REGULATING WHOLE EFFLUENT TOXICITY USING “PERCENT EFFECT” AS THE TEST ENDPOINT

Abstract

Traditionally, the No-Observed-Effect-Concentration (NOEC) or Inhibition Concentration (IC-25) are used to report the level of toxicity in effluent samples. These endpoints, and the methods used to calculate them, have several well known limitations that interfere with reasonable potential calculations, discharge limit development and permit compliance determinations. An alternative approach using “percent effect” as the test endpoint provides a more useful and intuitive method for implementing whole effluent toxicity limits in NPDES permits. The alternative procedure has been successfully applied in South Carolina.

I. Background

Most states have a narrative water quality criterion that prohibits the discharge of toxics in toxic amounts. States routinely use whole effluent toxicity (WET) tests to evaluate whether discharges are toxic or not.

An effluent is deemed non-toxic when it produces no statistically-significant negative effect on standard test organisms. The test procedure measures the effect of effluent-exposure on the rate of mortality, growth and reproduction.

In its simplest form the test can be run by comparing only two groups: effluent-exposed organisms and a control group exposed only to laboratory dilution water. The laboratory water is used to culture the test organisms and contains no toxic chemicals.

Most states, however, require that the toxicity test be performed using a range of effluent concentrations. In these tests, the laboratory dilution water is used to dilute the effluent. By comparing the effects at different concentrations, it is possible to confirm the presence of a dose-response relationship. A valid dose-response is an essential element for proving the existence of toxicity.
Testing a dilution series also allows dischargers and the permitting authority to be more precise about the true potential for toxicity in the effluent. A two-group test can only tell us whether the effluent passes or fails. A multi-concentration test can tell us how much dilution, if any, is required to assure that the discharge is not toxic.

The threshold at which the effluent no longer exhibits a statistically-significant negative effect is called the “No-Observed-Effect-Concentration (NOEC).” Most NPDES permits require that the NOEC be greater than or equal to the maximum permitted Instream Waste Concentration (IWC). The IWC is calculated using conservative estimates of the actual dilution available during relatively dry weather conditions.

In 1995, U.S. EPA promulgated standard methods for whole effluent toxicity testing. At that time, the agency recommended replacing the NOEC endpoint with the “inhibition concentration (IC).” The most common version of the new endpoint is the IC-25 or the effluent concentration likely to cause a 25% reduction in the rate of survival, growth or reproduction among test organisms. EPA guidance states that the IC-25 is functionally-equivalent to the NOEC.

The advantage to using the NOEC is that it includes a formal test for statistical significance; the IC-25 procedure does not. However, the NOEC is less precise and very sensitive to the specific dilution series used to test for toxicity. Some states remedy the known deficiencies by requiring both endpoints to fail before a sample is declared to be toxic.

It is important to note that both the NOEC and the IC-25 are expressed as percent of the original effluent sample. If the undiluted effluent causes no statistically-significant reduction in survival, growth or reproduction the NOEC equals 100% and the IC-25 is >100. If the undiluted effluent causes a statistically-significant reduction in survival, growth or reproduction but the negative effects are eliminated when the effluent is diluted in a 1:1 ratio using laboratory control water, then the NOEC equals 50% (reflecting the fact that the effluent is only half-strength). The IC-25 will necessarily fall somewhere between 50 and 100.

The standard dilution series used for most toxicity tests is shown in Figure 1. Usually, the test is performed using a control group, an undiluted effluent group, and four intermediate effluent concentrations (one of which is set to the IWC). The NOEC is calculated by comparing results from each of the test concentrations to the results observed for the control group. The lowest concentration at which a statistically-significant effect appears is called the Lowest-Observed-Effect-Concentration (LOEC). Theoretically, the NOEC is always a smaller concentration than the LOEC and the IC-25 should fall between the LOEC and the NOEC. In practice, this does not always occur.

<table>
<thead>
<tr>
<th>Control Group (0%)</th>
<th>6.25%</th>
<th>12.5%</th>
<th>25%</th>
<th>50%</th>
<th>Undiluted Effluent (100%)</th>
</tr>
</thead>
</table>

Table 1: Example of Standard Dilution Series for WET Testing

© 2001, Risk Sciences
II. Limitations

1) Toxicity Units

The most significant limitation to using the NOEC or IC-25 for expressing toxicity is that these measures are counter-intuitive. As toxicity increases, the NOEC and IC-25 decline. Higher toxicity requires greater dilution to mitigate the adverse effects. Therefore, an effluent with a NOEC of 25% is four times more toxic than an effluent where the NOEC equals 100%.

