

An application of adaptive cluster sampling for estimating total suspended sediment load

M. Arabkhedri, F. S. Lai, I. Noor-Akma and M. K. Mohamad-Roslan

ABSTRACT

Suspended sediment transport in river for a particular period is a timescale finite population. This population shows natural aggregation tendencies in sediment concentration particularly during floods. Adaptive cluster sampling (ACS) can be potentially conducted for sampling from this rare clustered population and estimating total load. To illustrate the performance of ACS in sediment estimation, a comparative study was carried out in the Gorgan-Rood River, Iran, with around a 5 year daily concentration record. The total sediment loads estimated by ACS were statistically compared to the observed load, estimations of selection at list time (SALT) and conventional sediment rating curve with and without correction factors. The results suggest that none of the sediment rating curves produced accurate estimates, while both ACS and SALT showed satisfactory results at a semi-weekly sampling frequency. The best estimation obtained by the rating curves did not show a percent error better than –40%; however, ACS and SALT underestimated the load at less than 5%. The results of this study suggest ACS could improve river monitoring programs.

Key words | adaptive cluster sampling, Gorgan-Rood River, sediment rating curve, selection at list time, suspended sediment load

M. Arabkhedri (corresponding author)

F. S. Lai

M. K. Mohamad-Roslan

Faculty of Forestry, University Putra Malaysia,
Serdang 43300, Selangor,
Malaysia

M. Arabkhedri (corresponding author)

Soil Conservation and Watershed Management
Research Institute,
PO Box 13445-1136,
Tehran,
Iran
E-mail: arabkhedri@scwmri.ac.ir

I. Noor-Akma

Institute for Mathematical Research (INSPERM),
University Putra Malaysia,
Serdang 43300, Selangor,
Malaysia

INTRODUCTION

Sediment transport in a river system and its deposition thereafter are two significant offsite effects of erosion from an upstream basin. Suspended sediment also carries various chemicals and pollutants that are transported into downstream water bodies. Most development projects in river basins need estimates of sediment output. Accurate estimation of suspended sediment yield has been a continuous challenge during recent decades.

Wash load usually constitutes the greatest part of suspended sediment in rivers. The amount of transported fine wash load particles is not sensitive to flow parameters (Vanoni 1975). Hence, it is often modelled using empirical relations (Asselman 2000) such as conventional sediment rating curve (CSRC). The existing gaps in sediment sample record have made CSRC an ordinary method for estimating missing values (Lai *et al.* 1995).

According to several studies, CSRCs most often significantly underestimate the long-term sediment transport rates as much by 50–60% or even more (Thomas 1985, 1988; Ferguson 1986; Koch & Smillie 1986; Walling & Webb 1988; Cohn *et al.* 1989; Asselman 2000; Cohn 2005).

Based on statistical considerations, Duan (1983) and Miller (1984) suggested that logarithmic transformation is the main cause of underestimation by regression models and recommended two widely used parametric and non-parametric correction factors (PCF and NPCF) to balance the bias. Ferguson (1986) reported satisfactory outcome of applying PCF to CSRC. In a follow-up paper, Koch & Smillie (1986) successfully applied NPCF to their sediment dataset. The application of these correction factors in many experiments was not very encouraging (Thomas 1985, 1988; Walling & Webb 1988; Crowder *et al.* 2007; Sadeghi *et al.*

doi: 10.2166/nh.2010.113

2008). The use of these correction factors could not entirely balance the existing bias and only improve the estimated load. This led some authors to claim that the bias associated with logarithmic transformation is not the prime cause of sediment rating curve inaccuracy (Walling & Webb 1988). Bias occurs as a consequence of higher samples taken during low flows, in contrast to the lower samples taken during high flows when most sediment is exported.

Extrapolation of sediment concentrations from sediment rating curves for flood flows also introduces uncertainties in load estimations. Both problems are due to the non-representativeness of sample sets. In this regard, Schleppe *et al.* (2006) suggested that the best way to obtain accurate and precise flux estimates is through probability sampling proportional to discharge.

A group of researchers have, however, suggested the use of non-linear regression in developing sediment rating curves (Horowitz 2003; Demissie *et al.* 2004; Schmidt & Morche 2006; Crowder *et al.* 2007; Sadeghi *et al.* 2008). Typically, the main shortcoming of the non-linear model is that residual errors are not identically distributed throughout the range of stream flow values. In addition, Thomas (1985) argued that the sediment rating curve could be termed non-statistical as sampling probabilities are not known; the estimators, therefore, cannot consider the probability structure, resulting in bias.

Survey sampling

In response to the shortcomings of sediment rating curve, the use of survey sampling for river sediment monitoring and estimating total suspended sediment were introduced in the 1980s. Survey sampling includes two parts: first, a sampling design describes the time schedule of sampling for river sediment studies and the second, the estimators, is used to estimate total load and its variance. These estimators are only relevant with the corresponding sampling designs and cannot be used with another sampling design.

Accurate estimates of total sediment load were reported by applying unequal probability survey sampling, including selection at list time (SALT) (Thomas 1985), time-stratified sampling (Thomas & Lewis 1993) and flow-stratified sampling (Thomas & Lewis 1995). Stratified sampling was

mostly used for the estimation of transported loads during storm hydrographs. Nevertheless, the SALT, which is a probability proportional to size sampling design, was also employed in the sampling of storm hydrographs and longer periods (up to several years). These sampling techniques were considered by Cohn (1995) as important innovations in estimating sediment transport from small watersheds. He recommended their uses in basins where the sediment rating curve is not valid. These procedures are only meaningful in basins instrumented with programmable sediment samplers.

