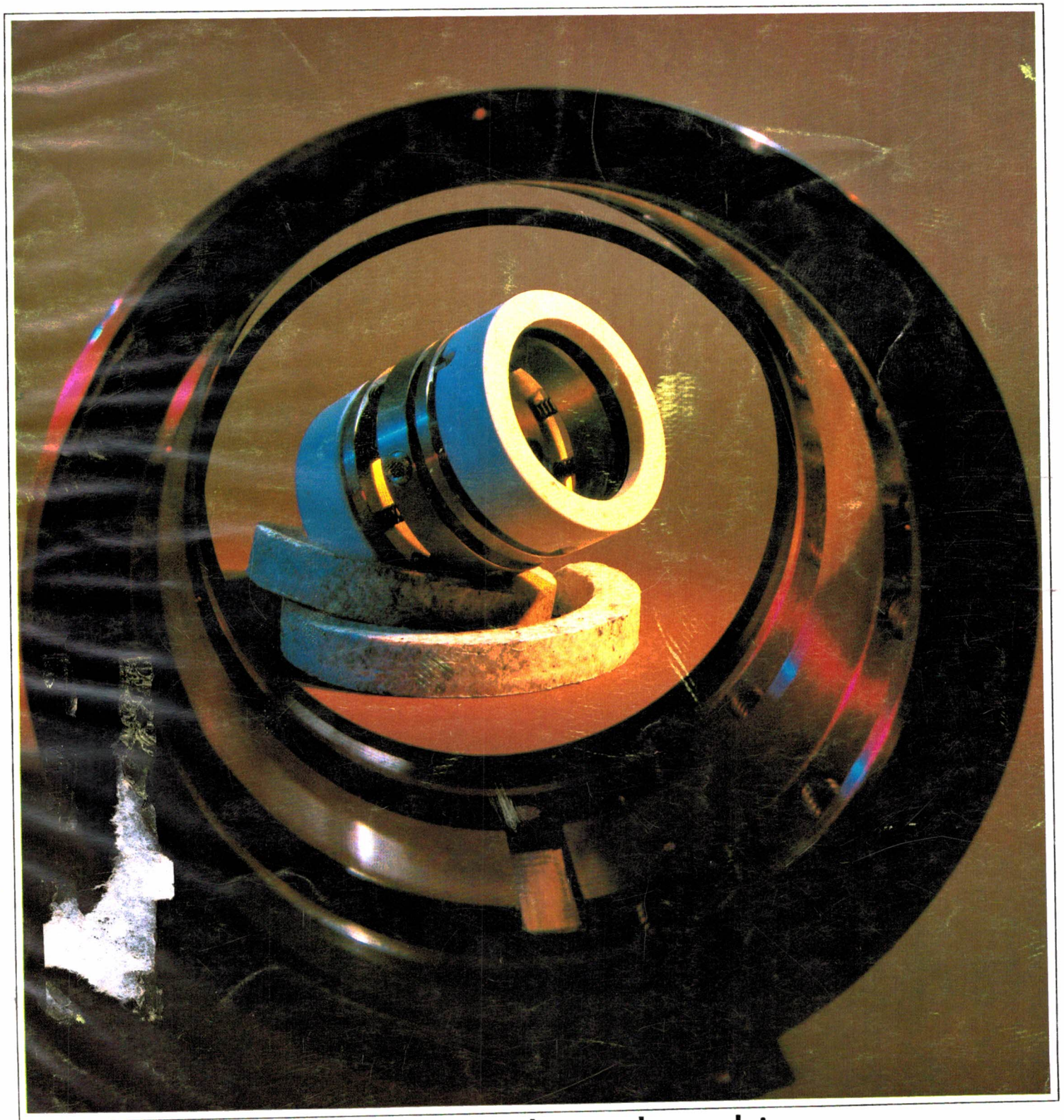
CHEMICAL ENGINEERING

A McGRAW-HILL PUBLICATION
OCTOBER 9, 1978



Prevent leakage with seals and packings

Measurement uncertainties in the pollutant-discharge permit system

A reported discharge having an associated uncertainty greater than 50% of its permit authorization is not unusual. Industry should determine its ability to measure effluent data, so as to recognize whether it is reporting an actual value or an accumulated error. Here are statistical procedures for doing this.

