

CHAPTER 2.9 – WET Test Variability

Whole effluent toxicity (WET) test variability is the result of a number of factors. The primary purposes of this chapter are to discuss potential sources of variability and to present ways for addressing and reducing that variability.

NOTICE: This document is intended solely as guidance, and does not contain any mandatory requirements except where requirements found in statute or administrative rule are referenced. This guidance does not establish or affect legal rights or obligations, and is not finally determinative of any of the issues addressed. This guidance does not create any rights enforceable by any party in litigation with the State of Wisconsin or the Department of Natural Resources. Any regulatory decisions made by the Department of Natural Resources in any matter addressed by this guidance will be made by applying the governing statutes and administrative rules to the relevant facts.

Types Of WET Variability

Whole effluent toxicity (WET) test variability has been discussed by a variety of groups in many different forums. This chapter is an attempt to generally discuss WET variability and ways to control it. More detailed guidance has been written by other groups, such as the Society of Toxicology and Chemistry (SETAC) “Expert Advisory Panel on WET Test Performance Evaluation and Data Interpretation” (see <http://www.setac.org/wetindex.html>). WET variability issues were first formally discussed at a “Pellston Workshop”, held September 16-21, 1995, at the University of Michigan Biological Station on Douglas Lake in Pellston, Michigan. This workshop included participation of experts from academia, industry, and government who were selected because of their experience and knowledge of WET test methods. The workshop provided a structured environment for the exchange of ideas and debate such that consensus positions could be derived and documented for some of the issues surrounding the science of WET testing. The participants at the 1995 Pellston Workshop categorized the different types of WET variability and highlighted ways that WET variability could be reduced (Grothe, et al., 1996). This group categorized WET variability into 3 types:

- **Intratest (within-test) variability.** Sources of intratest variability include the number of replicates, the number of organisms per replicate, and the sensitivity differences between organisms.
- **Intralab (within-lab) variability.** Intralab variability is that which is measured when tests are conducted under reasonably constant conditions in the same lab. Sources of intralab variability include those sources described for intratest variability, plus differences in: 1) test conditions (e.g., seasonal differences in dilution water & environmental conditions), 2) organism condition/health, and 3) analyst performance.
- **Interlab (between-lab) variability.** Interlab variability reflects the degree of precision that is measured when a sample is analyzed by multiple labs using the same methods. Variability measured between labs is a consequence of variability associated with both intratest and intralab variability factors, plus differences allowed within the test methods, technician training programs, sample and organism culturing/shipping effects, testing protocols, and testing facilities.

Pellston Workshop participants determined that both the regulatory and regulated communities could significantly influence factors that affect WET variability. They found that WET variability could be limited by controlling the factors that have the most influence:

- ❖ **Strict adherence to clearly specified methods.** Improper utilization of WET methods can have a substantial impact on variability. Since United States Environmental Protection Agency (EPA) methods must be written to apply to a wide range of different situations (e.g., different regions, climates, environments), the methods contain optional steps which can be used to fit different situations. In order to control WET variability, it is necessary to limit “optional” test methods that can cause differences in how tests are conducted.

- ❖ **Increasing analyst & regulator experience.** The experience and qualifications of the analyst performing the test will dictate how well culture and test methods are followed and the extent to which good judgment is exercised when issues arise in the process of conducting the test, analyzing data, and interpreting results. The issue of experience is of concern not only in relation to test results, but also in relation to the development and implementation of WET requirements by regulators. Although regulator experience does not directly influence WET variability, it is a key factor that determines how WET is implemented in a regulatory context.
- ❖ **Selection of quality labs.** Along with organism health (which is linked to lab quality), lab quality was considered by Pellston workshop participants to be one of the most important factors affecting test variability. Quality WET labs should be able to demonstrate a serious commitment to a quality assurance/control program that extends beyond analyst experience. Considerations such as an ongoing reference toxicant program, a review process for all toxicity test data and reports, a good sample custody tracking system that is always used, proper equipment maintenance, dilution water quality monitoring, facility maintenance, and attention to test organism health are all characteristics of a lab that is committed to generating quality data.

Edison Electric Institute & Western Coalition of Arid States, et al. vs. USEPA

Soon after the EPA promulgated its WET methods in October 1995, a lawsuit was filed by the electric power industry and some local governments and business groups that aimed to repeal the EPA rule requiring facilities to perform WET tests. The question before the court focused mostly on the chronic test methods, which the plaintiffs claimed were fallible and resulted in unacceptable levels of "false positives".

During negotiations surrounding the lawsuit, "evidence" was presented which allegedly showed the *"rate of false positives for some of these tests can be nearly 40%"*. These allegations have caused some concern among various permittees. It is very important to note that in Wisconsin's experience, a rate of false positives of 40% is highly unlikely, since WET tests conducted by Wisconsin permittees have only shown a failure rate of about 15% for acute and 19% for chronic from all of the WET tests done in the last ten years.

