Accuracy, Flow Duration Curve.

INTRODUCTION

Most water quality monitoring programs are conducted in accordance with continuous recording of flow discharges and discrete fixed-interval (similar to systematic) sediment sampling plans. Based on the collected records for a study period, the total sediment load is estimated using flow duration and sediment rating curves (Vanoni 1980). Several studies have shown that this method most often significantly underestimates the long-term sediment transport rates as much as 50 to 60 % and even more (Walling and Webb 1981, Thomas 1985, Ferguson 1986, Koch and Smillie 1986, Thomas

1988, Walling and Webb 1988, Cohn, Delong, Gilroy, Hirsch and Wells 1989, Asselman 2000, Cohn 2005).

In response to the shortcomings of sediment rating method, the concept of using sampling techniques to improve sediment load estimation was largely proposed by Thomas (1985). Essentially, the survey sampling approach comprises of two parts. The first pertains to the design of selecting sample from the study population. The second deals with a special estimator for estimating population parameters such as total (load).

Three approaches had been successfully employed in previous studies on this topic in the literature, including: Selection At List Time or SALT (Thomas 1985), time-stratified (Thomas and Lewis 1993), and flow-stratified (Thomas and Lewis 1995). With these designs, the inclusion probabilities of the sampling units from high flows are more than low flows. Cohn (1995) acknowledged these approaches as innovations in the river sediment estimation area.

The uses of these sampling designs are only restricted to the sediment gauging stations on rivers with programmable sampling devices. On the other hand, the common feature in the majority of water quality monitoring programs is manual sampling (Degens and Donohue 2002). Thus, another sampling technique is required for such gauging sites.

Adaptive Cluster Sampling

Adaptive cluster sampling (ACS) was introduced by Thompson (1992). This design is appropriate when the population is rare and highly clustered. In ACS, the investigator selects a sample set from the population in two steps. First, an initial sample set is taken; then for important elements (samples with high load for river sediment), adjacent units (neighbors) are chosen in the next step based on a threshold. With this design, a final sample set is produced which constitutes of several networks; some of them have only one unit (do not meet the condition) and some have more. The ACS procedure requires continuous symmetric sampling in the neighborhood until the measured variable drops below the threshold.

The application of ACS on spatial-scale populations such as wild animal, fish, forestry and soil was widely used in recent years. In these studies, higher efficiency of ACS was shown compared to conventional samplings when the study population is rare with high aggregation (Table 1). Arabkhedri *et al.* (2007) who conducted a study by the use of ACS on sediment load estimation in Gorgan-Rood river, Iran showed that it produces a better result compared to the conventional sediment rating curve.

TABLE 1: Some studies using adaptive cluster sampling

Study discipline	Reference
Surveying pacific hake larvae	Lo *et al.* (1997)
Estimating three rare trees in forest with different aggregation	Acharya *et al.* (2000)
Estimating fish populations with different aggregation	Su and Quinn (2003)
Estimating number of workers earning from different single-industries	Chaudhuri *et al.* (2005)
Mapping pollutant in soil	Juang *et al.* (2005)

Sediment concentration records are a temporal-scale population. For such a population, the use of symmetric neighborhood relation is not feasible in the field because the individual who samples cannot take preceding units. Therefore, in the case of temporal-scale population, an altered forward version of ACS is essential. The applicability of Symmetric Neighborhood Relation (SNR) was also reported problematic for some spatial-scale populations (Lo, *et al*. 1997, Salehi and Smith 2005).

Conducting Adaptive Cluster Sampling Design for River Sediment

For conducting an ACS in river suspended sediment as a temporal-scale study, some technical terms should be adapted (Table 2). In addition, ACS needs a condition or criteria, defined as a threshold for the variable of interest – sediment concentration values - as is the case of this study. However, sediment concentration is not known until the end of laboratory analysis. In this situation, discharge can be used as an auxiliary variable, because flow discharge is often used to correlate sediment, as in a sediment rating curve.

Figure 1 shows the effect of using the magnitude of discharge threshold (DT) on the number of additional samples (cluster size) during a storm hydrograph schematically. Each unit on the x-axis comprises a sampling unit or a time interval (refer to Table 2). The flow hydrograph exceeded two DTs -140 and 200 L/s.

TABLE 2: Adaptation of some survey sampling technical terms for river sediment study

No	Item	Definition
1	Population	A time continuous measurement of sediment concentration or load in a study period. Population is divided into N time intervals or sampling units.
2	Sampling unit	Each sampling unit refers to a short interval with relatively constant suspended sediment load. It could vary from several minutes in small flashy basins to even days in very large ones (Thomas 1985).
3	Observation unit	An observation unit is a representative small water sample collected in a gauging site during a sampling unit (time interval).
4	Variable(s) of interest	Water samples are analyzed in laboratory for measuring sediment concentration. The concentration may be used to compute corresponding load. These two values represent interested variables most of the time.
5	Auxiliary variable(s)	Other recorded variables such as water stage, discharge and turbidity during sampling from river are auxiliary variables.
6	Target population	It means a complete collection of sediment concentration or sediment load records from all sampling units during the study period.
7	Initial sample set	All n chosen observation units with a sampling design is called sample in text books. However, to prevent confusion with water sample, sample set substitutes this term in the current study. Fixed-interval sampling is suggested in this study because it is well known course for hydrologists.
8	Final sample set	The initial sample set and additional adaptive samples in the neighborhood constitute the final sample set.
9	Sample size	It refers to number of observations in corresponding initial or final sample set.

In Figure 1 if the initial sample set intersects the flow hydrograph on the 15th day (by intersection a , an initial sampling unit), it can be seen that the lower discharge threshold requires significantly more additional units than the higher threshold with respect to neighborhood relation. The higher threshold (200 L/s) has shorter base-time of a storm hydrograph leading to a smaller cluster size - 17 units compared to 31. If the initial sample set intersects the hydrograph on the 30th day (Intersection b), no additional unit would be chosen for the threshold 2; but for threshold 1, the cluster size is still 31 units.

