U.S. Department of Justice

Environment and Natural Resources Division

90-5-1-7-1318

Environmental Defense Section
P.O. Box 23986
Washington, DC 20026-3986

Telephone (202) 514-4767
Facsimile (202) 514-8865

March 17, 2005

VIA Hand Delivery

Mark L. Langer, Clerk
United States Court of Appeals for the
District of Columbia Circuit
United States Courthouse
3rd and Constitution Avenue, N.W.
Washington, D.C. 20001

Re: *Edison Electric Inst., et al. v. EPA,* Case No. 96-1062 (and consolidated cases)

Dear Mr. Langer:

Please find enclosed for filing in the above-referenced petition for review the original and five copies of the following document:

- **RESPONDENT EPA'S OPPOSITION TO WET COALITION AND WESTCAS'S PETITION FOR PANEL REHEARING**

Please file stamp the extra copy and return it to me in the enclosed return envelope.

Thank you in advance for your assistance.

Sincerely,

David S. Gualtieri
Counsel for Respondents

cc (by U.S. Mail): Counsel on attached Certificate of Service
Steve Sweeney, U.S. EPA OGC

ORAL ARGUMENT OCCURRED ON OCTOBER 15, 2004

UNITED STATES COURT OF APPEALS
FOR THE DISTRICT OF COLUMBIA CIRCUIT

EDISON ELECTRIC INSTITUTE, *et al.*, Petitioners, v. UNITED STATES ENVIRONMENTAL PROTECTION AGENCY, *et al.*, Respondents.	Docket No. 96-1062 (and consolidated cases)

RESPONDENT EPA'S OPPOSITION TO WET COALITION AND WESTCAS'S PETITION FOR PANEL REHEARING

TABLE OF CONTENTS

INTRODUCTION

On December 10, 2004, a unanimous panel of this Court denied the petitions for review of certain whole effluent toxicity ("WET") test methods promulgated by the United States Environmental Protection Agency ("EPA") pursuant to the Clean Water Act, 33 U.S.C. § 1251 *et seq.* ("CWA"). *Edison Electric Institute v. EPA*, 391 F.3d 1267 (D.C. Cir. 2004) (the "Opinion"). Petitioners Western Coalition of Arid States ("WESTCAS") and the WET Coalition (collectively, the "Petitioners")[1] now seek panel rehearing because, Petitioners claim, the Court erred in rejecting – and demonstrating the fallacy of – Petitioners' independent calculation of statistical variability of the WET test methods (known as a coefficient of variation or "CV"). The Rehearing Petition lacks merit and should be denied for numerous reasons, including that the Court did not, in fact, err.

First, Petitioners' methodology was fatally flawed, as Petitioners relied upon a dataset that EPA excluded from its CV calculations (for technical reasons explained in the record). In calculating their own CV, Petitioners contravened repeated and explicit EPA guidance that data derived from hypothesis testing (*i.e.*, NOEC data) not be used for such a calculation. Second, the CV calculations and arguments that Petitioners rely upon to support rehearing were raised for the first time either in their reply brief or in the instant Rehearing Petition, meaning that EPA never had the opportunity to address the numerous flaws in Petitioners' methodology. As a result, every argument Petitioners now raise is waived. Third, as confirmed by the attached affidavit from the EPA statistician responsible for such issues in the WET rulemaking, the Court did not err, and there is no factual basis to conclude that the Court made erroneous assumptions about the data as Petitioners allege. The Court's calculation was correct in the context for which it was provided and demonstrates that Petitioners' CV calculation was flawed and properly rejected. Finally, the issue on which Petitioners seek rehearing is contained in an explanatory footnote in the Opinion that is merely incidental to the Court's overall determination that the WET test methods are "precise."

---

[1] Under D.C. Circuit Rule 35(f), EPA need not respond to the amicus brief and affidavits attached to the motion of proposed *amici curiae*. Such briefs are allowed only by Court invitation. Regardless, the amicus brief and related affidavits contain many of the same flaws addressed herein and the Fox Declaration.

Even if Petitioners are correct that the Court erred, there are numerous other grounds upon which the Court determined that the WET test methods are precise. Thus, any error is harmless. The Rehearing Petition therefore should be denied.

