JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 45, No. 5 AMERICAN WATER RESOURCES ASSOCIATION October 2009

COMPARISON OF THREE PEBBLE COUNT PROTOCOLS (EMAP, PIBO, AND SFT) IN TWO MOUNTAIN GRAVEL-BED STREAMS[1]

Kristin Bunte, Steven R. Abt, John P. Potyondy, and Kurt W. Swingle[2]

ABSTRACT: Although the term "pebble count" is in widespread use, there is no standardized methodology used for the field application of this procedure. Each pebble count analysis is the product of several methodological choices, any of which are capable of influencing the final result. Because there are virtually countless variations on pebble count protocols, the question of how their results differ when applied to the same study reach is becoming increasingly important. This study compared three pebble count protocols: the reach-averaged Environmental Monitoring and Assessment Program (EMAP) protocol named after the EMAP developed by the Environmental Protection Agency, the habitat-unit specific U.S. Forest Service's PACFISH/INFISH Biological Opinion (PIBO) Effectiveness Monitoring Program protocol, and a data-intensive method developed by the authors named Sampling Frame and Template (SFT). When applied to the same study reaches, particle-size distributions varied among the three pebble count protocols because of differences in sample locations within a stream reach and along a transect, in particle selection, and particle-size determination. The EMAP protocol yielded considerably finer, and the PIBO protocol considerably coarser distributions than the SFT protocol in the pool-riffle study streams, suggesting that the data cannot be used interchangeably. Approximately half of the difference was due to sampling at different areas within the study reach (i.e., wetted width, riffles, and bankfull width) and at different locations within a transect. The other half was attributed to using different methods for particle selection from the bed, particle-size determination, and the use of wide, nonstandard size classes. Most of the differences in sampling outcomes could be eliminated by using simple field tools, by collecting a larger sample size, and by systematically sampling the entire bankfull channel and all geomorphic units within the reach.

(KEY TERMS: sediment; monitoring; fluvial processes; sampling technique; habitat evaluation; percent fines.)

Bunte, Kristin, Steven R. Abt, John P. Potyondy, and Kurt W. Swingle, 2009. Comparison of Three Pebble Count Protocols (EMAP, PIBO, and SFT) in Two Mountain Gravel-Bed Streams. *Journal of the American Water Resources Association* (JAWRA) 45(5):1209-1227. DOI: 10.1111/j.1752-1688.2009.00355.x

INTRODUCTION

Pebble counts are one of the most frequently used field methods to assess the size distribution of bed surface sediment in gravel-bed and cobble-bed streams. However, pebble counts exist in countless variations; more than 30 different procedures are used by the USDA Forest Service alone (Hilaire Bojonell, Hydrologist, USDA Forest Service, National Resources Information System, Corvallis, Oregon, personal communication, June 2008). In fact, nearly

[1]Paper No. JAWRA-08-0164-P of the *Journal of the American Water Resources Association* (JAWRA). Received August 29, 2008; accepted May 20, 2009. **Discussions are open until six months from print publication**.

[2]Respectively, Research Scientist (Bunte) and Professor (Abt), Engineering Research Center, Department of Civil Engineering, Colorado State University, 1320 Campus Delivery, Fort Collins, Colorado 80523; Program Manager/Hydrologist (Potyondy), Stream Systems Technology Center, Watershed, Fish, Wildlife, Air, & Rare Plants Staff, USDA Forest Service, Washington, D.C.; Environmental Scientist (Swingle), Boulder, Colorado (E-Mail/Bunte: kbunte@engr.colostate.edu).

every study that involves a pebble count uses its own adaptation to address the study aim, stream conditions, operator skill, and project budget.

To elucidate the variability among pebble counts, this study considers pebble counts to be composed of nine components: (1) the length of the sampling reach, (2) the sampling scheme (spatially integrated over the reach or segregated into different units), (3) the sampling pattern (transects, small-scale grids, or random), (4) the sampled portion of stream width, (5) spacing and number of particles collected per transect, (6) the sample size per reach, (7) identifying the particles to be extracted, (8) measuring particle size, and (9) particle-size analyses. For each component, there are several methodological options (Figure 1). For example, the component "particle-size measurements" includes the options of measuring the particle *b*-axis with a ruler, visually estimating the particle-*b*-axis size class, or passing a particle through a template (gravelometer). By selecting a specific option for each component based on study needs, a user takes an individualized pathway through the various pebble count components (Figure 1). Each pathway describes a different pebble count protocol.

Several studies have pointed out that doing pebble counts in a variety of different ways can lead to different results (Diplas and Lohani, 1997; Kondolf, 1997a,b; Whitacre *et al.*, 2007). More specifically, each methodological option selected for a pebble count component affects the sampling outcome, rendering a pebble count's sampling result a blend of the combined effects of all its methodological options. Pebble count users need to be aware of this variability. What adds to the problem is a lack of specific nomenclature to distinguish among different procedures. Typically, pebble count protocols are simply referred to as a "modified" Wolman pebble count after the originator of pebble counting (Wolman, 1954). Our investigation highlights the need to specifically describe the nature of any modifications introduced by a study.

The effects of some pebble count details have been addressed in depth, particularly the effects of sample size (e.g., Church *et al.*, 1987; Fripp and Diplas, 1993; Bevenger and King, 1995; Rice and Church, 1996; Petrie and Diplas, 2000). Hey and Thorne (1983) evaluated the benefits of using a template. Marcus *et al.* (1995) and Wohl *et al.* (1996) examined operator errors in measuring particle sizes and selecting particles from the bed, while Bunte and Abt (2001a,b) developed a sampling frame (SF) with grid intersections to minimize those errors. Recent studies have examined the combined effects of several components, such as natural variability among and within streams, operator variability, and sample sizes necessary for specific levels of accuracy. Roper *et al.* (2002) assessed the number of samples necessary to detect change. Larsen *et al.* (2004) analyzed how many years and stream sites need to be sampled within a region to detect an annual 1-2% change (i.e., a change of 15-30% over 10-20 years of monitoring). Olsen *et al.* (2005) assessed the proportion of variability attributable to stream heterogeneity *vs.* operator variability. Several studies evaluated the repeatability of sampling results for a specific protocol, e.g., Archer *et al.* (2004) for the Pacific Anadromous Fish Strategy (PACFISH) and the Inland Fish Strategy (INFISH) Biological Opinion (PIBO) protocol and Faustini and Kaufmann (2007) for the Environmental Monitoring and Assessment Program (EMAP) procedure. Whitacre *et al.* (2007) compared sampling results among different pebble count protocols (including EMAP and PIBO that are compared in this study) averaged over six different streams. While successful in pointing out differences among protocol results, these analyses do not show how selection of a specific methodological option contributes to differences in sampling results, and how results may be affected by stream type.

The overarching aim of the study presented here is to raise awareness of the fact that selection of different methodological options can have notable consequences on the results. Special emphasis is given to the effects of excluding morphological units from the sampling area, particle selection, and particle-size estimates. To demonstrate some of these cause-and-effect connections, the study compares three pebble count protocols applied to two streams. The study shows the overall effect that a specific sampling protocol has on the sampling result and demonstrates the effects of individual methodological options, specifically sampling location.

METHODS

One of the pebble count protocols selected for this study is the EMAP protocol named after the EMAP developed by the Environmental Protection Agency for physical habitat characterization (Kaufmann and Robison, 1998; Kaufmann *et al.*, 1999; USEPA, 2004; Peck *et al.*, 2006). The procedure includes a surface pebble count to monitor changes in the conditions of aquatic habitat and in the amount of silt and sand (<2 mm) supplied to a stream. The EMAP protocol was developed for rapid characterization of a reach and samples the wetted width of the channel. The protocol is applied in large national stream studies and uses many operators. The second protocol selected is the PIBO pebble count named after the PIBO Effectiveness Monitoring Program initiated in 1998 for long-term monitoring of aquatic and riparian resources in streams on Forest Service and Bureau of

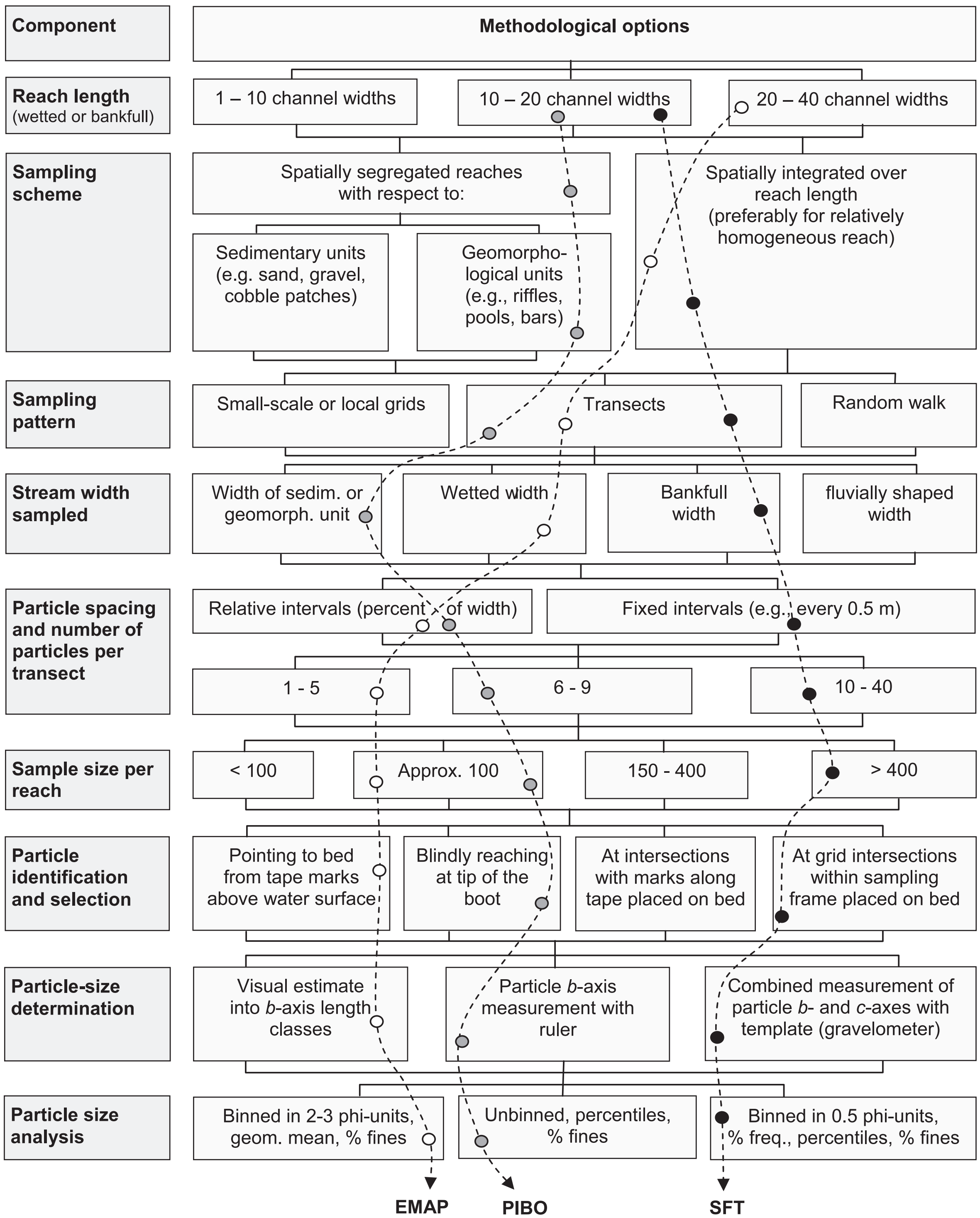

FIGURE 1. Components of Pebble Counts, Their Methodological Options, and Pathways Followed by Three Example Pebble Count Protocols: EMAP, PIBO, and SFT. The majority of methodological options refer to differences in sampling location.