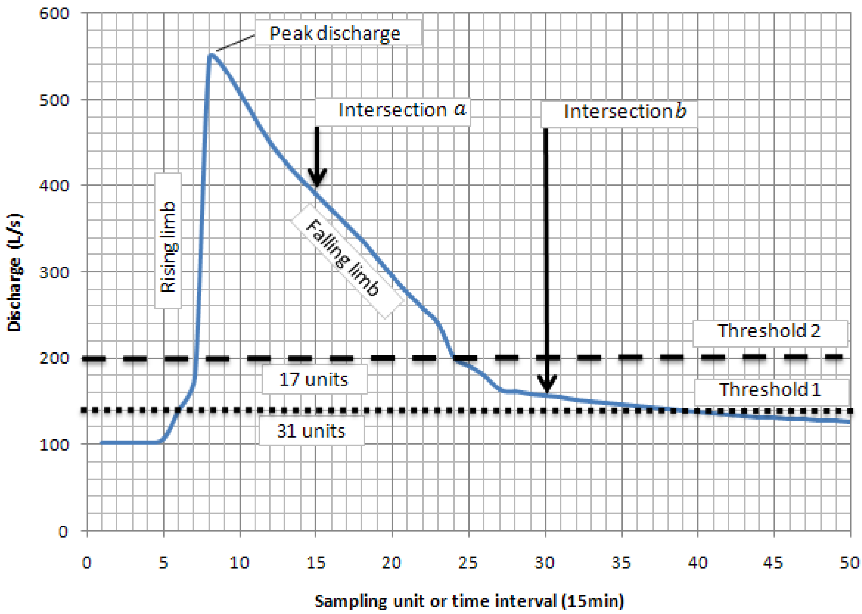

Figure 1: Effect of discharge threshold magnitude on cluster size in a storm hydrograph

Figure 2 is a flowchart showing the process of ACS with forward neighborhood relation (FNR). The person-in-charge for sampling must visit the gauging site, takes sediment sample, and records the instantaneous discharge (ID) in accordance with a predetermined time schedule. Based on discharge observation and a predetermined DT, he then makes a decision, whether he should continue taking the sample in the next sampling unit (i.e. next time period or adjacent neighbor) or not.

In other words, when the observed discharge exceeds the predetermined DT (Figure 2), he will take a sequence of samples in the subsequent time periods until the ID drops below the threshold. Otherwise, he will take the next sample according to the time schedule.

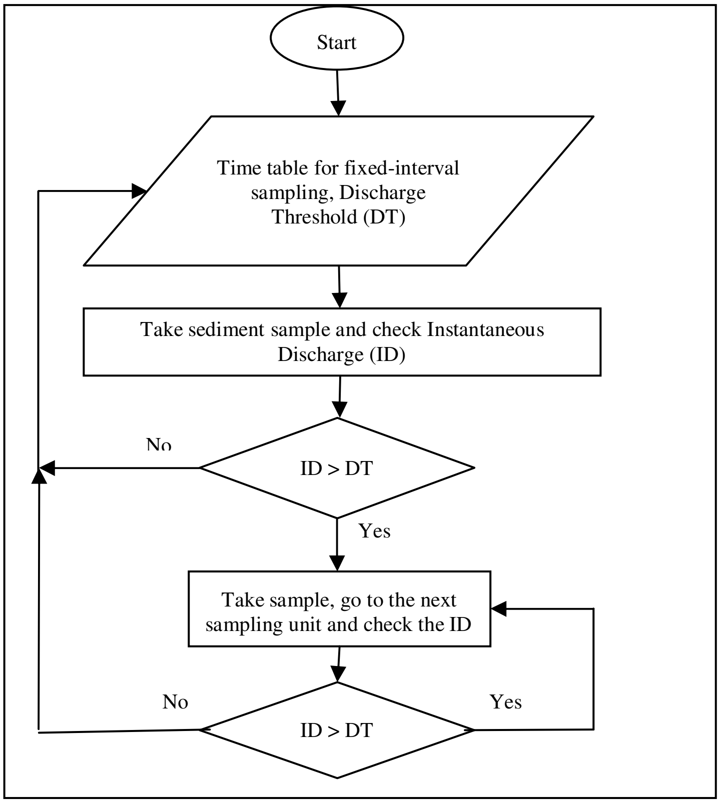

Figure 2: Flow chart of adaptive cluster sampling with forward neighborhood

Although the symmetric ACS is not applicable in river sediment field sampling, it can be simulated based on available continuous data for evaluation purpose. The first three steps of sampling with SNR are similar to the FNR. If the observed instantaneous discharge exceeds the discharge threshold, then additional samples are taken from preceding units as well as follow-up units.

Choosing Discharge Thresholds

Choosing an appropriate threshold by ranking the values of all initial samples had been suggested to manage the number of additional samples in spatial-scale studies (Thompson, 1996). As stressed earlier, this method is not possible in river sediment study, since discharges of forthcoming days, which is time-scale based, are not known. Therefore, this study has made use of the flow duration curve (FDC) to determine the threshold. FDC is a plot showing the percentage of time in historical record that flow in a stream is likely to equal or exceed any given magnitude.

This study was designed to evaluate the performance of ACS in river suspended sediment load sampling and estimating –as a temporal-scale problem- including the following aspects:

- how the ACS with SNR performs in sediment load estimation
- how the ACS by FNR works compared with the SNR
- the effect of DT magnitudes on load estimations
- the effect of sampling frequency or sample size

The study was performed in Sg. Pangsun (Figure 3), a small steep undisturbed upland catchment near Kuala Lumpur, which supplies water to a mini hydroelectric power station. The climate is equatorial characterized by high humidity, high rainfall and uniform annual temperature. The data used (including continuous records of river flow and sediment) had been collected in 1997 by Geoffery (1999). He had measured flow and sediment concentration using water level recorders and Total Suspended Solid Analyzer respectively for 8 months within every 15 minutes time interval.

SAMPLING SIMULATION, LOAD ESTIMATION AND COMPARISON

To extract required initial sample sets, calendar-based sampling from the entire population was simulated first under 10 different time schedules each in replicates of 50. In order to understand the performance of ACS, a wide range of sample sizes from 50 to around ¼ of population size were considered.

