9. Analysis

a. Analysis tools for dam removal

vi. Estimating Measurement Error

1.0 Rationale

For most dam removal or restoration monitoring projects, the emphasis is on changes in attributes, not the value of the attributes themselves. Inclusion of estimates of uncertainty is essential to certifying change. However, the values of attributes may have inherent meaning, such as spawning gravel sizes. In that case, quantifying the degree of uncertainty in measurements is valuable to establish whether or not the value truly falls within defined favorable versus unfavorable ranges. Without inclusion of uncertainty, any comparison of data lacks context. “There is no such thing as a perfect measurement” (Coleman and Steele 1999).

There are two types of error to consider: random and systematic. Random error is likely to be different for every measurement and can be attributed to the human component: interpolation, reaction time, etc. (Taylor 1997). Systematic error is consistent for a group of measurements taken the same way and it is usually a function of equipment: offsets due to temperature sensitivity, miscalibration, etc. Different approaches need to be utilized to calculate the two different kinds of error.

2.0 Definitions

<u>Error</u> is the difference between reality and our representation of it (Unwin 1995).

<u>Accuracy</u> is the closeness of results to values that are accepted to be true (Coleman and Steele 1999; Unwin 1995).

<u>Precision/Random error</u> is the difference between an observation and the average of a number of observations (Coleman and Steele 1999; Taylor 1997). Random errors lead to a normal or Gaussian frequency distribution of values around the average of a number of observations (Coleman and Steele 1999; Taylor 1997). For example, when reading a thermometer or tape measure, random error is introduced by the need to interpolate values between markings and from the angle at which the thermometer or tape measure is read. Or, if a stopwatch is used to time a velocity measurement, random error is introduced by the reaction time of the person stopping and starting the stopwatch resulting in slightly longer or shorter times.

<u>Bias/Fixed/Systematic error</u> is the difference between the average of a number of observations and the true value (Coleman and Steele 1999; Taylor 1997). Systematic error is consistent for all measurements of the same quantity taken the same way, and, therefore, cannot be detected by repeated measures (Coleman and Steele 1999; Taylor 1997). For example, if a thermometer used to measure water temperature was miscalibrated, it read 10 °C in ice water, then all measurements made with the thermometer would have a systematic error of +10 °C. Similarly, if a net used for macroinvertebrate samples had a small tear, there would be an unknown systematic error in the number of individuals, and types and number of species sampled due to loss of some individuals smaller than the tear.

Independence: Measurement errors are independent if the error associated with one measurement does not influence the error for another measurement – i.e. future errors cannot be predicted by prior errors (Taylor 1997). Random errors, by definition, are independent and systematic errors are not.

Error propagation is the transference of errors from a measured quantity to calculated result (Rabinovich 2000). For example, area is typically calculated as the product of width and depth. Therefore, the error associated with the calculated area is a combination of the error in the measurements of width and depth.

3.0 Methodology

3.1 Error estimation

There are three common methods used to estimate measurement error: 1) Repetition of measurements; 2) Comparison to results from measurements made using a different method; 3) Published or previously established values.

3.1.1 Repetition

One of the most popular ways to measure error for individual measurements is through repetition. To make this methodology successful, one must ensure that the same quantity is being measured each time, and systematic errors must be small (Taylor (1997) defines this as "smaller than required precision"). This can range in effort from a handful of repetitions to statistically rigorous quantities: sample sizes smaller than 31 result in underestimates of uncertainty using the sample standard deviation for a normal distribution unless the Student's t statistic is used as the multiplier (Dieck 2006).

There are three main ways to calculate and report error from repeated measurements: 1) Maximum difference; 2) Standard deviation; or 3) Probable error (PE) (Hubbard and Glasser 2005). If there are few repetitions, typically the error estimate (±) used is half the maximum difference between measurements, e.g. largest minus smallest divided by 2 (Downward 1995; Hubbard and Glasser 2005).

For numerous repeated measurements with random error and minimal systematic error, a normal or Gaussian distribution can typically be used as a good approximation (Taylor 1997). To check if the assumption of a normal distribution is acceptable, a Chi-squared test can be performed (see texts such as Taylor 1997 for details). By assuming a known frequency distribution, we can find the range (the mean ± the uncertainty) within which measured values will fall with a desired probability as a multiple of the standard deviation (Taylor 1997). For example, for a normal distribution there is a probability of approximately 95 % that an individual measurement is within 2 standard deviations of the true value (Taylor 1997).

$$\text{Sample standard deviation: } s = \sqrt{\frac{1}{N-1}\sum (x_i - \bar{x})^2}$$

An alternative way to report uncertainty is as the probable error (PE): there is a 50 % probability of a measurement being within ± PE of the true value (Taylor, 1997). The probable error is approximately equal to 0.67 times the standard deviation (Taylor, 1997).