David L. Russell, *Allied Chemical Corp. and* ***James J. Tiede,*** *Hazelton Laboratories America, Inc.*

☐ No matter how carefully we make a measurement, there will be some inaccuracy in it. The magnitude of accumulated inaccuracies associated with sampling, flow measurement and chemical analysis in determining the amount of a pollutant discharged by a plant is of great significance when the measurements are used to judge compliance with a regulation or a permit.

The National Pollutant Discharge Elimination System (NPDES) of the U.S. Environmental Protection Agency does not address the inherent inaccuracies in data measurement. During development of the "Effluent Guidelines," industrial efforts were directed toward the statistical evaluation of plant performance data, with little attention given to the overall accuracy of measurements.

Guideline estimates of plant performance have become "absolute" rules in NPDES permit negotiations, causing hardship to many industries. In one specific case where the Effluent Guidelines were strictly followed, a plant was required to limit its net discharge of a particular pollutant to 0.5 ppm, when the analytical accuracy of the chemical testing method employed was about 0.25 ppm.

To comply with existing and proposed permit limitations, industry needs to: (1) determine and state the limits on its ability to measure effluent data; (2) use these limits to set internal performance standards for pollution control; and (3) attempt to gain recognition of the magnitude of measurement uncertainties in current and future permit negotiations.

Part of a typical NPDES permit for a chemical plant is shown in Fig. 1. It describes effluent characteristics, discharge limitations and monitoring requirements. There are two discharge limits: the "daily average," defined as the daily average for the month (based upon a specified number of samples per month), and the daily maximum. Monitoring requirements specify the frequency of sampling and the type of sample.

An example demonstrating the magnitude of uncertainty associated with measuring the discharge of a

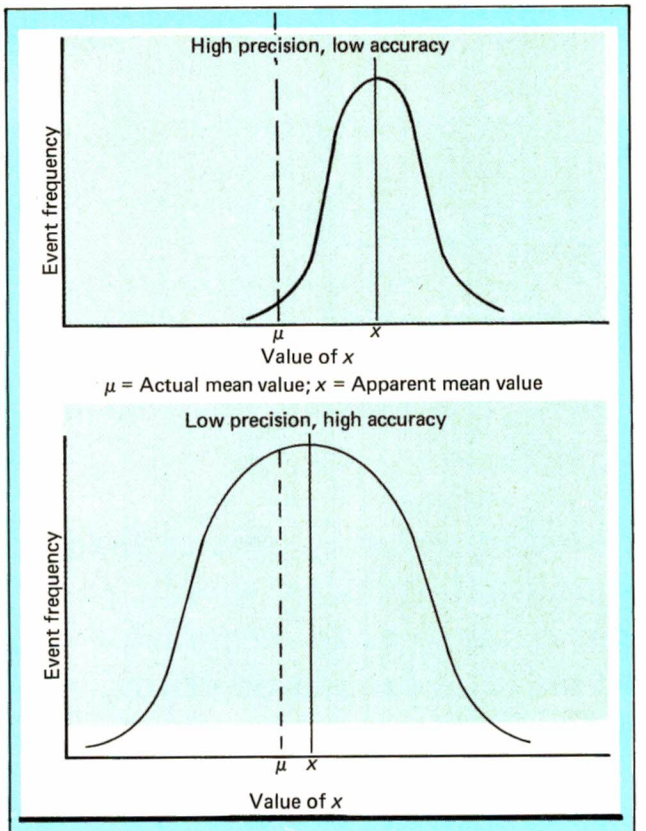
1. During the period beginning . . . . and lasting until . . . . , the permittee is authorized to discharge from . . . .