In spite of an obvious preference for survey sampling in producing unbiased sediment estimates, little attention has been paid to this method by the recent literature. The reason most likely relates to the difficulty and complexity of their use in sediment load estimation. In addition, these techniques are only relevant in a small number of basins equipped with a pumping sampler set. Most water quality monitoring programs, however, are carried out without the use of automatic samplers, as manual sampling is a common feature in majority of rivers (Degens & Donohue 2002).

Adaptive cluster sampling

One possibility of overcoming the problem of ill-equipped sediment stations discussed earlier is the use of adaptive cluster sampling (ACS) in collecting river water samples for sediment estimation where automatic samplers are not available. Thompson (1990) introduced this method in investigating rare clustered populations strictly in the statistical sense, but its application in specific problems such as forestry (Acharya *et al.* 2000; Talvitie *et al.* 2006), soil (Juang *et al.* 2005), social issues (Chaudhuri *et al.* 2005) and fishery (Bradbury 2000; Harris 2008) has been widely used during the last decade.

In a rare clustered population, the study variable in most sampling units is zero or negligible. However, it is substantial in the other sampling units (referred to as important units) because of heavy localization of the high valued units in certain parts. The main concept used in ACS is taking more samples from important units based on a neighbourhood relation.

Sediment transport is typical of a rare timescale clustered population. A study period can be divided into

N equal time intervals with almost constant discharge and sediment concentration in each period. Each time interval is a sampling unit. Sampling units with high sediment concentration show aggregation tendency and coincide mostly with flood events. For example, Walling (1994) reported that 95% of sediment load is carried over only 5% of the time and is negligible at other times for the Creedy basin in the UK. This means that there is natural clustering in sediment concentration populations particularly during floods. Markus & Demissie (2006) also reported that of 27 basins studied in the USA, 68% of the annual loads were transported by four highest floods each year.

Since ACS is new to the hydrological sciences, a basic form is briefly described using an illustrative example (Figure 1) to understand its context. In ACS, samples are taken in two steps. It also requires a threshold or boundary to determine important sampling units. First, an initial sample set (units shaded black in Figure 1) is taken. If the measured variable for an initial selected unit is equal to or exceeds the threshold (here assumed to be 25) then, as shown in Figure 1, the adjacent right and left units (symmetric relation) are taken in the next steps (shaded grey), until the study variable becomes less than the threshold. A set of samples including an initial sampled unit and all other related observed units is called a cluster. An edge unit is the observed neighbourhood of a previous selected unit that does not satisfy the condition. All units in a cluster without the edge units is termed a network of size x_k . Consequently, a final sample set is formed which constitutes the initial sample set together with all additional selected samples in the networks.

As stressed earlier, the sediment concentration population is a timescale population. For that reason, the classic ACS with symmetric neighbourhood relation (two sides) cannot be employed since the person in charge cannot sample those occurred events during the previous time intervals. Therefore, an altered forward neighbourhood relation (one side) was adapted in the current study instead. Sampling with a symmetric neighbourhood relation was reported to be unpractical in some cases or even impossible by some authors, who then suggested some remedies for their cases (Lo et al. 1997; Salehi & Smith 2005).

This paper describes the use of ACS design for the monitoring of suspended sediment to estimate total sediment load in rivers. In addition, comparisons are made between this method and four existing sediment load estimation approaches.

STUDY AREA AND AVAILABLE DATA

This study was performed in the Gorgan-Rood River at the Ghazaghly gauging site, located upstream of the Voshmgir dam, Golestan Province, Iran. This 7,062 km² river basin, situated in the north of Iran, originates from the Alborz Mountain and drains an average annual discharge of 12 m³ s⁻¹ into the Caspian Sea. The maximum precipitation in the mountainous area occurs in winter; however it occurs in late summer and early autumn in the lowlands. Loess without good vegetation covers the lower part of the basin and provides lots of sediment to the river system, particularly during summer rainstorms.

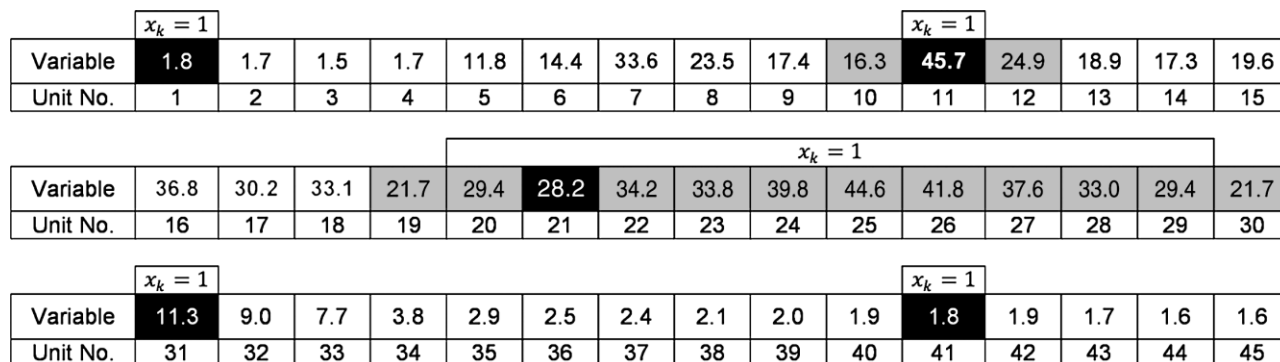


Figure 1 | Adaptive cluster sampling in a small string population.