EPA recommends translating the NOEC or IC-25 into “Toxicity Units” to restore the intuitive relationship between increasing pollution and rising numbers. Toxicity units are calculated by dividing 100 by the NOEC or the IC-25. Therefore, where the NOEC = 25%, the effluent toxicity may also be described as “4 toxicity units” (100 / 25). Where the IC-25 = 17%, the effluent can be described as “5.88 toxicity units” (100 / 17). In essence, toxicity units are a measure of the amount of dilution required to render a given effluent sample non-toxic (expressed as a factor of the original effluent sample).

The problem with using toxicity units is that a non-toxic effluent is described as having 1 toxicity unit not zero toxicity units. This is because the NOEC for non-toxic effluent is 100 and 100 / 100 = 1. The absence of a true zero value in toxicity nomenclature can cause difficulties when calculating whether a given discharge has reasonable potential to cause or contribute to instream toxicity.

For example, dischargers to ephemeral streams must demonstrate that their effluent is non-toxic at the end-of-pipe without benefit of dilution. In such cases, the effective permit limit is 1 toxicity unit. However, as described below, even a discharger who fails no WET tests in the period preceding permit renewal may appear to have reasonable potential for toxicity using the methods outlined in the Technical Support Document for Water Quality-based Toxics Control (TSD; 1991).

In the TSD, EPA recommends that reasonable potential be calculated by multiplying the highest recorded toxicity value (1 T.U.) a factor selected from Table 3-1 or 3-2 in the Technical Support Document (TSD). The lowest factor available in the Tables is 1.1. Therefore, when 1 T.U. is multiplied by the factor of 1.1 the discharger appears to have reasonable potential to exceed the permit limit of 1 toxicity unit despite the absence of any actual WET test failures in the historical record. Were the permit limit and non-failing tests both recorded as zero toxicity units, this problem would not occur.

Using 1 TU or 100 NOEC to represent zero toxicity can cause similar problems when calculating a long-term average for any given discharge. This is especially true when there is no dilution available and the permittee must meet WET limits at the end-of-pipe.
Another problem with using toxicity units is that they are a referenceless standard. Knowing the toxicity units for an undiluted sample does not tell us whether the sample is likely to be toxic after being discharged and diluted in the environment. Some states remedy this problem by altering the formula for calculating toxicity units to divide the IWC by the NOEC or IC-25. This approach expresses the potential for toxicity after accounting for available dilution. However, since dilution varies from location to location, the toxicity units are no longer directly comparable and cannot be added or subtracted from one another as might be done when developing wasteload allocations to implement a TMDL.

2) **Confirming Dose-Response**

The dose-response curve in biological assays is functionally equivalent to the calibration curve in chemical testing. Recent EPA guidance stresses the importance of using the dose-response curve to interpret WET test results, identify anomalous data, and minimize test variability.¹ Although a valid dose-response relationship is essential to corroborate the presence of toxicity, the standard methods for WET testing do not contain any formal procedure to confirm whether a valid dose-response relationship exists. A major deficiency in the use of NOEC/LOEC and IC-25 is that they are often calculated and reported without regard for the true dose-response relationship. And, EPA guidance provides no objective statistical test to confirm the presence of a valid dose-response.

3) **Definition of NOEC**

Some states define the NOEC as the highest concentration at which there is no statistically-significant adverse impact on standard test organisms. Other states define the NOEC as the highest concentration, less than the LOEC, at which there is no statistically-significant adverse impact during the WET test.

In theory, these definitions should mean the same thing because the NOEC is supposed to be less than the LOEC. In practice, this does not always happen. On occasion, the concentrations immediately above and below the LOEC can show no statistically-significant adverse effect. This is also known as an interrupted dose-response curve. When WET test results are used to assess reasonable potential or determine permit compliance the outcome may depend on which definition of NOEC is used to represent the potential for toxicity in the effluent.

¹ see Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the NPDES Program; EPA 833-R-00-003; June, 2000
In figure 2, the LOEC occurs at the 25% effluent concentration. However, some states would define the NOEC as 12.5% effluent and other states would define it at 100% effluent. According to EPA’s new guidance, the effect observed at 25% is probably a statistical anomaly.

**Figure 2: Defining NOEC with an Interrupted Dose-Response Relationship**

The complexity of the interpretation issue is illustrated in figure 3. Figure 3 differs from figure 2 only in the level of effect observed in undiluted effluent. However, viewed in context of all the other data, it now appears that the response observed at the 50% effluent concentration is the anomalous data. That interpretation is also consistent with EPA’s new guidance. But, note the relative subjectivity involved.