Such discharges shall be limited and monitored by the permittee as specified below:

Effluent characteristic	Discharge limitations kg/d or lb/d		Monitoring requirements	
	Daily avg	Daily max	Measurement frequency	Sample type
Flow-million gal/d	–	–	daily	continuous
Total suspended solids	–	–	1/wk	24-h composite
Ammonia (as N)	–	–	1/wk	24-h composite
Organic nitrogen	–	–	1/wk	24-h composite
Oil-grease	–	–	1/wk	grab
Temperature	–	–	daily max.	continuous
BODs	–	–	1/wk	24-h composite

2. The pH shall not be less than 6.0 nor greater than 9.0 and shall be monitored daily by means of a grab sample.
3. There shall be no discharge of floating solids or visible foam in other than trace amounts.
4. Samples taken in compliance with the monitoring requirements specified above shall be taken at the following location(s):

Typical NPDES effluent limitations and monitoring requirements **Fig. 1**



A comparison of precision and accuracy **Fig. 2**

pollutant will follow a brief introduction to the required statistical procedures.

Precision and accuracy

Precision is the ability to exactly duplicate the results of a single measurement. Accuracy is the ability to determine the exact or true value. The difference between the two is illustrated in Fig. 2. The precision and accuracy of various analytical instruments and test methods vary widely, depending on factors that are beyond the analyst's control.

Calculation of variance

Variance is the amount of spread or variation in a set of measurements. Small variance indicates precision in data measurement.

Mathematically, variance is the average of the sum of the squares of the differences from the average value in a set of observations.

Let the variance of X be denoted by Var (X), and let the covariance between two values Y and Z be denoted by Covar (Y, Z). For,

$$X = Y - Z \tag{1}$$

$$\text{Var}\ (X) = \text{Var}\ (Y) + \text{Var}\ (Z) - 2\ \text{Covar}\ (Y, Z) \tag{2}$$

If Y and Z are independent, Covar $(Y, Z) = 0$ [1]. Now, if

$$X = (YZ)$$

and Y and Z are independent, then:

$$\text{Var}\ (X) = \text{Var}\ (YZ)$$

Nomenclature

Var (A)	analytical variance
Var (B)	between-day variance
Var (W)	within-day variance
C	concentration, mg/L
D	discharge, lb/d
K	constant, lb/mg/million gal/L
Q	flowrate, million gal/d
S	standard deviation
e	number of days of sampling/mo
g	number of analyses on individual day
m	d/mo
n	analyses/sample
r	samples/d
u	subsamples/d
v	number of days subsamples are collected
x	number of days samples are collected

This can be approximated by a Taylor Series expansion of X as:

$$\text{Var}\ (X) = \left(\frac{dX}{dY}\right)^2_{\mu Z} \text{Var}\ (Y) + \left(\frac{dX}{dZ}\right)^2_{\mu Y} \text{Var}\ (Z) \tag{3}$$

where μ refers to the mean value of the variable given in the subscript.

Net basis calculations

Most NPDES permits are based on net values, an allowance over influent values. Representing Q as flowrate, C as concentration of a particular pollutant, and E and I as effluent and influent values, then the net discharge value, D_{NET}, of the pollutant is:

$$D_{NET} = Q_E\ C_E - Q_I\ C_I \tag{4}$$

If $\bar{X}_K$ and S^2_K denote the estimates of the mean and variance of variable K, then S^2_D, the estimate of the variance of D_{NET}, can be expressed by combining Eq. (2) and (3) as:

$$\text{Var}\ (D_{NET}) = S^2_D = \bar{X}^2_{C_E} \cdot S^2_{Q_E} + \bar{X}^2_{Q_E} \cdot S^2_{C_E} + \bar{X}^2_{C_I} \cdot S^2_{Q_I} + \bar{X}^2_{Q_I} \cdot S^2_{C_I} \tag{5}$$

All covariances are assumed to be equal to zero.

S_D, the square root of the variance, is an estimate of the standard deviation of D_{NET}.