The available data includes 1686 coincident discharge and concentration measurements conducted by the Ministry of Energy, Iran, from 1982 to 1987. Sampling was carried out every day. The daily mean discharge and a few hundred sediment concentration records are also available from 1973 to 1981. The maximum observed discharge is more than $200 \text{ m}^3 \text{ s}^{-1}$; however, the river may be almost dry for several days in summer every year. The sediment concentration during summer storms is usually high and, after the dry periods, may even exceed 50 g L^{-1} . The greatest observed daily load occurred after a short summer rain-storm (with a discharge of $32.5 \text{ m}^3 \text{ s}^{-1}$) in August 1986 and transferred 0.92×10^6 tonnes of suspended sediment. The smallest daily load for the same discharge during the study period was only about 5,000 tonnes per day, which occurred in April 1986. In comparison, the transported daily load by the maximum observed discharge is 147,000 tonnes and the minimum observed daily load was as low as 0.01 tonne. The total recorded load during the study period was 12.54×10^6 tonnes. Although the data are not recent, it is the best available sediment concentration record in Iran.

Data analysis

Sampling designs

In this study, three designs were used to extract sample sets: (1) ACS; (2) calendar-based and (3) SALT. Code written using MATLAB facilitated the simulation of samplings.

1. Adaptive sampling: For field sampling of the proposed ACS with a forward neighbourhood relation, a time schedule and a predetermined discharge threshold are required. The person in charge must visit the gauging site in accordance with the time schedule, take the sediment sample and record the instantaneous discharge. When the observed discharge exceeds the predetermined discharge threshold, he/she will take a sequence of samples in the subsequent time intervals until the instantaneous discharge drops below the threshold. Otherwise, the person will have to take the next sample according to the time schedule. Using the forward neighbourhood relation, he/she loses the sampling units before the initial chosen units (e.g. 20th day in Figure 1). With this method, no edge units are sampled because an auxiliary

variable is used (e.g. days 10, 12, 19 and 31 in Figure 1). We simulated the above ACS design by running MATLAB code for three initial sample sizes including 100, 200 and 400. It is noted that fixed-interval time schedules were used for selecting initial sample sets. Due to non-integer intervals, the real time intervals were not exactly equal for all units in the sample sets, causing one day variation between some consequent samples. The initial sample sets, therefore, were considered simple random sample sets.

Choosing an appropriate threshold by ranking the values of all initial samples was suggested in spatial-scale studies (Thompson 1996). This method is not possible in river sediment studies as discharges during following days, which are timescale based, are not known. This study has made use of the monthly flow duration curves using previous daily discharge records from 1973 to 1981 to define the threshold. The boundary between high and low flows (Table 1) was considered to be that which exceeded the flow 20% of the time within each month. These boundaries effectively separated high-load periods. Dickinson (1981) and Richards & Holloway (1987) suggested a daily flow which exceeds flow 15% and 20% of the year, respectively, for dividing high and low discharges. Table 2 shows the size of final sample sets taken by ACS for the three sampling frequencies.

2. Calendar-based sampling: For each sample size, an average time interval was calculated. The first sample was chosen randomly between one and the average time interval. The subsequent samples were then selected based on the fixed average time interval. By adding a very small random number (between ± 0.05) to time interval, we avoided taking the same sample sets for different replicates, even with the same first sample.
3. Selection at list time: Sampling simulation using the SALT design was adopted from Thomas (1985). In this study, we fitted a rating curve to the sediment and discharge records from 1973 to 1981 prior to the duration of the study:

$$(Q_s = 24.9 \times Q_w^{1.51}),$$

where Q_s and Q_w are sediment load (tonne day^{-1}) and discharge ($\text{m}^3 \text{ s}^{-1}$), respectively. This rating equation was

Table 1 | Average daily discharges (m^3s^{-1}) exceeded 20% of the time each month, based on a 9-year record before the study period

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
15.0	20.9	38.2	71.8	49.7	22.5	8.1	2.0	3.9	9.0	9.5	13.0

employed to make a preliminary estimate of the total sediment load, Z' , expected during the period to be monitored. It was completed by estimating total sediment load for 1686 comparable days during the 5-year period of 1976–1981. Then Z' was multiplied by a safety factor, W , to obtain a possible upper limit of sediment load in study period, Z^* :

$$Z^* = WZ'. \quad (1)$$

With this procedure, Z^* became sufficiently high to ensure that the real-time estimated sediment load using CSRC does not exceed Z^* . According to the suggestion of Thomas (1985), we chose $W = 10$ in this study because of the high sediment load variability expected in arid and semi-arid areas.

A column of cumulative loads was computed for the study period (obtained from the CSRC). For each simulation, a set of random numbers size n^* was generated between 1 and Z^* and sorted in ascending order. The method used to calculate n^* is described in Thomas (1985). The days on which the random numbers fall within the corresponding cumulative load values were selected as sampling units. Sometimes more than one random number fell within a day. Table 2 shows the size of selected sample sets by SALT.

As shown in Table 2, the average numbers of sample sets of three sampling frequencies (of 50 replicates) are about 120, 230 and 440. These sample sizes are labelled as small, medium and large based on their magnitudes, to prevent repetition using the values. The chosen average

sampling frequencies were approximately twice a month, once and twice a week. Table 2 shows that the number of records is not fixed among replicates selected by ACS and SALT.

