NACWA
A Clear Commitment to America's Waters

National Association of Clean Water Agencies

Whole Effluent Toxicity (WET) NPDES Permit Testing and Limitations for Public Agencies

White Paper

January 2006

TABLE OF CONTENTS

A.	Introduction and Background	3
B.	WET Testing Basics	5
C.	Permitting Considerations – General	8
D.	WET Permitting Approaches	18
E.	Additional Considerations	22
F.	TIE/TRE Requirements	24
G.	Conclusions	25

A. Introduction and Background

EPA's National Pollutant Discharge Elimination System (NPDES) program regulations require "whole effluent toxicity" (WET) effluent testing by publicly owned wastewater treatment facilities (POTWs) during each permit cycle. The regulations also require (with certain exceptions) that NPDES permits include limitations for WET, if WET data demonstrate that the facility has "reasonable potential" (RP) to cause or contribute to a receiving water excursion above either (1) numeric state water quality standards for WET or (2) the state's narrative ("no toxics in toxics amounts") water quality standard. These requirements are of substantial concern for clean water agencies because (1) WET permit limitations often pose difficult compliance challenges, (2) it is difficult and expensive to determine and/or correct the underlying cause of WET exceedances, (3) testing is expensive, and (4) like any permit requirement, once imposed, limitations are difficult to remove.

WET testing involves the exposure of cultured test organisms to varying concentrations of effluent, measurement of biological responses (mortality, growth, reproduction), and the calculation of an endpoint or effluent concentration at which the measured effect exceeds that of the control (a test performed on dilution water without effluent) or exceeds a predetermined measure of "toxicity." EPA and state NPDES agencies generally presume that an observed difference from control organism response indicates "toxicity" and from that extrapolate receiving water toxicity or a narrative water quality standard excursion. This is what leads to either numeric permit limits for WET or other permit conditions designed to confirm, determine the cause of, and eliminate "toxicity" in the effluent.

The National Association of Clean Water Agencies (NACWA), along with several state POTW organizations and industrial groups, challenged EPA's 2002 repromulgation of WET test methods in federal court. For strategic reasons the appeal focused on the most problematic sublethal endpoints within the chronic methods. NACWA believes particularly that for those methods and endpoints, the test methods are simply not accurate or precise enough for the NPDES program uses called for by EPA and state regulations, and they do not accurately predict receiving water conditions or "toxicity."[1] This is especially so where there are relatively low levels of "toxicity," as is often the case with POTW effluents. Although the appeal was aggressively pursued, the Court of Appeals for the District of Columbia Circuit upheld the test methods in December, 2004.[2] Nonetheless, the decision supported only the methods themselves and specifically did not support any particular NPDES permitting use of WET. The Court provided valuable input on proper and improper uses of WET methods. That input, the use of which this White Paper addresses, included:

- The Court's warning against the use of single WET test failures to bring enforcement actions. EPA's permitting system must account for the fact that sometimes a test will give a correct result, and sometimes the test will report (for example) twice the "true" level of toxicity.[3]

[1] Because NACWA believes that WET test results alone are inadequate to identify instream toxic conditions, this White Paper generally places the term "toxicity" in quotation marks.

[2] Edison Elec. Inst., NACWA, et al. v. EPA, et al., No. 96-1062 (D.C. Cir. Dec. 10, 2004) (rehearing denied 2005) (hereinafter "WET Court Decision") (Appendix A).

[3] Id. at 9.

- The validity of the methods does not imply the validity of any particular result.[4]
- Although more general Clean Water Act case law supports the near-finality of test results reported on a Discharge Monitoring Report (DMR), the Court noted that for WET methods nothing "forecloses consideration of the validity of a particular test result in an enforcement action."[5]
- States have the discretion to set toxicity thresholds to compensate for local conditions at the permitting stage. This may mitigate the lack of correlation between tests results and instream impact ("representativeness"), particularly at low levels of toxicity. EPA must establish representativeness in permitting.[6]
- Individual permits may be challenged if the clean water agency believes that regulation of toxicity is at a level posing minimal risk.[7]
- Permitting systems must account for the imprecision inherent in WET data.[8]

The development and presentation of the case by NACWA and the other petitioners also served to expose and better develop parts of EPA's rulemaking record. This provided valuable information and data on WET issues that clean water agencies will be able to use to their advantage in permit proceedings and in working with state agencies on WET programs.

Based on the WET Court Decision and the experience that NACWA members have gained in dealing with WET permitting among the states and EPA Regions, NACWA has assembled this White Paper. The White Paper includes guidance for working with NPDES permit agencies in developing (1) WET testing conditions, (2) endpoints for judging results, and (3) permit conditions that provide for (based on WET results) moving to less frequent routine testing, toxicity determination procedures, and numeric permit limits when justified. This White Paper also provides suggested NPDES permit language taken from permit language currently used by the states.

The difficulties of WET permitting are multiplied by the fact that there have been few if any court or administrative challenges brought by permittees over WET conditions or limits. Therefore, unlike some other NPDES permit issues, the definition of the correct or incorrect way to determine reasonable potential (RP) or to make other permit decisions is based directly on the (typically very general) regulations, EPA's very stringent WET permitting guidance, and the technical details that the clean water agency is able to develop. With the WET Court Decision and the remaining legal uncertainty about the methods themselves removed, this is likely to change. But, in the meantime, a clean water agency manager will need to carefully consider the effect of WET permit conditions.

As with any NPDES permit issue, success in obtaining an acceptable permit result depends on careful preparation and definition of the issues, a solid technical and regulatory approach that considers the clean water agency's specific conditions and needs, and building a full

---

[4] Id.

[5] Id.

[6] Id. at 12.

[7] Id. at 13.

[8] Id. at 8.

administrative record on which the NPDES agency may ideally make the right decision, or that will serve as a viable basis for judicial review of an improper NPDES agency decision. Although this White Paper is not intended as a litigation strategy, few difficult permit issues are successfully addressed without the detailed groundwork implied by the development of a full administrative record and the unstated possibility of review of an incorrect decision.

In an ideal NPDES permit case involving proposed WET limits, EPA or the state NPDES authority should be required to (1) prove representativeness of the specific WET endpoint(s) at the level of "toxicity" identified after instream dilution (because there are no numeric EPA water quality criteria and in most states no numeric water quality standards for WET); (2) factor WET test variabilities into any RP determination; and (3) design any WET limits in a manner that does not subject the clean water agency to undue liabilities for "false positive" results. The WET Court Decision provides support for this process. Realistically, in view of EPA's and the states' current approaches to WET permitting, achieving all of these results will be difficult in the short term. However, this White Paper provides guidance for achieving these objectives.

This White Paper provides general guidance. Because each case is unique, each must be evaluated based on its own merits and a consideration of all pertinent factors, and may need to consider factors not addressed in the White Paper. It does not provide specific legal or regulatory advice as to any individual NPDES permit, and in any such case the clean water agency manager may need to obtain case-specific legal and technical advice. The White Paper does not attempt to address EPA's Great Lakes Initiative WET procedures which are incorporated into regulation, or any specific state WET permitting guidance or program.

B. WET Testing Basics

This section provides an overview of the basics and additional details of the WET program.

1. Federal Requirements

EPA regulations impose few specific WET requirements for POTW NPDES permitting. At least four test events, using either acute (measuring short term biological impacts) or chronic (longer term – at least in relation to test organism life cycle) tests, are required for permit reissuance, with each test event on at least two different test species.[9] The regulations recommend acute tests if dilution at the edge of the mixing zone is greater than 1000:1. This extreme dilution implies little risk of chronic toxicity, and therefore focuses on acute near-field toxicity. Chronic tests are recommended if dilution is less than 100:1. If dilution is between 1000:1 and 100:1, either acute or chronic tests are appropriate. Accordingly, under the federal system or a comparable state program there is no specific regulatory basis for both acute and chronic testing.

