US Army Corps
of Engineers

Water Quality Technical Note AM-03
January 1998

# *Statistical Verification of Mean-Value Fixed Water Quality Monitor Sites in Flowing Waters*

*by Michael Vorwerk, Joe Carroll, and John Lemons*

## Purpose

This technical note describes a method for verifying the representativeness of mean-value and extreme-value water quality monitoring locations. Recommended techniques are illustrated using data collected with the total dissolved gas monitoring system on the Columbia and Snake Rivers. This technical note shows how statistical techniques can be applied to the design of monitoring systems to ensure that data collected are representative and thus scientifically defensible.

## Background

Water quality managers must carefully choose the locations for fixed water quality monitors, to ensure that the data they collect accurately reflect water quality conditions of the water of interest. Often, a monitor site will experience some spatial or temporal bias, and data collected there will not represent the release or river in question.

For rivers and hydroproject releases, bias may be the result of combined spill and generation releases (Lemons, Vorwerk, and Carroll 1996), releases into lacustrine tailwaters (Vorwerk and Carroll 1994), generation drawing water from a forebay with heterogeneities (Lemons and others 1996), point sources of pollution, or other processes (Vorwerk, Jabour, and Carroll 1996). A monitor system intake may be located in some portion of a flow and accurately measure its water quality, while not reflecting the quality of other portions (Figure 1). Thus, to provide usable data for operation, regulatory, or background monitoring needs, a manager must verify the representativeness of monitor sites with regard to the monitoring program goals.

This verification must include quantification of the spatial and temporal similarity between water quality data gathered at the monitor site and in the stream or river in question. Flowing water monitor systems can be designed to create temporal records of water quality information as either means or extreme values (Ward 1979). Different verification techniques are necessary for each of these designs. This technical note discusses the techniques necessary to verify mean-value monitor systems.

| Report Documentation Page | *Form Approved* *OMB No. 0704-0188* |
|---|---|



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| **JAN 1998** | | **-** |

| 4. TITLE AND SUBTITLE | |
|---|---|
| **Statistical Verification of Mean-Value Fixed Water Quality Monitor Sites in Flowing Waters** | 5a. CONTRACT NUMBER |
| | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | 5d. PROJECT NUMBER |
| | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Engineer Waterways Experiment Station,3909 Halls Ferry Road,Vicksburg,MS,39180-6199** | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT

**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT

**see report**

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | |
| **unclassified** | **unclassified** | **unclassified** | | **14** | |