Determination of the daily average

Now let us determine the magnitude of uncertainty in the calculation of the daily average.

Typically, the total daily variance is made up of three components: (1) analytical variance, (2) within-day variance, and (3) between-day variance.

A composite sample of r sub-samples/d taken at fixed time intervals is collected for analysis.

The stream being sampled is a once-through cooling water, with a relatively constant flowrate. A small amount of iron from plant production enters the cooling water from leaking heat exchangers and other minor sources. The feed water also contains some iron. The discharge permit is based on the *net* contribution of

iron, as may be the case for a plant that has quantities of a priority pollutant in its intake or raw water.

Since the plant's cooling-water losses are very small with respect to the flow, the net discharge, D, is equal to the flowrate, Q, times a constant, K, times the difference in concentration, C.

$$D = KQ(C_E - C_I) \quad (6)$$

If Var (A) is the analytical variance, Var (W) the variance within the day, and Var (B) the variance between the days, then the variance in D due to differences in measuring concentration, Var (C), is:

$$\text{Var}\ (C) = \frac{1}{2}\text{Var}\ (A) + \text{Var}\ (W) + \frac{(m-e)}{m}\text{Var}\ (B) \quad (7)$$

The analytical variance is halved because each sample is analyzed twice, and the multiplier on the between-day variance term is a correction for e days of sampling/mo of m days [5].

If a continuous, exactly flow-proportional, sampler were employed, the within-day variance, Var (W), would be deleted. However, if the sampler is not truly flow-proportional, the term must be retained, and the variance calculated.

The variance between the days, Var (B), is deleted for calculation of a particular-day estimate of variance, but must be retained for calculation of the variance of any nonparticular day [2].

Once Var (C_E) and Var (C_I) are calculated, and the variance in D due to inaccuracies in flow measurement, Var (D_Q), has been determined (as shown later), the total variance in D, Var (D_T), can be expressed as:

$$\text{Var}\ (D_T) = S_D^2 = K^2\,Q^2\,[\text{Var}\ (C_E) + \text{Var}\ (C_I)] + K^2\,[\text{Var}\ (Q) \cdot (\bar{C}_E^2 + \bar{C}_I^2)] \quad (8)$$

This is similar to Eq. (5), where $\bar{C}_E$ and $\bar{C}_I$ represent average values for the respective variables.

The standard deviation in the daily average is:

$$S_D \text{ avg.} = \sqrt{\frac{\text{Var}\ (D_T)}{e}} \quad (9)$$

Analytical variance—Var (A)

For a sample taken on day j, multiple analyses are made to determine the concentration C_{ji}, where i is the ith sample on day j. If n analyses are made on the sample, then:

$$\bar{C}_j = \frac{\sum_{i=1}^{n} C_{ji}}{n} = \frac{C_{ji} + C_{j2} + \cdots C_{jn}}{n} \quad (10)$$

If samples are collected for x days, then:

$$\text{Var}\ (A) = \frac{\sum_{j=1}^{x}\sum_{i=1}^{n}(C_{ji} - \bar{C}_j)^2}{x} \quad (11)$$

This is the estimate of analytical variance.

Estimates of the standard deviation associated with particular analysis methods have been published in "Standard Methods" [3], and "Methods for Chemical Analysis of Water and Wastes" [2]. Accuracy and precision data for some water and wastewater methods are shown in Table I.

If the tabular value for relative precision is P_M, and the average value of the measured parameter from analysis is $\bar{X}_M$, then the analytical variance is:

$$\text{Var}\ (A) = \left(\frac{P_M}{100\ \bar{X}_M}\right)^2 \quad (12)$$

Analytical precision and accuracy guidelines for selected water and wastewater parameter concentrations [3] — **Table I**