Other than RP requirements comparable to those for chemical-specific parameters, no other specific requirements are imposed. To the extent that a state program mirrors the federal program, there is no regulatory basis for WET requirements beyond those necessary for data generation prior to reapplication, at least in the absence of prior data showing effluent "toxicity."

However, many state programs have been far more creative and apply additional WET program requirements either through statute, regulation or guidance.

[9] 40 CFR 122.21(j)(5).

2. Acute/Chronic Testing

Although generally performed on the same test species, acute and chronic testing is performed pursuant to separate test protocols. Acute tests measure organism mortality based on relatively short term exposure. Because mortality is a readily distinguishable endpoint, and (at least in a relative sense) not subject to control comparison ambiguities, most clean water agency managers believe that they are not subject to substantial risk from acute tests of misidentification of effluent toxicity ("false positives" or Type I Error). Because mortality is a severe endpoint, determination and elimination of the cause of any such toxicity may be relatively straightforward in most cases, provided the toxicity is prolonged or consistent rather than an isolated incident.

By contrast, chronic testing focuses on (in addition to mortality) more subtle endpoints, typically growth and reproduction. Because of the nature of the endpoints the risk of misidentification of toxicity is substantially greater. As noted initially, the chronic methods sublethal endpoints were the focus of the WET litigation, particularly at low levels of "toxicity" typical for POTWs. They are also the focus of this White Paper.

3. Test Organisms/Species

For freshwater testing the usual test species are fathead minnows (*Pimephales promelas*, a small fish) and daphnia (*Ceriodaphnia dubia*, a small invertebrate "water flea"). Although the federally approved methods include other species, these two are used almost exclusively by EPA and state NPDES agencies for freshwater, for both acute and chronic testing. For estuarine and marine testing the methods include the mysid (*Americamysis bahia*, a small mysid shrimp), sheepshead minnow (*Cyprinodon variegatus*) and other organisms for both acute and chronic purposes.

In an unusual twist designed to allow West Coast states to use other (presumably more sensitive) acute and chronic estuarine and marine tests organisms, EPA's 2002 repromulgation of WET methods limited the estuarine and marine methods to use with Atlantic and Gulf Coast watersheds. This raises both additional difficulties and additional legal issues for West Coast estuarine and marine dischargers.

The organisms used for WET testing are laboratory cultured and subject to control tests, intended by EPA to define normal ranges of variation in their test endpoints. However, it is NACWA's view that the quality assurance/quality control (QA/QC) for the control tests is inadequate and needs to be significantly improved. The methods have minimal requirements for control performance, they do not address variability within and across laboratories, they do not have limits for intratest control variability, and there is no national standard for reference toxicant performance.

4. Test Endpoints

The WET test endpoints are typically mortality, growth and reproduction. Mortality is expressed, through metrics addressed below, as the number or percentage of organisms that die as the result of exposure to effluent. Growth and reproduction are expressed, also through metrics addressed below, as the reduction (compared to control tests) in weight gain of the organisms or reduction in the number of offspring.

Acute test results are most commonly expressed as either the "LC50" or the "NOAEC." The LC50 is the concentration of effluent within the test at which there is 50 percent mortality (half of the test organisms die). The No Observed Adverse Effect Concentration is the concentration of effluent within the test at which there is no statistically significant mortality compared to control tests.[10]

Chronic test results are most commonly expressed as either the "NOEC" or the "IC25." The No Observed Effect Concentration is the highest concentration of effluent within the test containers at which there is no statistically significant adverse effect (decrease in survival, growth or reproduction) compared to control tests. The IC25 (25 percent inhibition concentration) is the concentration of effluent within the test at which there is a 25 percent reduction in the measured effect compared to controls. Chronic results are occasionally expressed as a "LOEC." The Lowest Observed Effect Concentration is the lowest concentration of effluent at which there is an adverse effect, a less stringent endpoint than the NOEC.

5. Hypothesis Testing/Point Estimates

The metrics for expressing biological impact differ markedly from one another. "Hypothesis testing" uses a statistical test to determine (at any particular effluent test concentration) whether the response is different (less favorable) from the control. Hypothesis testing results in the use of only one test concentration from among the several effluent test concentrations used in any WET procedure to derive the test endpoint. The NOAEC, NOEC and LOEC are hypothesis testing endpoints.

Point estimates also use a statistical procedure, but use more of the WET test data to determine the point (the test effluent concentration) at which the response is equal to a specific target. Point estimate procedures are able to interpolate between responses to each tested concentration and have at least the potential of providing a more nearly "correct" result.

Because of the use of more of the test data, point estimates are a more technically rigorous approach, and the approach that poses less risk of false positives. The LC50 (acute) and the IC25 (chronic) are point estimates. The IC25 has the additional value of using the 25 percent point as a surrogate for the detection level concept used with chemical-specific pollutant identification. That is, rather than attempting to identify the point at which there is any difference from controls, the endpoint focuses on the more distinguishable 25 percent effect point. EPA also expresses a preference for the use of point estimates.

6. Test Method Flexibility

Because of the inherent difficulties in obtaining meaningful data from live test organisms, and in response to permittees' concerns, the methods themselves provide substantial flexibility to address difficulties that may lead to inaccurate results. EPA's 1996 memorandum on flexibility correctly notes that "[t]he test method manuals do not . . . strictly prescribe every aspect of

[10] The NOAEC endpoint was never field or laboratory validated by EPA. NOAEC data were not part of EPA's Interlaboratory Variability Study *infra*. Further, the acute test protocol references the LC50 endpoint only. Accordingly, there is a good argument that the NOAEC is not part of the adopted 40 CFR 136 Table 1A methods and should not be used for NPDES purposes.

method conduct."[11] Among other points, flexibility is provided in control of pH and resulting ammonia toxicity, temperature, hardness and test dilution concentrations.

For example, because (particularly for POTW effluents) ammonia is a common toxicant and because ammonia toxicity is highly related to pH, there is flexibility for the WET procedure to control artificial ammonia toxicity brought about by pH drift within the test. Similarly because toxicity can be influenced by water hardness (again particularly for POTW effluents that are expected to contain certain naturally occurring and added metals), provisions are made for taking test solution hardness into consideration.

C. Permitting Considerations - General

There are several WET testing considerations that affect most NPDES WET permit issues, irrespective of whether testing is monitoring for reapplication, compliance testing for numeric limits, or related to the identification and elimination of indications of "toxicity." This section provides guidance on those generic issues that may be helpful in protecting the clean water agency's interests in the NPDES permit process.

1. Frequency/Retests

WET testing is expensive and that is why most NPDES permit requirements are designed around quarterly testing (one per calendar quarter) at most. However, small numbers of data points subject the clean water agency more to the variabilities of the test procedures compared to more frequent chemical-specific testing, even if we assume that variabilities are comparable between WET and chemical methods. Although NPDES permits do not disallow the generation of additional data, there is an obvious tradeoff between costs and the value of additional data when the reliability of initial data is questionable. In critical situations, clean water agency managers often commission duplicate or repeat WET tests.

In a minor concession to the concerns of NPDES permittees, EPA has consistently stated that, in general, formal enforcement is not appropriate for single WET limit exceedances.[12] The WET Court Decision specifically endorsed this safeguard, noting that "WET tests will be wrong some of the time . . . ."[13] For chemical-specific testing, notwithstanding test variability, permittees do not generally consider that lab results that include an appropriate level of quality assurance/quality control (QA/QC) will misidentify a pollutant. However, EPA's data and the materials developed in the course of the WET litigation demonstrate that properly performed WET testing will frequently misidentify "toxicity." It is also important to note that the National Environmental Laboratory Accreditation Council (NELAC) recognizes that WET data are more variable than chemical-specific data.