Suspended sediment load estimating methods

Total sediment outputs for the study period were estimated with the following sediment estimation methods: (1) the modified Horvitz-Thompson estimator for ACS; (2) Hansen-Hurwitz estimator for SALT; and (3) CSRC, PCF and NPCF for calendar-based sample sets.

1. Modified Horvitz-Thompson estimator (ACS estimator):

Two estimators have been suggested for ACS by Thompson (1992). In this study, only the modified Horvitz-Thompson estimator was used because of its better performance (Salehi 2003). Since ACS is an unequal probability sampling, we need to calculate the probability first. As stated earlier, the initial sample sets were considered simple random samples. For a random initial sample set without replacement, assuming x_k is the number of units in the k th network, the probabilities of intersecting initial samples and networks are calculated from (Thompson 1992):

$$a_k = 1 - \binom{N-x_k}{n} / \binom{N}{n}. \quad (2)$$

The unbiased total sediment load is calculated by the modified Horvitz-Thompson estimator:

$$\hat{Y}_{ACS} = \sum_{k=1}^k \frac{y_k z_k}{a_k} \quad (3)$$

Table 2 | Size of sample sets taken by three sampling designs for three sampling frequencies

Sampling design	Sample size class	Adaptive cluster sampling (ACS)*			Selection at list time (SALT)		
		Small	Medium	Large	Small	Medium	Large
Size of sample set (50 replicates)	Minimum	109	217	427	107	207	417
	Maximum	136	242	455	130	246	464
	Average	117.9	227.6	440.1	118.6	228.3	440.5

*Calendar-based sample sets were chosen with the same sizes as those average numbers obtained for corresponding adaptive sample sets.

where \hat{Y}_{ACS} is estimator of total sediment load for ACS, k is the total number of distinct networks in the population and y_k^* is the total of y values in network k . For every network with one unit, $y_k^* = y$ and z_k is an indicator equal to 1 if any unit of the k th network is in the initial sample and equal to 0 otherwise.

2. Hansen-Hurwitz estimator (SALT estimator): First an estimate of sediment load, x_i , was obtained for discharge of each sampled unit using the available CSRC (explained earlier). The total estimated sediment load, X , was calculated by summing every estimated x_i . As SALT is an unequal probability sampling, we also need to find the probability (p_i) for each sampled unit. Assuming $p_i = x_i/X$, then \hat{Y}_{SALT} (the SALT estimate of total sediment load) can be obtained by the Hansen-Hurwitz estimator (Thomas 1985):

$$\hat{Y}_{SALT} = \frac{1}{n} \sum_{i=1}^N r_i \frac{y_i}{p_i}, \quad (4)$$

where y_i is the observed load during the study and r_i is the number of random values contained in the i th interval. The sample size n is given by

$$n = \sum_{i=1}^N r_i. \quad (5)$$

3. Sediment rating curves: Only calendar-based sample sets were used to estimate total sediment load with CSRC. To develop a CSRC, first, for each sample set, sediment load (dependent) and discharge (independent) variables were transformed to logarithm base 10. Then a linear rating curve was fitted to the data, using a least square technique. Coefficient of determination (R^2) and standard error were used to evaluate each rating curve. This equation was employed to estimate the missing daily sediment loads using daily discharge and then total sediment load for the study period.

Equations (6) and (7) below calculate PCF and NPCF. For calculating corrected estimates of sediment load, obtained correction factors then were multiplied by CSRC estimates:

$$PCF = 10^{1.1513\sigma_c^2}. \quad (6)$$

and

$$NPCF = 1/n \sum_{i=1}^n \exp(\hat{\epsilon}_i), \quad (7)$$

where σ_c^2 is the variance of the residuals from the regression relationship between load and flow (mean square error) in base-10 logarithms and $\hat{\epsilon}_i$ denotes the i th of n regression residuals between load and flow.

Estimation approaches, treatments and their comparison

Each estimation approach comprises two parts: sampling design and estimation method or estimator. As mentioned earlier, for survey sampling approaches such as SALT, each sampling design has a specific estimator. However, sediment rating curves can also be applied to calendar-based, ACS or SALT sample sets. We employed five estimation approaches in this study, as listed in Table 3. To avoid using long designations, a short name was considered for each approach.

Since sampling designs were simulated with three predetermined sizes of sample sets (110, 230 and 440), each estimation approach included three subgroups with respect to sampling frequency that were essentially the treatments. Each treatment comprised 50 replicates, which produced 750 separate estimations overall.

To compare different approaches and treatments, a number of statistics were calculated. Percent error was obtained for each estimated load to indicate over and underestimations. Average percent errors were also calculated to compare treatments to determine if they were

Table 3 | Sediment estimation approaches used in this study

Estimation approach: Sampling design and estimation method	Acronym
Adaptive cluster sampling (ACS) design and ACS estimator	ACS
Selection at list time (SALT) design and SALT estimator	SALT
Calendar-based sampling design and conventional sediment rating curve (CSRC)	CSRC
Calendar-based sampling design and parametric correction factor (PCF)	PCF
Calendar-based sampling design and non-parametric correction factor (NPCF)	NPCF

unbiased. A treatment is considered unbiased if the average percent error approaches zero.