Parameter and method	Relative precision, % (P_M)	Relative accuracy, %	Concentration
Hardness—EDTA	2.9	99.2	
Metals—atomic absorption			
Cadmium	21.6	91.8	50 μg/L
Chromium	47.8	83.7	110 μg/L
Copper	11.2	96.6	1,000 μg/L
Iron	16.5	99.4	300 μg/L
Lead	23.5	84.0	50 μg/L
Silver	17.5	89.4	550 μg/L
Zinc	8.2	99.6	500 μg/L
Nitrogen			
Ammonia (Nessler)	46.3	89.8	200 μg/L
	18.0	96.0	1,500 μg/L
Nitrate (Brucine)	66.7	92.4	50 μg/L
	5.5	94.0	1,000 μg/L
Organic (Nessler)	94.8	45.0	200 μg/L
	43.1	90.7	1,500 μg/L
Oxygen demand—biochemical	15	N/A	175 mg/L
—chemical	6.5	N/A	200 mg/L
Residue—total	52	N/A	
—suspended	33		15 mg/L
	10		242 mg/L
	0.76		1,707 mg/L

Iron concentration data for 2-day sampling period — **Table II**

Sample number	Concentration (mg/L) 1st day influent	1st day effluent	2nd day influent	2nd day effluent
1	0.40, 0.25 >0.32	0.35, 0.38 >0.37	0.22, 0.25 >0.23	0.60, 0.52 >0.56
2	0.25, 0.30 >0.27	0.84, 0.91 >0.88	0.30, 0.36 >0.33	1.34, 1.30 >1.32
3	0.34, 0.38 >0.36	0.56, 0.62 >0.59	0.34, 0.35 >0.35	1.00, 1.10 >1.05
4	0.45, 0.44 >0.45	0.77, 0.91 >0.84	2.60, 2.70 >2.65	4.45, 4.40 >4.42
5	0.30, 0.36 >0.33	0.52, 0.53 >0.53	1.00, 1.00 >1.00	0.45, 0.49 >0.47
6	0.56, 0.56 >0.56	0.57, 0.57 >0.57	0.88, 0.88 >0.88	0.46, 0.42 >0.44

C_I = 0.644 mg/L, C_E = 1.003 mg/L, n = 2

Summary of calculation of uncertainty for daily average — **Table III**

Step 1 — Calculate from influent and/or effluent data (Table II) Var (A), using Eq. (11): Var (A) = 0.002

Step 2 — Calculate from influent and effluent data:
W_1 and Var (W) using Eq. (13) (14) and (15)
W_1 influent = 0.4201 Var (W_1) = 0.0348
W_1 effluent = 1.1959 Var (W_E) = 0.0995

Step 3 — Calculate between-day variances from Eq. (16) and (17) (data base for calculations not shown)
B_1 influent = 0.1374 Var (B_1) = 0.0329
B effluent = 0.2820 Var (B_E) = 0.0405

Step 4 — Calculate Var (C_1) and Var (C_E) from Eq. (7) for appropriate sampling frequencies

Samples/mo, (e)	Var (C_1)	Var (C_E)	Var (C_I) + Var (C_E)
1	0.0676	0.1397	0.2073
4	0.0679	0.1360	0.2039
8	0.0647	0.1316	0.1963

Step 5 — Calculate standard deviation in the daily average from Eq. (8) and (9), where Q = 4.46, Var (Q) = 0.1563 and K = 8.34

$e = 1$ $S_D = [8.34^2 (4.46^2 (0.2073) + 0.1563 (0.644^2 + 1.003^2))]^{1/2} = 17.38$
$e = 4$ $S_D = [8.34^2 (4.46^2 (0.2039) + 0.1563 (1.4208))/4]^{1/2} = 8.62$
$e = 8$ $S_D = [8.34^2 (4.46^2 (0.1936) + 0.1563 (1.4208))/8]^{1/2} = 5.99$

Step 6 — Calculate confidence intervals for the daily average:

e	S	Degrees of freedom	Student's t	95% confidence interval
1	17.38	–	–	34.76*
4	8.62	3	3.182	27.43
8	5.99	7	2.365	14.17

*Assuming normal distribution

Variation within the day—Var (W)

Our sampler takes r individual (grab) samples/d and collects them as a single composite sample. To obtain an estimate of the within-day variance, Var (W), the single composite must be divided into a number of subsamples. To obtain a final result, two intermediate statistics, W_1 and W_2, must be calculated.