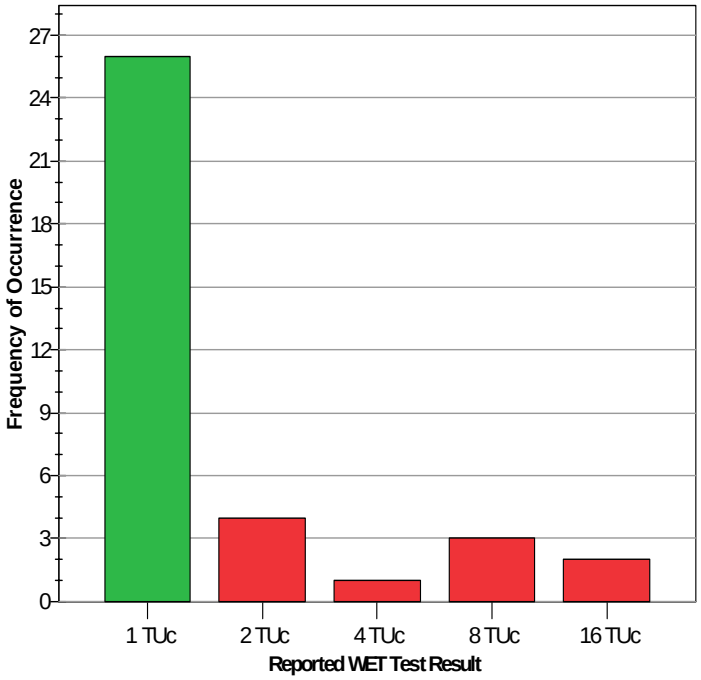
An important example of this effect is provided by EPA's own data generated in support of the methods. EPA's interlaboratory testing focused on relatively toxic samples, rather than samples with more dilute or minimal toxicity which might be expected from well operated

[11] Clairifications Regarding Flexibility in 40 CFR Part 136 Whole Effluent Toxicity Test Methods (EPA Apr. 10, 1996).
[12] 67 Fed. Reg. 69952, 68 (Nov. 19, 2002).

[13] WET Court Decision at 9.

POTWs. The one sample EPA tested with low toxicity had a 28 percent false positive rate[14] relative to the median response with *Ceriodaphnia* chronic results.[15] This means that (if this particular data set is representative) for a POTW with a quarterly two species test requirement (eight tests per year), on the average more than two tests per year will be expected to demonstrate "toxicity" when in fact none is present. This argues strongly for retest procedures before any finding of permit violation or before any mandatory switch to more frequent testing or TIE/TRE requirements. EPA's results are displayed graphically below.[16]

Figure 1
***Ceriodaphnia* Reproduction (NOEC) in EPA's Reference Toxicant Sample**



28% False Positive

A comparable example of variation in WET data, but on a more toxic sample, is shown by EPA's *Ceriodaphnia* reproduction data in effluent in Figure 3 *infra*. EPA's Interlaboratory Variability Study includes the comparable data for all of the repromulgated WET methods, and clean water agency managers may use those data to demonstrate the inherent variability in the methods that the agency proposes for the POTW permit. These data, and particularly graphical representations of the data, are useful in demonstrating that single WET test results should not

[14] This White Paper defines the false positive rate as the number of tests with a reported result above the central tendency of the data, divided by the total number of tests. EPA and some state NPDES agencies may use a false positive definition that does not consider all of the data above the point of central tendency to be false positives. However, using Figure 1as an example, the difference between 1 TUc (defined *infra)* and 2 TUc will typically be significant for a clean water agency, and NACWA considers that the definition used is appropriate.

[15] Final Report: Interlaboratory Variability Study of EPA Short-Term Chronic and Acute Whole Effluent Toxicity Test Methods, Vol. 1 ("Interlab Variability Study"), EPA 821-B-01-004 (Sept. 2004) Table 9.8, pp. 81-82.

[16] From Reply Brief of Petitioners at 26 in WET Litigation.

define NPDES permit violations and in advocating a weight-of-evidence approach to any judgements based on WET data.

2. Test Dilutions

Permit WET testing requirements frequently specify a particular "dilution series." For example, tested effluent concentrations may be 6.2, 12.5, 25, 50 and 100 percent (a "0.5" dilution series). While this may seem like a test design for identifying "toxicity" at whatever level it may occur, it is seldom an efficient test design. Rather, testing is more effectively centered on a dilution of concern (e.g., a mixing zone effluent concentration or a target concentration cited in a NPDES permit condition). That is, if a 25 percent effluent/75 percent receiving water concentration is the point at which "toxicity" or the absence of "toxicity" will be judged, test results at 100 percent are of no importance, and dilutions centered more closely on 25 percent increase the reliability of the procedures to represent potential for toxicity instream.

Although some state programs define what they consider an appropriate dilution series, these tend to be generic and ignore the site-specific nature of mixing and therefore the site-specific nature of toxicity determinations. The WET methods do not define or require specific dilution series, and the EPA's regulations do not otherwise require any particular series. Rather, test concentrations should be selected independently for each test based on the objective of the study, the expected range of toxicity, the receiving water concentration, and any available historic testing information on the effluent. Accordingly, the dilution series should be defined with regard to a particular endpoint, and both the clean water agency's and the NPDES agency's interests are served by that more specific approach. Using the 25 percent effluent example above, a 0.7 dilution series might focus on the concentration of interest and use test concentrations of 12.3, 17.5, 25, 35.7 and 51 percent. However, dilution factors much higher than 0.7 should not be used because small degrees of variance across tests dilutions may artificially result in irregular concentration-response curves or lower NOECs.

3. Data Quality Objectives

Like any laboratory program, WET testing should be approached holistically and the clean water agency manager and the laboratory should jointly determine the manner in which testing will be undertaken. Particularly when past WET results raise concerns or when the WET decision point stated in the permit or on which the NPDES agency focuses is very stringent, it is important that test problems be anticipated and guarded against. An experienced WET testing laboratory should consider and be able to specifically advise the clean water agency on method flexibility and on specific procedures that best serve to correctly identify/disprove indications of "toxicity" in a particular effluent-receiving water situation.

An assessment of concentration-response is critical. The classic toxicity response will demonstrate a consistently increasing response with increasing effluent concentration, and the laboratory should be prepared to question testing that does not identify such a response. The methods specifically provide for that level of QA/QC assessment. The methods do not specifically state that a permittee may invalidate a test purely on the basis of the concentration-response relationship. However, NACWA believes that, in the context of a full Data Quality Objectives program, the testing laboratory and the clean water agency should consider a test invalid if an adequate relationship is not present.

Data Quality Objectives must of course be implemented in a neutral manner based strictly on the merits of specific analyses. The laboratory must be prepared to disqualify data that are favorable to the clean water agency as well as data that are unfavorable. But, the laboratory should not hesitate to disqualify a WET test that does not demonstrate a proper concentration-response relationship and that otherwise does not meet proper Data Quality Objectives.

4. Expression of Results as "Toxicity Units"

EPA and generally the states express WET data not as a percent (percent of effluent in the test) having the identified biological effect as reported directly by the tests, but as "toxicity units" (TU). A TU is defined as 100 divided by the WET result expressed as percent effluent (e.g., a chronic test NOEC result of 80 percent is 100 ÷ 80, or 1.25 TUc). Among other claimed benefits, TUs are intended to avoid the counterintuitive feature of WET percent results where an increasing number indicates less "toxicity" (e.g. increasing NOEC values, up to a maximum of 100 percent, reflect increasing proportion of effluent in a test showing no observed effect).

However, the WET Court Decision raised serious concerns about the use of TUs. In response to a demonstration by NACWA and the other petitioners that WET results expressed as TUs displayed far more analytical variability than chemical-specific analytical methods, the Court concluded that the use of TUs to compute coefficients of variation (CV), an expression of variation in data from multiple tests on the same or different samples, gave a "grossly inflated result."[17]

Whether there is a substantial difference in CVs for WET data expressed as percent and as TUs is entirely dependent on the specific data set. It does appear that WET data that are generally reflective of low or nonexistent "toxicity," but with occasional outliers (a situation typical for POTWs), have the potential for the data expressed as TUs to exhibit relatively high CVs.