BACKGROUND

The underlying petitions for review challenged numerous aspects of the WET test methods, and the Court rejected each of those challenges. In short, the Court found that: EPA evaluated the WET test methods in accordance with appropriate criteria, 391 F.3d at 1269; the WET test methods are precise (*i.e.*, the methods do not exhibit excessive variability), *id.* at 1271; the WET test methods do not result in excessive "false positives," *id.* at 1272; EPA reasonably decided not to apply detection limits to WET tests, *id.* at 1273; EPA demonstrated the "availability and applicability" of the WET test methods, *id.*; and EPA demonstrated that WET tests can effectively predict instream impacts (*i.e.*, that they are "representative"). *Id.* at 1273-74. Because the Rehearing Petition concerns only the issue of WET method precision, we discuss below how EPA evaluated precision in the rulemaking, how Petitioners evaluated precision in their reply brief, and how Petitioners' methodology deviated from EPA guidance on the statistical methods for evaluating precision.

I. HOW EPA EVALUATED WET TEST METHOD PRECISION

Petitioners seek panel rehearing of the Court's disposition of one of many arguments that Petitioners raised with respect to the issue of WET test method precision.[2] EPA's Interlaboratory Study[3] demonstrated the reliability of 12 WET test methods by evaluating their measurement variability. The study generated dozens of datasets, as four types of samples (effluent, reference toxicant, blank, and receiving water) were evaluated for each test method. Interlaboratory Study at 8, J.A. 1151. EPA determined the variability of each WET test method by calculating its coefficient

---

[2] We refer the Court to EPA's merits brief ("EPA Br.") for a more comprehensive discussion of WET test method precision. *See* EPA Br. at 28-29, 49-52, 84-85.

[3] Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Whole Effluent Toxicity Test Methods, Vol. 1 (Sept. 2001) ("Interlaboratory Study"), J.A. 1123-1296.

of variation ("CV"),[4] the means of determining precision identified in the July 1998 Settlement Agreement. Settlement Agreement, ¶ 2 & Ex. B, ¶ 2, J.A. 687, 703-04; *see also* EPA Br. at 24-26, 28 (discussing purpose and scope of the Settlement Agreement).

In the Interlaboratory Study, EPA evaluated WET test method results generated by two statistical methods set forth in the WET test Methods Manual: point estimation and hypothesis testing. *See* EPA Br. at 16-22. The first, point estimation, compares, by plotting on a graph, biological effect (*e.g.*, number of *Ceriodaphnia* offspring produced) to, for instance, effluent concentration to determine an endpoint, such as the 25% Inhibition Concentration ("IC_{25}") – the concentration of effluent that would cause a 25 % reduction in a measurement such as reproduction or growth. Point estimation, therefore, enables the laboratory to determine the exact concentration at which a specified toxic effect occurs.

The second statistical method is hypothesis testing, wherein the laboratory determines, per statistical calculation, whether the hypothesis that the effluent, at a specified concentration, does not have a toxic effect can be rejected. The result is expressed as, for example, the No Observed Effect Concentration ("NOEC"). Hypothesis test results are limited to a finite set of discrete values corresponding to the effluent concentrations used (typically, and in the case of the Interlaboratory Study, 6.25%, 12.5%, 25%, 50%, and 100%). By contrast, WET test results based on point estimation can be any value above 0% up to 100%.[5]

The Interlaboratory Study and its design were subject to comment from Petitioners, as well as two rounds of intensive peer-review.[6] The "Study Design and Objectives" section of the

---

[4] The CV is "a standard statistical measure of the relative variation of a distribution or set of data, defined as the standard deviation divided by the mean." Interlaboratory Study at xvii, J.A. 1139.

[5] Although dischargers can use either method, EPA recommends point estimation. 67 Fed. Reg. at 69,957.

[6] EPA Response to Comments ("RTC") at 204-279, J.A. 1904-1972; EPA Response to Comments: Peer Review Report of the Interlaboratory Study (Sept. 2001), J.A. 1051-1122; Summary Report Peer Review of "Preliminary Report: Interlaboratory Variability Study of EPA Short-Term Chronic and Acute Whole Effluent Toxicity Test Methods" (Mar. 2001), J.A. 1038-50.

Interlaboratory Study states that one of its primary objectives was to "determin[e] CVs for LC_{50}s and IC_{25}s and ranges for NOEC values." Interlaboratory Study at 4, J.A. 1147. The study design did not call for EPA to convert LC_{50} and IC_{25} data to toxic units ("TUs") before calculating a CV. For NOEC values, EPA evaluated precision not by calculating a CV, but, rather, "by evaluating the range and distribution of NOEC values and the percentage of values falling within and beyond one concentration from the median." *Id.* at 78, J.A. 1221. EPA then compared the CVs from Interlaboratory Study with the CVs of approximately 200 other CWA test methods for chemical analytes in 40 C.F.R. Part 136. EPA found that the range of WET test method variability was comparable to, or better than, the variability of the chemical-specific methods.[7]