Land Management lands within the Upper Columbia River Basin (Kershner *et al.*, 2004; Henderson *et al.*, 2005). Pebble counts as part of the PIBO program are used to monitor the median (D_{50}) surface sediment size and the percent fines <6 mm of the surface sediment on riffles. The EMAP and PIBO pebble count protocols were selected for study because they are among the most widely used, with applications that involve hundreds of streams. These protocols were also selected because they represent major differences in sampling approaches [EMAP being reach-averaged and PIBO using a habitat-specific (riffle only) approach], and because they differ widely in their methodological details (Figure 1) many of which refer to various aspects of sampling locations. The EMAP and PIBO pebble count protocols were not developed for use at individual sites but rather to measure long-term change for a large number of streams within a region. However, the simplicity, speed, widespread use and reference in the literature make these protocols appealing for application to individual sites. The intent of this study is not to make an evaluation of whether the EMAP and PIBO protocols achieve their stated aims for long-term regional studies. This study compares sampling results from three pebble count protocols at two sites and illustrates some of the consequences of methodological choices associated with simple and rapid assessments.

The third protocol, named Sampling Frame and Template (SFT) (a SF and a template) for its major field tools is a rigorous, data-intensive approach and was developed by the authors to minimize operator error and sampling bias in field studies. The key features are: (1) identifying particles under intersections of thin elastic bands within the SF placed directly on the bed, (2) measuring particle sizes with a 0.5-phi template to eliminate operator error and achieve compatibility with sieve data, and (3) covering the entire reach in a systematic grid pattern the dimensions of which are determined by the study aim. This protocol, though time and data intensive, is thought to provide an accurate account of the reach-averaged particle-size distribution, while its detailed spatial record of particle sizes facilitate postsampling segregation by sampling location. It is intended to represent “best technical practice.”

Environmental Monitoring and Assessment Program Protocol

The EMAP protocol was developed for streams that are wadeable during seasonal low-flow periods, typically first to third order streams mapped as blue lines on 1:100,000 scale USGS topographic maps. The sampled reach length is 40 times the average wetted stream width. Five particles are collected along each of 21 transects spaced evenly at intervals of two wetted widths, yielding a sample size of 105 particles per reach. Particles are collected at 0, 25, 50, 75, and 100% of the wetted width; i.e., 40% of all counts fall onto the waterline which at the time of sampling is typically at a low to moderate flow stage. To identify particles for sampling, an operator spans a tape across the stream, sets a pointed meter stick onto the stream bottom at the respective stream width increments, and picks up the particle on which the stick rests. Particle *b*-axis lengths are visually estimated and tallied into seven size classes (Table 1) that span 2 phi units for boulders, cobbles, and coarse gravel, and 3 phi units for fine gravel (Faustini and Kaufmann, 2007). The pebble count protocol is part of a 13 data-element field assessment program that has been specifically designed to be completed by a field crew in half a day.

PACFISH/INFISH Biological Opinion Protocol

The PIBO program was developed for response reaches in unconstrained valley bottoms, with gradients of less than 2% (sometimes 3%), and bankfull widths of 1-15 m. The length of the sampling reach is 20 times the bankfull width (Henderson *et al.*, 2004). Sampling is spatially segregated and performed on four consecutive riffles or runs, each of which must be longer than 0.5 times the average bankfull stream width. A minimum of 25 particles is collected per riffle to yield at least 100 particles per reach. Each riffle is sampled along four transects perpendicular to the banks at 20, 40, 60, and 80% of the riffle length. At each transect, at least six to seven particles are selected at approximately even intervals across the bankfull stream width, starting at one step in from

TABLE 1. Size Classes Used for Particle-Size Categorization in the EMAP Protocol.

Size Class	Size Range (mm)	Size Range (phi)	Description
Large boulders	>1,000[1] to 4,000[1]	−9.97 to −11.97	Yard/meter stick to car size
Small boulders	>250[1] to 1,000[1]	−7.97 to −9.97	Basketball to yard/meter stick size
Cobbles	>64 to 250[1]	−6 to −7.97	Tennis ball to basketball size
Coarse gravel	>16 to 64	−4 to −6	Marble to tennis ball size
Fine gravel	> 2 to 16	−1 to −4	Ladybug to marble size
Sand	>0.06[1] to 2	4.06 to −1	Smaller than ladybug size, but visible as particles – gritty between fingers
Fines	0.001[1,2] to 0.06[1]	9.97 to 4.06	Silt, clay, muck – not gritty between fingers

Notes: EMAP, Environmental Monitoring and Assessment Program.
Combined from USEPA (2004) and Faustini and Kaufmann (2007). Categories such as wood and hardpan are not considered.
[1]Size class boundaries are slightly different from the Wentworth scale (Wentworth, 1922).
[2]Faustini and Kaufmann (2007) set a lower size boundary of 0.001 mm for the finest size class.

the bankfull width. Transect portions that extend into a pool are not sampled. Following instructions by Wolman (1954) and Leopold (1970), the operator paces a streambed area with a relatively homogeneous particle-size distribution, such as riffle, reaches down to the tip (or toe) of the boot and, with eyes averted, selects the first particle touched with a pointed finger. The particle *b*-axis length is measured to the nearest mm using a ruler. Similar to EMAP, PIBO pebble counts are part of a suite of field assessments that are designed to be accomplished in a one-day field visit.

The protocol described above was modified by the PIBO staff in 2004 (Archer *et al.*, 2006) to now collect five particles along 21 evenly spaced transects over at least the average bankfull width (Heitke *et al.*, 2007). Because this change is only reflected in unpublished project reports, the habitat-unit specific PIBO protocol was used as originally published (Henderson *et al.*, 2004) since it still represents the many monitoring groups that collect pebbles from specific habitat units (Johnson *et al.*, 2001). Consequently, all results obtained from the PIBO protocol in this study refer to the protocol version that exclusively sampled riffles.

Sampling Frame and Template Procedure

The SFT protocol was developed for coarse gravel-bed and cobble-bed streams, typically less than about 20 m wide. A SF is used to identify particles to be picked up from the bed (Bunte and Abt, 2001a,b) and a template (or gravelometer) (T) with square-hole openings that progress in 0.5 phi units is used for measuring particle size. The SF is a 0.6 by 0.6 m aluminum frame across which thin elastic bands are spanned to form a grid. Grid size is adjustable and may range from 0.05 to 0.5 m. The frame is placed on the bed at preset intervals along a tape at transects perpendicular to the banks and stretched from bankfull to bankfull across the stream. The operator selects the particle under the grid intersection (Bunte and Abt, 2001a,b), picks it up and measures its size using a template. The SFT protocol systematically samples more than 400 particles (as recommended by Rice and Church, 1996) over the bankfull width of the reach in a large-scale grid pattern, but is flexible in reach length, number and spacing of transects, and the number of particles sampled per transect and per placement of the SF. Typically, when used in coarse gravel and cobble-bed streams, elastic bands are spanned to yield four grid intersections spaced 0.3 m apart, and particles are collected under each of the four grid intersections. The 0.3 m spacing is chosen to avoid serial correlation that results from double counting particles or from sampling multiple particles within a cluster (Church *et al.*, 1987). The frame is placed at increments of 0.6-2 m along a transect, with either a side of the frame parallel to the tape or a corner pointing to it. The aim is to collect 10-40 particles along each of the 10-40 evenly spaced transects. Particle sizes are recorded in sequential order for each transect, noting in the field record the locations on the measuring tape of left and right bankfull as well as the waterlines along banks and mid-channel bars. Also documented are the geomorphological unit that a transect crosses, the transect number, and the spacing to adjacent transects. Such records retain spatial information of the sample locations and facilitate postsampling spatial segregation by units of similar sediment sizes or size distributions, geomorphological units, and whether a particle was collected within the wetted width, outside of it, or near the banks.

The SFT method was designed to reduce operator variability and bias resulting from particle selection and particle-size measurements, and a test of the SFT method in a separate study showed that the two operators (same as in this study) obtained virtually identical results (Bunte and Abt, 2001b). Particle identification under the grid points of elastic bands that touch the bed is unambiguous in low and calm flows. In swift (but wadeable) flows, or when the particle under a grid intersection cannot be seen, the operator places a finger at the grid intersection and points vertically down until a particle is touched. Advantages over sampling at the tip of the boot are that the particle to be selected is not determined by where an operator deems it safe to place a foot, and that an operator using the SF can stand or crouch more comfortably when pointing to a bed particle than when pointing downward to a particle at the tip of the boot. Other advantages of not sampling where the foot (or a pointing stick) is placed become evident when the SF is slid underneath woody debris, under overhanging bushes or undercut banks and when particles need to be selected by touch downward along the grid intersection in faster and deeper flow. Similarly, use of a 0.5-phi template (or a 0.25-phi template in well-sorted beds) effectively eliminates operator error in particle-size measurements. Templates are easy to use, readily available, and eliminate errors associated with incorrectly identifying the particle *b*-axis (Potyondy and Bunte, 2002). Using a template requires no more time than a ruler measurement. For the relatively few particles that are too embedded or too large to lift, operators measure the *b*-axes and *c*-axes to estimate the template hole size that will retain (or pass) the particle.

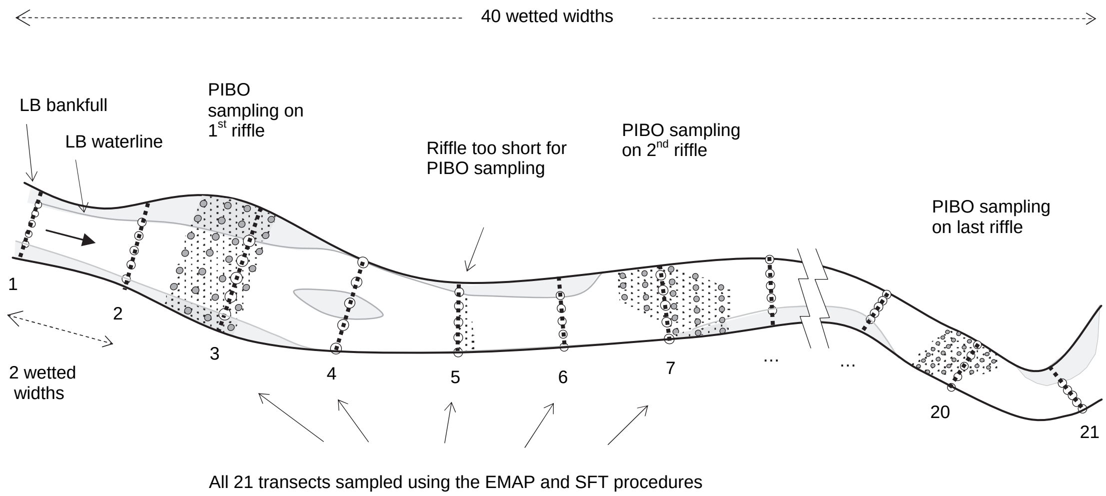

FIGURE 2. Schematic Presentation of a Stream Reach in Its Bankfull Width Showing Sampling Locations for the Three Procedures: EMAP (open circles), PIBO (gray circles), and SFT (dotted lines). SFT locations coincide with the EMAP transects, but extend over the bankfull width. Riffles are indicated as dotted areas, and exposed bars in light gray shading.