In 1998, EPA entered into a settlement agreement (see Attachment 1) to address some of the concerns raised by the parties in the lawsuit. EPA agreed to evaluate their WET methods and look for some ways to further control WET variability. EPA conducted an inter-lab study to evaluate the precision of 3 freshwater chronic tests, 4 marine chronic tests, and 5 acute tests. As a result of these studies, EPA published a series of documents related to WET variability (see <http://www.epa.gov/waterscience/WET/wetstudy.htm>) and made modifications to their WET test methods in November 2002 (see <http://www.epa.gov/OST/WET/>).

In addition to these EPA reports, earlier peer-reviewed data was published which discussed WET variability. The *"Technical Support Document for Water Quality-based Toxics Control"* (EPA/505/2-90-001) references published studies that show *"the precision of WET tests is similar to chemical-specific methods"* (pp. 11-12). In addition, interlab studies have been completed and published for the fathead minnow and *C. dubia* chronic tests (see references, below). These studies showed good reproducibility for these methods. Other researchers have also agreed that the precision of each of these methods is acceptable. Rue, et al, concluded that WET methods *"are comparable to accepted analytical methodologies"*. Another study by Grothe, et al, concluded that *"when comparing CVs for select toxicity test methods and commonly accepted analytical methods...the precision of both techniques is similar"*.

In December 2004, a federal appeals court ruled in EPA's favor, saying that the plaintiffs failed to prove that EPA's use of the WET tests are *"arbitrary, capricious, an abuse of discretion or otherwise not in accordance with the law."* Among the plaintiffs' contentions rejected by the court was that WET tests *"produce an unacceptably high number of false positives"*. The court noted that the plaintiffs had defined false positives more broadly than EPA, which claims that only 1-5% of WET tests detect toxicity when there is none. While the court

acknowledged that tests could find, on rare occasions, that toxicity was present when it was not, it pointed out that the plaintiffs failed to mention that there was the same likelihood that "*other permittees who should be violators may be deemed in compliance.*" The court decided that WET tests were no more variable than other permit-required tests, stating that "*perfection is not the standard against which we judge agency action. EPA's decision was informed by years of scientific studies, negotiation, and public notice and comment, and it represents the agency's expert judgment regarding the implementation of the aims of the Clean Water Act.*"

What Has Wisconsin Done To Control WET Variability?

Wisconsin has addressed many of the issues that have the potential to influence WET variability. The changes recommended at the Pellston workshop to improve WET variability have already been addressed in Wisconsin:

- **Strict adherence to clearly specified methods.** In addition to analyst experience and organism health, the following factors can affect WET variability: 1) sampling procedures (sample volume, type, storage conditions; frequency of composite sub-sampling), 2) sample holding time, 3) test duration, 4) deviations in feeding & environmental conditions (light, pH, temperature, DO, etc.), 5) dilution water, 6) number of concentrations and replicates tested, and 7) number of organisms per replicate. Each of these is addressed in EPA methods, but flexibility is allowed so states can make tests fit in specific situations. The more flexibility allowed in test methods, the higher the chance that tests will be done differently between labs or between tests, resulting in increased WET variability.

In order to control WET variability and improve the consistency of methods used by Wisconsin labs and permittees, the WDNR created the "*State of Wisconsin Aquatic Life Toxicity Testing Methods Manual, Edition 2*" (Methods Manual) and incorporated it by reference into ss. NR 149.22 and NR 219.04, Wis. Adm. Code. The Methods Manual contains specific procedures regarding testing and sampling procedures, types of tests, quality control/quality assurance procedures, etc., that labs must follow when performing WET tests for permit compliance. The WDNR also created the WET Program Guidance Document, in 1996, in order to provide guidance regarding issues that may effect test variability (e.g., proper sampling procedures, selection of a WET lab, use of representative data, etc.).

- **Increasing analyst & regulator experience.** In order to insure that WET tests are conducted by qualified, experienced analysts, section 3.17 of the Methods Manual sets forth minimum qualifications and training requirements that each lab must follow in order to maintain laboratory certification.

Wisconsin has an advantage over most states due to the dedication of a trained and experienced toxicologist to our WET program. The "Biomonitoring Coordinator" is responsible for the review and interpretation of WET tests and makes decisions when data is unusual or problems have occurred during testing. The Biomonitoring Coordinator also leads the "WET Team" which develops WET policy and guidance.