**Standard Form 298 (Rev. 8-98)**
Prescribed by ANSI Std Z39-18

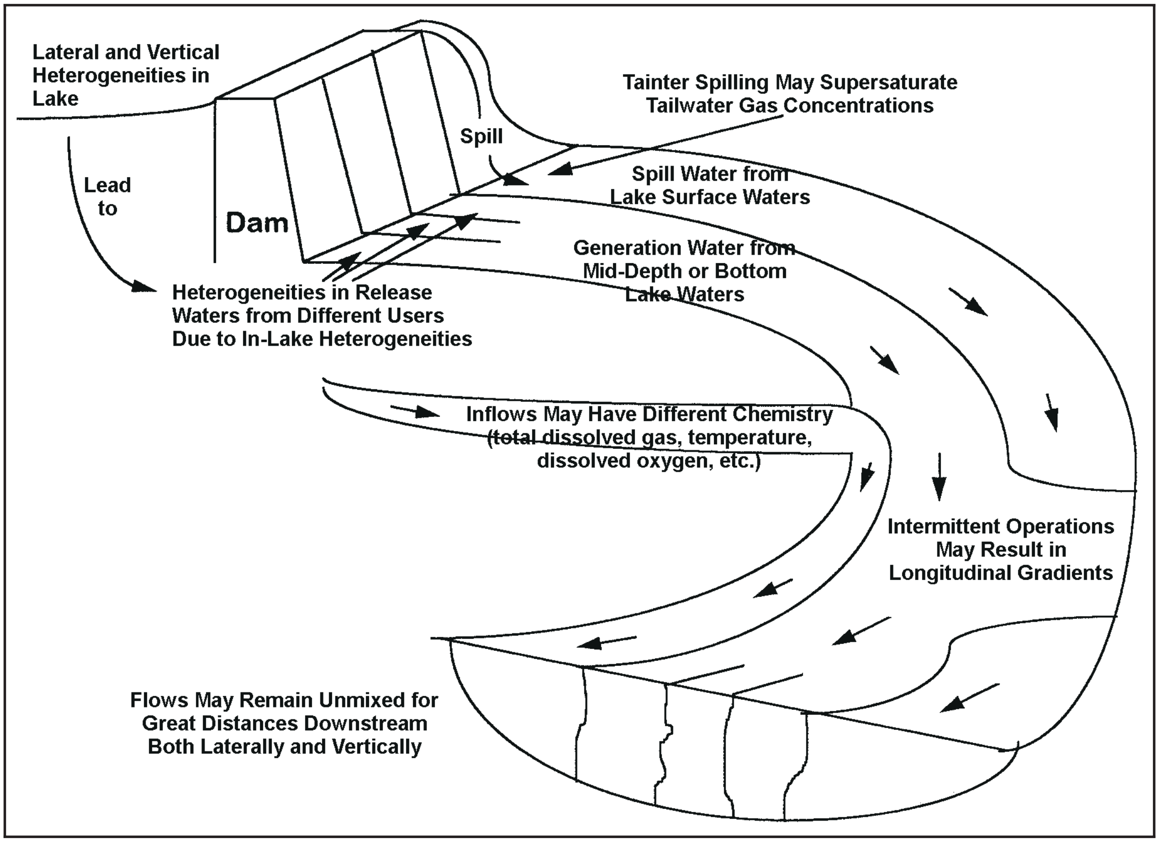



**Figure 1. Possible sources of heterogeneities in flowing water**

To obtain mean values of water quality parameters in flowing water, the analyst must have some knowledge of the mixing processes that are present. *In situ* data are needed for the verification. If the stream is turbulent and well mixed, it may be the case that any location can accurately represent the quality of the water. If the stream is not well mixed and has heterogeneities in water quality, the data must be flow weighted.

Flow-weighted data allow one to calculate the mass transport of parameters through the cross section of the stream in time. Some examples of flow-weighting include temporal quantification of dissolved oxygen mass or average dissolved oxygen concentration moving down a river, a record of average total dissolved gas saturation, mass transport of nutrients, or a record of average temperature. The important aspect is that the value of the parameter of interest is averaged across the area of the channel cross section with respect to velocity.

Any verification must be both qualitative and quantitative. This technical note describes approaches for statistically quantifying and verifying the adequacy of monitoring sites for measuring the average water quality at river transect. Total dissolved gas data collected from the Columbia and Snake Rivers are used to illustrate these techniques. The statistical methods provided will allow users with a basic knowledge of statistics to design and implement studies to verify the representativeness of their own monitor locations. A review of statistics with water quality applications can be found in Gaugush (1986). It should be noted that, although this technical note is based on the use of automated fixed water quality monitors, the procedure described can be applied to manual monitoring as well.

## Approach for Mean Data

### Data Collection and Preparation

The basic approach to verifying the representativeness of a monitor site is to compare matched pairs of observations from the monitor and averaged from the flow (Figure 2). These pairs must be taken over as many different times, flow conditions, and water quality variations as possible.

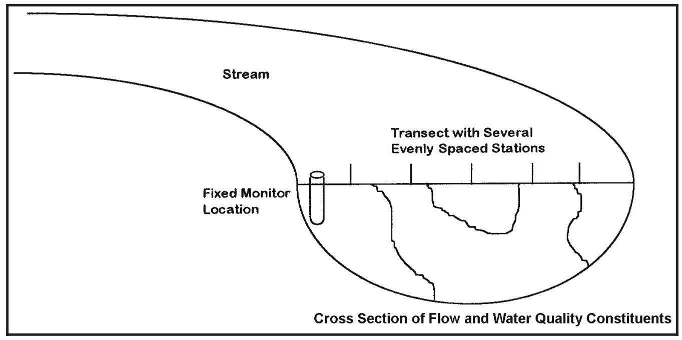



**Figure 2. Cross section of flow with evenly spaced sample stations along a transect**

The observations in the flow must be distributed so they adequately describe water quality conditions across the stream. For wider streams and rivers or for more highly variable water quality conditions, more sample locations are necessary. The sample values from the stream are averaged with an area-weighted average. If velocities vary greatly in the stream cross section, the data averaging must also be flow-weighted. The next section provides details on this weighting.

In practice, data are often limited, and the only available option is averaging the transect data with a simple arithmetic average, and then carrying out the statistical comparison. However, if the stations are not evenly spaced or if the water column has lateral or vertical heterogeneities in water quality or velocity, then a flow-weighted average should be calculated.

### Flow-Weighted Data

The following method can be used to calculate a flow-weighted average. For each sample station, *i*, and depth, *z*, with velocity $U_{i,z}$ and water quality parameter value $P_{i,z}$, assign an area $A_{i,z}$ that the information gathered at that location represents (Figure 3). The area can be difficult to calculate and is most often approximated from depth soundings, maps, surveying techniques, global positioning equipment, and "best-guess."