Field Study to Compare the Three Pebble Count Protocols

The three protocols differ in almost every pebble count component (see Figure 1), thus slight adjustments were necessary to better compare study results. A common reach length had to be selected for all protocols, and sampling locations had to be as similar as possible without compromising the characteristics of any of the procedures. In accordance with the EMAP protocol, reach length for the study was taken as 40 wetted widths, and all qualifying riffles (6 and 7 in the two study streams) within the reach were sampled using the PIBO protocol (Figure 2). The locations of the SFT transects coincided with the EMAP transects but extended over the bankfull width. In this study, the SF had two intersections spaced 0.5 m apart (Figure 3), and the frame was moved in 1-m increments such that two particles were collected from every 0.5-m increment of stream width, yielding a total of 19-40 particles per transect depending on the bankfull width. The two operators took turns sampling and recording with the SFT method (the close similarity of their sampling results had been established) (Bunte and Abt, 2001b). All PIBO and almost all of the EMAP samples were collected by one operator only. Operator field time varied among the three sampling protocols. Excluding reach reconnaissance and surveying, SFT took 6-8 h and yielded 461 and 598 particles per site; EMAP took 3-4 h for 105 and 120 particles; while sampling the six and seven PIBO riffles took 1.5-2 h for 170 and 201 particles. The time that operators spent in the SFT and EMAP protocols for wading between transects and setting up the tape was identical and amounted to about 2-3 h in the 320-m long study reaches. Thus, the SFT and PIBO protocols were faster per particle than the EMAP protocol.

FIGURE 3. Sampling Frame as Used in This Study: One Frame Corner Pointed Toward the Tape, and Samples Were Collected at the Left and Right Grid Intersections Spaced 0.5 m Apart. The operator is passing a retrieved particle through a template hole. A thin veneer of mud covered particles in areas of tranquil flow at Willow Creek.

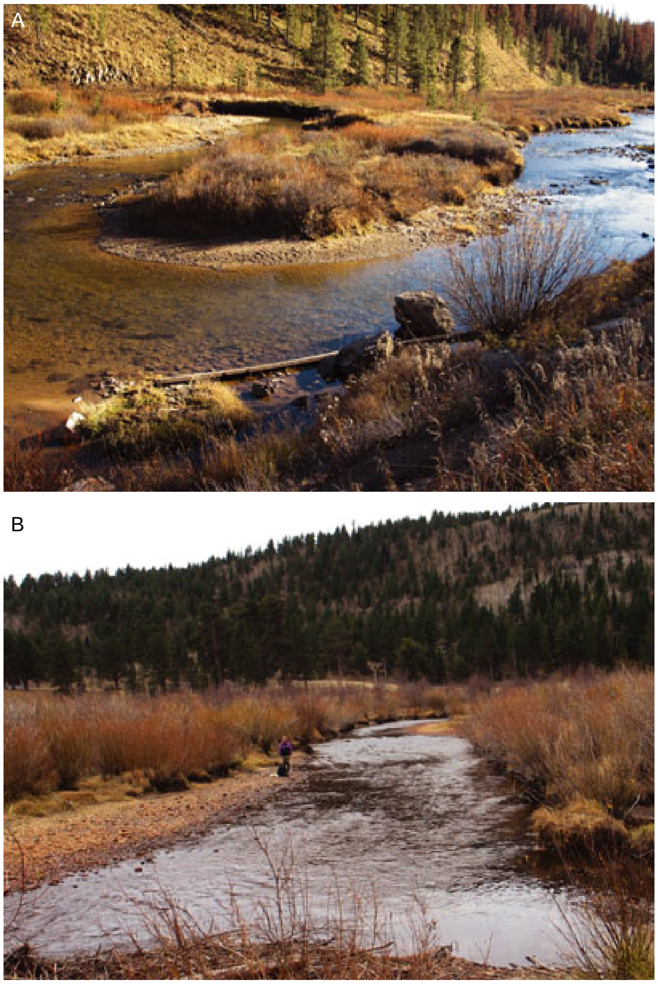

FIGURE 4. (A) Photo of the Willow Creek Study Site Showing One of the Incised Meander Bends. Flow is toward the right. (B) Photo of the study site at the North St. Vrain Creek showing alternate bars in the downstream part of the reach. Flow is toward the viewer.

Because the PIBO protocol is meant for response reaches (Montgomery and Buffington, 1997, 1998) in gravel-bed streams, while the EMAP protocol encompasses a variety of stream types, two sites with pool-riffle morphologies were selected for study: Willow Creek about 32 km NNW of Granby, and the North St. Vrain Creek about 32 km NW of Boulder, both in north central Colorado. The Willow Creek study reach has four tight meander bends (Figure 4A) and is incised up to 1.5 m into a floodplain about 60 m wide covered by willows, grass, shrubs, and scattered pine trees. Partial confinement by valley walls had to be tolerated because gravel-bed pool-riffle streams that are publicly accessible and wadeable are not common in north central Colorado. Watershed lithology is mainly sandstone and conglomerates with some volcanic materials. Many particles are platy in shape. The North St. Vrain Creek takes an irregular meandering course through a large glacial outwash valley, covered by willow thickets. The study reach is also incised up to a meter into a vegetated floodplain (as is typical of Colorado mountain streams) but unconfined by valley walls (Figure 4B). Most particles are granitic and ellipsoidal in shape. Both study streams had exposed gravel bars and similar bed material particle-size distributions with D_{50} sizes of 37 and 42 mm, respectively (Table 2).

Data Analysis

Postsampling segregation of the systematic SFT data were used to show that different stream locations harbor different particle sizes. Because the SFT protocol minimizes operator error with respect to particle identification and measurements, systematically samples the reach over the bankfull width, and collects a large sample size (461 and 598 per reach in the two study streams), results from the SFT protocol were considered the best available approximation of the true particle-size distribution within the entire reach and specified portions of it. The study compared particle-size distributions among the three sampling protocols, specifically the D_{50} and D_{84} particle sizes, the percent fines <5.6 mm, and the percent silt and sand (<2 mm). Note that this study computes D_{50} and not D_{gm}, the geometric mean particle size typically presented in EMAP studies. Instead of the percent fines <6 mm reported in the PIBO procedure, this study reports fines <5.6 mm because the 5.6 mm size break aligns with a 0.5 phi size class and was simpler to compute.

TABLE 2. Characteristics of the Two Sampling Reaches Selected for This Study.

Stream	Stream Gradient (m/m)	Site Elevation (m)	Bankfull Width[1] (m)	Wetted Width[2] (m)	Wetted Width[1] (m)	Ratio w_{wwet}/w_{bkf}	Reach Length (m)	Surface D_{16}, D_{50}, D_{84} Size (mm)[3]
Willow Cr.	0.0038	2,668	6.8-16.0 (9.5)	8	3.6-10.2 (6.8)	0.71	320	4.4, 37, 103
North St. Vrain Cr.	0.0055	2,535	11.0-20.3 (14.4)	8	5.9-15.4 (10.7)	0.74	320	3.7, 42, 88

[1]Number in parentheses is mean width, averaged over 21 transects.
[2]Estimated from five transects near reach center and rounded to nearest m.
[3]Computed from SFT protocol.

RESULTS

Comparison of Results From the Three Sampling Protocols

Particle-size distributions obtained from the three pebble count procedures varied widely, but the variability followed a similar pattern in both study streams. Sampling results among the protocols differed mostly at the fine end of the bed material size distribution. The EMAP protocol indicated that both study streams had 32% fines smaller than 5.6 mm, while PIBO indicated 7.1 and 6.5% fines <5.6 mm (Figure 5), a greater than fourfold difference between the two protocols. The SFT protocol indicated 17 and 18% fines <5.6 mm. The D_{50} particle sizes obtained by the PIBO protocol in the two study streams (42 and 51 mm) were nearly twice as large as those from EMAP (24 and 26 mm), while the SFT results again took middle values (37 and 42 mm). The bed material D_{84} sizes were least different among the procedures, reaching 103 and 128 mm for the two streams according to the EMAP, 86 and 95 mm according to the PIBO protocol, and 103 and 88 mm according to SFT. A higher percentage of fines <6 mm and smaller D_{50} sizes for the EMAP compared with the PIBO protocol were also reported by Whitacre *et al.* (2007), although the difference in sampling results between the two protocols was less pronounced in that study, perhaps due to the effects of averaging study results over multiple streams.

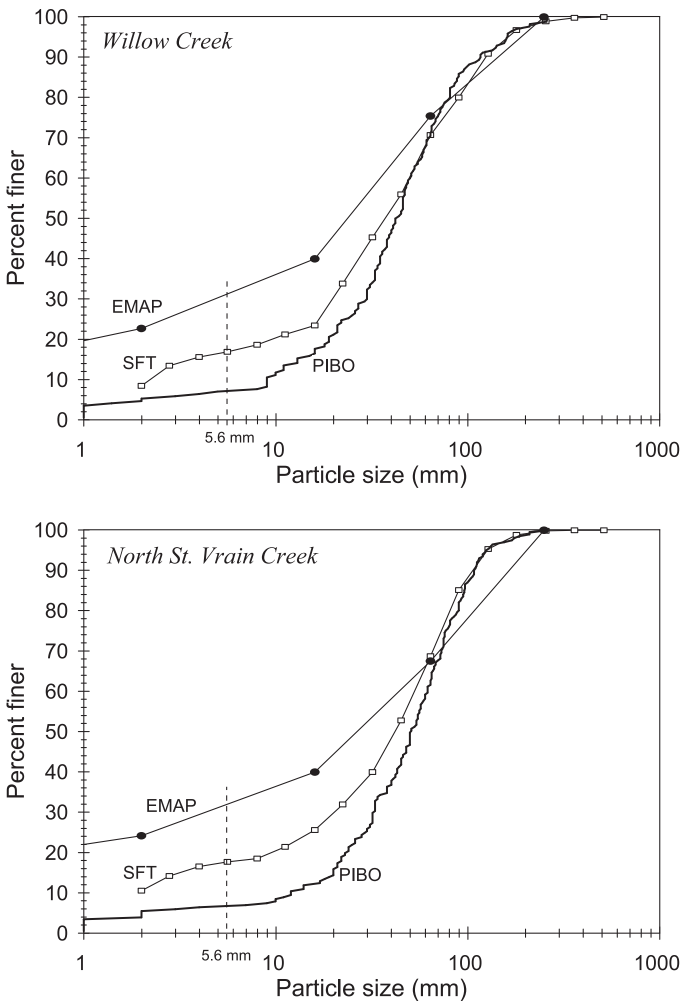

FIGURE 5. Particle-Size Distributions Obtained From the Three Pebble Count Procedures at Willow Creek (top) and North St. Vrain Creek (bottom). The cut-off point for fines <5.6 mm is indicated by the dashed line.

Effects of Sampled Streambed Locations

Riffle and Runs Coarser Than the Reach. Segregating data from the SFT protocol into different geomorphological units showed for both study streams that exposed bars had by far the finest particle-size distributions in the two reaches with 30% fines <5.6 mm (Figure 6), pools had about 20% fines <5.6 mm, while inundated riffles had only about 10%. The geomorphological units in pool-riffle streams clearly harbored different sediments: inundated riffles consisted mostly of gravel and cobbles, and exposed bars mostly of fines <5.6 mm and gravel. In inundated pools, more than 50% of all particles were cobbles and fines <5.6 mm. Consequently, sampling on different geomorphological units in the pool-riffle study streams produces different results. If exposed bars that store much of the fines <5.6 mm are not included in a pebble count, the amount of fines <5.6 mm in a reach may be substantially underestimated.