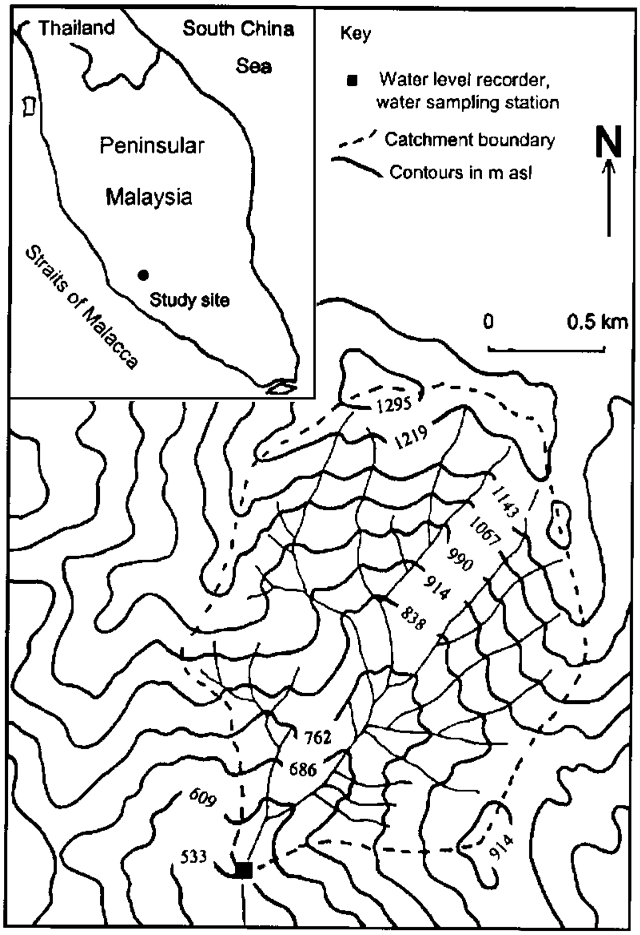

Figure 3: Location of study site (With small changes after Lai and Detphachanh 2006)

We chose 3 different DTs for the evaluation of ACS using the FDC of study river (Figure 4). Prior to choosing these DTs, several trials were carried out to understand the behavior of different levels. These trials demonstrated that the ratio of additional adaptive samples to the corresponding initial sample size vary from one sample set to another, but decreases by increasing the size of initial sample size. Therefore, first a "reference" DT was chosen using the FDC. The selection of this DT was a little subjective, however it was chosen in a manner that it never produces additional adaptive samples more than around twice the corresponding initial sample size. A DT exceeded 2.5% of time showed appropriate results. In the next step, two new DTs were also considered at twice and half the occurrence probability of the reference DT respectively (i.e. discharge exceeded 5 and 1.25% of time). Table 3 shows the selected DTs. They are called high, medium (or reference discharge) and low in this paper in accordance with their magnitudes. Table 3 also shows all possible additional units for the selected DTs.

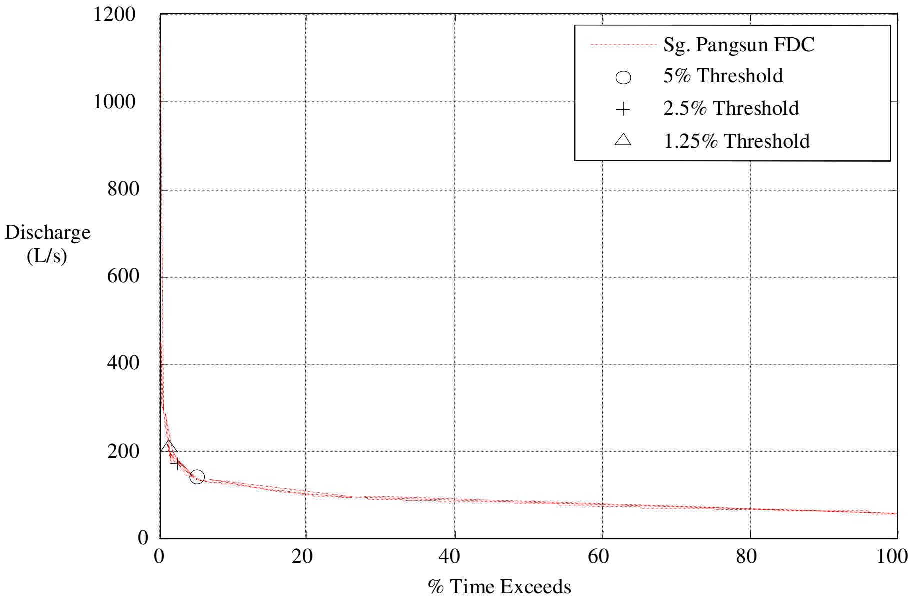

Figure 4: Flow duration curve

TABLE 3: Discharges exceeding selected thresholds using flow duration curve

Discharge frequency	Designation	Discharge (L/s)	Number of clusters	Number of additional samples
FDC 5	Low DT	142.5	19	1175
FDC 2.5	Medium/Reference DT	170.7	16	587
FDC 1.25	High DT	208.6	14	294

Example: FDC5= Flow duration Curve exceeds 5% of time, DT= Discharge Threshold,

Figure 5 shows the combination of parameters influencing the adopted ACS. The intersection of two neighborhood relationships (symmetric and forward) and the three DTs produced six combinations. All sample sets that were the same regarding these two parameters were grouped under a "scenario" in this study. As shown in Figure 5, each scenario was simulated with ten different sizes of initial sample sets creating 60 new combinations, which were essentially the "treatments". The simulation for treatments each in replicates of 50 produced 3000 "final adaptive sample sets" which eventually resulted in 3000 "estimates" with the Horvitz-Thompson estimator. The suitability of this estimator compared to the Hansen-Hurwitz was reported by Salehi (2003).