The variance of replicates $V[\hat{Y}]$ and coefficient of variation (CV) were calculated to study the precision of different estimation treatments. The most precise treatment shows a very small amount of variance; the CV illustrates that the replicates estimate considerably close loads.

Finally, to decide which treatment works better, the normalized root mean square deviation $\text{NRMSD}[\hat{Y}]$ was obtained:

$$\text{NRMSD}[\hat{Y}] = \sqrt{V[\hat{Y}] + (\text{Bias}[\hat{Y}])^2/Y}. \quad (8)$$

where $\text{Bias}[\hat{Q}_s]$ is calculated:

$$\text{Bias}[\hat{Y}] = \frac{\sum_{i=1}^{50} \hat{Q}_s}{50} - Q_s, \quad (9)$$

where \hat{Q}_s and Q_s in Equation (9) are the estimated and observed total sediment load, respectively. The smallest $\text{NRMSD}[\hat{Y}]$ denotes the most accurate treatment.

RESULTS

The results, including mean estimated load, standard deviation, CV, percentage of error for maximum, minimum and mean estimates and NRMSD are shown in Table 4. The comparison of ACS and SALT with respect to the mean error shows that ACS always slightly underestimates the mean load by -15% to -4% , while the estimates of SALT are almost unbiased with mean errors ranging from -7% to $+2.2\%$. The two approaches provide a similar percent error for the large sample size. The two survey sampling approaches also produce almost similar results with regard to the range of percent errors. While percent errors of individual estimates for large sample sets of SALT vary from -28 to $+35$, ACS shows narrower ranges. The largest error for SALT exceeds 100% compared to 65% for ACS.

Use of the CSRC, however, drastically underestimates the sediment load by more than 71%. The mean estimations for CSRC treatments are almost constant at about 3 million tonnes, approximately one-quarter of the observed load. Application of the two correction factors

Table 4 | Mean estimated loads for study approaches with some corresponding descriptive statistics

Estimation Approach	Mean estimated load (t)	Standard deviation (t)	Coefficient of variation (%)	% Error for estimations			
				Maximum	Minimum	Mean	NRMSD
Small sample size = 120							
ACS	10,630,179	4,020,785	37.8	65.3	-58.0	-15.3	0.35
SALT	11,574,591	3,015,862	26.1	60.1	-43.4	-7.7	0.25
CSRC	2,993,263	271,232	9.1	-70.9	-80.0	-76.1	0.76
PCF	7,631,879	816,759	10.7	-25.5	-53.6	-39.2	0.40
NPCF	19,642,455	8,158,387	41.5	243.7	-47.1	56.6	0.86
Medium sample size = 230							
ACS	11,117,111	2,825,871	25.4	53.6	-44.0	-11.4	0.25
SALT	12,822,539	3,461,289	27.0	98.1	-35.9	2.2	0.28
CSRC	2,962,293	97,028	3.3	-74.4	-78.6	-76.4	0.76
PCF	7,509,623	604,973	8.1	-32.7	-50.4	-40.1	0.40
NPCF	21,508,511	7,029,676	32.7	173.0	-21.2	71.5	0.91
Large sample size = 440							
ACS	11,952,291	1,307,695	10.9	13.2	-27.0	-4.7	0.11
SALT	12,035,096	1,605,437	13.3	35.2	-28.1	-4.1	0.13
CSRC	2,982,137	60,568	2.0	-75.4	-77.3	-76.2	0.76
PCF	7,530,822	283,010	3.8	-35.2	-44.7	-40.0	0.40
NPCF	20,473,038	2,816,191	13.8	105.2	25.0	63.2	0.67

yields contradictory results. The average biases for PCF treatments are about -40% , demonstrating about 30% improvement compared to CSRC. However, the average biases for NPCF treatments show large overestimations of more than 56% . Despite these overestimations, the percent errors of NPCF individual estimates have too many fluctuations, sometimes even becoming negative. PCF does not produce any estimates greater than the observed load, indicating a systematic underestimation as for CSRC. It can be seen from the data in Table 4 that the best individual estimate for PCF has a -25% error.

From Table 4, strong evidence of positive correlation between standard deviation and average estimated load can be understood. That is, estimation treatments with large mean estimated loads report much more variation and a wide range of estimations. Standard deviations and CVs of both survey sampling approaches are less than NPCF only, but greater than the two other sediment rating curve approaches. ACS shows a fairly reasonable standard deviation with the large sample sets. Its CVs, however, vary from 11% for large sample size to 39% for smaller sample sizes. The calculated CVs for the medium and large sample sizes treatments of ACS are smaller than the analogous values of SALT. For the small size, however, ACS produced a larger CV.

Comparing ACS and SALT using NRMSD yields similar results as the CV. For the medium and large sample sizes, ACS shows smaller values (0.25 and 0.11) compared to the SALT. The NRMSD of SALT for the small to large sample sets ranges from 0.28 to 0.13 , respectively.

Table 4 indicates that CVs of three CSRC treatments range from $2-9\%$. These values are very low compared to other treatments. Despite precise results, the amounts of

NRMSDs range from 0.69 to 0.76 demonstrating the second lowest accuracy. Using normal distribution rules, it is statistically predicted that at least 99.9% of all possible estimates of CSRC would be less than one-third of the observed load, revealing that it cannot provide any estimates in Gorgan-Rood with an error less than -69% .

Regardless of the large improvement of PCF on percent error compared to the CSRC, its variance is a little higher than the latter. The highest CV of PCF estimates is about 11% , which occurs for the small sample set. It gradually drops to 4% for large sample sizes. This correction factor also substantially improves the accuracy (NRMSD = 0.40) compared to the analogous value for CSRC (0.76).