Examples from the literature:

Downward (1995) delineated channel boundaries from aerial photographs and maps in GIS. In order to establish the uncertainty in digitizing those boundaries, Downward (1995) chose a boundary location from one map to repeatedly digitize. 'Slithers' were created between a "'control' boundary line" and each digitized boundary line; error was calculated as the standard deviation for the assumed normal frequency distribution of mean individual slither displacement (slither area/arc length) (Downward 1995).

3.1.2 Comparison:

Another method used to estimate error is comparison to results from a more accurate or better established technique. Systematic and random errors are measured together using this technique. This is a common method in remote sensing where the coordinates for locations on images or photos are checked against maps or images of the same landscape with known errors. In the case of sediment sampling, volumetric samples are typically the standard, so calculated values from other measurement techniques can be compared to results from volumetric sampling, such as bulk samples (Fripp and Diplas 1993).

Examples from the literature:

Simonson and Lyons (1995) compared values for species abundance, richness, and assemblage structure for 9 small (< 8 m wide) streams in Wisconsin from a single catch per effort (CPE) tow-barge electrofishing pass versus multiple electrofishing tow-barge passes with block-nets and removal (REM). Both types of sampling were performed on each stream, with CPE upstream of REM sampling for 4 of the streams, and REM upstream of CPE for 5 for the streams (Simonson and Lyons 1995). The Wilcoxon's signed-rank test, Spearman's rank correlation, and a similarity index were used to compare values calculated from the CPE sampling and the removal sampling (Simonson and Lyons 1995).

3.1.3 Established values:

Finally, there may already be error estimates for standard methods or equipment. In that case, as long as the same protocols are observed and conditions apply, finding and using published error estimates may be the best use of time and resources.

Example Sources of Error Estimates

Category	Methodology	Citation
Sediment	Pebble Count	(Fripp and Diplas 1993; Green 2003; Wohl et al. 1996; Wolman 1954)
	Bulk and Freeze Core Sampling	(Ferguson and Paola 1997; Zimmermann et al. 2005)
Geomorphology	GPS and DEM	(Brasington et al. 2000; Cheng and Granata 2007)

	Total station	(Downward 1995)
Macroinvertebrates	Surber	(Li et al. 2001)
Physical Habitat	USFS: AREMP, PIBO; EPA: EMAP	(Whitacre et al. 2007)
Fish	Electrofishing	(Reid et al. 2009; Rosenberger and Dunham 2005; Simonson and Lyons 1995)
Hydrology	Stage, Discharge, Rating curves	(Carter and Anderson 1963; Rantz and others 1982; Sauer and Meyer 1992)

<u>3.2 Error propagation</u>

Once individual errors for measurements are estimated, there is the issue of error propagation. Error propagation is the combination of multiple types of error for a given end value. The maximum error for a quantity is always the sum of the individual errors – in the worst case scenario where all errors lead to either overestimation or underestimation of the quantity of interest (Taylor 1997). If errors are independent and random, there is a 50 % probability that the individual errors will be opposite (some underestimating and some overestimating) and at least partially cancel each other (Taylor 1997). Therefore, the combined error is the quadratic sum of the individual errors (Taylor 1997).

Modified from Taylor (1997):
If a quantity, q, is the sum and/or difference of multiple terms: $q = (x + \cdots + z) - (u + \cdots + w)$, then the uncertainty in the value of q, δq, is:

$$\delta q \begin{cases} = \sqrt{(\delta x)^2 + \ldots + (\delta z)^2 + (\delta u)^2 + \ldots + (\delta w)^2} \\ \leq \delta x + \ldots + \delta z + \delta u + \ldots \delta w \end{cases}$$

If a quantity, q, is the product and/or quotient of multiple terms: $q = \dfrac{x \times \cdots \times z}{u \times \cdots \times w}$, then:

$$\frac{\delta q}{|q|} \begin{cases} = \sqrt{\left(\dfrac{\delta x}{x}\right)^2 + \cdots + \left(\dfrac{\delta z}{z}\right)^2 + \left(\dfrac{\delta u}{u}\right)^2 + \cdots \left(\dfrac{\delta w}{w}\right)^2} \\ \leq \dfrac{\delta x}{|x|} + \cdots + \dfrac{\delta z}{|z|} + \dfrac{\delta u}{|u|} + \cdots \dfrac{\delta w}{|w|} \end{cases}$$

Example: Error in channel width from coordinates

If channel width is measured from coordinates, taken for example with a total station, instead of with a tape measure, then the error in the channel width is a function of the error in the coordinates for each end point. To establish the random error in each end point (assuming systematic error is either smaller than precision or consistent for both end points), the coordinates for each point were taken 32 times. Assuming a normal frequency distribution, the error in x and y coordinates for each end point can be calculated as the standard deviations for each.

End point	Error/standard deviation in x (δx)	Error/standard deviation in y (δy)
River left (RL)	0.15 m	0.09
River right (RR)	0.1 m	0.12

Since the random errors which create a frequency distribution of x and y coordinates for the end points are due largely to decision-making and rod placement by the surveyor, they can be assumed to be random and independent for each direction and each end point. Based on that assumption, the quadratic sum of the individual errors for each direction and both endpoints can be calculated as the total error for the channel width.