Suppose we collect u subsamples per day for v days, and C_{ab} is the bth subsample collected on the ath day, then $\bar{C}_a$ is calculated according to Eq. (10), and

$$W_1 = \frac{\sum_{a=1}^{v} \sum_{b=1}^{u} (C_{ab} - \bar{C}_a)^2}{v\ (u - 1^*)} \qquad (13)$$

*Degrees of freedom

W_1 contains Var (A), which must be removed. Therefore, W_2 is defined as:

$$W_2 = \frac{(W_1 - \text{Var}\ (A))}{n} \qquad (14)$$

Finally, since a 24-h composite sample is composed of r individual samples, and since we have taken r/u individual samples for each subsample, then:

$$\text{Var}\ (W) = \frac{W_2}{u} \qquad (15)$$

Calculation of uncertainty for influent and effluent daily values for a single day — **Table IV**

Using previously developed data from the first day of sampling, estimate the standard deviation of the discharge.

C_I = 0.383 mg/L for the first day of sampling
C_E = 0.638 mg/L for the first day of sampling
Var (C_I) = Var (W_I) + Var (A)/2 = 0.0358
Var (C_E) = Var (W_E) + Var (A)/2 = 0.1005
where:
Q = 4.46 million gal/d, and Var (Q) = 0.1563,
$S^2 = 8.34^2 (4.46^2 (0.0358 + 0.1005) + 0.1563 (0.383^2 + 0.638^2))$
S^2 = 194.60,
S = 13.95 lb/d

Assuming a normal probability distribution, the true mean net discharge for this particular day has a 95% probability of being in the range of 9.49 lb/d ± 2 (13.95), or between −18.41 and +37.39 lb/d.

Variation between days—Var (B)

If the average concentration of a particular pollutant on an individual day is C_i, and $\bar{C}$ is the mean value of g

analysis determinations on individual days, then B_1, the variance of the concentration between those days, is expressed as:

$$B_1 = \frac{\sum_{i=1}^{g}(C_i - \bar{C})^2}{g - 1^*} \quad (16)$$

*Degrees of freedom

B_1 contains analytical and within-day variances, which must be removed, so Var (B) is:

$$\text{Var } (B) = \frac{B_1 - \text{Var } (A) - \text{n Var } (W)}{n} \quad (17)$$

Variance of flow measurement—Var (D_Q)

A carefully installed flowmeter may have an inaccuracy of about 1.5% of the flow. In an open channel, the inaccuracy may approach 20%. The most frequently overlooked source of open-channel flowmetering error is failure to observe energy and boundary conditions for the installation. Such oversights can greatly change the head-discharge relationship upon which most open-channel flow measurement is based.

A recent example points this out: A surface-level measuring device that relied on a weighted bobbing line to determine depth of flow was installed in an open channel. A theoretical calibration curve, relating depth of flow to discharge, was computed. However, the upstream flow conditions, energy relationships and water velocity in the channel were not considered when installing the meter.

When the flowmeter was found incapable of measuring less than a 25% change in flow, conditions of installation were examined. A steep upstream slope permitted the channel water velocity to reach 20 ft/s. At that speed, the depth-discharge relationship for the channel was found substantially different from the theoretical calibration curve.