The CV and NRMSD of NPCF are almost greater than other study estimation approaches. The highest estimate is more than twice that of the observed load. However, the smallest estimate shows 21% underestimation. Overall, the NPCF did not improve the CSRC estimates in Gorgan-Rood, since it has produced almost the same amount of bias (only positive).

Table 4 also shows the relationship between sample size and calculated statistics. For example, from the data in this table, it is apparent that sample size does not affect NRMSD in the case of CSRC and PCF. However, for ACS, there is a significant negative correlation between these two parameters i.e. NRMSD decreases with increasing sample size. SALT and NPCF, in comparison, show the lowest accuracies for the medium sample size.

The better performance of survey sampling compared to the sediment rating curve approaches can easily be seen in the results presented in Table 5. While none of the CSRC estimations demonstrate a percent error less than 50% , all estimates of PCF yield errors of between $30-50\%$.

Table 5 | Percent of estimates placed in the range of different percent errors

Estimation approach	Sample size	CSRC			PCF			NPCF			SALT			ACS		
		120	230	440	120	230	440	120	230	440	120	230	440	120	230	440
Percent error	< 50	0	0	0	96	98	100	54	44	34	98	94	100	88	100	100
	< 40	0	0	0	58	50	48	52	38	26	94	90	100	62	94	100
	< 30	0	0	0	4	0	0	38	22	4	70	82	98	40	64	100
	< 20	0	0	0	0	0	0	32	14	0	46	60	86	22	40	92
	< 10	0	0	0	0	0	0	16	8	0	24	28	46	8	14	54
	< 5	0	0	0	0	0	0	8	2	0	8	8	16	4	2	30

In comparison, NPCF shows a relatively uniform distribution of estimation errors in different categories, but 34–54% of estimates are classified in >50% error category (Table 5). Surprisingly in the case of NPCF, the increasing sample size has an inverse effect on the accuracy. All estimations made by ACS and SALT with the large sample sets fall within the class of less than 30% and 40%. This table shows the superiority of SALT for the small sample sets, but ACS performed better for large sample sets.

DISCUSSION

The results showed that the mean estimates obtained by ACS and SALT were comparatively close to the observed value; the sediment rating curve approaches produced substantial under or overestimations, however. These findings are expected, because both designs are inherently unbiased. Nevertheless, ACS showed a tendency to underestimate the sediment load. The underestimation was most probably a result of the missing important sampling units taken during the first halves of the storm hydrographs by the forward neighbourhood relation (compared to the symmetric procedure). It is noted that in basins with a clockwise hysteresis loop, a higher rate of sediment loads are transported during the rising limbs and the peaks (Gomi *et al.* 2005), which coincide with the first halves of storm hydrographs.

As indicated in Table 4, the underestimation with the adopted ACS is greater when the sample size became smaller. The reason for this underestimation is explained in the example shown in Figure 2. When the sampling interval is weekly, the high flow part (greater than the discharge threshold) of storm hydrograph is intersected at days 9, 16 and 23. Additional samples with forward neighbourhood relation will therefore be taken from day 9 to 25. For the biweekly interval, however, the high flow part is only intersected at day 16, which leads to an additional 9 days of samples from day 16 to 25. Using the sediment graph in Figure 2, it is apparent that the majority of the load occurred before day 16 and a very low amount at the end part of the storm hydrograph. In other words, the chance of taking additional samples from the first parts of storm hydrographs increases when the sample size increases. Choosing more

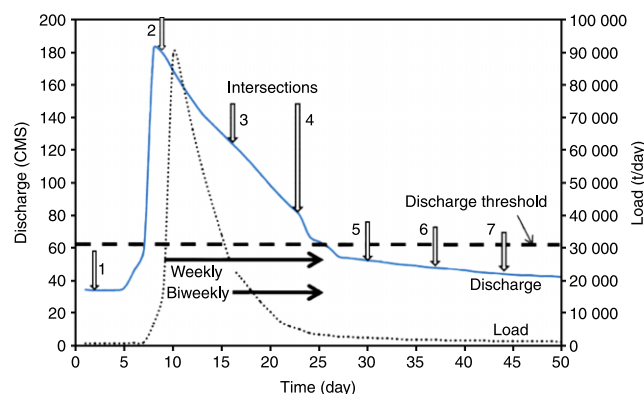


Figure 2 | The effect of initial sampling frequencies (weekly and biweekly) on the number of intersections during storm hydrograph and associated additional samples.

additional samples during the first part of storm hydrographs will eventually lead to a smaller amount of underestimation. In Gorgan-Rood, the average underestimation was negligible when about a quarter of the population had been sampled.