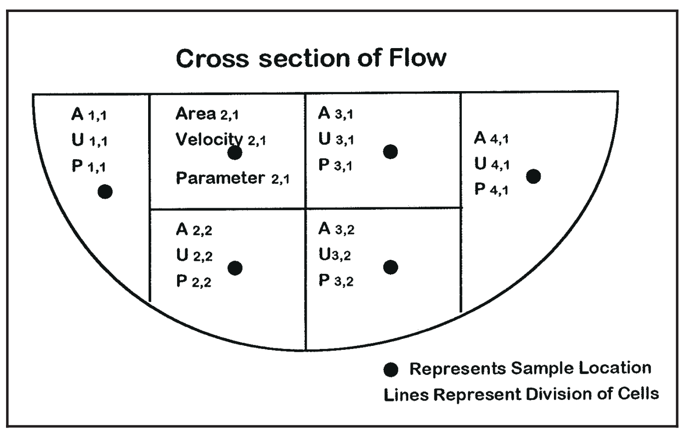



**Figure 3. Hypothetical sample scheme for flow-weighting data**

The transect flow-weighted average of the parameter *P* can then be expressed as

$$\overline{P} = \frac{\sum_{all\,i} \sum_{z=0}^{z\,max} A_{i,z} U_{i,z} P_{i,z}}{\sum_{all\,i} \sum_{z=0}^{z\,max} A_{i,z} U_{i,z}} \qquad (1)$$

Averaging should be carried out for each sampling time. Again, only in the most carefully designed and executed studies is such information available. More typically, an analyst may have three to seven lateral measurements along a transect to compare with fixed monitor information. In this case, the analysis can be performed, but the analyst must be aware that those limited data lessen the weight that may be given to any conclusions.

### Statistical Comparison

At this point, the verification data set should contain *n* pairs of data ($X_{m,j}$ , $X_{s,j}$), each containing a monitor observation $X_{mj}$ and an average stream value $X_{sj}$ for time *j*, where *m* and *s* indicate that the observation came from the monitor or stream, respectively. Next, one tests the relationship between the two locations using a paired t-test (following Hines and Montgomery 1980). This test assumes that the samples each come from a normally distributed, independent distribution. However, moderate departures from normality should not adversely affect the analysis (Pollard 1977). The difference between each pair of observations, $D_j = X_{mj} - X_{sj}$ , should come from a normally distributed independent distribution.

To verify that the data come from a normally distributed population, either of two methods can be used. The easiest method is to plot the data on normal probability paper or use a statistics software package to generate a normal probability graph. A second method is to use a quantitative test such as the Kolmogorov-Smirnov test or Lillefore’s test. Further details of these tests can be found in Hines and Montgomery (1980) and Pollard (1977). Within this technical note, normal plots are used; these were generated using SPSS (SPSS, Inc., Chicago, IL), a statistical analysis software package.

Once it has been determined that the data come from a normal or nearly normal distribution, one can begin the comparison by stating the hypotheses. The null hypothesis is that the mean of the differences between pairs, $\mu_D$ , is zero. This implies that monitor value agrees with stream values and is representative. The alternative hypothesis is that the mean of the differences is not zero; that is, the monitor values do not agree with stream values and are not representative. This is stated as follows:

$$H_0 : \mu_D = 0 \tag{2}$$

$$H_1 : \mu_D \neq 0 \tag{3}$$

These hypotheses are tested with the following statistic:

$$t_0 = \frac{\overline{D}}{S_D / \sqrt{n}}, \tag{4}$$

where

$$\overline{D} = \frac{\sum_{j=1}^{n} D_j}{n}, \tag{5}$$

$$S_D^2 = \frac{\sum_{j=1}^{n} D_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} D_j\right)^2}{n-1}, \tag{6}$$

and we reject $H_0$ if $t_0 > t_{\alpha/2, n-1}$ or if $t_0 < t_{\alpha/2, n-1}$. The confidence level, $\alpha$, is typically taken to be 0.05 and is the type I error, or the probability of rejecting $H_0$ when $H_0$ is true. If $H_0$ is rejected, we conclude that the fixed monitor system does not represent the water quality of the stream at the $\alpha$ confidence level. If $H_0$ is not rejected, we conclude that "we have not found sufficient evidence to reject $H_0$" (Hines and Montgomery 1980). This may be because the monitor site accurately represents the stream or because the sample size (that is, the number of comparisons) is so small that not enough data are available to make the stronger conclusion to reject $H_0$. So, for verification, we need a large enough sample size to minimize the type II error (that is, the probability of accepting $H_0$ when $H_0$ is false).

Similarly, one-sided hypotheses can be tested as follows:

$$H_0{:}\mu_D \le 0,\ H_1{:}\mu_D > 0, \text{ reject } H_0 \text{ if } t_0 > t_{\alpha,n\text{-}1} \tag{7}$$

$$H_0{:}\mu_D \ge 0,\ H_1{:}\mu_D < 0, \text{ reject } H_0 \text{ if } t_0 < t_{\alpha,n\text{-}1} \tag{8}$$

### Determining the Power of the Test

The rejection of the null hypothesis is considered a "strong" conclusion because we control the type I error (choice of $\alpha$), or the probability of rejecting $H_0$ when $H_0$ is true. On the other hand, the acceptance of the null hypothesis is considered to be a "weak" conclusion, because we do not control the type II error ($\beta$), or the probability of accepting $H_0$ when $H_0$ is false.