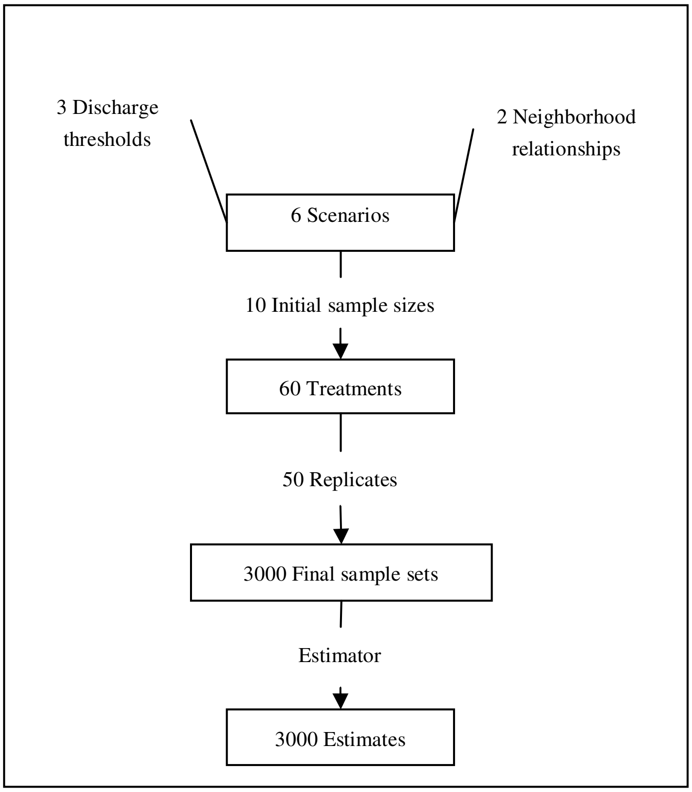

Figure 5: Concepts of scenarios, treatments and estimations

Thompson (1992) presented the modified Horvitz-Thompson estimator for ACS. If the initial sample set size n is taken by random sampling without replacement from a population size N and x_k denotes the number of units in the k^{th} network, then the probabilities of intersecting initial sample set and networks can be calculated by:

$$a_k = 1 - \binom{N - x_k}{n} \Big/ \binom{N}{n} \tag{1}$$

The total can be calculated by applying the modified Horvitz-Thompson estimator to the final sample set:

$$\hat{T}_{HT} = \sum_{k=1}^{k} \frac{y_k^* z_k}{a_k} \tag{2}$$

where,

$\hat{T}_{HT}$ is the estimator of total

k is total number of distinct networks in the population

y_k^* is total of y values in the network k. For every network with one unit $y_k^* = y$

z_k is an indicator equal one if any unit of the k^{th} network is in the initial sample, equal zero otherwise.

Computer codes written in MATLAB software were developed for extracting several thousand intended sample sets and estimating their corresponding total sediment load. Knowing the total observed load and total estimated sediment loads, treatments were compared with respect to the unbiasedness, precision and accuracy using the average percent error, coefficient of variation (CV), and normalized root mean square deviation (NRMSD) respectively. The NRMSD incorporates the effect of variance and bias into a single measure. It is calculated as:

$$NRMSD\left[\hat{Q}_s\right] = \sqrt{V\left[\hat{Q}_s\right] + \left(Bias\left[\hat{Q}_s\right]\right)^2} \Big/ Q_s \tag{3}$$

where $Bias\left[\hat{Q}_s\right]$ and $V\left[\hat{Q}_s\right]$ are the bias and variance respectively and Q_s is the observed load. Treatments with small $NRMSD\left[\hat{Q}_s\right]$ were considered more accurate. Zamyadi *et al.* (2007) have recently employed root mean square error to compare different strategies of load estimation in Canada.

RESULTS

Sediment loads were estimated for 6 scenarios which are given in the following two subsections with respect to the neighborhood relation:

Adaptive Cluster Sampling with Symmetric Neighborhood Relation

Table 4 shows some experimental results for the treatments with SNR. From column 3, it is apparent that the number of clusters is a function of DT. For example, the average number of clusters for the largest initial sample (5000) of low DT (FDC5) is only 4763 about 169 less than the analogous number of clusters for the high DT (FDC1.25). This is attributed to the higher possibility of merging nearby clusters in the low DT case.

The fifth column of Table 4 and Figure 6A demonstrate that the discharge thresholds significantly affect the number of additional adaptive samples. These results show a sharp rising trend of additional adaptive samples for smaller initial sample sets (size 50 to 500) followed by a gentle rising for larger sample sets. In fact, several important units in the large size sample sets belong to initial samples rather than the additional adaptive samples. The number of additional adaptive samples in each low level discharge is twice of the next higher one, because of their occurrence probabilities.

TABLE 4: Descriptive statistics for scenarios with SNR*

Scenario	ISS	NC	Mean FSS	NAS	MeanEr (%)	CV	NRMSD
SNR-FDC5	50	50.00	779.04	729.04	26.53	16.41	0.34
	150	146.98	1021.32	874.34	6.64	7.54	0.10
	300	292.04	1252.24	960.20	5.05	5.28	0.08
	500	484.52	1518.18	1033.66	5.12	3.63	0.06
	1000	964.22	2086.22	1122.00	4.12	0.37	0.04
	1500	1439.68	2577.36	1137.68	1.84	0.31	0.02
	2000	1915.80	3061.96	1146.16	0.98	0.22	0.01
	3000	2866.46	4019.78	1153.32	0.29	0.21	0.00
	4000	3814.68	4969.12	1154.44	0.05	0.15	0.00
	5000	4762.86	5918.22	1155.36	-0.06	0.11	0.00
SNR-FDC2.5	50	49.98	154.32	104.34	2.80	42.41	0.44
	150	149.78	464.22	314.44	17.77	10.55	0.22
	300	298.38	717.68	419.30	9.89	6.34	0.12
	500	495.32	952.88	457.56	4.44	8.79	0.10
	1000	986.28	1515.04	528.76	6.71	1.02	0.07
	1500	1475.88	2033.12	557.24	4.53	0.55	0.05
	2000	1963.78	2531.16	567.38	3.10	0.29	0.03
	3000	2937.54	3507.16	569.62	1.72	0.22	0.02
	4000	3916.80	4487.20	570.40	1.32	0.20	0.01
	5000	4886.02	5456.40	570.38	1.07	0.12	0.01
SNR-FDC1.25	50	50.00	62.40	12.40	-24.77	46.77	0.43
	150	150.00	219.68	69.68	14.42	28.09	0.35
	300	299.84	411.40	111.56	0.56	23.41	0.24
	500	499.82	709.58	209.76	24.65	13.24	0.30
	1000	995.82	1236.86	241.04	10.29	2.98	0.11
	1500	1491.12	1754.22	263.10	6.59	0.61	0.07
	2000	1986.70	2256.00	269.30	3.69	0.41	0.04
	3000	2967.34	3245.40	278.06	1.36	0.29	0.01
	4000	3963.56	4242.78	279.22	0.53	0.25	0.01
	5000	4932.34	5211.66	279.32	0.12	0.25	0.00