EPA's recommended procedures for dealing with WET data (and the procedures of many state NPDES agencies) are entirely dependent on use of TUs and the use of CVs based on TU data. However, TUs are not recognized in the WET methods themselves or in EPA regulations. Accordingly, the clean water agency manager should consider challenging the use of TUs in an appropriate circumstance. RP determinations, calculation of permit limits and any other procedures that involve the comparison and manipulation of WET data should be performed only on data expressed as percent. That approach is strongly supported by the WET Court Decision observations about the use of TUs.

The only use of WET data expressed as TUs should be as a final step after the calculation procedures, and only to take advantage of the claimed attributes of the TU metric in making increasing numeric data consistent with increasing "toxicity." The corresponding percent data should also be expressed, and any further calculations should use the percent data.

5. Data Interpretation with Hypothesis Tests

As noted above, EPA generally recommends the use of point estimates (LC50, IC25) rather than hypothesis tests (NOAEC, NOEC). However, EPA also states that both approaches produce the same result. This is generally not the case, as demonstrated by the comparisons in figures two and three and figures four and five below. To the extent that a state procedure or a NPDES permit

[17] WET Court Decision fn. 4.

requirement mandates the use of hypothesis test results, the clean water agency manager ideally should also develop and report the appropriate point estimate results.

Although QA/QC is always important when generating WET data, hypothesis test data presents an even more critical need for full QA/QC within the context of a Data Quality Objectives approach. The laboratory should pay particular attention for hypothesis tests to the PMSD limits (a QA/QC procedure specified by the WET test methods) for the chronic methods and to making sure that control test response is correct and representative. Control response is particularly important for chronic WET tests where minimum levels of performance are set by the methods (weights, fecundity, number of juveniles produced) but variability in performance within and between tests is not addressed. Although complex, control response considerations may be the difference between a WET test that exceeds a NPDES permit limit or otherwise impacts the clean water agency, and a test that can be demonstrated to atypically demonstrate "toxicity."[18]

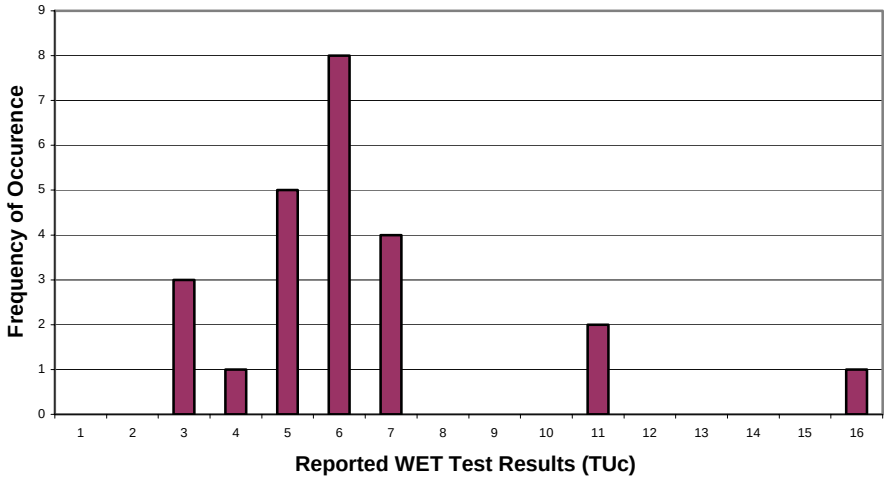
6. Alternate Endpoint Approaches

As noted above, WET acute results are typically expressed as either the LC50 or the NOAEC, and chronic results as either the IC25 or the NOEC. EPA expresses a preference for the "point estimate" LC50 and IC25, both of which make better statistical use of the data generated by the testing. The availability of different endpoints underscores the fact that there is no "correct" result in WET testing. An IC25 and a NOEC from the same test data set, both calculated correctly, will typically produce different results, and series of tests will produce different distributions of results.

This point is illustrated by considering EPA's chronic C*eriodaphnia* reproduction data.[19] Figures 2 and 3 below, respectively, display the calculated IC25 and NOEC data from EPA's interlaboratory results on an effluent sample intended to be moderately toxic.

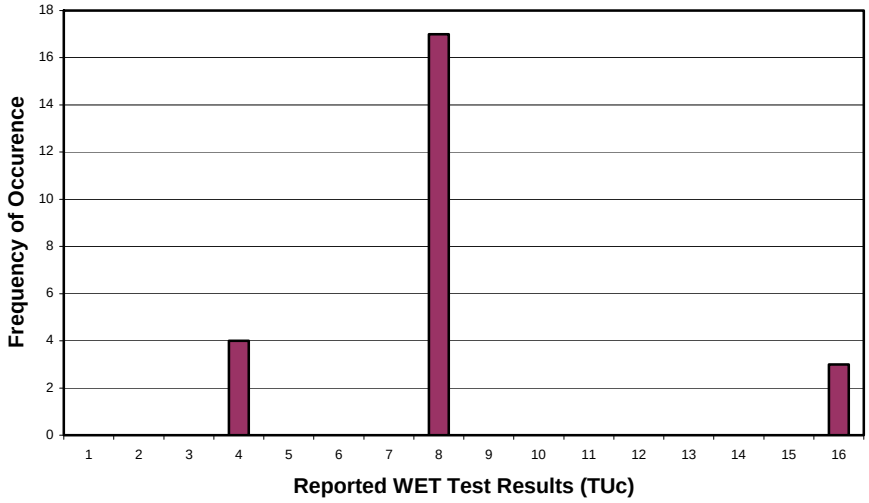
[18] For example, a chronic mysid test requires that the average dry weights of control organisms at the end of a test must be greater than 200 micrograms per individual. However, a testing laboratory may normally produce organisms that are 300 plus or minus 30 micrograms per individual. If a mysid test is conducted by this laboratory and the average weight of control organisms falls outside this range, a substantial case can be presented that the test is unrepresentative. Similarly, the CV for the controls may normally be 20 plus or minus 5 percent and the CV for the latest test 5 percent. Because hypothesis test endpoints are driven partially by intratest variability, this latest test may atypically predict "toxicity" where normally it would not.
[19] Interlab Variability Study Table 9.9, pp. 83-84.

Figure 2
***Ceriodaphnia* Reproduction IC25 in EPA's Effluent Sample**



29% False Positive

Figure 3
***Ceriodaphnia* Reproduction NOEC in EPA's Effluent Sample**



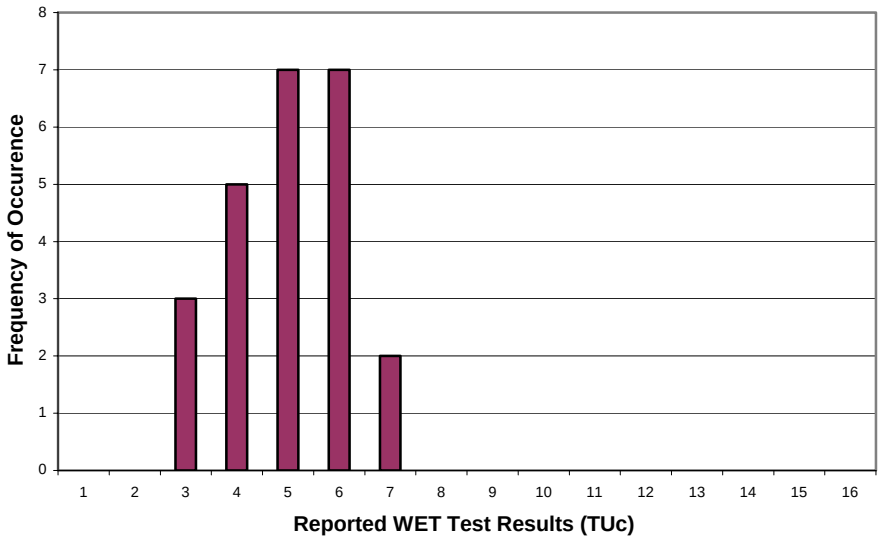
12% False Positive

First, the median, central tendency or "correct" result is different when expressed as IC25 and NOEC. In any particular permit case the distinction in the *Ceriodaphnia* example between 6 (IC25) and 8 (NOEC) Toxicity Units may be critical in determining RP or in determining compliance with a numeric permit limit. EPA's Interlaboratory Variability Study data for other test organisms and other biological effects can be readily displayed and will typically show similar comparisons. It could be observed that IC25 generally produces a "better" (less "toxicity") result. However, this is largely due to the manner in which the IC25 uses the available data, as contrasted with the NOEC identification of only the (highest effluent) concentration with no biological effect. In other words, this is largely due to a more rigorous identification of the point at which there is no "toxicity." The use of appropriate test dilutions should eliminate or moderate any such result disparity.