These illustrations also show the statistical inefficiency of using NOEC-LOEC to assess toxicity. Although many organisms may be exposed to several different effluent concentrations, the NOEC-LOEC calculations are made by comparing the biological response from a single treatment group to the control group. The IC-25 is also calculated by comparing only 2 test groups (usually the NOEC and the LOEC). Thus, much of the statistical power available is wasted in the system.
4) **Method Detection Level (MDL)**

Another limitation of using NOEC/LOEC or the IC-25 to represent toxicity potential is the assumption that the test method provides equally reliable results throughout the entire dynamic range. Unlike chemical testing, there is no Method Detection Level (MDL) to define the threshold at which WET test procedures can no longer accurately distinguish the presence or absence of toxicity in a sample. All reported results are presumed to be equally reliable because there is no way to develop an MDL for WET based on the NOEC or IC-25 endpoints.

The absence of an MDL is most keenly felt when identical split samples are subjected to separate WET tests (at the same or different laboratories) and the results are incongruent. One test indicates the sample is toxic and the other test indicates the sample is not toxic. Without an MDL, both results are presumed valid despite the obvious paradox. With and MDL it is more likely that the apparent inconsistency can be resolved.
5) **Software Problems**

EPA has noted some special limitations of the IC-25 and the software used to estimate this endpoint. On occasion, the software is unable to generate appropriate confidence levels for the given test data. Usually, this occurs when the statistical algorithms are unable to converge around a single solution. This is a strong indicator of a poor or uncertain dose-response relationship. EPA warns that the confidence levels should not be used until the software errors can be corrected. In the meantime, it is impossible to evaluate the validity of the IC-25 estimate without such confidence levels.

EPA has also warned that techniques used to “smooth” WET test data prior to calculating the IC-25 can introduce an unintended bias into the analysis. Once again, the agency recommends that the data be graphed and carefully reviewed before affirming the presence of toxicity.

Additional bias may be introduced into the analysis by the manner in which EPA’s software uses Monte Carlo resampling techniques to estimate the IC-25. Assuming the sample is truly non-toxic, Figure 4 shows the estimated IC-25 range from the Monte Carlo resampling technique. Note that the mean is 100 as expected.

![Figure 4: Range of IC-25 Responses from a Non-Toxic Sample](image)

---

2 Method Guidance and Recommendations for Whole Effluent Toxicity Testing; EPA-821-B-00-004; July, 2000

By design, EPA’s software discards all estimates that are higher than the highest tested concentration (usually 100). This is done because, logically, values greater than 100% effluent are deemed “non-sensical.” The IC-25 is calculated as the mean of all retained data (see Figure 5).

Figure 5: Range of IC-25 Responses After Censoring Certain Data

Note that the censoring process eliminated 37 of the original 80 data points. Note also that the mean of all retained data is approximately 93 because the mean of values less than 100 can never equal 100. For dischargers that must demonstrate compliance with toxicity limits at end-of-pipe, the data handling routines built into the software can cause artificial non-compliance.

The censoring problem is most pronounced as the effluent moves closer and closer to the threshold definition of a non-toxic sample (e.g. IC-25 = 100 or NOEC = 100).
III. Alternate Approach: Percent Effect

All of the limitations described above result from the use of NOEC/LOEC and IC-25 to describe the toxicity of effluent samples. However, these limitations are not inherent to toxicity testing only to the choice of endpoints used to represent toxicity.

Expressing toxicity as a percentage of the initial effluent sample is counter-intuitive and causes certain anomalies when analyzing and interpreting data. A better approach would be to express toxicity as the percent effect at the permitted instream waste concentration.

The percent effect is strongly related to statistical endpoints already used in toxicity testing. For example, the IC-25 is the effluent concentration which is believed to cause a 25% reduction in growth or reproduction. The LC-50 is the effluent concentration which is believed to cause a 50% reduction in the rate of survival among test organisms.

Just as the IC-25 or LC-50 look for the concentration at which a specific level of biological impact occurs, the percent effect approach quantifies the level of impact at a given concentration. The percent effect is estimated using sophisticated non-linear regression models to calculate the relationship between effluent concentration and demonstrated impact (see figure 6).