*SNR=Symmetric neighborhood relation, FDC5 & 2.5 & 1.25 are discharges exceed 5 & 2.5 & 1.25 % of time in flow duration curve, ISS= Initial sample size, NC=Number of cluster, MeanFSS=Average final sample size, NAS= Number of additional adaptive samples, CV= Coefficient of variation, MeanEr=Average error, NRMSD=Normalized root mean square deviation

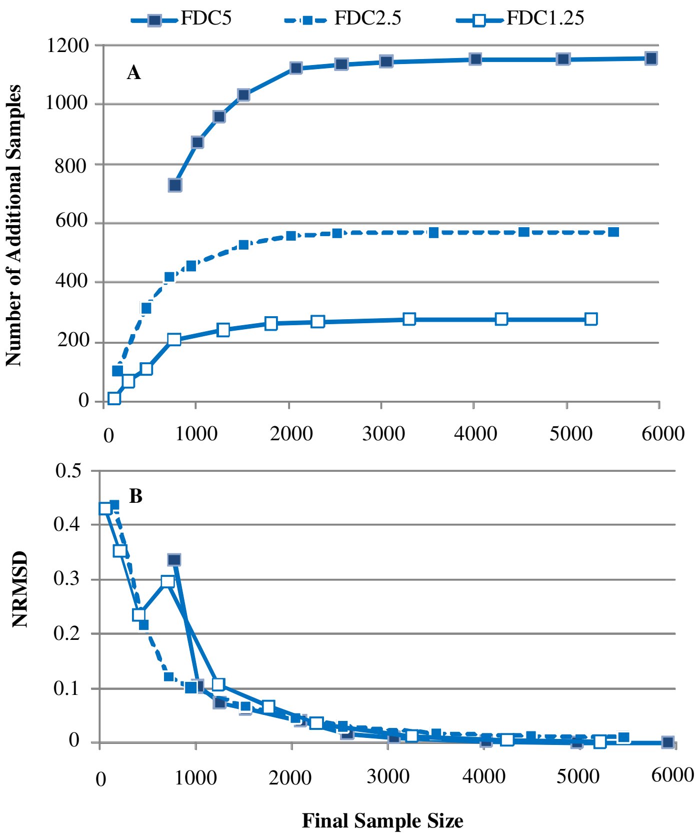

Figure 6: Comparison of SNR scenarios with respect to number of additional samples and NRMSD

As shown in column 6, Table 4, ACS-SNR mostly overestimates the mean load for small and medium (up to 1500 or 2000) initial sample sizes. It then gets almost unbiased for large and very large initial sample sizes (>2000). The trend of mean percent error for the low DT is quite steady, starts from a high value gradually decreases to unbiased. Whereas, irregularities are shown in mean percent error trend for the UCS-YFDC1.25 and UCS-YFDC2.5 scenarios (the high and medium DTs). Note that the smallest final sample size for the low DT is 780 compared to 154 and 62 for the other scenarios.

Regarding the CV, all scenarios show similar trends; start from a high value for the smallest initial sample size, gradually drop to less than one for the largest initial sample size (Column 7, Table 4). Although, the low DT shows better performance regarding the CV compared to other scenarios for similar initial samples; the relation is inverse considering final adaptive sample size.

It is apparent from subplot B of Figure 6 that NRMSDs of estimations mostly show a negative correlation with final sample size except an irregularity for the FDC1.25. They became almost zero for large sample sizes.

NRMSD for final sample sizes smaller than 500 is not better than 0.22, however, a few good results (NRMSD≈0.10) are obtained for sample sets with a size around 1000.

Adaptive Cluster Sampling with Forward Neighborhood Relation

Table 5 presents the results obtained from the simulations of ACS with FNR. Number of clusters for all treatments in this table are only one or two units less than the analogous values in SNR scenarios (Table 5) revealing that FNR does not affect the number of clusters very much.

TABLE 5: Descriptive statistics obtained for scenarios with FNR*

Scenario	ISS	NC	Mean FSS	NAS	MeanEr (%)	CV	NRMSD
FNR-FDC5	50	50.00	380.38	330.38	-35.90	21.99	0.39
	150	147.00	817.94	670.94	-30.07	15.52	0.32
	300	291.92	1084.70	792.78	-21.19	4.29	0.21
	500	484.70	1349.18	864.48	-18.38	7.96	0.19
	1000	964.18	1944.92	980.74	-13.85	5.62	0.15
	1500	1441.88	2472.08	1030.20	-10.87	3.78	0.11
	2000	1916.42	2974.96	1058.54	-8.16	3.00	0.09
	3000	2865.20	3957.44	1092.24	-4.50	1.59	0.05
	4000	3821.46	4930.26	1108.80	-2.85	0.95	0.03
	5000	4761.62	5882.10	1120.48	-1.85	0.59	0.02
FNR-FDC2.5	50	50.00	98.28	48.28	-30.26	36.06	0.39
	150	149.92	319.94	170.02	-22.65	21.00	0.28
	300	298.42	573.30	274.88	-14.15	5.32	0.15
	500	495.38	841.58	346.20	-15.66	9.20	0.17
	1000	986.10	1415.06	428.96	-12.07	5.00	0.13
	1500	1474.26	1938.06	463.80	-10.54	4.45	0.11
	2000	1962.92	2452.80	489.88	-7.32	3.42	0.08
	3000	2932.52	3451.84	519.32	-4.84	1.50	0.05
	4000	3903.10	4436.26	533.16	-3.37	0.93	0.03
	5000	4886.44	5428.28	541.84	-2.57	0.75	0.03

*FNR=Forward neighborhood relation, FDC5 & 2.5 & 1.25 are discharges exceed 5 & 2.5 & 1.25 % of time in flow duration curve, ISS= Initial sample size, NC=Number of cluster, MeanFSS=Average final sample size, NAS= Number of additional adaptive samples, CV= Coefficient of variation, MeanEr=Average error, NRMSD=Normalized root mean square deviation