Thus, to determine the meaning of our conclusion when we accept the hypothesis that a monitor represents the flow, we must determine the type II error. For the monitor location to be acceptable, the type II error must be acceptably small.

To estimate the type II error, or $\beta$, a statistic $d$ is calculated, and with $\alpha$ and $n$, $\beta$ can be determined from operating characteristic charts available in statistics books (Hines and Montgomery 1980, p 604). Using Equations 5 and 6, we calculate $d$ as follows:

$$d = \frac{\left|\overline{D}\right|}{S_D} \tag{9}$$

Once $\beta$ is found, the probability of correctly accepting $H_0$ is the power, namely $P = 1 - \beta$. Because we want only to correctly accept $H_0$, we desire the power to be as close to 1 as possible. The question then becomes, What's good enough?

Since we typically choose $\alpha$ to be 0.05, it seems reasonable to attempt to hold $\beta$ to a similar probability. However, because we have no direct control over $\beta$, probabilities less than 0.2 are probably sufficient. Thus, we consider comparisons with the power greater than 0.8 to be acceptable.

If we are designing a verification study, pilot studies, such as the one described in examples 1 and 2, provide *a priori* knowledge of $\overline{D}$ and $S_D$. This information can be used to design the verification study with a sample size large enough to ensure that the power is as great as desired. This is accomplished through increasing the sample size until the desired value for $\beta$ is achieved on the operating characteristic curve.

## Example 1: Columbia River Camas/Washougal Station—Hand Calculation

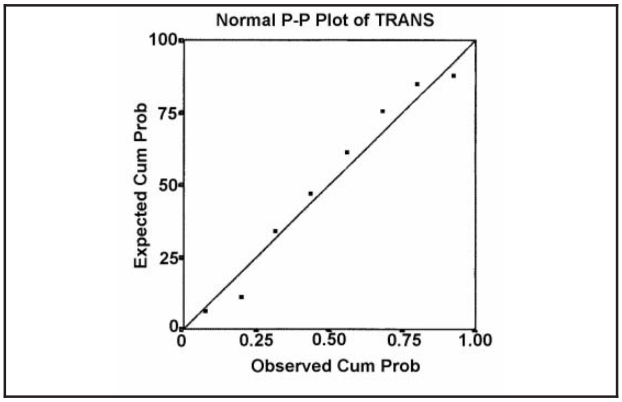
The following example illustrates this method with data from the Camas/Washougal total dissolved gas monitoring station (CWMW) on the Columbia River. To assist smolt in their downstream migration, the U.S. Army Corps of Engineers spills surface water from projects on the Columbia and Snake Rivers. This spillage causes air to be driven into the water column to depths where it causes gases in the water column to be supersaturated with respect to surface saturation. This supersaturation can be detrimental to fish, so the Corps monitors spill gas concentrations in the rivers. Thus, this system is designed to determine the extreme total dissolved gas concentrations resulting from spilling water. This information is used for compliance and in project operations.

To determine if these monitors could be used to determine the flux of total dissolved gas in the river, the statistical verification studies presented in this technical note were carried out. The verification is based on comparing monitor data with data collected at eight transects near the CWMW monitor site (river mile 122) on 3 days (Table 1). The stations on the transects were approximately evenly spaced, so the data for each transect were simply averaged together to obtain an average total dissolved gas concentration at that transect.

**Table 1**
**Average Total Dissolved Gas as Percent Saturation, Columbia River Transects and Camas/Washougal Monitoring Station Fixed Monitor**

| Date | Transect Mile | Percent Saturation | | No. Samples |
|---|---|---|---|---|
| | | Transect Average | Monitor | |
| 18 May 95 | 119.9 | 115.1 | 113.4 | 5 |
| 25 May 95 | 121.2 | 118.1 | 115.5 | 5 |
| 25 May 95 | 121.6 | 119.0 | 117.3 | 5 |
| 25 May 95 | 122.1 | 119.4 | 118.5 | 5 |
| 25 May 95 | 119.9 | 117.0 | 113.4 | 7 |
| 27 Jul 95 | 121.2 | 112.1 | 109.8 | 32 |
| 27 Jul 95 | 121.6 | 116.0 | 111.9 | 15 |
| 27 Jul 95 | 122.1 | 112.9 | 109.5 | 15 |

Figures 4 and 5 show normal probability plots of the transect and fixed monitor system data, respectively. Ideally, the data would be randomly distributed along the normal distribution line, with points close to and on either side of the line. Though the transect data in Figure 4 do not appear to be completely random about the normal line, they are sufficiently normal for this