II. PETITIONERS' REPLY BRIEF TABLE RELIES ON AN *EXCLUDED* DATASET

The Rehearing Petition is based exclusively on arguments raised in Petitioners' reply brief relating to their own, independent calculation of a CV for a single dataset in the Interlaboratory Study: the seventh column of Table 9.8, Results for *Ceriodaphnia* chronic test performed on reference toxicant samples, Interlaboratory Study at 81-82, J.A. 1224-25. These data are NOEC data. Petitioners used their own methodology to calculate a CV from these data and presented their result (179.3%) in the third row of Table 2 in their reply brief. Reply Br. at 25 ("Reply Brief Table"). Except for a general reference to Appendix E of EPA's 1991 Technical Support Document ("TSD"), Reply Br. at 25, n.28, Petitioners did not explain in their reply brief how this figure was derived.

As noted, EPA did not calculate CVs for *any* NOEC values in the Interlaboratory Study. Additionally, EPA stated at numerous points in the Interlaboratory Study that it did not calculate a CV for the *Ceriodaphnia* reference toxicant dataset because "this sample type failed to produce toxicity that could be definitively measured within the test concentration range," due to problems with preparation of the samples. Interlaboratory Study at 77, J.A. 1220; *see also id.* at 31, 66, J.A.

---

[7] Memorandum from M. Kelly, EPA Engineering and Analysis Division (October 16, 2002) ("Comparison Memo"), J.A. 1814-23; *see also* EPA Br. at 50-51; 391 F.3d at 1271, n.4 (citing Comparison Memo).

1174, 1209; EPA Br. at 84-85; RTC at 255-56, J.A. 1951-52. During the rulemaking, neither Petitioners nor peer reviewers objected in their comments to EPA's decision not to compute a CV for this dataset. Nor did Petitioners raise this issue in their opening brief.

III. THE REPLY BRIEF TABLE RELIES ON "TOXIC UNITS," NOT WET TEST RESULTS, CONTRARY TO THE DESIGN OF THE INTERLABORATORY STUDY

Although Petitioners' reply brief did not explain how they derived a CV of 179.3%, it is now clear that they converted the WET test results to toxic units ("TUs") before calculating that CV. Reh'g Pet. at 5; Tr. at 52 (same). Toxic units are not WET test results; rather, they are the *inverse* of the WET test results multiplied by 100. *See* EPA Br. at 23 (explaining how TUs are calculated); 391 F.3d at 1270-71. Conversion of WET test results to TUs simplifies the expression of concentration-based toxicity measurements by providing values that correlate to the level of toxicity. TSD at 6, J.A. 384. For example, it is easier to comprehend that a sample with 8 TUs is more toxic than a sample with 1 TU than it is to comprehend that a sample with an IC_{25} of 12.5% is more toxic than a sample with an IC_{25} of 100%. Most permitting authorities express permit limits and require the reporting of toxicity testing in the form of TUs. *See* EPA Br. at 23; TSD at 6, J.A. 384.

The Interlaboratory Study – the design and results of which were subject to public comment and peer review – used only WET test results to calculate CVs. Interlaboratory Study at 4, 66, 149, J.A. 1147, 1209, 1292. EPA did not convert Interlaboratory Study test results to TUs. *Id.* EPA received no objection to its method of calculating CVs from commenters, including Petitioners, or from peer reviewers. *See supra* at 3. Additionally, the WET test methods themselves, which are at issue in this litigation, do not even address or mention the concept of TUs, let alone to assess test method variability. *See* Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms (4th ed. Oct. 2002) (the "Methods Manual"), J.A. 1554-1813. EPA has never used TUs to calculate CVs to evaluate *test method* variability (*i.e.*, method precision), and no EPA document discusses the use of TUs in this context. To the extent EPA guidance and documents in the administrative record discuss using TUs to calculate CVs, those

documents are limited to specific implementation situations. For instance, Appendix E of the 1991 TSD, which Petitioners claim to have applied to their CV calculation, Reh'g Pet. at 8-11, is entitled "Lognormal Distribution and Permit Limit Derivations" and states that: "This appendix provides supporting information for the statistical methodology used in *permit limit calculations*." TSD at E-1 (emphasis added).[8/] It does not purport to apply to assessing test method variability. The section of EPA's Variability Guidance that Petitioners cite also is inapplicable to the assessment of test method variability.[9/] That document discusses how permitting authorities may calculate CVs using TU values "to make a reasonable potential determination[10/] and to calculate permit limits," Variability Guidance at 6-1, J.A. 859, a process that evaluates a different type of data for a wholly different purpose. *See* Declaration of John Fox ("Fox Decl.") at ¶¶ 21-22 (attached as Exhibit 2).