Error in x (δx) = $\sqrt{\delta_{RL}^2 + \delta_{RR}^2} = \sqrt{0.15^2 + 0.09^2} = 0.17$ (versus 0.24 for the simple sum)

Error in y (δy) = $\sqrt{\delta_{RL}^2 + \delta_{RR}^2} = \sqrt{0.1^2 + 0.12^2} = 0.16$ (versus 0.22 for the simple sum)

Error in channel width (δw) = $\sqrt{\delta^2 + \delta^2} = \sqrt{0.17^2 + 0.16^2} = 0.2$ m (versus 0.4 for simple sum)

4.0 Additional Information

Taylor (1997) has a very coherent and applicable textbook on error analysis for physical measurements: An introduction to error analysis: the study of uncertainties in physical measurements. Sausalito, Calif., University Science Books.

The International Organization for Standardization (1995) has created the Guide to the Expression of Uncertainty in Measurement which has the official definitions of terms, worldwide standard for methods to determine and express uncertainty, and detailed explanations of the methods themselves, the reasons for the methods, and examples of the methods being put to use.

5.0 References

Brasington, J., Rumsby, B. T., and McVey, R. A. (2000). "Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey." *Earth Surface Processes and Landforms*, 25(9), 973-990.

Carter, R. W., and Anderson, I. E. (1963). "Accuracy of current-meter measurements." *American Society of Civil Engineers Journal*, 105-115.

Cheng, F., and Granata, T. (2007). "Sediment transport and channel adjustments associated with dam removal: Field observations." *Water Resources Research*, 43(W03444), 14.

Coleman, H. W., and Steele, W. G. (1999). *Experimentation and uncertainty analysis for engineers*, Wiley, New York.

Dieck, R. H. (2006). *Measurement uncertainty: methods and applications*, Instrument Society of America, Research Triangle Park, N.C.

Downward, S. R. (1995). "Information from topographic survey." Changing river channels, A. M. Gurnell and G. E. Petts, eds., Wiley, New York, 303-323.

Ferguson, R. I., and Paola, C. (1997). "Bias and precision of percentiles of bulk grain size distributions." *Earth Surface Processes and Landforms*, 22(11), 1061-1077.

Fripp, J. B., and Diplas, P. (1993). "Surface sampling in gravel streams." *Journal of Hydraulic Engineering*, 119(4), 473-490.

Green, J. C. (2003). "The precision of sampling grain-size percentiles using the Wolman method." *Earth Surface Processes and Landforms*, 28(9), 979-991.

Hubbard, B., and Glasser, N. F. (2005). *Field techniques in glaciology and glacial geomorphology*, John Wiley & Sons, Chichester, West Sussex, England ; Hoboken, NJ.

International Organization for Standardization. (1995). *Guide to the expression of uncertainty in measurement*, International Organization for Standardization, Switzerland.

Li, J., Herlihy, A., Gerth, W., Kaufmann, P., Gregory, S., Urquhart, S., and Larsen, D. P. (2001). "Variability in stream macroinvertebrates at multiple spatial scales." *Freshwater Biology*, 46(1), 87-97.

Rabinovich, S. G. (2000). *Measurement errors and uncertainties: theory and practice*, AIP Press, New York.

Rantz, S. E., and others. (1982). "Measurement and computation of streamflow: volume 1. Measurement of stage and discharge." *2175*, United States Geological Survey, Washington, D. C.

Reid, S. M., Yunker, G., and Jones, N. E. (2009). "Evaluation of single-pass backpack electric fishing for stream fish community monitoring." *Fisheries Management and Ecology*, 16(1), 1-9.

Rosenberger, A. E., and Dunham, J. B. (2005). "Validation of abundance estimates from mark–recapture and removal techniques for rainbow trout captured by electrofishing in small streams." *North American Journal of Fisheries Management*, 25(4), 1395-1410.

Sauer, V. B., and Meyer, R. W. (1992). "Determination of error in individual discharge measurements." U. S. Geological Survey, Norcross, Georgia.

Simonson, T. D., and Lyons, J. (1995). "Comparison of catch per effort and removal procedures for sampling stream fish assemblages." *North American Journal of Fisheries Management*, 15(2), 419-427.

Taylor, J. R. (1997). *An introduction to error analysis: the study of uncertainties in physical measurements*, University Science Books, Sausalito, Calif.

Unwin, D. J. (1995). "Geographical information systems and the problem of 'error and uncertainty'." *Progress in Human Geography*, 19(4), 549-558.

Whitacre, H. W., Roper, B. B., and Kershner, J. L. (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., Anthony, D. J., Madsen, S. W., and Thompson, D. M. (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.

Zimmermann, A., Coulombe-Pontbriand, M., and Lapointe, M. (2005). "Biases of submerged bulk and freeze-core samples." *Earth Surface Processes and Landforms*, 30(11), 1405-1417.