TABLE 5 (continued): Descriptive statistics obtained for scenarios with FNR*

Scenario	ISS	NC	Mean FSS	NAS	MeanEr (%)	CV	NRMSD
FNR-FDC1.25	50	50.00	56.54	6.54	-28.99	46.41	0.44
	150	149.96	177.62	27.66	-28.34	20.69	0.32
	300	300.00	366.66	66.66	-22.64	16.25	0.26
	500	499.48	616.84	117.36	-19.65	14.46	0.23
	1000	996.56	1159.28	162.72	-20.54	13.70	0.23
	1500	1492.42	1686.70	194.28	-15.53	6.39	0.16
	2000	1987.32	2200.84	213.52	-12.73	8.97	0.15
	3000	2976.62	3212.16	235.54	-7.15	4.76	0.08
	4000	3957.56	4206.10	248.54	-4.09	1.68	0.04
	5000	4943.76	5198.44	254.68	-3.02	1.25	0.03

*FNR=Forward neighborhood relation, FDC5 & 2.5 & 1.25 are discharges exceed 5 & 2.5 & 1.25 % of time in flow duration curve, ISS= Initial sample size, NC=Number of cluster, MeanFSS=Average final sample size, NAS= Number of additional adaptive samples, CV= Coefficient of variation, MeanEr=Average error, NRMSD=Normalized root mean square deviation

Figure 7A shows the number of additional adaptive samples for ACS-FNR scenarios. It can be compared to Figure 6A of SNR that shows similar trends for different analogous scenarios. The number of additional samples for the smallest initial sample set of FNR scenarios are almost half of their analogous values with SNR. However, for the largest sample size, the former requires about 90% of additional samples compared to the SNR.

Column 6, Table 5 presents percent of mean error for estimated loads. It underestimates the load by -30 to -40 % with the initial sample set size 50. Mean error shows an ascending trend for all scenarios. The mean percent error for largest sample size is almost near zero.

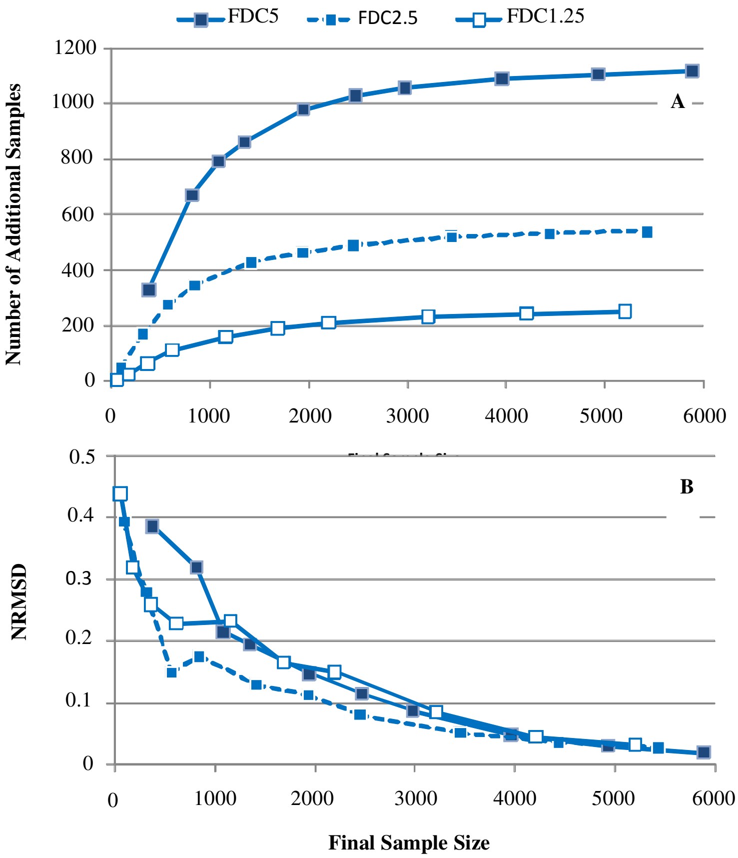

Figure 7: Comparison of FNR scenarios with respect to number of additional samples and NRMSD

As shown in Column 7, Table 5, the trend of coefficient of variation for all three scenarios are similar. The CV values are near zero for very large sample sizes.

The last column of Table 5 provides the value of NRMSD and Subplot B of Figure 7 compares scenarios graphically. NRMSD shows negative correlation with final sample size, however like as SNR. Nevertheless, the amounts of calculated NRMSDs for FNR are a little higher than their analogous values for SNR. For example, as Table 5 represents the minimum NRMSD with SNR is zero corrected to two decimal places. In comparison, the least NRMSD of FNR is not better than 0.03.

As it is apparent from Figure 7B, surprisingly the amount of NRMSD for estimates of FDC2.5 is the least and better than FDC1.25 and FDC5 respectively. This finding is different with the result obtained from SNR scenarios.

DISCUSSION

So far, a number of scenarios and treatments on the use of ACS for sediment load estimation were compared under two SNR and FNR categories. The scenarios and treatments differed by the input parameters including neighborhood relation, discharge threshold, and initial sample size. The significance of related input parameters to achieve accurate sediment load is evaluated in the next three subsections.

The Effect of Initial Sample Size

Looking at the results of SNR scenarios show large variations and irregularities of mean percent error or overestimation for treatments with small and even medium initial sample sizes. The irregularities included instabilities in trends of mean percent errors for initial sample sizes $\leq$1500. The estimates for large ($\geq$2000) initial sample sets showed relatively small percent error (<5) and very low CV (even <1 in several treatments). This result is consistent with Thompson (1992) who found ACS is more effective with high rate sampling of initial sample (even 25 or 50% of population). Therefore, taking an initial sample set with size at least 10-15 % of the concerned population would be necessary to achieve relatively accurate estimates.

This small amount of biasedness could be attributed to the sampling design adopted in this study where the discharge threshold -as an auxiliary variable - was used instead of sediment concentration threshold (variable of interest). It is expected that the estimations made by the adopted ACS would be unbiased relative to discharge and not necessarily to sediment load.