Second, the lesser spread of data points with the NOEC endpoint reflects the less intense use of the test data generated. Because the NOEC uses only the greatest single test concentration at which there is no adverse impact on the organisms, reported results are artificially limited to a few test concentrations. By contrast, the IC25 does an interpolation between data points, using more of the data, and will produce a more precise estimate of the effect concentration. Thru the same effect, the NOEC may generally demonstrate less data spread and a lower coefficient of variation, suggesting less analytical variability.

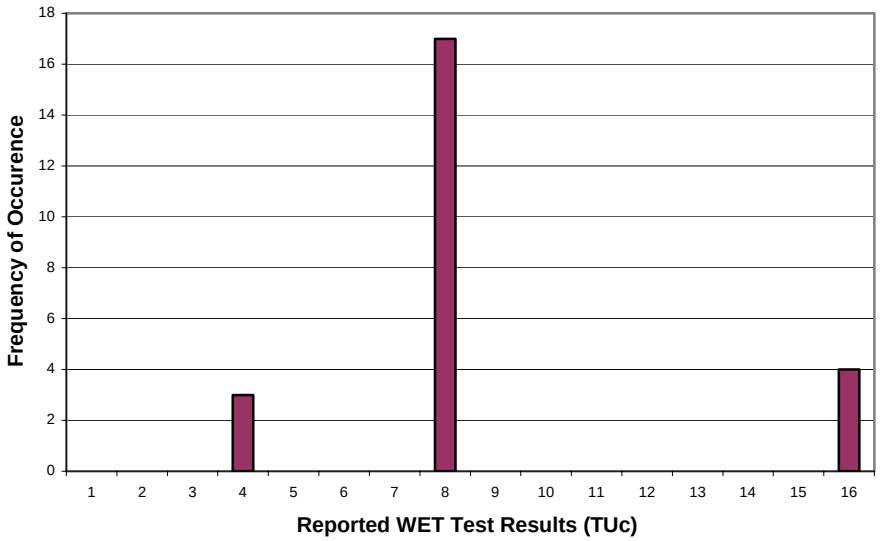
Figures 4 and 5 below illustrate similar IC25/NOEC effects for fathead minnow chronic reproduction tests.

Figure 4
Fathead Minnow Growth IC25 in EPA's Effluent Sample



8% False Positive

Figure 5
Fathead Minnow Growth NOEC in EPA's Effluent Sample



17% False Positive

Overall, NACWA believes that the endpoint comparison illustrates that (1) WET data are defined by the process from which the data are derived, rather than representing any true measure of pollutant; (2) only one endpoint and its CV and any other statistical measures should be used in any particular NPDES procedure; (3) different endpoints should not be compared, averaged or otherwise used together; and (4) irrespective of the endpoint chosen any single WET result is suspect.

In light of concerns that many clean water agency managers have expressed with use of the standard WET test endpoints, there have been calls for the development of other endpoints that could address some of these problems. In particular, some attention has focused on the Percent Effect (PE) endpoint. PE focuses testing on the instream waste concentration (IWC), and compares biological response at the IWC with a predetermined Regulatory Level (RL). The PE is the percent difference in the biological measure (survival, growth, reproduction) from the control. The procedure uses the RL as a detection level surrogate – the percent effect (difference from controls) that the WET result must exceed to be considered real toxicity. For example, a South Carolina PE approach at one point used a 40 percent RL for single WET tests.

The current use of PE or any other approach to WET endpoints is likely to raise substantial objections from EPA and perhaps from the affected state NPDES agency. The PE approach is also believed to require substantial developmental work to define a mathematical protocol for determining the PE for individual tests and for determining and justifying the RL. It is also assumed, but unproven, that PE would provide substantial benefits for NPDES permittees. However, the IC25 has some of the advantages of the PE including better use of test data and a detection level surrogate. The use of IC25, including the flexibility to focus test concentrations on the WET limit or decision point, represents the best current approach to WET permitting involving the more problematic sublethal chronic tests.

7. WET (Instream Toxicity) Criteria

Unlike most chemical-specific pollutants of interest, there are no EPA Clean Water Act section 304 water quality criteria for WET, and few if any states have adopted numeric criteria for WET within their state water quality standards. However, EPA's Technical Support Document for Water Quality-Based Toxics Control[20] (TSD) recommends a "magnitude" for WET of 1.0 TUc and 0.3 TUa, and most state NPDES agencies use those numbers. These numbers are then used as if they were instream water quality criteria, driving RP and permit limit calculations.

The chronic number is based on the belief that there should be no "toxicity" in the receiving water outside of the mixing zone.

The acute "magnitude" is based on the same idea, but because the LC50 is the point at which there is 50 percent mortality, EPA applies a theoretical 0.3 correction factor to estimate a negligible (one percent) mortality level. The basis for the 0.3 correction is questionable,[21] and a relatively small amount of laboratory work in a specific NPDES permit case may produce a

[20] EPA/505/2-90-001 (1991).

[21] EPA's TSD data show that 90 percent of the time use of the 0.3 factor overstates "toxicity."

justification for a less severe acute criterion.[22] The acute criterion is particularly onerous for many clean water agencies because, by definition, it is not a criterion that can be shown to be achieved instream, even with perfect LC50 results, unless the discharge has available dilution (for acute mix purposes) of at least 3.3:1.

Further, if acute data are expressed as the NOAEC endpoint, the 0.3 conversion has no applicability. This is because, by definition, the NOAEC identifies the effluent concentration having no mortality (unlike the LC50 which purports to identify the test concentration with 50 percent test organism mortality), and there is no correction to be made. Therefore, an acute instream number expressed as NOAEC should be no more stringent that 1.0 TUa.

In the absence of EPA or state WET criteria, state NPDES agencies should not attempt to use the 1.0 TUc and 0.3 TUa numbers as if they were binding criteria. Instead, state permit determinations of whether a POTW causes or contributes to impairment of the state's narrative criterion should be based on the totality of the evidence. That is, other data showing that the narrative criteria are protected or that instream beneficial uses are not impaired should negate any determination based solely on a calculated exceedance of the EPA numbers.

However, in response to this point EPA or the state is likely to cite EPA's policy of "independent applicability." The policy, which is not found in either the Clean Water Act or in regulations, provides that WET data, chemical-specific data and instream biological assessments should all be considered independently, and that any one (or two) types of data should not overrule an adverse indication in any other data element. As applied to the WET program, independent applicability would disallow a clean water agency demonstration based on benthic testing and aquatic life studies that, despite adverse WET test results, instream beneficial uses are not being impaired.[23] [24]

In 1998 EPA issued an Advance Notice of Proposed Rulemaking in which it discussed its "current thinking" on the relative merits of independent applicability and "weight of evidence" approaches, and it specifically invited comments on alternative approaches to the use of chemical, toxicity and biological methods in determining reasonable potential.[25] No final action was ever taken on that proposed rulemaking. Further, although not disavowing independent applicability, EPA's 2006 Integrated 303(d) Report Guidance recommends that higher quality data may be weighted more favorably in water quality standards attainment determinations.[26] Although the federal NPDES program regulations arguably support the independent applicability of numeric chemical-specific standards (and the 1998 Advance Notice expressed concerns about weight of

---

[22] For example, a series of acute WET tests may be designed to determine the LC50 and the LC1 on a particular effluent. The correction would be the (mean, 90th percentile or other appropriate statistic) of the ratio of the LC1 percent test result to the LC50 percent test result.