PIBO Results Coarser Than SFT Sampling on Riffles. Extending over the bankfull width, PIBO riffle transects included portions of exposed bars. Sampling close to the PIBO riffle transects with the SFT protocol indicated that the streambed area crossed by these transects was still coarser than the reach average. At Willow Creek, SFT-measured riffle transects had fewer fines <5.6 mm (11%) than the reach (17%), but more medium gravel to cobbles. At the North St. Vrain Creek, the difference between riffles and the reach was less, perhaps because the reach was more free-formed, and riffle-pool units were less affected by local hydraulic conditions in narrow bends than at Willow Creek.

Because the PIBO protocol sampled from streambed areas that contained fewer fines <5.6 mm than the reach as a whole, a coarse size distribution was expected from the PIBO protocol. However, the PIBO results were even coarser than expected from

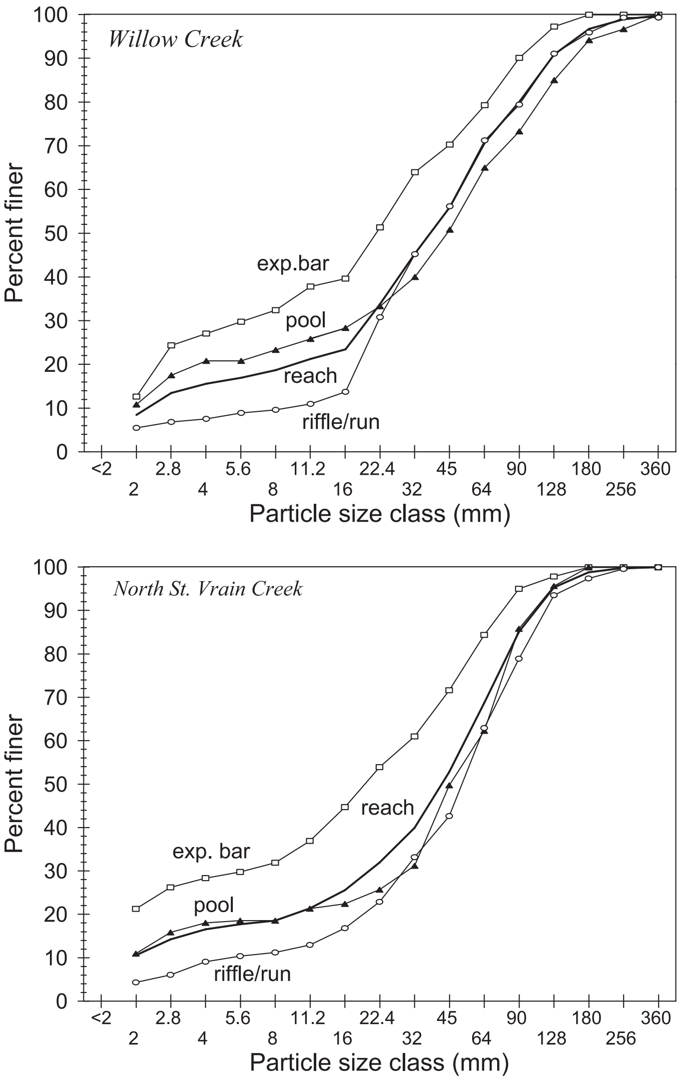

FIGURE 6. Particle-Size Distributions Sampled From Exposed Bars, Pools, Riffle/Run, and the Overall Reach at Willow Creek (top) and the North St. Vrain Creek (bottom) Using the SFT Protocol.

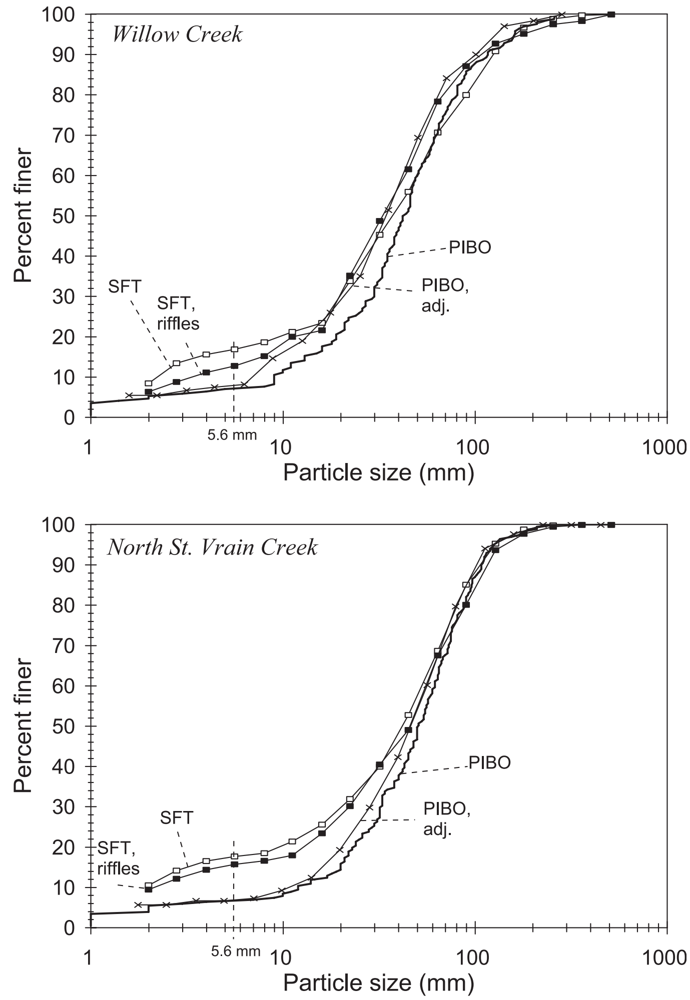

FIGURE 7. Particle-Size Distributions the Reach-Averaged SFT Procedure (SFT), and the SFT Procedure on Riffle Transects Close to Those Sampled by the PIBO Procedure (SFT riffle), From the PIBO Procedure (PIBO), and From the PIBO Procedure Adjusted for Both Heel-to-Toe Sampling as Well as for Ruler Measurements (PIBO adj.) at Willow Creek (top) and North St. Vrain Creek (bottom). The cut-off point for fines <5.6 mm is indicated by the dashed line.

sampling on riffles alone. Compared with the SFT samples collected close to the PIBO samples on riffles, PIBO's D_{50} sizes were 11 and 27% coarser while the D_{84} sizes were 15% finer and 8% coarser at the two study streams. The percent fines <5.6 mm was approximately halved (i.e., decreased by 45 and 59%) (Figure 7). Having sampled in similar streambed areas, the observed differences suggest that PIBO's methods of particle identification and size measurements must contribute to the coarseness of the PIBO results (see sections "Effects of Selecting Particles at the Tip of the Boot" and "Effects of Ruler *vs*. Template Measurements").

EMAP Results Finer Than SFT Sampling Within the Wetted Width. Segregating the SFT data into wetted width and bankfull width indicated that the wetted width had a coarser bed than the bankfull reach because the wetted width does not include the relatively fine bar sediment (Figure 8). Consequently, a relatively coarse size distribution should be expected from the EMAP protocol that samples only within the (relatively coarse) wetted width. Instead, the EMAP protocol produced a particularly fine-grained particle-size distribution with 23 and 24% silt and sand (<2 mm) at the two study streams, whereas the SFT protocol indicated 7 and 8% silt and sand within the wetted width.

The reason for the high percentage of silt and sand collected by the EMAP protocol may be found in the combined effects of stream morphology and collecting 40% of all particles at the waterline. Both study streams were incised into floodplain sediments rich in

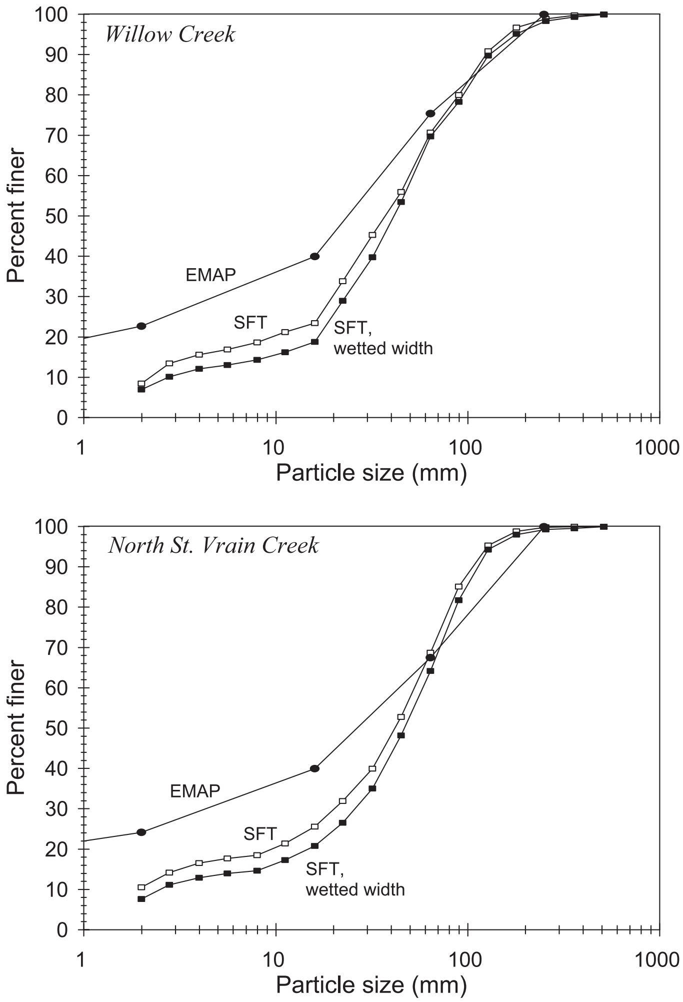

FIGURE 8. Particle-Size Distributions From EMAP and SFT Procedures and SFT Procedure on Wetted Stream Width Only at Willow Creek (top) and North St. Vrain Creek (bottom).