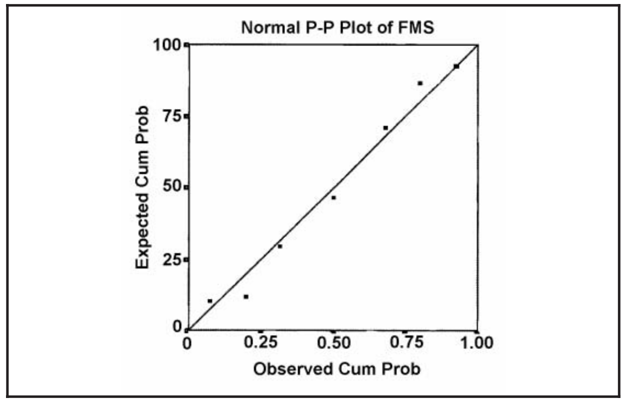




**Figure 4. Normal probability plot of transect data. (Straight line plots the normal distribution; square symbols are the observed data.)**

**Figure 5. Normal probability plot of fixed monitor station data. (Straight line plots the normal distribution; square symbols are the observed data.)**

analysis. The data essentially fit the normal distribution line, but show a trend to be above the line for higher cumulative probabilities and below the line for lower cumulative probabilities. We conclude that the data are approximately normally distributed.

Figure 5 suggests that the fixed monitor data are also normally distributed. Note that the data in Figure 5 are somewhat more randomly distributed on each side of the normal line, with fewer "runs" or continual observations on one side or the other of the normal line. This graphically based determination is subjective. To lessen subjectivity, tests as discussed above can be used (Kolmogorov-Smirnov, Lillefore's, etc.), but the analyst is often forced to use whatever data are available.

Because the data were collected for another study and not specifically for monitor verification, the transect locations did not coincide exactly with the monitor location. For our comparison, all transects that were within 3.5 km of the fixed monitor station were selected. The number of samples varied with transect mile and date. The May samples had five or seven evenly spaced measurements at a constant depth of 4.6 m. July samples had multiple depths and five to seven sample locations. The calculations of the differences, the square of the differences, and the totals of the two sites are depicted in Table 2.

**Table 2**
**Differences (D), Squares of Differences ($D^2$), and Totals for Data Specified in Table 1 (Sample Size, n = 8)**

| Date | Transect Mile | D | $D^2$ |
|---|---|---|---|
| 18 May 95 | 119.9 | 1.7 | 2.9 |
| 25 May 95 | 121.2 | 2.6 | 6.8 |
| 25 May 95 | 121.6 | 1.7 | 2.9 |
| 25 May 95 | 122.1 | 0.9 | 0.8 |
| 25 May 95 | 119.9 | 3.6 | 13.0 |
| 27 Jul 95 | 121.2 | 2.3 | 5.3 |
| 27 Jul 95 | 121.6 | 4.1 | 16.8 |
| 27 Jul 95 | 122.1 | 3.4 | 11.6 |
| **Total** | | **20.3** | **60.1** |

Figure 6 is a normal probability plot of the differences between the transect and fixed monitor data pairs. Though the data show some tendency to be lower than the normal plot for low probabilities and higher than the normal plot for high probabilities, the data appear to be approximately normally distributed.

Equations 2 and 3 were used to test whether the data collected at the fixed monitor site represent the water quality within the river. First, the parameters necessary for the test statistic were calculated.

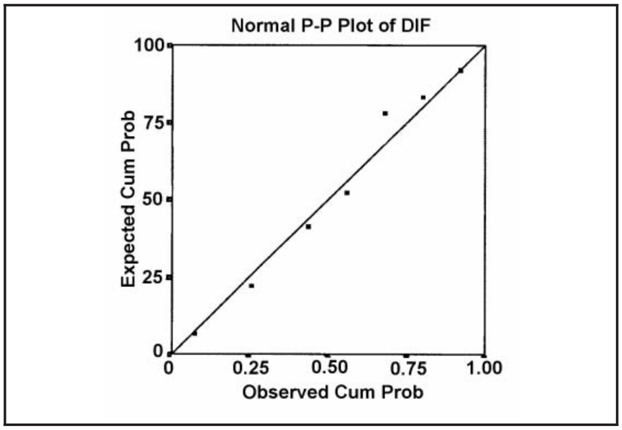



**Figure 6. Normal probability plot of the differences between transect and fixed monitor station pairs of observations. (Straight line plots the normal distribution; square symbols are the differences.)**

The mean difference (Equation 5) was

$$\overline{D} = \frac{\sum_{j=1}^{n} D_j}{n} = \frac{20.1}{8} = 2.5 \tag{10}$$

The variance was estimated using Equation 6:

$$S_D^2 = \frac{\sum_{j=1}^{n} D_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} D_j\right)^2}{n-1} = \frac{60.1 - \frac{1}{8}(20.1)^2}{8-1} = 1.4 \tag{11}$$

The test statistic, $t_0$, was then calculated using Equation 4:

$$t_0 = \frac{\overline{D}}{S_D / \sqrt{n}} = \frac{2.5}{1.2 / \sqrt{8}} = 5.9 \tag{12}$$

Next, the test statistic calculated in Equation 12 was compared with $t_{\alpha/2, n-1}$. This value can be found in various statistics books in the Students' *t* table or *t* distribution table (Hines and Montgomery 1980, p 596). For $\alpha = 0.05$ (our choice) and $\upsilon = n - 1 = 7$ (determined by the sample size of 8), the value of $t_{\alpha/2, n-1} = t_{0.025,7} = 2.365$ (from tables). Then, since

$$5.9 = t_0 > t_{\alpha/2, n-1} = t_{0.025,7} = 2.365 \tag{13}$$

we rejected $H_0$ and concluded that the difference between the transect values and the fixed monitor station values was not zero. The fixed monitor did not adequately represent the water quality in the river at this location.