[23] See TSD Ch. 1.

[24] As to an independent applicability argument by EPA or a state NPDES agency, it is important to note that the federal regulations on WET are themselves inconsistent with a strict independent applicability approach. In addition to WET data, "other information" must inform a decision on RP, and appropriate chemical-specific NPDES permit limits may be used in place of WET limits. 40 CFR 122.44(d)(1)(v).

[25] 63 Fed. Reg. 36742, 803 (July 7, 1998).

[26] Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act at IV.K (EPA July 29, 2005) (draft).

evidence primarily in that context), there is no firm legal basis for an application of the policy of independent applicability that would prevent other lines of evidence from counteracting WET data. Accordingly, clean water agency managers should not hesitate to present a demonstration that, despite an exceedance of the EPA recommended WET magnitudes, instream beneficial uses are not being impaired, and therefore the narrative water quality standard is not exceeded.

8. Instream Mixing/Dilution

Like chemical-specific pollutants, EPA and most state programs accept the concept of mixing zones within which WET numeric provisions do not apply. This is appropriate because WET effects are clearly concentration-dependent. EPA's Technical Support Document emphasizes this by noting "a discharger's chance of being charged incorrectly with causing instream toxicity is low if and only if dilution in the receiving water is considered."[27]

Because mixing will readily provide relief in terms of the stringency of WET limits, target values or other provisions, an evaluation of receiving water mixing should be an initial step in any clean water agency manager's efforts to address WET.[28] State programs frequently provide default mixing assumptions or protocols for defining acute and/or chronic mixing. Although these default procedures are typically very conservative, the clean water agency also has the option of using a "CORMIX" (a frequently used mixing protocol) or other standard or individualized mixing programs to better define mixing.

Discharges to receiving waters with tidal effects often are subject to substantially more mixing than what would be implied by steady state flow statistics. That is, the movement produced by the tides induces additional mixing. Although tidal mixing is more difficult to model than free flowing streams, modeling or empirical mixing studies are frequently worth the additional effort.

In a difficult case of application of WET limits, it may also be to the clean water agency's advantage to consider an instream diffuser in order to artificially induce additional mixing. Although this approach is seldom used for WET purposes, such use is fully consistent with frequent use of diffusers for chemical-specific purposes and with the structure of the NPDES program. Any such approach would involve design and construction costs. However, in the long term additional mixing may provide a more permanent and more reliable solution to WET results than treatment, pretreatment, incoming wastes modification or other compliance strategies. Any dilution gained also helps the clean water agency manager deal with chemical-specific water quality standards issues.

The clean water agency manager may also consider the application of Monte Carlo or other dynamic modeling procedures to assist in dilution demonstrations.[29] A Monte Carlo procedure uses a simulation involving variations in receiving stream flow, effluent volume, level of "toxicity" and any other identifiable variables affecting the occurrence of instream "toxicity." The procedure avoids the simultaneous use of multiple worst case variables, and predicts a more realistic profile of pollutant concentration or effect. Monte Carlo procedures are used with chemical-specific pollutants to examine instream impacts in conjunction with the exposure and recurrence intervals

[27] TSD Box 1-3 (emphasis in original).

[28] See TSD 2.1.1 & 5.2.2.

[29] See TSD 5.3.2.

associated with EPA's criteria and most state water quality standards. Although seldom used for WET purposes, Monte Carlo procedures have the potential to address instream impacts in a manner that is more accurate than the standard critical low flow dilution assumptions.

All of these approaches to mixing are designed to correctly reflect instream exposure to effluents. They are not devices that avoid any NPDES program requirement.

D. WET Permitting Approaches

The states' NPDES approaches to WET conditions vary from simple reissuance application requirements, to immediately effective numeric limits for WET, to complex Toxicity Identification Evaluation/Toxicity Reduction Evaluation (TIE/TRE) approaches. Each approach includes particular risks for clean water agencies. However, the WET Court Decision and the experiences of other clean water agencies provide opportunities for avoiding or improving on NPDES permit conditions that the states might otherwise impose.

1. Reapplication Monitoring Only

The only WET monitoring specifically required by the federal NPDES regulations for POTWs is for an application for permit reissuance. Reissuance monitoring should be four events, two species each event, for either acute or chronic toxicity. Further, the federal regulations do not provide a basis for a reopener provision in the event of WET data that the agency may see as justifying additional requirements. Rather, the regulations provide an exclusive list of reasons for permit modification by the agency, which do not include additional WET (or chemical-specific) data.[30] Although there is a more generic reopener provision, it is only for sludge regulation changes and other specific purposes. Accordingly, under the federal regulations, and with the possible exception of substantial new indirect dischargers or other fundamental changes inconsistent with the clean water agency's most recent permit application information, the clean water agency manager should oppose any attempt to include a permit reopener predicated on the results of WET testing. Instead, as with chemical-specific data, the five-year reapplication process is the NPDES agency's opportunity to consider RP.

Even for permits that impose only reapplication monitoring requirements, the clean water agency manager should carefully choose and work with its WET laboratory to define Data Quality Objectives and to make proper use of test method flexibility so that results reflect to the extent practical instream conditions and avoid anomalous indications of "toxicity." The clean water agency manager should work with the NPDES agency in an attempt to focus WET testing and reporting on test dilutions reflecting exposure assumptions and on the use of point estimates (LC50 or IC25).

Appendix B includes examples of appropriate reapplication monitoring requirements. As with Appendices C and D, NACWA does not consider all of the various WET provisions in the permit examples to be desirable examples for clean water agencies.

2. Routine WET Monitoring and TIE/TRE

Some state NPDES agencies impose more intensive or more routine WET monitoring, pursuant to specific state programs, because of a generic concern that treats WET data differently

[30] 40 CFR 122.62(a).

from chemical-specific data, or in response to existing POTW WET data raising concerns about "toxicity." In any such case, the clean water agency manager should address the same issues outlined immediately above for reapplication monitoring permit conditions.

As noted above, in the absence of a state regulatory provision, a permit requirement for routine WET monitoring does not have a strong regulatory basis, at least in programs similar to the federal program. However, the states often include routine monitoring in cases where existing WET data raise concerns, and it is clearly in the clean water agency's interests to have routine monitoring conditions rather than (1) numeric limits for WET or (2) a substantial debate with the NPDES agency over RP.

In any case of routine monitoring conditions, the NPDES agency obviously anticipates some use of the data if it should be unfavorable, and the clean water agency manager should work to craft appropriate WET program provisions. Those provisions would ideally include a critical WET value or target number based on averaging of multiple test results or the exceedance of critical values in at least two consecutive tests, a retest provision in the event of unfavorable results, and progression to a phased TIE/TRE program rather than numeric WET limits.

The WET Court Decision specifically supports a clean water agency's efforts in the following ways.

- The Court's observation that state NPDES agencies have the discretion to set toxicity thresholds to compensate for local conditions supports the setting of WET target numbers or thresholds that incorporate available mixing. Mixing evaluations should be undertaken when the clean water agency manager anticipates routine WET monitoring conditions because such permit conditions will typically lead to either TIE/TRE procedures or more immediate WET limits, and it will be more difficult to address mixing after the thresholds for those steps have been already established. In particular, later mixing work has missed the opportunity to focus previous test dilution series on the more correct endpoint, and has therefore missed the opportunity to maximize the reliability of the resulting data base.
- The Court's warning against the use of any single WET test result should inform the manner in which the permit handles early results before there is a substantial data set. Decisions on advancing to TIE/TRE procedures or numeric limits should only be made after there is a data set adequate to characterize the data given the variabilities in results demonstrated by EPA's interlaboratory data. EPA recommends that RP and limit determinations be based on at least ten tests.
- Building on the WET Court Decision that states that "nothing forcloses consideration of the validity of a particular test result in an enforcement action," it is also true that nothing should forclose consideration of the validity of test results applied to any permit purpose. This is where the clean water agency manager's advance work with its WET testing laboratory and the establishment of Data Quality Objectives may be critical.
- TUs should be used only as a final expression of permit target numbers. Any calculations should be performed on WET data expressed as percent.