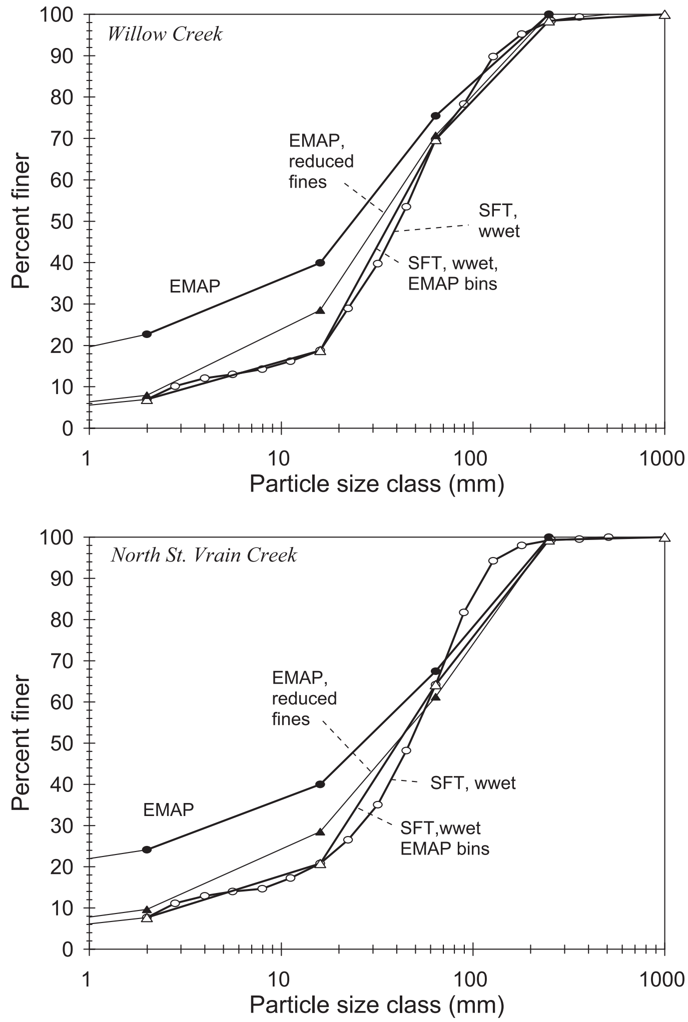

FIGURE 9. Particle-Size Distributions From the EMAP Protocol (black circles), the EMAP Protocol With a Reduced Percentage of Fines <2 mm to Show the Waterline Effect (black triangles), the SFT Procedure Collected Within Wetted Width (open circles), and the Wetted Width-SFT Results Binned in Size Classes 2-3 phi Units Wide (open triangles) for Willow Creek (top) and the North St. Vrain (bottom).

silt and sand (as is typical of Colorado mountain streams), thus the waterline ran along banks cut into fine sediment for parts of the reach (see Figures 4a and 4b). Of the total 23 and 24% silt and sand indicated by the EMAP protocol at the two study streams, only about one quarter (5 and 6%) were collected in the three central locations at 25, 50, and 75% of the wetted stream width, whereas three quarters of the percent silt and sand (18%) were added by sampling at the two waterline locations. By comparison, for the SFT protocol that sampled in 0.5 m width increments and collected 10 and 13% of its particles near waterlines, silt and sand amounted to 7 and 8% within the wetted width and to 5% when particles collected close to the waterline were excluded from the count. To visualize the effect of disproportionately sampling fines at the waterline and its contribution to the fineness of EMAP's sampling results, EMAP's 23 and 24% silt and sand are reduced to 7 and 8% as indicated by the SFT protocol and replotted in Figure 9.

Effects of Particle Selection and Size Measurements

Effect of Binning into 2-3 phi Units Wide Size Classes. Methodological differences in particle selection, size measurements, and analysis can also substantially affect the outcome of pebble counts. To estimate the effects of binning, the SFT samples collected within the wetted width and binned in 0.5 phi units were rebinned using the wide EMAP size classes. In the two study streams, EMAP binning lowered the D_{40} sizes by 13 and 20%, the D_{50} sizes by 10 and

13% from 41 and 47 mm (0.5 phi bins) to 37 and 41 mm, and increased the D_{84} particle size by about 44 and 73% from 109 and 97 mm to 157 and 168 mm, respectively (Figure 9). A similar effect of binning into EMAP size classes was described by USFS (2004). Faustini and Kaufmann (2007) explain the effects of binning into wide size classes: the center of the size distribution becomes finer, and the coarse end becomes coarser when the sampled material has a tail toward fines (typical of mountain gravel-bed streams). The reverse happens in distributions with a coarse tail (typical of sand or fine gravel beds with some coarser gravel): the center becomes coarser and the fine end finer. In the two mountain gravel-bed study streams, wide binning explained the coarseness of the EMAP result at the coarse end of the size distributions (particularly at the North St. Vrain Creek) as well as finer particle sizes in the center of the distribution.

Taken together, the effects of sampling 40% of all particles at the waterline and binning into wide size classes explained a large portion of the difference between the EMAP and the SFT sampling results in the two study streams. This is visualized in Figure 9 where the size distribution of SFT, limited to sampling the wetted width and binned in EMAP classes, is much closer to the EMAP size distribution adjusted for the effect of oversampling silt and sand along the waterline than the results from the original EMAP and SFT protocols within the wetted width.

Effects of Visual Particle-Size Estimates and Particle Selection. Despite having moved closer together, some difference remains in the size distribution obtained from SFT limited to the wetted width and binned in EMAP classes and the EMAP size distribution adjusted for the effect of oversampling silt and sand along the waterline: the adjusted EMAP protocol has more fine gravel. Comparison of the percent frequency of particles from the EMAP protocol (that after adjustment has approximately the same amount of fines <2 mm as SFT) to those sampled with the SFT procedure within the wetted width (= same stream locations) and rebinned into EMAP size classes (= same binning) showed that EMAP had 47 and 20% more particles in the size class 2-16 mm than SFT in the two study streams. By contrast, there were 30 and 37% fewer particles in the size class 16-64 mm, as well as 14 and 7% fewer particles in the size class 64-250 mm (Figure 10). Particles <2 mm were not involved in the comparison. The most likely explanation for the large quantity of 2-16 mm gravel and the small quantity of 16-64 mm gravel and cobbles (64-250 mm) in the EMAP samples is that operators erred in their visual estimates of the 16-mm border and placed particles into the 2-16 mm class that should have been in the size class 16-64 mm. The same error of operators favoring the 2-16 mm size class over the 16-64 mm class in visual estimates of particle *b*-axis sizes was observed in USFS (2004). The difficulty of accurately identifying particles when operators set the EMAP meter stick on the bed, particularly in deep and swift flow may have contributed to the difference between the protocol results. Knowing that large particles make larger targets for a moving stick than small particles, operators may have overcorrected when the stick pointed between two particles and may have selected the smaller one.

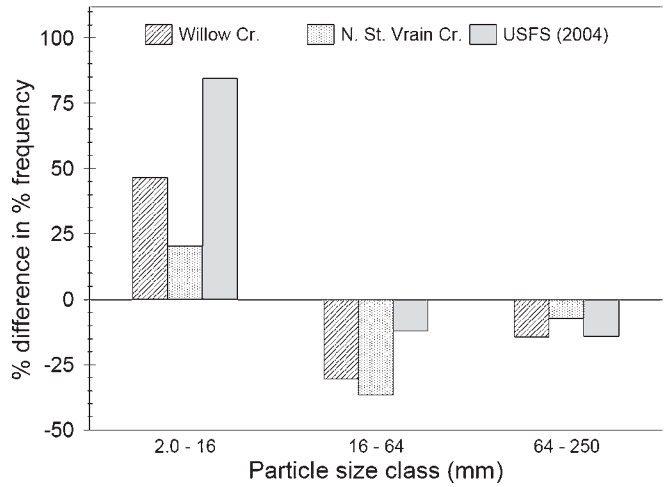

FIGURE 10. Percent Difference in the Percent Frequency of Particles Collected in Three EMAP Size Classes (2-16, 16-64, and 64-250 mm) Between EMAP Samples and SFT Samples Collected in the Wetted Width and Rebinned in EMAP Size Classes. Particles <2 mm were not included in the comparison.

Taken together, the methodological options employed by the EMAP protocol (sampling the wetted width, collecting 40% of particles at the waterline, binning into 2-3 phi wide size classes, particle selection with a pointing stick, and visual size estimates) largely explain the difference in sampling results between the EMAP and SFT procedures at the two study streams. Sampling the wetted width made the distribution lack sand and gravel. The lack of sand was more than compensated by an abundance of silt and sand contributed to the sample from floodplain sediment by collecting 40% of all particles at the waterline. Wide binning made the D_{50} size finer and the D_{84} coarser, while a relative abundance of small gravel likely stems from operators either misclassifying coarse gravel as fine gravel or by selecting fine over coarse gravel.

Effects of Selecting Particles at the Tip of the Boot. The effects that collecting particles at the tip of the boot (PIBO) and measuring the particle *b*-axis length with a ruler have in comparison to using a SF and template can be estimated. Bunte and Abt (2001a,b) found that operators avoid stepping onto large cobbles and boulders that appear slippery or protrude from the streambed in order not to risk loosing their footing. This causes undersampling of cobbles, particularly by operators with small feet. When the tip of the boot is placed above a gap between two large gravel particles, Marcus *et al.* (1995) observed that operators are more likely to touch a particle along its side and select it (incorrect) rather than bending further down to point at a particle seated deeply in the bed (correct). Touching particles sideways undersamples fines. The combined effects of pacing and selecting a particle at the tip of the boot can result in bias against small gravel and cobbles, while favoring mid-sized and coarse gravels. Based on results by Bunte and Abt (2001a,b) as well as unpublished data by the authors, both biases may amount to a factor of 1.1 in the percent frequency of cobbles or small gravel. However, underrepresenting both distribution tails while overrepresenting the distribution center does not necessarily exert a notable effect on the D_{50} particle size in coarse gravel and cobble-bed streams. At the two study streams, riffles did not contain many large cobbles (>128 mm), and avoiding stepping on them was not a critical issue. A moderate presence of cobbles may be a reason why PIBO and SFT on riffles do not differ much in their percent of large cobbles. The difference in particle-size measurements (ruler *vs.* template) likely caused a larger effect.

Effects of Ruler *vs.* Template Measurements. There are three main differences between ruler *b*-axis measurements and particle-size measurements with a template. One concerns measurement accuracy. The particle *b*-axis of an unevenly shaped three-dimensional object cannot be measured as accurately with a ruler as it can with calipers or a template (compare the diameter of a coffee mug measured with a ruler held behind it with the measured diameter across the top of the mug). Similarly, identifying the *b*-axis of an unevenly shaped three-dimensional object is error prone particularly for angular and "odd"-shaped particles (Marcus *et al.*, 1995). Errors in particle *b*-axis measurements may cancel out if they are randomly distributed, but systematic visual errors will not. Ruler *b*-axes measurements are more error prone for large and rounded or odd-shaped particles than for small, flat, and even-shaped particles.

A second difference is that ruler *b*-axis measurements made on particles with a somewhat flattened shape indicate a coarser particle size than template measurements. The reason is that ellipsoidal and platy particles with a *b*-axis larger than the diameter of a sieve can pass diagonally through a square-hole sieve opening $D_{(\mathrm{sieve})}$. The difference in reported particle size causes error and confusion among studies.

The third difference is that ruler measurements provide information on the *b*-axis length while template measurements provide information of the sieved particle size which corresponds to particle volume. The usefulness of either one of the two measurements might vary with study objectives. Typically, studies concerning fluvial geomorphology and sediment transport are more likely to require information on particle volume and mass rather than on *b*-axis length because transportability is commonly related to particle mass unless specific shape effects are considered (e.g., Gintz *et al.*, 1996). Sieving with square-hole sieves or templates takes into account the combined size of two axes (*b* and *c*), and those measurements represent particle mass better than the *b*-axis length.

Ruler and template measurements of particle size are different approaches that yield different results. For relatively small and flattened particle shapes, operator error in measuring the *b*-axis size may be small, but increases for round, angular, and odd-shaped particles. Particle-measurement errors could be entirely eliminated by passing the particle through a template, which not only provides continuity with sieve data, but takes no more time than measuring the particle *b*-axis with a ruler and is less prone to operator error.