The results also indicated a very high CV for the ACS with the small sample sets, which can be explained by considering the nature of flow and sediment load duration curves in the Gorgan-Rood. Figure 3 reveals that 45% of the load has been carried only during 1% of time. By neglecting the sample from this period, ACS would almost be changed to a calendar-based sampling with relatively high CV. The high CV value for small sample size is in agreement with Thompson (1992) who showed by example that the variance of estimations obtained by the modified Horvitz-Thompson

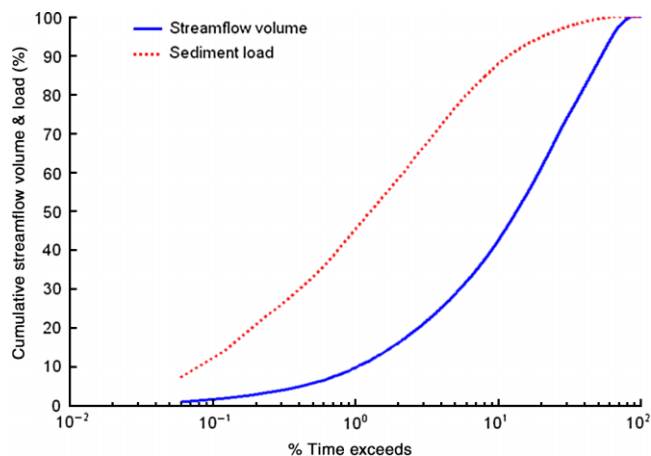


Figure 3 | Cumulative percentage of stream flow and sediment load duration curves in Gorgan-Rood River, Iran.

estimator exceeds two-thirds of the analogous value for simple random sampling when the size of sample sets are less than 10% of population. However, this ratio is less than 1/15 for very large sample sets (equal to 50% of the population size).

Despite the good results of SALT in this study, the outcomes are not as encouraging as the two previous studies carried out by Thomas (1985) and Thomas & Lewis (1993). There are several possible explanations for contradictory findings. First, in choosing a representative sample set, SALT needs an appropriate auxiliary variable highly correlated with concentration (Thomas & Lewis 1993). However, the correlation between concentration and the only available auxiliary continuous record (discharge) was very poor (less than 10%). It is difficult to model sediment load estimation in rivers such as Gorgan-Rood which have sophisticated variations in daily load level, as explained before. Good results in previous studies can be explained by their analyzed datasets. Thomas (1985) used sediment records synthesized using CSRC for two years (not true records), and the dataset of Thomas & Lewis (1993) involved only five storm events.

CONCLUSION

ACS is able to produce relatively accurate sediment load estimation as well as SALT but much better than rating curve approaches in the Gorgan-Rood when the sample size is large enough. The most significant advantage of SALT and ACS designs compared to the calendar-based method is that the two former designs select more samples during high flows. Another important advantage of survey sampling approaches compared to the sediment rating curves is that the former gives a valid estimate of variance of the estimated total loads, therefore provides an estimate of the error present in estimating total suspended sediment load. Although both ACS and SALT take more samples during high flows, the use of SALT is restricted in most rivers where the gauging site is not equipped with an automatic sediment sampler.

Despite the better accuracy of ACS compared to other estimation approaches, this study showed that it considerably underestimated the sediment load in Gorgan-Rood

when the sample size is less than 15% of the population. In other words, the adopted ACS with forward neighbourhood relation is inherently biased because of several unsampled units during the rising limbs of storm hydrographs. As discussed earlier, this bias can be largely eliminated by taking large samples. Developing an unbiased estimator by considering the nature of storm hydrographs and sediment graphs may be a necessary next step for further studies.

Testing of this method in a variety of conditions is suggested. It is expected that ACS will yield more accurate load estimation for rivers with fairly long storm hydrographs. In such rivers, there would be a greater chance to take samples from all events, which leads to relatively unbiased estimates with smaller variations.

ACKNOWLEDGEMENTS

The authors thank two unknown reviewers for their useful comments and suggestions in the earlier version of this article.

REFERENCES

- Acharya, B., Bhattarai, G., de Gier, A. & Stein, A. 2000 Systematic adaptive cluster sampling for the assessment of rare tree species in Nepal. *For. Ecol. Manage.* **137**(1–3), 65–73.
- Asselman, N. E. M. 2000 Fitting and interpretation of sediment rating curves. *J. Hydrol.* **234**(3–4), 228–248.
- Bradbury, A. 2000 Stock assessment and management of red sea urchins (*Strongylocentrotus franciscanus*) in Washington. *J. Shellfish Res.* **19**, 618–619.
- Chaudhuri, A., Bose, M. & Dihidar, K. 2005 Sample-size-restrictive adaptive sampling: an application in estimating localized elements. *J. Stat. Plan. Inference* **134**(1), 254–267.
- Cohn, T. A. 1995 Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers. *Rev. Geophys.* **33**, 1117–1124.
- Cohn, T. A. 2005 Estimating contaminant loads in rivers: An application of adjusted maximum likelihood to type 1 censored data. *Water Resour. Res.* **41**(7), 1–13.
- Cohn, T. A., Delong, L., Gilroy, E., Hirsch, R. & Wells, D. 1989 Estimating constituent loads. *Water Resour. Res.* **25**(5), 937–942.
- Crowder, D. W., Demissie, M. & Markus, M. 2007 The accuracy of sediment loads when log-transformation produces nonlinear sediment load-discharge relationships. *J. Hydrol.* **336**(3–4), 250–268.