Appendix C includes examples of appropriate routine WET monitoring provisions, though the other WET provisions are not necessarily desirable for clean water agencies.

3. Reasonable Potential Determination and Numeric WET Permit Limits

The least desirable result for a clean water agency is an NPDES permit agency determination that reasonable potential exists for the effluent to cause or contribute to an excursion beyond the state's narrative water quality standards. In any such instance the clean water agency manager's first task is to determine whether, in fact, RP exists. EPA's TSD RP procedures (which Great Lakes Initiative states and some other states use) are very restrictive, better designed for chemical-specific RP (the accuracy of chemical-specific data being determinable and CVs being more properly determinable), and would in the majority of cases without substantial instream dilution bring about a finding of RP. However, the federal regulations and many states' regulations or procedures involve a wider inquiry. The agency must consider the variability of the pollutant, the sensitivity of the WET test species, instream dilution,[31] and any other available information.[32] Instream dilution has been addressed above. In an appropriate case, the clean water agency manager should consider instream mixing evaluations or steps to modify instream dilution through a diffuser to avoid or reduce the adverse impact of WET limits. The clean water agency manager and the public should understand that both natural and induced mixing are consistent with the NPDES program, and any potential adverse effects within mixing zones are also addressed by the program. The use of mixing processes is not an avoidance of Clean Water Act responsibilities.

In terms of pollutant variability, the WET Court Decision warned of the inappropriateness of the use of single test failures for enforcement. The same principle should apply to other uses of single tests showing unusual results. In any case where there are intermittent adverse WET data, the clean water agency manager has the opportunity to evaluate its WET data in the context of EPA's interlaboratory variability data for the test species and test endpoint. EPA claims that its WET methods exhibit interlaboratory variability (CV) between 11 and 44 percent. Although NACWA believes that these CVs are systematically understated, test species with higher variabilities present a more compelling case that one or two test exceedances within a permit cycle should not be the principal basis for a RP determination. EPA's claimed CVs are provided in its rulemaking record.[33]

In terms of sensitivity of the WET test species, the WET Court Decision noted that state NPDES agencies may set "toxicity" thresholds to mitigate the lack of correlation between test results and true instream impact (representativeness). NACWA believes that representativeness, or the lack thereof, is a principal disconnect between WET testing and its use as if the data were chemical-specific data. This is particularly true for the chronic sublethal endpoints, and at relatively low levels of "toxicity" (at or near 1.0 TUc). Generally, NACWA believes that EPA does not have data that demonstrate representativeness in the absence of lethal conditions. Rather, EPA's conclusions of representativeness for the sublethal endpoints involve an assumption based on data from tests where there was lethality. Although the burden should be on the permit agency to demonstrate representativeness as to any particular RP determination, the agency is likely to rely on general EPA claims regarding its methods. In a strict legal sense such general claims, in the absence of specific data, should not be seen as supporting an RP determination. However, in a

[31] 40 CFR 122.44(d)(1)(ii).

[32] Id. 122.44(d)((1)(v).

[33] Interlab Variability Study Tables 9.65 & 9.66, pp. 150-51.

practical sense, the clean water agency manager will need to advance the issue by raising specific doubts about representativeness.

EPA's WET rulemaking and the briefs of NACWA and the other petitioners can provide additional guidance for a demonstration concerning representativeness. For example, if a claimed RP determination were to be made using the chronic marine test methods, the clean water agency manager might take the initiative and himself cite EPA's support for representativeness. The briefing materials provide a cogent argument as to how that EPA position is without basis.[34] At this point the burden should shift back to the agency to produce specific data showing representativeness of chronic sublethal effects. NACWA found in its research that such data do not exist. EPA has stated that "no single marine case study has been designed with the goal to comprehensively evaluate [the representativeness] relationship." Also, "[w]e have discovered no case studies in the scientific literature that describe a detailed analysis of the toxicity of an effluent discharging to the estuarine or marine environments or that also describe a corresponding impact on the water column and benthic communities of the receiving system."[35] Pointedly, "WET tests originally were not designed to predict receiving system impacts."[36]

Similarly, data do not show the representativeness of the freshwater chronic sublethal endpoints. Helpful documentation is available in the WET litigation materials. However, those materials do not fully develop the freshwater issue, and additional research into EPA's Comprehensive Effluent Testing Program documents will be necessary. The CETP documents from the 1980s report on investigations of representativeness conducted on several waterways. For some of the work no WET test toxicity was identified, and the testing proved nothing.[37] For other work a relationship was purportedly identified between effluent WET test results and WET test results on receiving water samples.[38] However, this merely demonstrates that the researchers were able to identify points in the receiving waters where the appropriate concentration of effluent remained, and that the distinction between laboratory dilution water and ambient water did not make a substantial difference. In the Back River Maryland study a relationship between WET test results and instream biological quality was not identifiable.[39] Other studies purported to identify a specific relationship between WET test results on effluent and instream biological impact.[40] Such identifiable effects occurred only where there was test organism mortality, and that they prove nothing about representativeness for relatively low toxicity manifested in chronic sublethal

---

[34] See, e.g. Reply Brief of Petitioners at 34 in WET Litigation.

[35] Whole Effluent Toxicity Testing: An Evaluation of Methods and Receiving Stream Impacts, Society of Environmental Toxicology and Chemistry (SETAC) section 10.3 ("Predicting Receiving Stream Impacts from Effluent Toxicity: A Marine Perspective" 1996) (the "Pellston Conference") (a "Discussion Initiation Paper" by EPA staff on behalf of EPA).

[36] Id.

[37] Validity of Effluent and Ambient Toxicity Tests for Predicting Biological Impact, Scippo Creek, Circleville, Ohio EPA 600/3-85/004 (June 1985).

[38] Validity of Effluent and Ambient Toxicity Tests for Predicting Biological Impact, Back River, Baltimore Harbor, Maryland EPA 600/8-86/001 (July 1986).

[39] Id.

[40] See, e.g. Validity of Effluent and Ambient Toxicity Tests for Predicting Biological Impact, Skeleton Creek, Enid, Oklahoma EPA 600/8-86/002 (Mar. 1986).

endpoints. However, it is this point that appears to require further research into the specific test data.

As noted above, EPA's regulations also require the consideration of any other available information as part of an RP determination. Other information may include favorable bioassessment data, favorable results of scans for chemical-specific pollutants included in the clean water agency's reapplication data, fisheries assessments from the state's environmental or fisheries resource agencies, and any other data supporting the attainment of designated uses in the receiving waters. As discussed above, EPA's policy of independent applicability should not prevent the consideration of such additional water quality data. In the absence of numeric criteria for WET within the state's water quality standards, and in the absence of a state regulation disallowing a weight of evidence approach, a demonstration that benthic metrics and water column biological integrity are not impaired (in critical low flow conditions) should offset WET data.

TUs should be used only as a final expression of permit limits or target numbers. Any calculations should be performed on WET data expressed as percent.

Appendix D includes examples of appropriate WET permit limit provisions, though the other WET provisions are not necessarily desirable for clean water agencies.

E. Additional Considerations

1. Compliance Schedules

Permit compliance schedules, whether for TIE/TRE programs or numeric WET limits should provide sufficient time for the required tasks. Under federal law and most states' programs a compliance schedule within a permit may extend at least to the end of the permit term. Given the difficulty of WET programs it would not be unreasonable in many cases to receive all or the majority of a five year permit term to complete a TIE/TRE program or to achieve compliance with numeric WET limits.