Ruler-measured and template-measured particle sizes can be made comparable by the use of conversion factors (Church *et al.*, 1987; Shirazi *et al.*, 2009), but this requires information on particle shape. The ratio D_{sieve}/b increases with the degree of particle platyness (ratio *c*-axis to *b*-axis lengths) according to $D_{\mathrm{sieve}}/b = 1/\sqrt{2}[1 + (c/b)^2]^{0.5}$ (Church *et al.*, 1987). The ratio D_{sieve}/b can then be expressed as a change in the sieve size of a particle. Conversion is complicated by the fact that the ratio of *c*-axis to *b*-axis may change among particles of different size, of different lithology, and different transport history, thus a conversion factor does not achieve accurate conversion for all particles (Shirazi *et al.*, 2009). Conversion is needed not only for compatibility among studies, but also for comparison with data from sieved volumetric samples, e.g., when the sizes of surface and subsurface sediment are compared.

To assess the difference between ruler-measured and template-measured particle sizes in the study streams, an ellipsoidal particle shape with a *c*/*b*-axis ratio of 0.75 is assumed for particles at the North St. Vrain Creek, thus the ratio $D_{(\mathrm{sieve})}/b = 0.88$. For Willow Creek with mostly platy particles, an estimated average *c/b* ratio of 0.5 leads to $D_{(sieve)}/b = 0.79$. Particles that vary in their *b*-axes (in mm) by factors of 0.88 and 0.79, respectively, vary by approximately 0.18 and 0.34 units in their phi sizes. Particles measured with a ruler can thus be estimated to be about 0.18 phi units coarser than if they had been measured with a 0.5 phi opening template at the North St. Vrain Creek, and 0.34 phi units coarser at Willow Creek. To show the effects of particle selection and ruler measurements on the sampling outcome, PIBO results were adjusted for favoring mid-sized gravel (see section "Effects of Selecting Particles at the Tip of the Boot") as well as for the effects of ruler *b*-axis measurements (i.e., particle size for all computed cumulative percent frequencies were decreased by 0.18 and 0.34 phi units, respectively, for the two study streams). The adjustments generally moved PIBO's size distributions closer to those collected by SFT on riffles (Figure 7).

DISCUSSION

The study indicates that methodological options employed by the EMAP and PIBO protocols caused significant differences in the sampling results from the two study streams. Some methodological options caused underrepresentation of fines <5.6 mm, others overrepresented the amount of silt and sand (<2 mm), some resulted in overly high percentages of cobbles and others overly low. In PIBO's case, the selected methodological options generally caused a coarsening of the results. In EMAP's case, one methodological option (sampling 40% of all particles at the waterline) caused oversampling fines (if the water line runs along cut banks). These effects were fortuitously offset by another methodological option (not sampling on exposed bars where fines are stored) that caused undersampling of fines. A pebble count protocol cannot rely on offsetting biases, because the direction and magnitude of sampling differences introduced by methodological options vary with stream type and site conditions. Results from the study illustrate that sampling results are highly protocol-dependent and that results from different protocols cannot be used interchangeably and must be carefully interpreted. For example, a fourfold difference in the measured percentage of fines <5.6 mm as obtained in this study could easily lead to contrasting conclusions about the impairment of a reach. A twofold difference in the surface D_{50} size when used in the Shields equation causes a twofold difference in the computed critical flow depth for incipient motion which typically means an approximately fivefold difference in critical discharge. A twofold variability in the D_{50} size may also cause estimated bed-load transport rates computed from bed-load equations to vary by orders of magnitude.

Effects of Sampling Location in Various Stream Types

The sampling approaches used in EMAP and PIBO pebble count protocols are not well suited to characterize a reach as a whole in several stream types. While sampling exclusively within the wetted width or on riffles allows evaluation of these biologically important areas, these sampling locations are not well suited for evaluating change in the percent of fines supplied to the reach. The largely unsampled streambed areas – bars in the case of EMAP and pools in the case of PIBO – can cover large portions of the reach and tend to contain more fines than the sampled streambed areas. In mountain gravel-bed streams, bed load is comprised mostly of sand and fine to medium gravel particles that travel mainly over submerged bars due to secondary flows (e.g., Bridge and Jarvis, 1976, 1982; Dietrich and Whiting, 1989; Anthony and Harvey, 1991; Julien and Anthony, 2002; Bunte *et al.*, 2006). EMAP and PIBO's sampling schemes, that exclude bars to a large extent, therefore neglect information on the most frequently transported bed-load sizes. To characterize fines (silt to very fine gravel) or monitor their change, sampling should include (or even focus on) streambed areas where fines are preferentially stored i.e., on bars, in eddy deposits, in pools, and in the subsurface sediment. The relatively coarse bed surface sediment within the wetted width and on riffles is neither particularly indicative of the sediment sizes transported during normal high-flow events nor is it particularly indicative of moderate changes in the amount of silt/sand and fine gravel supplied to a reach.

The degree to which sampling only portions of the reach may exclude certain bed material sizes and introduce biases against small or large particles compared with whole-reach sampling is not fixed for a specified pebble count protocol but varies with stream morphology. In the pool-riffle study streams, the wetted width had significantly fewer fines <5.6 mm than the bankfull reach because it largely excluded exposed bars where fines were stored. Wetted width sampling would probably cause less bias against fines in braided streams because the EMAP protocol samples over unvegetated mid-channel bars where fines are typically more plentiful than in the wetted low-flow channel. Plane-bed and step-pool reaches likely show the least effect from sampling only the wetted width because lateral bars are rare. Also, the bed between the wetted and the bankfull width is typically small in area and not much finer than the bed material within the wetted width. Incised reaches are the most problematic stream morphology to sample with the EMAP protocol. Collecting 40% of all particles along the waterline samples an unduly large amount of the sediment exposed on cut banks. Sediment on cut banks is likely of a different size than the currently active streambed sediment and, when cut into floodplain deposits, it is likely finer. Much of EMAP's oversampling of silt and sand that occurs in streams incised into a floodplain could be avoided by not sampling to the water's edge when collecting only a few particles per transect. The influence of the bank particles could also be de-emphasized by sampling a larger number of particles (20 or more) per transect in even spacing over the bankfull width.

Stream type also affects the outcome of riffle sampling. While riffles and runs are typically coarser than the reach in many stream types, rock fall from adjacent hillslopes and exhumed boulders may produce a coarser bed between riffles than on riffles. Stream type and variability among riffles also affects sampling precision. Riffle formation can be affected by a variety of processes, and riffles therefore have different shapes and particle-size distributions, particularly in streams where the formation of riffle-pool units is influenced by local hydraulic conditions around narrow bends or large woody debris. Differences in riffle shapes can generate operator error in defining the areal extent of a riffle, thus the riffle area sampled by different operators may not be identical and cause variability in the sampling result. At Willow Creek with its narrow bends and some adjacent hillslopes, particle-size distributions among individual riffles within the study reach differed mainly at their coarse ends (by a factor of 2-3 in the D_{50} to D_{84} sizes). At the North St. Vrain Creek, riffles differed mainly at their fine ends (the percentage of fines <5.6 mm ranged from 1 to 15%) (Figure 11). The precision obtained from riffle sampling with a preset sample size is higher when riffles within the sampled reach have similar particle sizes. This condition is most likely found in pool-riffle streams disconnected from direct hillslope influence and without forcing flow around narrow bends or large woody debris. By contrast, study streams with gradients of up to 2 and 3% that typically have plane-bed morphology with isolated, often forced, riffles and pools (Montgomery and Buffington, 1997) encompass a wide variability among the riffles within a reach. Uncertainty associated with riffle sampling might be reduced if a protocol description pointed out sources of variability among riffles, making a stronger case for the desirability of sampling riffles of similar type, and sampling over more than four riffles.

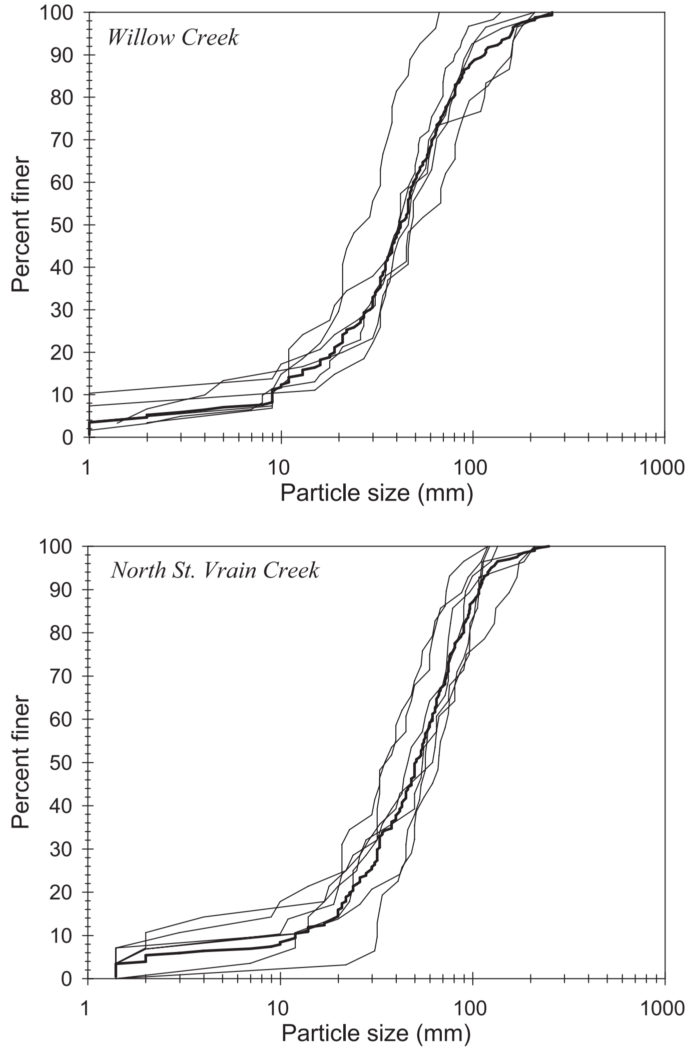

FIGURE 11. Particle-Size Distributions Obtained From PIBO Procedure on Six and Seven Individual, Consecutive Riffles at Willow Creek (top) and North St. Vrain Creek (bottom). The thick line shows the average overall riffles in a reach.

Effects of Sample Size

Sample sizes in the EMAP and PIBO pebble count protocols are small and typically not able to adequately characterize a reach or detect a change when applied to studies in individual coarse gravel and cobble-bed streams. The bed material size distribution in these streams is typically poorly sorted, ranging from sand to cobbles or boulders, and is skewed toward a tail of fines. A sample size of 100 may suffice to characterize the D_{50} to D_{84} particles in those streams to within 0.25 phi units with 95% confidence levels, but in order to accurately evaluate the tails of the distribution, a larger sample size (Green, 2003), typically about 400 particles is needed for errors less than 0.2 and 0.1 phi units for the D_{16} and D_{95} sizes (Rice and Church, 1996). Bevenger and King (1995) suggest that several hundred particles are needed to detect change in user-defined percent fines in individual streams between two visits (e.g., for a stream with 10% fines, approximately 300 particles are needed to detect a 10% change in fines, with Type I and II errors both set at 0.05). For multiple streams, Archer *et al.* (2004) concluded that over 200 particles needed to be sampled using the PIBO protocol to detect a 10% change in the D_{50} particle size, accounting for both observer difference and heterogeneity between streams with Type I and Type II errors set to 0.1, and almost 800 particles were needed to detect a 10% change in the percent fines <6 mm. These examples show that a 100-particle sample size is generally not suitable for differentiating between reaches or detecting change between visits at specific sites. Sample sizes by EMAP and PIBO are meant to detect long-term changes in the percent silt and sand (<2 mm) and percent fines <6 mm when a large number of sites are monitored over many years. For example, the EMAP protocol claims that monitoring the same 30-50 sites over 13-23 years permits detecting an average 1-2% per year change in the percent silt and sand <2 mm (Type I error <0.05 and Type II error <0.2) (Larsen *et al.*, 2004), and that 18-32 years would be required when randomly selecting sites from a set within a region. Note that by this time the size of tens of thousands of particles will have been estimated.