Figure 6: Example of Non-Linear Regression to Estimate Percent Effect
The data illustrated in Figure 6 shows that control organisms are expected to produce approximately 26 offspring per female. A 25% reduction from that value would be approximately 19.5 offspring per female. Figure 6 shows that the IC-25 is approximately 70% effluent.

Assuming that the dischargers permitted instream waste concentration was 80% effluent, we can also determine the percent effect at that concentration. Figure 6 indicates we should expect approximately 17 offspring per female exposed to an 80% effluent concentration. That would be 9 fewer offspring than was estimated for the control group. And, this corresponds to a 34.6% reduction in reproduction.

Figure 6 is also depicted with the upper and lower 95% confidence intervals. Note that at the 80% effluent concentration the estimated reproduction ranges from a low of 8 offspring per female to a high of 24 offspring per female. This corresponds to a 95% confidence band that ranges from –69% to –7.7% of control performance.

Thus, while it is clear that the percent effect is negative, we cannot say with statistical certainty that the adverse effect is greater than 25%. This is partly because there is some statistical uncertainty about what the true control performance is. Figure 6 shows that the 95% confidence interval ranges from 24 to 28 offspring per female (26 is the median estimate of control performance).

The specific formula used to generate figure 6 is typical of the non-linear equations routinely relied on to evaluate dose-response data in pharmetoxicology. There are many other equations that are equally well suited to this purpose and all will generate roughly the same fitted curve (see D.J. Finney, Statistics in Bioassay, 1987).

Recently, most major statistical software packages began including modules to perform such curvefit analysis. These packages also contain a module for estimating regression coefficients using the Generalized Linear Model (GLM).

The primary advantage to using GLM is that the data need no longer meet traditional assumptions in order to infer a valid a percent effect. Specifically, it is not necessary for the data to be normally-distributed. Nor is it necessary for the variance to be heterogeneous. These improvements make the GLM technique more robust than traditional regression models.

GLM and curvefitting software provide ideal tools for predicting and managing whole effluent toxicity in the NPDES permitting program. There are significant methodological advantages and substantial implementation advantages to using these tools to regulate toxicity based on the percent effect rather than the effluent concentration.
IV. Methodological Advantages

1) It Is Possible to Develop an MDL for WET Based on Percent Effect

Using data from EPA’s recent interlaboratory variability study, it is possible to estimate the level of percent effect that is likely to occur by chance (see figure 7). The chart shown in figure 7 was generated using interlaboratory reproduction data for Ceriodaphnia that were only exposed to laboratory control water.

Figure 7:

Estimated MDL Using Percent Inhibition

By repeatedly sampling two random groups of ten organisms and comparing the mean reproduction for those groups, we can calculate the frequency with which apparent adverse impacts occur for reasons unrelated to effluent toxicity. For example, figure 7 shows that we can be 90% certain that the effect is real when the observed inhibition is greater than 25%.
Federal regulations define the MDL as the level at which we can be 99% certain that the test correctly distinguishes between the presence and absence of any given pollutant parameter (see 40 CFR 136.2). Figure 7 shows that an inhibition greater than 40% is unlikely to occur by chance more than 1% of the time (e.g. 99% confidence). Therefore, if the MDL for toxicity were calculated in the same manner as it is for chemical tests, inhibitions less than 40% would be reported as “not detected” or ND.

Similar graphs can be prepared for any combination of species, method, and biological endpoint (survival, growth, reproduction) provided there is a large body of data describing organisms performance in non-toxic control water.

2) **Dose-Response Confirmation Is Automatically Integrated Into the Analysis.**

Because the percent effect is derived from a regression equation, dose-response analysis is automatically built into the process. Moreover, there are common and familiar tests for evaluating the goodness-of-fit for each equation (see figure 8). This helps reduce the subjectivity inherent to interpreting dose-response relationships.

<table>
<thead>
<tr>
<th>Table 8: Statistical Test for Dose-Response Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Dose Response Equation</td>
</tr>
<tr>
<td>( r^2 ) Coef Det</td>
</tr>
<tr>
<td>Source</td>
</tr>
<tr>
<td>Regr</td>
</tr>
<tr>
<td>Error</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>Lack Fit</td>
</tr>
<tr>
<td>Pure Err</td>
</tr>
</tbody>
</table>

© 2001, Risk Sciences
The results reported in figure 8 coincide with the graph shown in figure 6. This output from a statistical package indicates that the logistic dose-response equation is highly statistically-significant (p<0.00000). The equation appears to explain 31-35% of the variance in reproduction based on the effluent concentration the organisms were exposed to. Although the explained variance may seem, at first glance, to be relatively low it is consistent with the level of explanatory power expected from this test method (see WERF Report entitled WET: Accounting for Variance, 1999).