IV. EPA GUIDANCE "STRONGLY RECOMMENDS" THAT ONLY POINT ESTIMATE – NOT HYPOTHESIS TEST – VALUES BE USED TO GENERATE CVs

Even for the limited permitting purposes for which EPA recommends using TUs to calculate a CV, EPA's guidance is explicit that TUs be derived only from point estimate (*e.g.*, IC_{25}) values, and not hypothesis testing (*e.g.*, NOEC) values. However, Petitioners calculated their own CV from NOEC data. The EPA guidance upon which they rely, Pet. Reh'g at 5, when read in context and in its entirety, advises against just that:

> Section 6.1 discusses use of the CV of sample measurements of toxicity to make a reasonable potential determination or to calculate permit limits. Two points must be understood: (1) this CV is to be calculated using toxic unit (TU) values (USEPA 1991a) (see Section 6.2) and (2) *EPA strongly recommends* that point estimates *(not NOEC or LOEC values)* be used to calculate the TU values (USEPA 1994a, 1994b).

Variability Guidance at 6-1, J.A. 859 (emphasis added). EPA explained that the use of point estimate data was strongly recommended because "[p]oint estimates have less analytical variability

---

[8/] Appendix E, attached as Exhibit 1, is not included in the Joint Appendix, nor did Petitioners provide it.

[9/] Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System (June 2000) ("Variability Guidance"), J.A. 818-873.

[10/] This step in the permitting process is discussed in EPA's merits brief. EPA Br. at 10, 68-69.

than NOECs . . .[and] [p]oint estimates make the best use of the WET test data for purposes of estimating the CV. . . ." *Id.* at 6-4, J.A. 862.

In the 1991 TSD, EPA similarly warned against computing a CV based on NOEC values:

> Since the IC is a point estimate, a CV can be calculated. A CV cannot be calculated if hypothesis testing is used because results are only available for the effluent concentrations used. For this reason, estimates of test precision cannot be calculated for NOECs derived by hypothesis testing.

TSD at 5-6, J.A. 383-84; *see also id.* at 12, Box 1-5, J.A. 390 (explaining why CV only calculated for point estimate data and NOEC evaluated, instead, by mean and range). Consistent with this guidance, EPA's Interlaboratory Study computed CVs from point estimates (LC_{50} and IC_{25}) *only*. Interlaboratory Study at 4, 66, 149, J.A. 1147, 1209, 1292. NOEC data were evaluated differently by analyzing the range and distribution of NOEC values. *Id.* at 78, J.A. 1221.

ARGUMENT

I. PETITIONERS' CV METHODOLOGY IS FATALLY FLAWED

The Court appropriately rejected a fatally flawed argument in Petitioners' reply brief. Petitioners now seek to rehabilitate that argument and introduce additional evidence they did not raise before. By applying their own methodology to generate a CV from data in the Interlaboratory Study, Petitioners argued that the *Ceriodaphnia* reference toxicant dataset exhibited far more variability than the other CVs reported in the Interlaboratory Study. Petitioners' CV of 179.3% was not calculated by EPA, was not reported in the Interlaboratory Study, appears *nowhere* in the administrative record, and relies upon a methodology contrary to all relevant guidance.[11] Thus, their very attempt to generate a CV in these circumstances was inappropriate and was rightly rejected.

---

[11] Petitioners' argument that it was the Court that introduced and relied upon "extra-record" data is ironic. Reh'g Pet. at 11-13. Rather, it was *Petitioners* who attempted to introduce extra-record evidence in their reply brief, based on their own, independent calculation. The Court merely identified obvious errors in those calculations. Thus, *Safe Food & Fertilizer v. EPA*, 365 F.3d 46 (D.C. Cir. 2004), wherein the Court cited a risk assessment that was not part of the administrative record to support the challenged EPA rule, is distinguishable. Here, the Court performed a simple calculation, based on data already in the administrative record, to show that Petitioners' extra-record calculation was unsupportable.

A. Petitioners Relied on an *Excluded* Dataset.

In the Interlaboratory Study, EPA declined to calculate a CV from the *Ceriodaphnia* reference toxicant data upon which Petitioners rely because those samples failed to produce toxicity that could be definitively measured due to problems with preparation of the samples. Interlaboratory Study at 77, J.A. 1220; *see also id.* at 31, 66, J.A. 1174, 1209. At oral argument, Judge Randolph alluded to EPA's decision not to compute a CV for these data. Tr. at 29 ("EPA right around there made some sort of a comment that it's invalid to do this kind of a calculation"). Moreover, neither comments by Petitioners nor peer reviewers during the rulemaking called into question EPA's decision not to compute a CV for these data. Thus, from the outset, Petitioners relied upon a wholly inappropriate dataset that was properly excluded by EPA.