A fourfold increase in sample size (from 100 to 400) can be achieved with less than a fourfold increase in field time. In this field study, determining the EMAP transect locations, installing the tape across the stream, and wading between sample locations took more time than actually selecting and measuring the five particles. While wading across a transect, operators could easily collect 10 or 20 particles to double or quadruple the sample size without expending a comparable increase in time and effort. Sampling six instead of four riffles in the PIBO protocol with 25 particles each adds about 30-45 min to the field time, while increasing sample size to 150 particles. Sampling twice the number of particles on each riffle would take even less time per sampled particle and increase sample size to 300.

Effects of Particle Selection and Particle-Size Measurements

The methodological options employed by EMAP and PIBO protocols in particle selection and measurement can have undesirable effects on sampling results. Operators in this field study were not convinced that pointing with a stick through the water column provided an accurate identification of the particle to be selected in deep or swift flow, particularly not for small and medium gravel particles adjacent to cobbles. PIBO's particle selection at the tip of the boot favors the selection of mid-sizes gravel particles and can underrepresent cobbles and fines (Marcus *et al.*, 1995; Bunte and Abt, 2001a,b). Compared with effects from other methodological options, the effect of PIBO's particle selection was small in the two gravel-bed study streams, but would be larger on riffles rich in cobbles and boulders.

Visual estimates of particle *b*-axis size classes differed from those obtained from template-measured particle sizes, even though operators in this study used locations on their hands and fingers to help visualize the 16, 64, and 250 mm size class boundaries in the field. Visual estimates appear to have increased the amount of 2-16 mm gravel and decreased the amount of 16-64 mm gravel, a tendency also observed in USFS (2004). This effect, perhaps enhanced by the effects from particle selection, explains the abundance of 2-16 mm gravel obtained by the EMAP pebble count. Ruler measurement of the particle *b*-axis rendered the entire distribution too coarse in comparison to template-measured particle sizes, particularly at Willow Creek where particles are of ellipsoidal and platy shape. However, knowing particle shape, this discrepancy is largely correctable. Binning particles into size classes 2-3 phi units wide, when compared with 0.5 phi wide, caused the central part of the distribution to be too fine (by 13 and 20% for the D_{40} size) and particularly the coarse part to be too coarse (by 44 and 73% for the D_{84} size) in the study streams where bed material has a tail toward fines. Other binning effects occur in beds with a tail toward coarse particles.

The EMAP and PIBO protocols have selected methodological options that make field work fast. Collecting five particles at 0, 25, 50, 75, and 100% of the wetted width at the tip of a meter stick is time saving for the EMAP protocol because flow depth is measured with the stick at those locations as well. However, the convenience of this methodological option brings about several problems. Using a meter stick to point at a particle through the water column (instead of selecting the particle immediately below a grid intersection) likely introduces error, particularly in deep and swift flows. Sampling only the wetted width which largely excludes lateral bars undersamples fines <5.6 mm, while collecting 40% of the particles at the waterline in incised streams oversamples particle sizes that make up a steep or cut bank, typically silt and sand <2 mm. The claimed convenience of visually estimating the particle size class may lead to random operator error or bias if misidentifications occur systematically. Binning into size classes 2-3 phi units wide introduces additional biases that vary depending on the percentile in question and whether the bed sediment has a tail of fine or coarse particles.

The PIBO's time-saving practice of determining the location from which to extract a particle by pacing and selecting a particle at the tip of the boot may cause undersampling of fines as well as cobbles, particularly in coarse gravel and cobble beds where stepping location may be affected by where operators consider it is safe to step (Bunte and Abt, 2001b). Measuring particle *b*-axis sizes with a ruler takes no less time than passing a particle through a template, but can introduce operator error (Marcus *et al.*, 1995) and renders the entire distribution coarser, typically by about 0.15-0.35 phi units depending on particle flatness. Better wadeability and not having to span a tape across transects make PIBO's sampling exclusively on riffles convenient and attractive for field studies other than those focusing solely on characterizing riffle habitat. However, excluding fines <5.6 mm that have accumulated in pools and on many bars poses problems if riffle sampling is used with the aim to characterize the reach and its amount of fines. Finally, a rapidly obtained sample size of just over 100 particles per reach comes at a cost of sample accuracy and precision, particularly if the sampling aim was to identify changes in the percent of fines at a site which typically requires sampling of several hundred particles.

CONCLUSION AND RECOMMENDATIONS

Sampling results were greatly different among the three pebble count protocols EMAP, PIBO, and SFT. The percent fines <5.6 mm differed by a factor of 4-5 and the D_{50} sizes by a factor of 2 in the two study streams. These results suggest that pebble count results are protocol-specific and cannot be used interchangeably. The present study does not evaluate the suitability of the EMAP and PIBO protocols for their intended purposes. The EMAP and PIBO pebble count protocols are broad-scale probability based sampling designs that purposefully select methodological options to minimize field time. However, EMAP and PIBO's rapid field techniques may be adopted by others for use in fluvial geomorphology or sedimentation studies at individual sites. This study explored how methodological options used by each protocol influence the sampling result. Compared with the data-intensive SFT method designed for individual stream studies, the methodological options used by the EMAP and PIBO protocols result in important differences in the measured bed material size distributions. The magnitudes and directions of the differences vary among stream types and site conditions. In combination, the observed individual differences may either offset or reinforce each other, depending on the specific conditions at a study site. These errors can be avoided or minimized by applying more rigorous field techniques.

Effects of Sampling Location

Sampling only within the wetted width and largely excluding exposed bars results in a bias against fines (as bars are typically finer than the wetted width bed), and this bias becomes pronounced when exposed bars cover a large portion of the reach area such as in pool-riffle streams. Sampling a wetted width that may fluctuate between visits depending on flow stage also makes it difficult to compare results between visits, and this is typically not suitable for individual stream studies. By contrast, tying sampling locations to a specific geomorphic surface such as the bankfull width provides consistent reference points between visits. Collection of 40% of all particles along the waterline overrepresents fine particles in streams that are incised into fine-grained floodplain sediment and overrepresents coarse particles in streams incised into coarse sediments. Both, the bias against fines when sampling the wetted width and the bias toward fines from sampling at the waterline along cut banks is expected to be less in streams with plane-bed or step-pool morphologies because bars are scarce or small there and cut banks are less likely to coincide with the low-flow waterline. For the PIBO protocol, sampling exclusively on riffles limits the analysis to riffle sediment. The riffle particle-size distribution may be coarser than the reach because the excluded bars and pools typically harbor finer sediment. The resulting underrepresentation of finer sediment likely persists in braided and plane-bed streams but might reverse when gravel riffles develop in boulder-strewn streams or those that receive coarse hillslope particles. The waterline problem does not occur to the same extent in PIBO's sampling of about 7 particles per riffle transect because sampling starts a step away from the water line and because the low-flow waterline on riffles is much less likely to touch cut banks. In summary, sampling location – although not very influential in studies of benthic invertebrate (Herbst and Silldorff, 2006; Rehn *et al.*, 2007) – matters greatly in stream studies that analyze particle-size distributions.

Systematically sampling the reach with even-spaced transects and even-spaced increments over the bankfull width ensures that particle-size information from all streambed areas are included in the analysis and avoids concentrating 40% of all samples at the water line. Recording particle sizes in sequence (i.e., individually for each transect, from bank to bank instead of tallying tick marks, and noting major geomorphological, sedimentary, or habitat units as well as flow stages) permits postsampling spatial segregation. Distinguishing particle-size distributions among geomorphological units and monitoring change in particle sizes among units can be a powerful tool for fluvial analysis.

Particle Selection

Pointing with a stick through the water column is not the most accurate tool for particle identification, particularly not in deep and swift flows. Similarly, sampling inundated beds at the tip of the boot, eyes averted, may undersample fines and cobbles in coarse gravel and cobble beds when operators avoid stepping on slippery large rocks. Both particle selection problems can be largely avoided when particles are selected under grid intersections of elastic bands within a SF placed on the bed. On poorly visible beds, the operator points downward at grid intersections, and the error in particle selection may approach that of a Wolman pebble count.

Particle-Size Estimates

Visually estimating *b*-axis size-class boundaries can cause operator error. Operators in this study may have overestimated the amount of small gravel particles when visually estimating the 16-mm boundary at the expense of large gravel. Using a ruler to measure a particle's *b*-axis coarsens the distribution compared with sieve-measured or template-measured data. The difference between ruler and template measurements increases with particle flatness. The bias is largely correctable (Shirazi *et al.*, 2009), particularly if particle shape is known (Church *et al.*, 1987), and correction is necessary if ruler measurements are to be compared with sieve data. Using a template – which takes no more time than a ruler measurement – eliminates operator error in particle-size estimates. Templates provide a particle-size estimate that more closely reflects particle mass – often the preferred particle characteristic in studies of fluvial geomorphology and sedimentation – and permit direct comparison with sieved data.

Compared with binning in 0.5 phi sizes, binning into 2-3 phi wide size classes as is done by EMAP moderately lowered the D_{50} sizes, and notably increased the D_{84} sizes (by 44-73%) in the study streams where the bed material sediment had a tail of fines, typical of mountain gravel-bed streams. The effects shift in streambeds with a tail toward coarse particles, typical of sand-bedded streams that contain some gravel. Use of a template with 0.5 phi openings eliminates this binning error. In well-sorted bed material, a template with a 0.25 phi size opening may be more appropriate.

Sample Size

Sample size is specific to a study aim and site conditions. In individual stream studies, sampling 100 particles may suffice to characterize a well-sorted facies. However, in streambeds that are poorly sorted (encompass sand to large cobbles) and skewed (typically with a tail of fines), sampling 100 particles does not provide results sufficiently precise for many applications, such as monitoring change in the percentage of fines or accurate input for bed-load computations.

ACKNOWLEDGMENTS

We thank John Faustini for sharing a spreadsheet that clarified the computation of particle-size distribution for the EMAP procedure. This study was funded by the USDA Forest Service, Stream System Technology Center, Fort Collins, Colorado.

LITERATURE CITED

Anthony, D.J. and M.D. Harvey, 1991. Stage-Dependent Cross-Section Adjustments in a Meandering Reach of Fall River, Colorado. Geomorphology 4:187-203.

Archer, E.J., A. Birnie, B. Bouwes, J. Heitke, and G. Kliewer, 2006. Summary Report 2001-2006, Effectiveness Monitoring Program for Streams and Riparian Areas Within the Upper Columbia River Basin Region 1. National Forest Service 2001-2006 Annual Summary Report. Unpublished paper, 257 pp. http://www.fs.fed.us/biology/fishecology/new.html#pibo_reports, *accessed* August 5, 2008.

Archer, E.K., B.B. Roper, R.C. Henderson, N. Bouwes, S.C. Mellison, and J.L. Kershner, 2004. Testing Common Stream Sampling Methods for Broad-Scale, Long-Term Monitoring. General Technical Report RMRS-GTR-122. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, 15 p.

Bevenger, G.S. and R.M. King, 1995. A Pebble Count Procedure for Accessing Watershed Cumulative Effects. USDA, Forest Service, Rocky Mountain Forest and Range Experiment Station, Research Paper RM-RP-319. Fort Collins, Colorado, 17 pp.