Figure 8 also indicates that the 3-parameter regression model may not be the best fit for the data. The third parameter adds very little new explanatory power to the equation (p=0.11398). Nevertheless, the results indicate that the equation can provide a statistically-significant and reliable estimate of the probable percent effect at any given effluent concentration. A 2-parameter regression produces approximately the same curve with a higher F-value; but, it is a less rigorous dose-response model.


The GLM technique will always generate confidence limits for each of the parameters in the regression equation. It is not dependent on Monte Carlo resampling or “bootstrapping” to accomplish this task. Current software that rely on resampling and bootstrapping are often unable to consistently generate confidence limits for reasons related to the underlying program code. EPA recently recommended that the confidence limits from such software not be used until the program code can be repaired.

In addition, the GLM technique will generate true confidence limits. That is the width of the confidence limits can vary across the entire dynamic test range based on the amount of data available and the stability of that data at any given point in the dilution series. This is well illustrated in figure 6.

Because there is considerably more data available at the low end of the dilution series (0% to 40%), the confidence limits are much narrower in this part of the range. Because there is no data in the zone between 50% and 100% effluent, the confidence limits grow wider to reflect the increasing uncertainty.

If WET test data is used to evaluate reasonable potential or permit compliance, knowing the true confidence limits can significantly reduce the risk of error during the decision-making process. Most important, the GLM technique can overcome the most common logical mistake made with regard to the use of IC-25 estimates. The lack of confidence limits on the printout was frequently misinterpreted to mean that the confidence bounds were very narrow or non-existent. In fact, for most cases, the opposite is true. The lack of confidence intervals often indicates that the IC-25 estimate itself is unstable.
4) **Estimating Percent Effect Using GLM Makes the WET Test More Sensitive.**

Unlike the IC-25 procedures, the GLM technique uses all of the available data to establish the correlation between effluent concentration and percent effect. Consequently, the procedure makes more efficient use of the available statistical power to identify the threshold for toxicity.

Noisy data often confuses an IC-25 analysis. It is necessary to “smooth” the data before any conclusion can be inferred. The GLM procedure minimizes the number of transformations necessary to analyze data without losing any of the information that is useful for quantifying uncertainty and developing appropriate confidence limits.

In fact, all of the software problems that plague the IC-25 procedure are easily overcome when estimating percent effect using a GLM technique. There is no longer any need to censor data in order to “make sense” of the results. There is no artificial threshold at which results cannot be logically interpreted. When the impacts are negative, they are considered inhibitions. When they are positive they are considered stimulations. Undiluted effluent can have a range of response that is either positive or negative. If the bootstrapping procedure were applied to such data, the range of response surrounding zero effect would likely average to zero effect. This is a significant improvement vs. the IC-25 procedure.

5) **Percent Effect Is More Intuitive and Easier to Interpret**

Unlike NOEC or IC-25, the percent effect endpoint increases as toxicity increases. In addition, the absence of toxicity is expressed as “zero” rather than 1 toxicity unit. These simple conventions make it easier to understand and interpret toxicity test results. It also makes it easier to identify correlations between WET test results and the concentration of various chemical pollutants in the effluent.

One significant advantage is that the two primary measures of test sensitivity (PMSD and reference toxicant results) can be expressed using the same measure of performance. Reference toxicant tests have long been used to demonstrate that WET tests were adequately sensitive. But, the results of most reference toxicant tests is expressed as a specific concentration of the toxicant. Knowing that the salt solution (NaCl) is toxic at 1 g/L does not provide an obvious measure of just how sensitive the test is. Knowing the range of observed responses for a given salt concentration provides more useful information for evaluating the relative sensitivity of each test.
In the recently released guidance, EPA recommended using the Percent Minimum Significant Difference (PMSD) to determine whether a given test was adequately sensitive. The PMSD is the smallest difference that the test can detect and identify as statistically-significant. EPA recommends that most toxicity tests should be able to identify a 20% inhibition as statistically-significant. The specific threshold varies for different test endpoints. The PMSD is a natural extension of the percent-effect approach.

V. Implementation Advantages

1) WET Tests Can Be Calibrated to Site-Specific Water Quality Conditions.

Whole effluent toxicity tests must be run under standardized conditions that may or may not reflect the natural quality of receiving waters. Some states have very soft waters naturally. The lack of hardness in these streams can cause stress on test organisms that are cultured in much higher hardness. Drinking water taken from these streams may not be able to pass a toxicity test before it is used and released as treated wastewater.