One major weakness of the ACS design is the ineffective use of a small initial sample size for a population when the aggregation and localization is significant. In fact, high CV of treatments with small initial sample size is because of extreme aggregations of study sediment population. For example, 45 % of load was transported during only one percent of the time. Consequently, the probability of intersecting these very important rare storm flow events by a small initial sample set is low leading to high variance and CV values in every 50 replicates.

Effect of Discharge Threshold

The number of additional samples becomes restricted by the use of high discharge thresholds. However, the study has also demonstrated that choosing low discharge threshold does not necessarily lead to an accurate estimation although the latter would involve more samples.

Based on these findings, choosing a threshold that the additional samples do not exceed 25-30% of initial sample size could be satisfactory. The FDC may be helpful to find an appropriate threshold in this instance.

The number of additional samples also is affected by the following properties of the study population: i) the length of storm flow events and ii) the length of each sampling unit. For rivers with long storm hydrographs and daily sampling unit, flow at a higher percentage of time can be selected. In practice, simulation of two or three discharge levels using the FDC of study river will help to decide on a threshold.

The Effect of Forward Neighborhood Relation

The ACS estimates for FNR scenarios for initial samples with size smaller than 1000-1500 showed significant underestimations, whereas the present errors decreased to -2 to -13 % for larger initial sample sizes.

The lower estimations for FNR scenarios are probably due to fewer samples taken during the first halves of the storm hydrographs particularly, the rising limb compared to the SNR as shown in Figure 8. The horizontal axis in this figure is real time. If the initial sample set intersects the storm hydrograph only at 19:45 hrs, the SNR will extract 17 samples, while only 9 additional samples are identified by FNR using the discharge threshold 200 L/s.

Most rivers usually carry more sediment load during the rising limb of a single storm than the corresponding discharge during the falling limb due to hysteresis effects which is called clockwise hysteresis (Vanoni 1980, Williams 1989, Gomi, Moore and Hassan 2005). Therefore, the first half of a hydrograph that coincides with the rising limb and sampling units around the peak have a considerable role in load transfer rather than the end half. This can be seen in Figure 8, which is an actual hydrograph of Sg. Pangsun dated 8^{th} and 9^{th} November 1997. The amount of load transported by the first 8 sampling units or time intervals of storm hydrograph (from 17:45 hrs to 19:45 hrs) is 1450 kg, while for the remaining 9 sampling units, the load is only 142 kg.

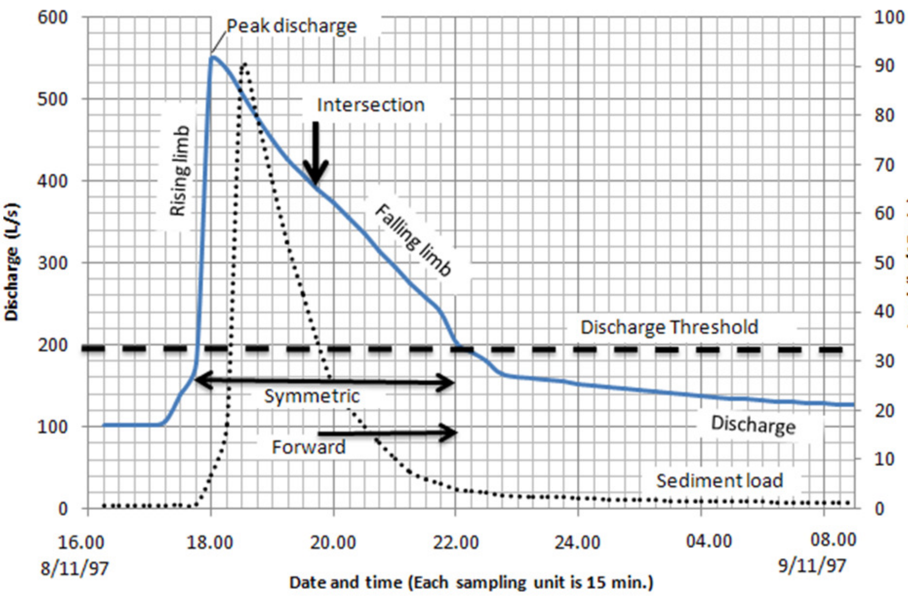

Figure 8: Comparison of SNR and FNR on the number of additional samples; shown by arrows' length

The underestimation with forward neighborhood relation also depends on the length of rising and falling limbs. Table 6 shows the percentages of rising and falling limbs for storm hydrographs higher than the study discharge thresholds. The lengths of average rising limb are less than 15 % of base time. Therefore, by substituting FNR instead of SNR, smaller amounts of load are anticipated.

TABLE 6: Average percentages of rising and falling limbs in Sg.Pangsun

Discharge thresholds*	Rising limb (%)	Falling limb (%)
YFDC5	10	90
YFDC2.5	11	89
YFDC1.25	14	86

*The values for discharge thresholds have been given in Table 3

The effects of FNR on less additional samples from the first half particularly occur for short storm hydrographs, when only one initial sample intersects the event hydrograph. A large initial sample set, which intersects a prolonged storm hydrograph in many sampling units, would increase the chance of including samples from the early half and consequently increases the number of additional adaptive samples. Therefore, it is expected that the both neighborhood relations (symmetric and forward) produce closer estimates by increasing the size of the initial sample.

CONCLUSION

This study showed relatively reliable performances for the ACS either with the SNR or with FNR when large enough initial sample sets would have been taken. In fact, an intensive initial systematic sampling ensures identifying the important additional samples. As a rule, taking an initial sample set larger than 10 % of the expected population size is suggested.

The current study placed much emphasis on the performance of ACS. However, the cost of sampling is an important consideration. The cost of taking samples with the ACS is probably less than calendar-based with respect to the travelling costs because the person-in-charge can stay on gauging site to take additional samples if so decided by ACS. This sampling design can also be recommended for remote gauging sites on rivers that are not easily accessible to save cost.

In spite of the advantages offered by ACS, an important limitation needs to be considered. The number of taken additional samples in the ACS differs from year to year due to the stochastic nature of discharge and sediment. This will affect the number of final samples taken, while the budget allocated for sediment sampling is usually fixed each year.