B. Petitioners' CV Calculation Ignored EPA Guidance on the Use of NOEC Data

Petitioners' reliance on NOEC data in the Reply Brief Table contravenes consistent and straightforward EPA guidance on the calculation of CVs and is inconsistent with the manner in which CVs were calculated in the Interlaboratory Study. In referencing the EPA Variability Guidance on this point, Petitioners misleadingly omitted from their quotation EPA's admonition that, for those limited circumstances when a CV is derived from TU values, "EPA strongly recommends that point estimates (not NOEC or LOEC values) be used to calculate the TU values." Variability Guidance at 6-1, J.A. 859. Nor did Petitioners inform the Court of EPA's basis for this recommendation – "[p]oint estimates have less analytical variability than NOECs." *Id.* at 6-4, J.A. 862.[12/] Petitioners' reliance on the 1991 TSD, Reh'g Pet. at 6, is also misplaced, as EPA therein

---

[12/] Petitioners' reliance on EPA's Draft National Whole Effluent Toxicity Implementation Guidance, EPA 832-B-04-003 (Dec. 28, 2004), is equally unavailing. *See* Pet. for Reh'g at 6, n.7. This document (pertinent excerpts of which are attached as Exhibit 3) is a *draft* document that post-dates the 2002 WET Rule, is not part of the administrative record, and, therefore, need not be addressed by the Court. Regardless, like the other documents discussed herein, in this document EPA *again* admonishes that in calculating a facility-specific CV *for permitting purposes*, "[r]egulatory authorities should calculate the facility-specific CV *using point estimate techniques* to determine the need for and derive a permit limit, even if the permit compliance monitoring test results will be determined using hypothesis test procedures . . . [p]oint estimates make the best use of the WET test data for purposes of estimating the CV. . . ." *Id.* at 23 (emphasis added).

stated in the clearest terms: "A CV cannot be calculated if hypothesis testing [*i.e.*, NOEC] is used because results are only available for the effluent concentrations used." TSD at 5, J.A. 383; *see also id.* at 12, Box 1-5, J.A. 390. EPA's Methods Manual similarly instructs testing laboratories not to calculate CVs based on NOEC data, Methods Manual at 15-16, 39, J.A. 1583-84, 1607, and to use point estimation to report WET test results to facilitate the calculation of CVs. *Id.* at 39, J.A. 1607.

Further evidence that NOEC data are not to be used to calculate CVs is the fact that the Interlaboratory Study – *the* record document on the topic of WET method precision – calculated CVs only for LC_{50} and IC_{25} data and did not calculate CVs for NOEC data. Interlaboratory Study at 4, 66, 149-51, J.A. 1147, 1203, 1292-94.[13] Instead, EPA used a different, more appropriate analytical tool to evaluate NOEC data. *Id.* at 78, J.A. 1221. Thus, the administrative record does not support Petitioners' use of NOEC data to compute their CV, and the Court correctly rejected it.

C. Petitioners Improperly Converted the Data at Issue to Toxic Units.

The Court also correctly held that Petitioners erred by using TUs, as opposed to WET test results, to calculate a CV for the *Ceriodaphnia* reference toxicant data. As discussed *infra*, the Court correctly noted that doing so with respect to this dataset resulted in a predictably inflated CV. Fox Decl. at ¶¶ 8(d), 12, 14 & 15. In fact, combined with their improper use of NOEC data, conversion of this particular dataset to TUs, further inflated the CV. *See id.* at ¶¶ 12, 15-16. No EPA guidance recommends the use of TUs to compute a CV for the purpose of assessing *test method* variability. *Id.* at ¶ 23. Petitioners' claim that EPA recommends the use of TUs for a broad array of purposes, including to assess test method variability, *see* Reh'g Pet. at 5-6, is erroneous, misleading, and mischaracterizes the cited documents. Rather, EPA has discussed using TU values for CV calculations only in narrow and specific circumstances, such as deriving permit limits, *see supra* at 5-6, and not to assess test method variability. The only record documents discussing the

---

[13] At oral argument, Judge Randolph identified the problem with relying on NOEC, as opposed to point estimate (*e.g.*, IC_{25}), data. Tr. at 29-30 (noting predefined limits of NOEC data and concurring with EPA counsel's explanation that NOEC data, unlike IC_{25} data, are not continuous).

methodology for calculating a CV to evaluate method variability pertain to the Interlaboratory Study, and we have already demonstrated that the design and execution of that study computed CVs based solely on WET test results, not TUs. *Supra* at 3-4.