We tested the hypothesis that the transect % TDG values (total dissolved gas as percent saturation) were greater than the fixed monitor % TDG values using Equation 7. We hypothesized that $H_0{:}\mu_D \leq 0$ with alternative $H_1{:}\mu_D > 0$. We rejected $H_0$ if $t_0 > t_{\alpha,n-1}$. Again, using $\alpha = 0.05$ and $\upsilon = n - 1 = 7$ (determined by the sample size of 8), the value of $t_{\alpha,n-1} = t_{0.05,7} = 1.895$.

Also, since $5.9 = t_0 > t_{\alpha,n-1} = t_{0.05,7} = 1.895$, we rejected $H_0$ and concluded that the difference between the transect values and the fixed monitor station values was greater than zero. The fixed monitor consistently recorded total dissolved gas percent saturation values that were less than the average of those actually present in the river at this location during this study. Thus, we conclude that the fixed monitor system does not accurately represent the flux of total dissolved gas in the river.

To avoid tedious hand calculations, software packages are useful for calculating the paired t-test statistics for the data sets. Two commonly used packages are SPSS (SPSS Inc., Chicago IL) and SAS (SAS Institute, Inc., Cary, NC).

## Example 2: Columbia and Snake Rivers Fixed Monitoring System

The technique employed in the above example can be used to look at an entire monitoring system. Though the fixed monitoring system is designed to determine extreme concentrations of total dissolved gas in spill waters, here we explore the potential of each station for use in monitoring the average total dissolved gas concentration in the river. The system consists of monitors at 26 sites. As in example 1, these fixed monitor sites were compared with transect data collected during 1995.

Again, because the transect study was designed to aid modelers and not strictly to verify the fixed monitor system, adequate data were not available for each location. The analysis shown here was intended only to provide insight into the representativeness of the monitoring system. Details, such as verifying normality, have been omitted. The results presented here might best be used to design future, more rigorous verification studies.

### Data Collection

Transects within 3.5 km of each fixed monitor site were used for comparisons to fixed monitor data. This created a larger data set than if only transects that were adjacent to the monitor sites were used. Larger data sets reduce the type II error, that is, the probability of accepting $H_0$ when $H_0$ is false. The paired test requires at least two pairs of data for each site. This constraint eliminated three stations, leaving 23 for further possible analysis.

### Results

Because of the number of comparisons that were desired, SPSS was used to analyze the data. A paired t-test was run on each of the 23 fixed monitor sites and their comparable transect data. The results of these analyses are shown in Table 3.