- Degens, B. P. & 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.
- Demissie, M., Xia, R., Keefer, L. & Bhowmik, N. 2004 The sediment budget of the Illinois river, Champaign, Illinois: Illinois State Water Survey, Contract Report 2004–13, 51.
- Dickinson, W. T. 1981 Accuracy and precision of suspended sediment loads. *Proceedings of Erosion and Sediment Transport Measurements Florence, Italy*. IAHS, Wallingford, Oxfordshire, pp. 195–202, IAHS Publication 133.
- Duan, N. 1985 Smearing estimate: a nonparametric retransformation method. *J. Am. Stat. Assoc.* **78**, 605–610.
- Ferguson, R. 1986 River loads underestimated by rating curves. *Water Resour. Res.* **22**(1), 74–76.
- Gomi, T., Moore, R. D. & Hassan, M. W. 2005 Suspended sediment dynamics of small forest streams of the Pacific Northwest. *J. Am. Water Resour. Assoc.* **41**, 877–898.
- Harris, J. 2008 Evaluation of adaptive cluster sampling for remotely-operated underwater vehicle surveys of inshore rockfish (*Sebastes* spp.), *School of Resource and Environmental Management*, Simon Fraser University, p. 73.
- Horowitz, A. J. 2003 An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrol. Process.* **17**, 3387–3409.
- Juang, K. W., Lee, D. Y. & Teng, Y. L. 2005 Adaptive sampling based on the cumulative distribution function of order statistics to delineate heavy-metal contaminated soils using kriging. *Environ. Pollut.* **138**(2), 268–277.
- Koch, R. & Smillie, G. 1986 Comment on River loads underestimated by rating curves by R.I. Ferguson. *Water Resour. Res.* **22**(13), 2121–2122.
- Lai, F. S., Lee, M. J. & Rizal, S. M. 1995 Changes in sediment discharge resulting from commercial logging in the Sungai Lawing basin, Selangor, Malaysia. In *Proceedings of Symposium on Effect of Scale on Interpretation and Management of Sediment and Water Quality*, Boulder, USA. IAHS, Wallingford, Oxfordshire, pp. 55–62, IAHS Publication 226.
- Lo, N. C., Griffith, H. D. & Hunter, J. R. 1997 Using a restricted adaptive cluster sampling to estimate Pacific hake larval abundance. *California Cooperative Oceanic Fisheries Investigations*. Report 37, pp. 160–174.
- Markus, M. & Demissie, M. 2006 Predictability of annual sediment loads based on flood events. *J. Hydrol. Eng.* **11**(4), 354–361.
- Miller, D. M. 1984 Reducing transformation bias in curve fitting. *Am. Stat.* **38**(2), 124–126.
- Richards, R. P. & Holloway, J. 1987 Monte Carlo studies of sampling strategies for estimating tributary loads. *Water Resour. Res.* **23**(10), 1939–1948.
- Sadeghi, S. H. R., Mizuyama, T., Miyata, S., Gomi, T., Kosugi, K., Fukushima, T., Mizugaki, S. & Onda, Y. 2008 Development, evaluation and interpretation of sediment rating curves for a Japanese small mountainous reforested watershed. *Geoderma* **144**(1–2), 198–211.
- Salehi, M. M. 2005 Comparison between Hansen-Hurwitz and Horvitz-Thompson estimators for adaptive cluster sampling. *Environ. Ecol. Stat.* **10**, 115–127.
- Salehi, M. & Smith, D. R. 2005 Two-stage sequential sampling: a neighborhood-free adaptive sampling procedure. *J. Agric. Biol. Environ. Stat.* **10**(1), 84–103.
- Schleppi, P., Waldner, P. A. & Fritschi, B. 2006 Accuracy and precision of different sampling strategies and flux integration methods for runoff water: comparisons based on measurements of the electrical conductivity. *Hydrol. Process.* **20**(2), 395–410.
- Schmidt, K. H. & Morche, D. 2006 Sediment output and effective discharge in two small high mountain catchments in the Bavarian Alps Germany. *Geomorphology* **80**(1–2), 131–145.
- Talvitie, M., Leino, O. & Holopainen, M. 2006 Inventory of sparse forest populations using adaptive cluster sampling. *Silva Fennica* **40**(1), 101–108.
- Thomas, R. 1985 Estimating total suspended sediment yield with probability sampling. *Water Resour. Res.* **21**(9), 1381–1388.
- Thomas, R. B. 1988 Monitoring baseline suspended sediment in forested basins: the effects of sampling on suspended sediment rating curves. *Hydrol. Sci.* **33**(5), 499–514.
- Thomas, R. B. & Lewis, J. 1995 A comparison of selection at list time and time-stratified sampling for estimating suspended sediment loads. *Water Resour. Res.* **29**(4), 1247–1256.
- Thomas, R. B. & Lewis, J. 1995 An evaluation of flow-stratified sampling for estimating suspended sediment loads. *J. Hydrol.* **170**(1–4), 27–45.
- Thompson, S. K. 1990 Adaptive cluster sampling. *J. Am. Stat. Assoc.* **85**, 1050–1059.
- Thompson, S. K. 1992 *Sampling*. Wiley, New York, p. 343.
- Thompson, S. K. 1996 Adaptive cluster sampling based on order statistics. *Environmetrics* **7**(2), 123–133.
- Vanoni, V. A. 1975 Sedimentation Engineering. *ASCE Manuals and Reports on Engineering Practices No. 54*, 745.
- Walling, D. E. 1994 Measuring sediment yield from river basins. In: Lal, R. (ed.) *Soil Erosion Research Methods*. Soil and Water Conservation Society Publication, Ankeny, Iowa, pp. 39–80.
- Walling, D. E. & Webb, B. W. 1988 The reliability of rating curve estimates of suspended sediment yield: some further comments. *Sediment Budgets, Porto Alegre, Brazil*. IAHS, Wallingford, Oxfordshire, pp. 337–350, IAHS Publication 174.

First received 29 November 2008; accepted in revised form 15 July 2009. Available online December 2009