Additionally, there are significant differences between calculating a CV for permitting purposes and for assessing method variability, not the least of which is the fact that CVs calculated for permitting purposes are based on *one facility's discharge* monitoring data collected *over a period of time*, which can be assumed to have a lognormal distribution. Fox Decl. at ¶ 21. By contrast, the Interlaboratory Study data at issue here are results that *multiple laboratories* obtained when they tested the *same reference toxicant sample. Id.* at ¶ 22. Those data could be presumed to have a normal distribution. *Id.* Thus, by converting this particular dataset to TUs, Petitioners erred by relying on guidance that addresses wholly distinct analytical scenarios.

Petitioners also err by arguing – for the first time in the instant petition – that EPA should have converted WET test results to TUs before computing the CVs to determine the "mass" of toxicity in a sample and to facilitate a "mass-based . . .to . . .mass-based" comparison of WET test methods and chemical-specific methods. *See* Reh'g Pet. at 6-7.[14] As the Court recognized, toxicity is a method-defined analyte, 391 F.3d 1269-70, and WET tests measure the *effect* of a sample on test organisms. *Id.* at 1268-69; EPA Br. at 45. Thus, whether expressed as a WET test result or its inverse multiplied by 100 (*i.e.,* TU), the "mass" of toxicity simply is not measured in the way that the mass of a chemical analyte can be determined using EPA chemical-specific test methods. *See* Fox Decl. at ¶ 25. Thus, Petitioners establish no reason why TUs, rather than WET test results, should be used to calculate CVs or to compare the variability of WET test methods to other EPA-

[14] The portion of the TSD that Petitioners rely on for this point is both inapplicable and taken out of context. Reh'g Pet. at 6. The development of regulatory waste load allocations relies on mass balance principles to calculate necessary dilutions. Toxic units are used in that process because, like the expression of permit limits, *see supra* at 5, the related modeling requires values that increase when the percent of effluent in the receiving stream increases. *See* TSD at 85. Ex. 1; Fox Decl. at 25. These considerations are irrelevant to the evaluation of method variability. *See* Fox Decl. at 25.

approved CWA test methods. *See* Fox Decl. at ¶¶ 24-25.

Petitioners' claims about the arithmetic averaging of disparate WET test results, Reh'g Pet. at 6, n.6, reveal a similar lack of understanding about the assessment of WET test method variability. Whole effluent toxicity is the *aggregate* toxic effect of an effluent. It cannot be assumed that mixing two samples with different test results will produce a composite sample with a test result that is the arithmetic average of the two. Fox Decl. at ¶ 26. Moreover, because the concentration-response curve of toxicity is not necessarily a straight line, the toxic concentration that is measured when two different toxic samples (of equal volume) are physically combined will not necessarily lie at the midpoint between the concentration-response curves associated with the two original samples. *Id.* Thus, EPA appropriately used WET test results to assess variability.

II. PETITIONERS' ARGUMENTS ARE WAIVED

In their reply brief and the instant Rehearing Petition, Petitioners proffer a completely new approach to computing CVs to evaluate test method variability. Despite ample opportunity to do so during a lengthy rulemaking and settlement negotiations, *see* EPA Br. at 5-7 (discussing opportunities for public comment), Petitioners have *never* proposed that EPA calculate CVs in the way they now advocate. Nor did Petitioners present this novel methodology to EPA or the Court in their opening merits brief. While Petitioners did assert CV values for the *Ceriodaphnia* reference toxicant data of 0.17, 0.9, and 1.7 in their opening brief, they did not explain the methodology they applied nor did they indicate that their CVs were based on TUs, as opposed to WET test results.[15] Nor did Petitioners cite a single record document in support of these CVs. *See* Pet. Br. at 31, 38

---

[15] Petitioners' claim that they raised in their opening brief the issue of how CVs should be calculated, including the use of TUs, is incorrect. *See* Reh'g Pet. at 10, 11, 12. In their opening brief, Petitioners baldly asserted that "[t]he CV's associated with these data are 0.9 and 0.17." Pet. Br. at 31. Petitioners did not identify the source of the CV calculation, the method for calculating those CVs, or whether TUs were used. Petitioners' additional reference to page 38, n.27 of their brief is similarly unavailing, as their assertion that the data "had CVs from 0.9 to 1.7" is also unadorned by *any* explanation of its basis. Rather than speculate as to how Petitioners derived these figures, EPA simply pointed out in response that the referenced dataset had been excluded for CV purposes because of an error in sample preparation. EPA Br. at 84-85.

n.27. Petitioners did not reveal their CV of 179.3% or their underlying methodology until their reply brief, and then only vaguely. Reply Br. at 25 & n.28. For instance, not until oral argument did Petitioners specify that the CV was calculated from WET test results converted to TUs. Tr. at 52.