Bridge, J.S. and J. Jarvis, 1976. Flow and Sedimentary Processes in the Meandering River South Esk, Glen Clova, Scotland. Earth Surface Processes and Landforms 1:303-336.

Bridge, J.S. and J. Jarvis, 1982. The Dynamics of a River Bend: A Study in Flow and Sedimentary Processes. Sedimentology 29:499-543.

Bunte, K. and S.R. Abt, 2001a. Sampling Surface and Subsurface Particle-Size Distributions in Wadable Gravel- and Cobble-Bed Streams for Analysis in Sediment Transport, Hydraulics, and

Streambed Monitoring. USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, General Technical Report RMRS-GTR-74, 428 pp. http://www.fs.fed.us/rm/pubs/rmrs_gtr74.html, *accessed* July 22, 2009.

Bunte, K. and S.R. Abt, 2001b. Sampling Frame for Improving Pebble Count Accuracy in Coarse Gravel-Bed Streams. Journal of the American Water Resources Association 37(4):1001-1014.

Bunte, K., J.P. Potyondy, S.R. Abt, and K.W. Swingle, 2006. Path of Gravel Movement in a Coarse Stream Channel. Proceedings of the Eighth Federal Interagency Sedimentation Conference, April 2-6, 2006, Reno, NV, Session 1B-2, 9 pp. [CD ROM]. http://pubs.usgs.gov/misc_reports/FISC_1947-2006/1st-7thFISCs-CD/8thFISC/Session%201B-2_Kbunte.pdf, *accessed* July 22, 2009.

Church, M., D.G. McLean, and J.F. Walcott, 1987. River Bed Gravels: Sampling and Analysis. *In*: Sediment Transport in Gravel-Bed Rivers, C.R. Thorne, J.C. Bathurst, and R.D. Hey (Editors). John Wiley and Sons, Chichester, pp. 43-87.

Dietrich, W.E and P. Whiting, 1989. Boundary Shear Stress and Sediment Transport in River Meanders of Sand and Gravels. *In*: River Meandering, S. Ikeda and G. Parker (Editors), Water Resources Monograph 12, American Geophysical Union, Washington, D.C., pp. 1-50.

Diplas, P. and V.K. Lohani, 1997. Discussion of: "Application of the Pebble Count: Notes on Purpose, Method, and Variants," by G. Mathias Kondolf. Journal of the American Water Resources Association 33(6):1397-1399.

Faustini, J.M. and P.R. Kaufmann, 2007. Adequacy of Visually Classified Particle Count Statistics From Regional Stream Habitat Surveys. Journal of the American Water Resources Association 43(5):1293-1315.

Fripp, J.B. and P. Diplas, 1993. Surface Sampling in Gravel Streams. Journal of Hydraulic Engineering 119(4):473-490.

Gintz, D., M.A. Hassan, and K.-H. Schmidt, 1996. Frequency and Magnitude of Bedload Transport in a Mountain River. Earth Surface Processes and Landforms 21:433-445.

Green, J.C., 2003. The Precision of Sampling Grain-Size Percentiles Using the Wolman Method. Earth Surface Processes and Landforms 28:979-991, doi: 10.1002/esp.513.

Heitke, J.D., E.J. Archer, D.D. Dugaw, B.A. Bouwes, E.A. Archer, R.C. Henderson, and J.L. Kershner, 2007. Effectiveness Monitoring for Streams and Riparian Areas: Sampling Protocol for Stream Channel Attributes, 75 p. http://www.fs.fed.us/biology/fishecology/emp, *accessed* August, 2008.

Henderson, R.C., E.K. Archer, B.A. Bouwes, M.S. Coles-Ritchie, and J.L. Kershner, 2005. PACFISH/INFISH Biological Opinion (PIBO): Effectiveness Monitoring Program Seven-Year Status Report 1998 Through 2004. *Gen. Tech. Rep.* RMRS-GTR-162. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, 16 p.

Henderson, R.C., E.K. Archer, and J.L. Kershner. 2004. Effectiveness Monitoring for Streams and Riparian Areas Within the Upper Columbia River Basin: Sampling Protocols for Integrator Reaches Steam Channel Parameters – 2001. *In*: J.L. Kershner, E.K. Archer, M. Coles-Ritchie, E.R. Cowley, R.C. Henderson, K. Kratz, C.M. Quimby, D.L. Turner, L.C. Ulmer, and M.R. Vinson, Guide to Effective Monitoring of Aquatic and Riparian Resources. *Gen. Tech. Rep.*, RMRS-GTR-121. U.S. Department of Agriculture, Rocky Mountain Research Station, Fort Collins, Colorado, 57 pp.

Herbst, D.B. and E.L. Silldorff, 2006. Comparison of the Performance of Different Bioassessment Methods: Similar Evaluation of Biotic Integrity From Separate Programs and Procedures. Journal of the North American Benthological Society 25(2):513-530.

Hey, R.D. and C.R. Thorne, 1983. Accuracy of Surface Samples From Gravel Bed Material. Journal of Hydraulic Engineering, 109(6):842-851.

Johnson, D.H., N. Pittman, E. Wilder, J.A. Silver, R.W. Plotnikoff, B.C. Mason, K.K. Jones, P. Roger, T.A. O'Neil, and C. Barrett, 2001. Inventory and Monitoring of Salmon Habitat in the Pacific Northwest – Directory and Synthesis of Protocols for Management/Research and Volunteers in Washington, Oregon, Idaho, Montana, and British Columbia. Washington Department of Fish and Wildlife, Olympia, Washington., 212 pp.

Julien, P.Y. and D.J. Anthony, 2002. Bed Load Motion and Grain Sorting in a Meandering Stream. Journal of Hydraulic Research 40(2):125-133.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck, 1999. Quantifying Physical Habitat in Wadeable Streams. EPA/620/R-99/003. U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P.R. and E.G. Robison, 1998. Physical Habitat Characterization. *In*: Environmental Monitoring and Assessment Program c Surface Waters: Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams. J.M. Lazorchak, D.J. Klemm and D.V. Peck, (Editors). EPA/620/R-94/004F. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., pp. 77-118.

Kershner, J.L., E.K. Archer, M. Coles-Ritchie, E.R. Cowley, R.C. Henderson, K. Kratz, C.M. Quimby, D.L. Turner, L.C. Ulmer, and M.R. Vinson, 2004. Guide to Effective Monitoring of Aquatic and Riparian Resources. *Gen. Tech. Rep.*, RMRS-GTR-121. U.S. Department of Agriculture, Rocky Mountain Research Station, Fort Collins, Colorado, 57 pp.

Kondolf, G.M., 1997a. Application of the Pebble Count: Notes on Purpose, Method, and Variants. Journal of the American Water Resources Association 33(1):79-87.

Kondolf, G.M., 1997b. Reply to Discussion by Panayiotis Diplas and Vinod K. Lohani: "Application of the Pebble Count: Notes on Purpose, Method, and Variants." Journal of the American Water Resources Association 33(6):1395-1396.

Larsen, D.P, P.R. Kaufmann, T.M. Kincaid, and N.S. Urquhart, 2004. Detecting Persistent Change in the Habitat of Salmon-Bearing Streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Science 61:283-291, doi: 10.1139/F03-157.

Leopold, L.B., 1970. An Improved Method for Size Distribution in Stream-Bed Gravel. Water Resources Research 6:1357-1366.

Marcus, W.A., S. Ladd, J. Stock, and J. Stoughton, 1995. Pebble Counts and the Role of User-Dependent Bias in Documenting Sediment Size Distributions. Water Resources Research 31(10):2625-2631.

Montgomery, D.R. and J.M. Buffington, 1997. Channel-Reach Morphology in Mountain Drainage Basins. Geological Society of America Bulletin 109(5):596-611.

Montgomery, D.R. and J.M. Buffington, 1998. Channel Processes, Classification, and Response. *In*: River Ecology and Management. Lessons From the Pacific Coastal Ecoregion. R.J. Naiman and R.E. Bilby (Editors), 13-42, Springer Verlag, New York, NY, 705 pp.

Olsen, D.S., B.B. Roper, J.L. Kershner, R. Henderson, and E. Archer, 2005. Sources of Variability in Conducting Pebble Counts: Their Potential Influence on the Results of Stream Monitoring Programs. Journal of the American Water Resources Association 41(5):1225-1236.

Peck, D.V., A.T. Herlihy, B.H. Hill, R.M. Hughes, P.R. Kaufmann, D.J. Klemm, J.M. Lazorchak, F.H. McCormick, S.A. Peterson, P.L. Ringold, T. Magee, and M. Cappaert, 2006. Environmental Monitoring and Assessment Program-Surface Waters Western Pilot Study: Field Operations Manual for Wadeable Streams. EPA/620/R-06/003. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C.

Petrie, J. and P. Diplas, 2000. Statistical Approach to Sediment Sampling Accuracy. Water Resources Research 36(2):597-605.

Potyondy, J.P. and K. Bunte, 2002. Sampling With the US SAH-97 Hand-Held Particle-Size Analyzer. Federal Interagency Sedimentation Project, Waterways Experiment Station, Vicksburg, Mississippi, 6 pp. http://fisp.wes.army.mil/Instructions%20US_SAH-97_040412.pdf, *accessed* July 22, 2009.

Rehn, A.C., P.R. Ode, and C.P. Hawkins, 2007. Comparisons of Targeted-Riffle and Reach-Wide Benthic Macroinvertebrate Samples: Implications for Data Sharing in Stream Condition Assessments. Journal of the North American Benthological Society 26(2):332-348.

Rice, S. and M. Church, 1996. Sampling Surficial Fluvial Gravels: The Precision of Size Distribution Percentile Estimates. Journal of Sedimentary Research 66(3):654-665.

Roper, B.B., J.L. Kershner, E. Archer, R. Henderson, and N. Bouwes, 2002. An Evaluation of Physical Stream Habitat Attributes Used to Monitor Streams. Journal of the American Water Resources Association 38(6):1637-1646.

Shirazi, M.A., J.M. Faustini, and P.R. Kaufmann, 2009. Streambed Gravel Sampling and Frequency Base Conversion: A Solution to Dataset Sharing. Water Resources Research 45, W01414, doi:10.1029/2007WR006151.

USEPA (U.S. Environmental Protection Agency), 2004. Wadeable Stream Assessment: Field Operations Manual EPA 841-B-04-004. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C., 106 pp.

USFS (U.S. Forest Service), 2004. Stream Notes (April 2004). Ask Doc Hydro. Visual Estimates and Pebble Counts. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Stream Systems Technology Center, Fort Collins, Colorado. http://www.stream.fs.fed.us/news/, *accessed* August, 2008.

Wentworth, C.K., 1922. A Scale of Grade and Class Terms for Clastic Sediments. Journal of Geology 30:377-392.

Whitacre, H.W., B.B. Roper, and J.L. Kershner, 2007. A Comparison of Protocols and Observer Precision for Measuring Physical Stream Attributes. Journal of the American Water Resources Association 43(4):923-937.

Wohl, E.E., D.J. Anthony, S.W. Madsen, and D.M. Thompson, 1996. A Comparison of Surface Sampling Methods for Coarse Fluvial Sediments. Water Resources Research, 32(10):3219-3226.

Wolman, M.G., 1954. A Method of Sampling Coarse River-bed Material. Transactions, American Geophysical Union 35(6):951-956.