Attempts to use the natural receiving water as the dilution water during a toxicity test often causes control organisms perform inadequately thereby failing to meet EPA’s minimum test acceptance criteria. The only alternative is to use a laboratory dilution water which may or may not accurately represent the true baseline stress condition for ambient receiving waters.

Using GLM to quantify the percent effect, it is possible to determine whether a specific effluent discharge cause or contributes to toxicity when the receiving waters are naturally stressful to the standard test organism. This is best accomplished by running two tests side-by-side.

The first test uses a sample of upstream receiving water diluted by laboratory control water to estimate the baseline inhibition effect of the local water supply. The second test is performed on effluent also diluted by laboratory control water. If the percent effect observed in the second test is less than or equal to that observed in the first test, then the effluent is not causing or contributing to an exceedence of the water quality criteria prohibiting the discharge of toxics in toxic amounts because there is no NET negative effect on the standard test organism.

Over time, it may be possible to accurately predict the percent effect that is likely to result from local water based on a large number of tests. Thereafter, it is probably no longer necessary to run side-by-side tests each time. The percent effect in effluent can be compared to a table value for the local receiving waters.
This approach also works well where groundwater is the primary source of municipal and industrial supply. The ionic structure of groundwater often fails to support uninhibited reproduction and growth among standard test organisms. The same is true where toxicity testing is used to evaluate the quality of stormwater runoff. Rain is naturally low in pH and several key ions. Rainwater will often cause a toxicity test to fail even if no other chemical pollutants are added to the runoff. Therefore, it is essential to quantify the natural level of background effect before attempting to estimate the contribution of toxicity added by human activities.

2) Reasonable Potential Analyses Are More Reliable Based on Percent Effect.

Because percent effect can be characterized by a true “zero,” the mathematical calculations used to estimate reasonable potential are no longer biased by the units of measure. In addition, the development of an MDL for WET tests can eliminate the spurious results which tended to inflate reasonable potential calculations.

Perhaps most important, it will be easier to distinguish method variability from effluent variability when evaluating reasonable potential. Known levels of method variability can be subtracted from observed levels of test variability to calculate the net level of effluent variability. This is made possible because all of the variability is expressed in the same units of measure: percent effect.

In the absence of such a partitioning technique all of the observed test variability will be assumed to represent variation in effluent quality. No consideration will be given to the amount of variability contributed by the measurement system itself. Thus, reasonable potential will continue to be overestimated for WET test results. Shifting to percent effect can significantly reduce this systemic bias.

3) Percent Effect Can Provide the Basis for More Flexible WET Permit Limits.

In South Carolina, where the percent effect endpoint is already employed, permit limits for WET are expressed as a “test maximum” and a monthly (or quarterly) average (see attachment). The maximum limit prohibits a discharger from releasing effluent that causes more than a 40% inhibition at the IWC. The monthly average prohibits a discharger from releasing effluent where the mean of all WET tests performed during the monitoring period is greater than a 25% inhibition. If only a single WET test is conducted during the monitoring period, it must meet both the maximum and the average limit.

The advantage to this approach is that it encourages additional WET testing for borderline cases. For example, if the first WET test shows a 30% inhibition, the discharger has every incentive to run another WET test in order to reduce the average below 25%.
The ability to average test results is also useful for determining compliance based on identical split samples. Where one test passes (0% inhibition) and another fails (38% inhibition) the average of both tests (19%) may indicate overall compliance with WET limits in the permit.

**Conclusions**

Using the percent effect endpoint is not a panacea for eliminating analytical variability from whole effluent toxicity testing. Rather, it improves our ability to manage such variability.

EPA relies on whole effluent toxicity testing to identify pollutants that were toxic below the threshold of detection for standard chemical analyses. EPA also believes that WET testing can be used to identify pollutants that generate synergistic toxicity. In both cases, it is not necessary to know the cause of toxicity to be able to regulate its effects.

If the central focus of toxicity testing is to prevent adverse impacts to the environment, then it makes sense that the primary measure of success should be the “percent effect.” Translating the estimated effect back into the effluent concentration believed to be responsible for the effect merely adds unnecessary complexity and serves to obscure the issue we care most about.

Regulating toxicity based on the percent effect is more logical, more intuitive, more robust, and more stable (statistically). If the high powered desktop computers were available when toxicity testing was first developed, it is likely that the developers would have chosen “percent effect” as a superior strategy for managing toxics in the environment.