Thus, EPA had no opportunity to respond to the inappropriateness of Petitioners' methodology, and the Court appears to have observed as much. Tr. at 31. Any argument not raised in Petitioners' opening brief, or raised only summarily without explanation or reasoning, is waived if it is first raised in the reply brief. *E.g.*, *City of Waukesha v. EPA*, 320 F.3d 228, 251, n.22 (D.C. Cir. 2003).[16] The Court has sharply criticized this type of "sandbagging" and refused to consider arguments raised under closely analogous conditions.[17] Nor may a party raise new and additional matters for the first time in a petition for rehearing.[18] Thus, the Court could have declined even to consider Petitioners' CVs and arguments based on the Reply Brief Table.[19]

III. THE COURT DID NOT ERR

As the foregoing shows, Petitioners' CV calculation was fundamentally flawed and contrary to all applicable EPA guidance. Moreover, the basic statistical analysis the Court undertook to demonstrate the distorting effects of Petitioners' CV calculation was correct, consistent with the methods EPA applied in the Interlaboratory Study, and conforms to relevant EPA guidance. Petitioners assert that the Court erred by criticizing Petitioners' use of TUs to calculate a CV of

---

[16] At oral argument, EPA asserted that Petitioners' arguments are waived. Tr. at 31.

[17] *Board of Regents of the Univ. of Washington v. EPA*, 86 F.3d 1214, 1221 (D.C. Cir. 1996) (raising "obscure" argument in opening brief in "conclusory fashion and without visible support" but then, in reply brief, raising new, specific objection on same issue, constitutes "sandbagging") (citations omitted).

[18] *See, e.g., American Policyholders Ins. Co. v. Nyacol Prods., Inc.*, 989 F.2d 1256, 1264 (1st Cir. 1993); *Utahns for Better Transp. v. U.S. Dept. of Transp.*, 319 F.3d 1207, 1210 (10th Cir. 2003); *Holley v. Seminole County School Dist.*, 763 F.2d 399, 401 (5th Cir.1985) ("thorny issues" cannot be presented for the first time in petition for rehearing); 20A Moore's Federal Practice, § 340.12 (3d ed. 2004).

[19] Petitioners' attempt to raise issues on rehearing in this Court that have not first been formally presented to, evaluated by, and acted upon by EPA also violates the requirements of *Oljato Chapter of the Navajo Tribe v. Train*, 515 F.2d 654, 660, 666-67 (D.C. Cir. 1975).

179.3%, Reh'g Pet. at 1, which, the Court found, produced a "grossly inflated result." 391 F.3d at 1271, n.4. The Court appears to have performed its calculations merely to illustrate the straightforward proposition that when the values in the particular dataset – which is characterized by many "high" values and few "small" values – are mathematically inverted and multiplied by 100, *i.e.*, converted to TUs, a statistical calculation based on the manipulated data will produce a "grossly inflated result."[20] We have already demonstrated that the Court's criticism was on point and is supported by the administrative record. *Supra* at 9-11; *see also* Fox Decl. ¶¶ 8(d), 12, 14-15, 21-23. Moreover, the Court's calculation of the two CVs was accurate and effectively demonstrated the distorting effects of first converting this particular dataset to TUs. *Id.* at ¶¶ 8(d), 9-12, 15-16.

Petitioners also wrongly argue that the Court erred in assuming that the data in the *Ceriodaphnia* reference toxicant dataset at issue were normally distributed. *See* Reh'g Pet. at 1-2, 7-8. The Court neither stated nor gave any basis to conclude that it assumed that the data were normally distributed. 391 F.3d at 1271, n.4. Regardless, the Court's CV calculations (0.43 and 1.47) can be arrived at without making *any* assumptions about the data's distribution. Fox Decl. at ¶¶ 9-12, 17-18. More importantly, it was Petitioners who erred yet again by assuming that the NOEC dataset at issue fit a delta-lognormal distribution and by applying the delta-lognormal procedures of Appendix E of the 1991 TSD. *Id.* at ¶¶ 19-20. Because NOEC data are limited to a finite set of discrete values, they cannot be assumed to fit any distribution, *id.*, and Petitioners offer no evidence to support their contention that the data are delta-lognormally distributed. *Id.* at ¶ 22. Thus, Petitioners have not identified a single error in the Opinion.