**Table 3**
**Verification of Fixed Monitor Station Location with Transect Data**

| Station | FMS Mean* | Trans. Mean* | Dif. Mean* | Dif. s.d.* | T Value | d.f. | 2-Tail Sig. | Relationship** |
|---|---|---|---|---|---|---|---|---|
| BON@ | 108.5 | 111.3 | –2.8 | 0.4 | –16.1 | 3 | 0.001 | FMS > Transect |
| CWMW@ | 113.7 | 116.2 | –2.6 | 1.1 | –6.6 | 7 | 0.000 | Transect > FMS |
| HPKW# | 113.7 | 116.0 | –2.3 | 4.0 | –0.8 | 1 | 0.565 | Accept Null Hypoth. |
| IDSB@ | 126.8 | 120.9 | 5.9 | 6.8 | 3.1 | 12 | 0.009 | FMS > Transect |
| IDSW@ | 126.9 | 120.9 | 6.1 | 6.0 | 3.7 | 12 | 0.003 | FMS > Transect |
| IHR# | 111.8 | 111.9 | –0.1 | 0.5 | –0.3 | 2 | 0.794 | Accept Null Hypoth. |
| JDA@ | 107.8 | 106.0 | 1.8 | 0.2 | 15.2 | 1 | 0.042 | FMS > Transect |
| JHAW@ | 109.8 | 106.5 | 3.3 | 3.7 | 2.4 | 6 | 0.053 | FMS > Transect |
| KLAW# | 109.8 | 110.2 | –0.5 | 0.5 | –2.0 | 3 | 0.152 | Accept Null Hypoth. |
| LGNW | 109.3 | 108.9 | 0.4 | 1.8 | 1.0 | 13 | 0.362 | Accept Null Hypoth. |
| LGS@ | 107.3 | 108.1 | –0.7 | 0.0 | –33.7 | 1 | 0.019 | Transect > FMS |
| LGSW | 110.7 | 113.7 | –3.0 | 7.8 | –1.5 | 13 | 0.169 | Accept Null Hypoth. |
| LMNW@ | 117.6 | 113.9 | 3.8 | 0.8 | 13.7 | 7 | 0.000 | FMS > Transect |
| MCPW@ | 117.9 | 115.4 | 2.4 | 2.6 | 3.7 | 15 | 0.002 | FMS > Transect |
| MCQO# | 113.9 | 112.7 | 1.1 | 1.4 | 1.6 | 3 | 0.202 | Accept Null Hypoth. |
| MCQW# | 112.0 | 112.7 | –0.7 | 2.2 | –0.6 | 3 | 0.569 | Accept Null Hypoth. |
| SKAW@ | 112.9 | 114.1 | –1.2 | 1.2 | –2.8 | 7 | 0.026 | Transect > FMS |
| TDA# | 106.0 | 106.2 | –0.1 | 0.8 | –0.2 | 1 | 0.852 | Accept Null Hypoth. |
| TDAB# | 105.8 | 106.2 | –0.3 | 0.7 | –0.7 | 1 | 0.621 | Accept Null Hypoth. |
| TDTO@ | 112.0 | 115.5 | –3.5 | 1.3 | –8.6 | 9 | 0.000 | Transect > FMS |
| WANO# | 106.5 | 106.6 | –0.1 | 0.2 | –0.5 | 2 | 0.682 | Accept Null Hypoth. |
| WRNB | 113.5 | 114.0 | –0.5 | 0.9 | –1.4 | 6 | 0.227 | Accept Null Hypoth. |
| WRNO | 114.4 | 114.0 | 0.5 | 0.8 | 1.7 | 6 | 0.147 | Accept Null Hypoth. |
| AGGR. FILE | 114.8 | 113.9 | 0.9 | 4.6 | 2.6 | 156 | 0.012 | FMS > Transect |

* Variable is total dissolved gas percent saturation.
** Decision made at alpha = 0.05 significance level.
@ Additional study recommended.
\# Additional data collection recommended.

The “Relationship” column was created by comparing the “T value” column ($t_0$) with values from a Students’ $t$ table using the degrees of freedom in the “d.f.” column. First, we tested to see if the difference was zero. If this was not rejected, we labeled the “Relationship” column “Accept Null Hypoth.”

If the null hypothesis was rejected, Equations 7 and 8 were used with the appropriate values from the Students’ $t$ table to determine whether the transect data were greater or lesser than the fixed monitor station (FMS) data. These results were labeled in the “Relationship” column as “Transect > FMS” or “FMS > Transect,” respectively.

For 11 of the 23 stations, the statistical tests rejected the hypothesis that the FMS and transect data were equal. This means that data collected at these FMS sites did not reflect the water quality conditions occurring across the river.

These stations, which had nonequivalent FMS and transect comparisons, are marked with an ampersand. It is recommended that further analysis be conducted on these stations to determine if the fixed monitor system needs to be moved, modified, or increased in scope. It is possible that the differences detected occur uniformly, allowing a simple addition or subtraction from the FMS data to then accurately represent river conditions. If the variance is large temporally or spatially, these stations should be relocated. To ensure the validity of these conclusions, it is generally accepted that a sample size of at least seven is necessary.

At the remaining 12 stations, the null hypothesis that the FMS and transect data were equal was accepted. This may be because the FMS adequately represents the transect, or simply because the limited data did not provide sufficient evidence to reject the null hypothesis. Thus, further analysis is needed to determine whether the monitors represent the flow.

### Determining the Power of the Test

Using Equation 9, we calculated the statistic $d$ for each station where the null hypothesis was accepted. These results are shown in Table 4. The table shows that in no case is the power greater than 0.32. Thus, we conclude that in each case where the conclusion of the test was to accept the null hypothesis (fixed monitor data represents water quality conditions in the river), there are insufficient data to make a reasonable statistical decision.

We next calculated the necessary sample size for each of these 12 stations to obtain the desired target power of 0.8. These values are shown in Table 5. With the exception of stations KLAW and MCQO, the sample sizes are somewhat unrealistic. This occurs because of the relationships between the sample means and standard deviations.

From Equation 9, $d = \frac{|\overline{D}|}{S_D}$. The power of the test relies on this relationship, in addition to the sample size $n$ . In these other stations, the variance is so large compared with the mean that sample sizes are not reasonable. This implies that the fixed monitors are not located in such a way that their values change uniformly with the flow values. Thus, a first step at improving these monitors would be to place them in locations experiencing more uniform changes with flow and to increase the number of fixed monitor locations across the flow.