- **Selection of quality labs.** In order to insure labs are of the highest quality and are able to demonstrate a serious commitment to a quality assurance/control program, the WDNR, under state statutes, certifies labs to perform different types of environmental analysis. In order for a lab to apply for certification for acute and chronic WET testing, the lab must submit a completed application and a quality assurance plan to the lab certification program and pass an on-site evaluation. WET labs must have an ongoing reference toxicant program, a review process for all test data and reporting, a good sample custody system, proper equipment maintenance, dilution water quality monitoring, facility maintenance, and attention to test organism health, and make other demonstrations of good lab practices in order to pass an audit.

References:

Fathead Minnow 7-day Test: Round Robin Study, Intra- and Interlab Study to Determine the Reproducibility of the Seven-day Fathead Minnow Larval Survival and Growth Test, August 1988, American Petroleum Institute, API Publication No. 4468.

Precision of the EPA Seven-day Ceriodaphnia dubia Survival and Reproduction Test, Intra- and Interlab Study, May 1990, National Council of the Paper Industry for Air and Stream Improvement, Inc, NCASI Technical Bulletin No. 588.

A Review of Inter- and Intralab Effluent Toxicity Test Method Variability, 1988, Rue, W.J., J.A. Fava, and D.R. Grothe, Aquatic Toxicology and Hazard Assessment: 10th Volume, ASTM STP 971.

A Perspective on Biological Assessments, 1990, Grothe, D.R., R.A. Kimerle, and C.D. Malloch, Water Environment and Technology.

Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving Stream Impacts, 1996, Grothe, D.R., Dickson, K.L. and Reed-Judkins, D.K. (eds.), SETAC Special Publication, SETAC Press.

Attachment 1: WET Methods Lawsuit Settlement

Edison Electric Institute & Western Coalition of Arid States ("WestCAS") v. EPA

The following is a summary of the lawsuit settlement, provided by EPA in 1998:

1. Parties are Edison Electric ("Utilities") and Western Coalition of Arid States ("WestCAS").
2. The Settlement Agreement stays the litigation pending completion of EPA tasks. After EPA completes the tasks, the parties will dismiss the litigation. The settlement is NOT a consent decree; EPA has no independently enforceable obligation to complete the tasks.
3. The Final Rule, published in November 1995, standardized testing procedures to measure whole effluent toxicity using seventeen different procedures/species: 4 freshwater chronic; 6 marine chronic; and 7 acute.

Studies

4. Under the settlement, EPA will conduct multilab studies to evaluate precision of 3 of the freshwater chronic tests, 4 of marine chronic tests, and 5 of the acute tests. The design of the studies, and the results of the studies, will be subject to a formalized peer review. The studies will begin in the autumn of 1998 and the peer review on the study results should be completed by the summer of 2000.
5. After review of the study results and other information, EPA will propose to ratify or withdraw each of the 17 methods. EPA will propose that action in September 2000 and take final action in September 2001.
6. EPA is confident that the studies will formally verify the Agency's decision to standardize the test procedures for compliance monitoring of WET.

Methodology Requirements and Recommendations

7. EPA will take additional actions to modify the test method manuals. Some provisions of the test method manuals are mandatory in order to assure consistency in measurement; some are discretionary in order to optimize successful test completion.
8. In January 2000, EPA will propose rulemaking to convert some of the discretionary provisions into mandatory ones, specifically, the *Ceriodaphnia* reproduction test will now require randomization and blocking-by-parentage and each of the test methods will require the development of a valid concentration-response as a prerequisite for a "valid" test result. EPA will also propose standardized procedures to control pH shift in test chambers.
9. Also in January 2000, EPA will prepare and distribute guidance and recommendations regarding adjustment to statistical error rate assumptions, confidence intervals, available dilution waters, and permutations of "valid" concentration-responses.
10. In January 2001, EPA will revise the test method manuals to incorporate the new requirements and recommendations.
11. EPA believes that the existing test method manuals provide sufficient direction and flexibility to accomplish the revisions that would be incorporated into the January 2001 edition, but that the revisions will improve the test methods nonetheless.

Analytic Variability Guidance

12. Variability in test results arises from a variety of sources, including effluent variability, analytic variability, and analyst variability, among others. Procedures to implement test procedures in NPDES permitting do not currently account for variability in a comprehensive way. Technology-based standards, i.e., Effluent guidelines, account for effluent variability and analytic variability after direct measurement of actual discharges. Procedures for water quality-based permit writing account for effluent variability, but not necessarily analytic variability or analyst variability.
13. Under the settlement, EPA will prepare and distribute guidance on how to take analytic variability into account in determining the need for and in the derivation of an effluent limitation for whole effluent toxicity. The guidance will be subject to a formalized peer review and will be available in April 2000. Ultimately, the guidance may either supplement or supplant guidance to permit writers contained in EPA's Technical Support Document for Water Quality-Based Toxics Control.