Choosing a sampling design for a gauging station depends on the nature of data collected, in particular the advantages and limitations. The most significant advantage of SALT and ACS designs compared to the calendar-based is that the two former methods work better as they select more samples during high flows. Therefore, they are suitable for estimating suspended sediment load since sediment is carried mostly during floods. Although, both the ACS and SALT take more samples during high flows, the use of the latter is restricted in most rivers, where the gauging site was not equipped with automatic sediment sampler.

REFERENCES

Acharya, B., Bhattarai, G., de Gier, A., and Stein, A. 2000. Systematic Adaptive Cluster Sampling for the Assessment of Rare Tree Species in Nepal, *Forest Ecology and Management*, **137**, 65-73.

Arabkhedri, M., Lai, F. S., Noor-Akma, I., and Mohamad-Roslan, M. K. 2007. Estimating Suspended Sediment Load Using the Adaptive Cluster Sampling Strategy and Design Approach, *Rivers' 07- the 2nd international conference on managing rivers in the 21th century: solutions towards sustainable river basins*, 6.

Asselman, N. E. M. 2000. Fitting and Interpretation of Sediment Rating Curves, *Journal of Hydrology*, **234,** 228-248.

Chaudhuri, A., Bose, M., and Dihidar, K. 2005. Sample-Size-Restrictive Adaptive Sampling: An Application in Estimating Localized Elements, *Journal of Statistical Planning and Inference*, 134, 254-267.

Cohn, T. A. 1995. Recent Advances in Statistical Methods for the Estimation of Sediment and Nutrient Transport in Rivers, *Reviews of Geophysics*, **33**, 1117–1124.

Cohn, T. A. 2005. Estimating Contaminant Loads in Rivers: An Application of Adjusted Maximum Likelihood to Type 1 Censored Data, *Water Resources Research*, **41**, 1-13.

Cohn, T. A., Delong, L., Gilroy, E., Hirsch, R., and Wells, D. 1989. Estimating Constituent Loads, *Water Resources Research*, **25**, 937-942.

Degens, B. P., and Donohue, R. D. 2002. *Sampling Mass Loads in Rivers – a Review of Approaches for Identifying, Evaluating and Minimising Estimation Errors*, Water and Rivers Commission, Perth, WA.

Ferguson, R. 1986. River Loads Underestimated by Rating Curves, *Water Resources Research*, **22**, 74-76.

Geoffery, J. G. 1999. *Suspended Sediment Yield Analysis Using a Continuous Monitoring System in the Sg. Pangsun, Ulu Langat, Selangor*, Universiti Putra Malaysia.

Gomi, T., Moore, R. D., and Hassan, M. W. 2005. Suspended Sediment Dynamics of Small Forest Streams of the Pacific Northwest, *Journal of the American Water Resources Association*, **41**, 877-898.

Juang, K. W., Lee, D. Y., and Teng, Y. L. 2005. Adaptive Sampling Based on the Cumulative Distribution Function of Order Statistics to Delineate Heavy-Metal Contaminated Soils Using Kriging, *Environmental Pollution*, **138**, 268-277.

Koch, R., and Smillie, G. 1986. Comment On "River Loads Underestimated by Rating Curves" By R. I. Ferguson, *Water Resources Research*, **22**, 2121-2122.

Lai, F. S., and Detphachanh, S. 2006. Sediment Production During an Unusually Dry Year in the Steep Forested Sungai Pangsun Catchment, Selangor, Peninsular Malaysia, *Forest Ecology and Management*, **224**, 157-165.

Lo, N. C., Griffith, H. D., and Hunter, J. R. 1997. Using a Restricted Adaptive Cluster Sampling to Estimate Pacific Hake Larval Abundance, *California Cooperative Oceanic Fisheries Investigations. Reports* **37**, 160–174.

Salehi, M., and Smith, D. R. 2005. Two-Stage Sequential Sampling: A Neighborhood-Free Adaptive Sampling Procedure, *Journal of Agricultural, Biological, and Environmental Statistics*, **10**, 84–103.

Salehi, M. M. 2003. Comparison between Hansen-Hurwitz and Horvitz-Thompson Estimators for Adaptive Cluster Sampling, *Environmental and Ecological Statistics*, **10**, 115-127.

Su, Z., and Quinn, T. J. 2003. Estimating and Efficiency for Adaptive Sampling with Order Statistics and Stopping Rule *Environmental and Ecological Statistics*, **10**, 17-41.

Thomas, R. 1985. Estimating Total Suspended Sediment Yield with Probability Sampling, *Water Resources Research*, **21**, 1381-1388.

Thomas, R. B. 1988. Monitoring Baseline Suspended Sediment in Forested Basins: The Effects of Sampling on Suspended Sediment Rating Curves, *Hydrological Sciences*, **33**, 499-514.

Thomas, R. B., and Lewis, J. 1993. A Comparison of Selection at List Time and Time-Stratified Sampling for Estimating Suspended Sediment Loads, *Water Resources Research*, **29**, 1247-1256.

Thomas, R. B., and Lewis, J. 1995, An Evaluation of Flow-Stratified Sampling for Estimating Suspended Sediment Loads, *Journal of Hydrology*, **170**, 27-45.

Thompson, S. K. 1992. *Sampling*, New York: Wiley.

Thompson, S. K. 1996. Adaptive Cluster Sampling Based on Order Statistics, *Environmetrics*, **7**, 123-133.

Vanoni, V. A. (ed.) 1980. *Sedimentation Engineering*, A.S.C.E.

Walling, D. E., and Webb, B. W. 1981. The Reliability of Suspended Sediment Load Data, *Erosion and Sediment Transport Measurements*, **133**, 177 – 194.

Walling, D. E., and Webb, B. W. 1988, The Reliability of Rating Curve Estimates of Suspended Sediment Yield: Some Further Comments, *Sediment Budgets*, **174**, 337 – 350.

Williams, G. P. 1989. Sediment Concentration Versus Water Discharge During Single Hydrologic Events in Rivers, *Journal of Hydrology*, **111**, 89-106.

Zamyadi, A., Gallichand, J., and Duchemin, M. 2007. Comparison of Methods for Estimating Sediment and Nitrogen Loads from a Small Agricultural Watershed, *Canadian Biosystems Engineering*, **49**, 1-27 to 21-36.