**Table 4**
**Calculation of Parameters Needed to Determine *d*, β, and the Power of the Test**

| Station | $\overline{D}$ | $S_D$ | $d = \frac{\lvert\overline{D}\rvert}{S_D}$ | *n* | β from Table | Power |
|---|---|---|---|---|---|---|
| HPKW | –2.3 | 4.0 | 0.58 | 2 | 0.94 | 0.06 |
| IHR | –0.1 | 0.5 | 0.20 | 3 | 0.96 | 0.04 |
| KLAW | –0.5 | 0.5 | 1.0 | 4 | 0.74 | 0.26 |
| LGNW | 0.4 | 1.8 | 0.22 | 14 | 0.90 | 0.10 |
| LGSW | –3.0 | 7.8 | 0.38 | 14 | 0.76 | 0.24 |
| MCQO | 1.1 | 1.4 | 0.79 | 4 | 0.79 | 0.21 |
| MCQW | –0.7 | 2.2 | 0.32 | 4 | 0.93 | 0.07 |
| TDA | –0.1 | 0.8 | 0.13 | 2 | 0.97 | 0.03 |
| TDAB | –0.3 | 0.7 | 0.43 | 2 | 0.95 | 0.05 |
| WANO | –0.1 | 0.2 | 0.50 | 3 | 0.92 | 0.08 |
| WRNB | –0.5 | 0.9 | 0.56 | 7 | 0.72 | 0.28 |
| WRNO | 0.5 | 0.8 | 0.63 | 7 | 0.68 | 0.32 |

**Table 5**
**Determination of Sample Size Needed to Obtain Desired Power of 0.8**

| Station | $d = \frac{\lvert\overline{D}\rvert}{S_D}$ | *n* |
|---|---|---|
| HPKW | 0.58 | 28 |
| IHR | 0.20 | 300 |
| KLAW | 1.0 | 10 |
| LGNW | 0.22 | 300 |
| LGSW | 0.38 | 75 |
| MCQO | 0.79 | 15 |
| MCQW | 0.32 | 75 |
| TDA | 0.13 | 400 |
| TDAB | 0.43 | 50 |
| WANO | 0.50 | 32 |
| WRNB | 0.56 | 30 |
| WRNO | 0.63 | 25 |

## Conclusions

This technical note has demonstrated statistical techniques for verifying the representativeness of fixed monitoring systems that monitor mean values of parameters in flowing water. These techniques were illustrated with data collected on the Columbia and Snake Rivers. Based on the criteria detailed in this technical note, a preliminary analysis of the 1995 Columbia and Snake Rivers fixed monitor system data set revealed that none of the fixed monitor systems accurately represented the average river total dissolved gas concentrations. This demonstration was, however, based on limited transect data, which were not specifically collected for the purposes of monitor site verification.

These examples given in this technical note illustrate use of the statistical approach to eliminate the subjectiveness involved in determining whether a monitoring station accurately represents the water quality in a river. The information presented can be used to guide managers to the most problematic locations, so improvements can be made on a “worst-case first” basis. Additionally, pilot studies similar to the ones used to collect the data used in this technical note can be used to help design verification studies to control the power of the test, obtaining the desired trust in the results.

Many other factors, such as cost, ease of accessibility, and equipment availability, contribute to the difficulties in monitor system design and installation. The cost of an intensive analysis like the ones described above may be prohibitive to many water quality managers. However, the ideas presented herein should make the manager more aware of the difficulties involved in collecting representative data and improve the final system design.

## References


Gaugush, R. F., tech. ed. (1986). “Statistical methods for reservoir water quality investigations,” Instruction Report E-86-2, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Hines, W. W., and Montgomery, D. C. (1980). *Probability and statistics in engineering and management science*. John Wiley and Sons, New York, 291-93.

Lemons, J. W., Vorwerk, M. C., and Carroll, J. H. (1996). “Remote monitoring of total dissolved gas: Design, installation, and verification of remote monitoring systems,” *Water Quality ’96, Proceedings of the 11th Seminar*, Miscellaneous Paper W-96-1, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Lemons, J. W., Vorwerk, M. C., Jabour, W. E., and Carroll, J. H. (1996). “Remote downstream monitoring of Savannah River hydropower releases,” Miscellaneous Paper EL-96-5, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Pollard, J. H. (1977). *A handbook of numerical and statistical techniques*. Cambridge University Press, Cambridge, 177-79.

Vorwerk, M. C., and Carroll, J. H. (1994). “Implications of reservoir release and tailwater monitor placement,” *Lake and Reservoir Management* 9(1),172-78.

Vorwerk, M. C., Jabour, W. E., and Carroll, J. H. (1996). "Evaluation of methods for in situ monitoring of releases from hydropower projects," Water Quality Technical Note AM-01, U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Ward, R. C. (1979). "Regulatory water quality monitoring: A systems perspective," *Water Resources Bulletin* 15(2), 369-81.

## Point of Contact

For additional information, contact Dr. Michael C. Vorwerk (kmvor@gorge.net), Mr. John W. Lemons, or Mr. Joe H. Carroll, (541) 298-6656, at The Dalles Dam, U.S. Army Engineer Waterways Experiment Station.

WQRP

WATER QUALITY RESEARCH PROGRAM