The following is a calculation of Var (D_Q):

A particular flowmeter installation on a 90°V notch weir has a head discharge relationship of [4]:

$$Q = 1.62\ H^{2.5} \quad (18)$$

where Q is in million gal/d, and H is in ft. The uncertainty in the reported discharge, Var (D_Q), is actually a combination of the uncertainty in measurement of H and the precision of the flowmetering system, and is expressed as:

$$\text{Var } (D_Q) = \left(\frac{dQ}{dH}\right)^2 S_H^2 + \text{Var (System)} \quad (19)$$

When $H = 1.5$, $S_H = \frac{5}{8}$ in. (0.05 ft) and $Q = 4.46$ million gal/d, then:

$$\text{Var } (D_Q) = 1.62^2\ (2.5\ H^{1.5})^2(0.05)^2 + \text{Var (System)}$$

The variance in the system is generally stated in terms of the precision of mechanical and electrical components, which in this case is estimated as 3% of the flow. Thus:

$$\text{Var } (D_Q) = 0.1384 + [(0.03)(4.46)]^2 = 0.1563 \quad (20)$$

and: $S_Q = [\text{Var } (D_Q)]^{1/2} = 0.395$ or about 8.9%

The estimate of the standard deviation in the measurement of Q is 0.395 million gal/d.

Practical example of analysis of variance

Table II shows concentration data for a two-day sampling period. Table III summarizes the calculation of uncertainty for the daily average, based on the data in Table II and the example just given.

If the NPDES permit-limit daily average is 30 lb/d, and compliance monitoring is done once per month, the standard deviation is ±17.38 lb/d, more than half the daily average. The 95% confidence interval for the daily average is 30 ± 34.8 lb/d, assuming normal distribution.

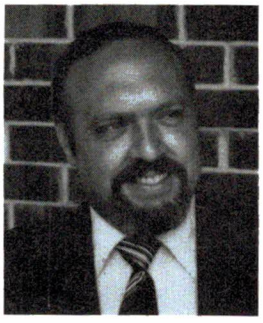
When monitoring is done more frequently, the uncertainty in the average decreases. At 4 samples/mo, the standard deviation is ±8.62 lb/d; at 8 samples/mo, it is 5.99 lb/d. The 95% confidence intervals for 4 and 8 samples/mo are 30 ± 27.43 and 30 ± 14.17 lb/d, respectively.

For our example, the amount of iron discharged during a particular day (a number required by the NPDES permit daily maximum) has a measurement variance of 13.95 lb/d, as calculated in Table IV. For a net discharge of 9.49 lb/d as calculated from Eq. (6), the 95% confidence interval is 9.49 ± 27.9 lb/d, or between −18.41 and 37.29 lb/d. The negative value is possible because this is a net discharge calculation.

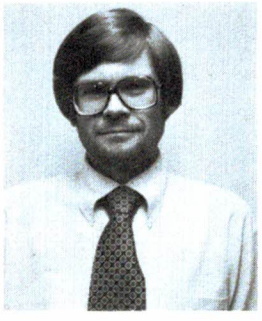
References

1. Ostle, B. and Mensing, R. W., "Statistics In Research," Iowa State Univ. Press, Ames, Iowa, 1975.
2. Methods for Chemical Analysis of Water and Wastes, EPA, USEPA-1625/6-74-003a, 1974.
3. "Standard Methods," 14th ed., American Public Health Assn., New York, 1975.
4. Chow, V. T., "Open Channel Hydraulics," McGraw-Hill, New York, 1959.
5. Cochrane, W. G., "Sampling Techniques," 3rd ed., John Wiley & Sons, New York, 1977.

The authors



David L. Russell is a senior environmental engineer for the Specialty Chemicals Div. of Allied Chemical Corp. at B and A Works, Wilmington Turnpike, Marcus Hook, PA 19061. Prior to joining Allied, he was a private consultant for nine years. He received a B.S. in Civil Engineering from the University of Illinois (Urbana) and holds an M.S. in Environmental Engineering from West Virginia University. Mr. Russell is a registered engineer in Ill. and Pa.



James J. Tiede is a senior biostatistician for Hazelton Laboratories America, Inc., 9200 Leesburg Turnpike, Vienna, VA 22180. At the time this article was written, he was a statistical scientist for the Specialty Chemicals Division of Allied Chemicals Corp., Buffalo, N.Y. Dr. Tiede received his Ph.D. in statistical science from State University of New York at Buffalo and a B.S. in mathematics from St. John Fisher College, Rochester, N.Y.