NPDES permit provisions imposing TIE/TRE requirements and eventual WET limits are best drafted to move progressively through each step, rather than providing within the permit the eventual numeric limit. For example, an appropriate compliance schedule would require workplan submission and allow a defined period to complete the TIE/TRE. Based on the program findings and conclusions, the NPDES agency would then determine whether numeric WET limits are required and, if so, what those limits will be. This avoids the argument that compliance with the limit must be within the five year permit term. It also preserves for a later time the clean water agency's arguments against WET limits or over the specific numbers. If, for example, the clean water agency's program is successful in identifying and removing an offending wastestream that had caused the "toxicity," no numeric limits should be seen as necessary. That is, with the source of the problem gone, there is no RP. Alternately, if the NPDES agency sees the problem as being addressed through new treatment processes or changes to existing processes, the agency will have a good argument that, although the "toxicity" is now gone, numeric limits are necessary to require and maintain the increased level of treatment.

2. Removing WET Limits and Permit Conditions

The removal of numeric WET limits and other permit conditions should be handled in a similar manner to chemical-specific limits and corresponding permit conditions. Numeric limits subject to a compliance schedule may be removed through permit modification prior to the date the limits become effective without violating "antibacksliding" restrictions.[41] Although EPA and state NPDES agencies will typically take the position that limits may not be removed after becoming effective, that is an overbroad reading of antibacksliding. Water quality-based limits may be revised or removed as long as (1) for waters attaining state standards, antidegradation requirements are met, and (2) for waters not yet attaining state standards, the cumulative effect of wasteload allocations provides for standards attainment.[42] Also, limits may be changed or removed if justified by alterations to the POTW, new information or other specified causes.[43] Although these provisions are complex and often poorly understood, they allow changes in or removal of permit conditions in many appropriate circumstances.

3. What Constitutes a Violation?

An exceedance of a currently effective numeric WET permit limit constitutes a permit violation and a violation of federal and state law. However, WET triggers and other conditions short of traditional limits should be carefully drafted so that the permit requires WET testing, TIE/TRE procedures and other management provisions in a way that is binding and that allows the agency to properly enforce the permit, without prematurely characterizing specific WET test results as violations. WET testing should typically be on a specified schedule and require the reporting of results with the Discharge Monitoring Report for the month of the test. A failure to test and report as required is a permit violation. Because of the difficulties inherent in WET testing, retests will frequently be required in order to obtain a test that meets the clean water agency's Data Quality Objectives and QA/QC criteria. The permit should ideally reflect this aspect of WET testing and state that, in the event of such testing difficulties, a retest as soon as practicable is in compliance with the testing requirement. Even in the absence of such a permit provision, NACWA believes that the record adequately identifies the difficulties with WET testing, and the clean water agency manager would have a viable impossibility or other defense to a charge of noncompliance based on a WET test properly rejected because of QA/QC problems.

A separate issue from retests because of Data Quality Objectives failure is retests based on WET data variability, and the fact that at least occasionally a WET test will show "toxicity" when there is no effluent toxicity.[44] NACWA believes that the record justifies a permit retest provision for any final NPDES permit numeric WET limits. However, in low dilution situations (with WET limits of 1.0 TUc or only slightly higher) a retest averaging provision is generally of little value - because no number of 1.0 TUc results will average a greater than 1.0 TUc result down to 1.0. In

[41] National Whole Effluent Toxicity (WET) Implementation Guidance Under the NPDES Program 5.3.2, EPA 832-B-04-003 (Nov. 2004) (draft) (EPA Implementation Guidance).

[42] 33 U.S.C. 1313(d)(4) & 1342 (o)(1).

[43] Id. 1342 (o)(2).

[44] WET Court Decision at 9.

recognition of this, EPA's WET Guidance approves of the use of median rather than average value in such situations, although with some substantial restrictions.[45]

TIE/TRE procedures are typically step-by-step analytical exercises where the eventual result is unknown until completion, and where the path to completion itself is unknown. Therefore, a properly constructed permit condition should require an approvable facility-specific TIE/TRE work plan by a specific date, and should require that the work plan itself include specific enforceable milestones allowing the NPDES agency to properly enforce the permit requirements. A permit violation would result if the clean water agency failed to develop the work plan on time or failed to implement and report to the NPDES agency on the specific work plan steps on time. The clean water agency manager's task is to develop the work plan in a manner that satisfies the NPDES agency's need for enforceable conditions on an identifiable timeline, which gives itself and its consultants time to perform the TIE/TRE procedures, and that where practical bases its requirements on elapsed time after the prior required NPDES agency approval rather than on a specific date.

F. TIE/TRE Requirements

1. Timeframe

NPDES permit requirements for TIE/TRE procedures should ideally not specify a date for completion. Rather, the procedure is a step-by-step process where an identification of the pollutant or condition leading to the indication of effluent "toxicity" may be identified at an initial point, or initial results may rule out a particular conclusion and require subsequent analytical steps. There is ample EPA guidance on performance of TIE/TRE procedures to establish during the permit process either an open-ended program or a program that is likely to provide enough time for completion based on past experience. Generally, NACWA believes that a permit should not anticipate TIE/TRE completion in less than two years from initiation.

2. TIE/TRE Goal

The single goal of a TIE/TRE process should be to identify the pollutant, combination of pollutants or other factors causing the WET testing indications of "toxicity," and to identify a treatment or management approach to remove or correct the cause in a manner that either (1) results in future acceptable WET test results or (2) identifies the cause of the adverse WET test results as a factor other than effluent "toxicity."

3. Degree and Variability of Toxicity

TIE procedures are only effective if sufficient toxicity is available over a number of consecutive tests. The clean water agency manager should strive for language defining these conditions if the NPDES permit agency includes language in a permit that specifically deals with this level of decision making, rather than including a permit requirement for the submission of a TIE plan that will address such details. For example, the clean water agency manager could propose that the TIE not be initiated until a pattern of significant toxicity is measured (e.g. two consecutive tests exceeding the permit decision trigger by more than 20 percent). This approach will help ensure that false positives will not drive the process. Also, there should be language addressing specifically when a TIE has been completed (e.g. two or three consecutive tests meeting

[45] EPA Implementation Guidance at 5.2.4.

the permit decision trigger). This latter language will help ensure that a TIE does not continue indefinitely.

G. Conclusions

This White Paper has presented guidance on a number of issues that may be useful for a clean water agency manager in addressing WET testing and NPDES permitting in a specific POTW permit situation. Some of the more important issues addressed in detail are listed below. Achieving an acceptable result for the clean water agency NPDES permit in a specific case may involve utilizing one or more of these issues to properly interpret and use WET data, to formulate appropriate permit conditions and convince the NPDES agency of the appropriateness such conditions, and to build an administrative record to support a correct NPDES agency decision.

- Distinctions among WET test endpoints, the metrics for expressing endpoints (e.g. the chronic NOEC and IC25), and EPA's recommendations for point estimates (e.g. IC25).
- Test method flexibility.
- QA/QC and Data Quality Objectives, and in particular concentration-response relationships.
- Frequency of WET testing, retests, and EPA and WET Court Decision statements concerning the use of single test results in enforcement.
- WET test dilutions.
- EPA data demonstrating false positive rates for the WET test methods, at least on moderately "toxic" samples.
- Use of WET data expressed as Toxicity Units, and the use of data expressed as percent (rather than TU) in any statistical or other manipulations of the data.
- Differences in WET data results between hypothesis tests and point estimates.
- Use of WET instream "toxicity" criteria.
- Instream mixing and dilution.
- Proper expression of NPDES permit requirements for reapplication and routine WET monitoring.
- Triggers for TIE/TRE requirements, and effective TIE/TRE permit conditions.
- Representativeness, or the degree to which WET data accurately predict instream toxicity.
- The use of state narrative water quality standards, RP, and NPDES permit numeric limits for WET.