---

[20] That the Court also noted that a CV of 0.43 was within the range of CVs reported in the Comparison Memo, 391 F.3d at 1271, n.4, is immaterial and did not generate improper extra-record evidence. *See* Reh'g Pet. at 11. In fact, it is unclear whether this observation factored into the Court's overall determination that the WET test methods are precise. Even if it did, it would have been incidental to the Court's overall determination, which was supported by EPA's affirmative demonstration of precision. *See infra.*

IV. IF THE COURT ERRED, ANY ERROR WAS HARMLESS

Petitioners exaggerate by arguing that the Court's refutation of the CV they calculated "determined the outcome of this matter." Reh'g Pet. at 14. To the contrary, the Court's rejection of Petitioners' methodology was not central to its determination of the precision issue, let alone its resolution of the entire case and its rejection of every one of Petitioners' other challenges. In short, an alleged technical error in an explanatory footnote that merely demonstrates why Petitioners' invented methodology was flawed and misleading cannot justify rehearing where the Court identified several other reasons to support its conclusion that WET tests are precise.

In its analysis of the precision issue, the Court found that EPA's methodology for assessing the variability of the WET test methods was appropriate and consistent with EPA past practice. 391 F.3d at 1270. The Court appears to have relied heavily upon the fact that "[t]he record contains extensive raw data, from the main EPA Interlaboratory Study and other privately commissioned studies, regarding the variability of WET toxicity measurements." *Id.* The Court then relied upon the analysis in the Comparison Memo, *id.* at 1271, n.4, and accepted EPA's argument that "these WET test methods exhibit a degree of precision compatible with numerous chemical-specific tests already in use." *Id.* at 1271. There is no dispute that the CVs in the Interlaboratory Study (as opposed to Petitioners' exaggerated results using their own methodology) are well within the range of CVs reported in the Comparison Memo, which range the Court endorsed. *Id.* at 1271, n.4.

The Court further explained that there will be "some degree of variation" with any measurement system, *id.* at 1272, and that "[n]o matter how narrow the error band, or how precise the test, there always will be some measurements on the high end of the range, and some on the low." *Id.* at 1271. The Court observed that EPA further mitigated the variability inherent in the methods by assuming a "false negative" rate that gives dischargers the benefit of the doubt 20% of the time (*i.e.*, by generating "no toxicity" WET test results when toxicity is, in fact, present). *Id.* at

1272.[21] Finally, the Court held that, in addition to being correct, *id.* at 1271, EPA's conclusions regarding precision were within its technical expertise and were entitled to deference. *Id.* at 1270.

Thus, the Court appears to have relied upon EPA's affirmative demonstration of WET test method precision, based on the totality of the administrative record. While the Court may have found further support in its rejection of Petitioners' inflated CV, those considerations appear to have been subordinate to EPA's affirmative demonstration. Indeed, the very calculation and discussion that Petitioners now challenge appears in a footnote and is set off by the phrase "For example," suggesting the comparative weight given to this finding. *Id.* at 1271, n.4. While Petitioners' arguments based on their own inflated CV may have been "at the heart of petitioners' claims of extreme variability," *id.* at 1271, the Court's rejection of those arguments was *not* at the heart of the Court's decision. In short, the Court had ample alternative grounds to determine that the WET test methods are precise and would have reached the same result even if it erred as Petitioners allege.

CONCLUSION

For the foregoing reasons, the Rehearing Petition should be denied.

Respectfully submitted,

THOMAS L. SANSONETTI
Assistant Attorney General

DAVID S. GUALTIERI
CHRISTINA BECHAK PARASCANDOLA
Environmental Defense Section
Environment & Natural Resources Division
United States Department of Justice
P.O. Box 23986
Washington, DC 20026-3986
(202) 514-4767/

Dated: March 17, 2005

Of Counsel:

STEPHEN J. SWEENEY
Senior Attorney
Office of General Counsel
Water Law Office (2355A)
U.S. Environmental Protection Agency
1200 Pennsylvania Ave., NW
Washington, DC 20460

[21] We note that, even though the *Ceriodaphnia* reference toxicant samples tested in the dataset at issue were, in fact, moderately toxic, 26 of the 36 laboratories (72%) validly measured "no toxicity" in this sample. Interlaboratory Study at 81-82, J.A. 1224-25.

CERTIFICATE OF SERVICE

I hereby certify that I caused a true and correct copy of the foregoing RESPONDENT EPA's OPPOSITION TO WET COALITION's AND WESTCAS'S PETITION FOR PANEL REHEARING to be served this 17th day of March 2005, by first class mail, on the following:

James N. Christman
Penny A. Shamblin
Andrea W. Wortzel
Hunton & Williams
Riverfront Plaza, East Tower
951 E. Byrd Street
Richmond, VA 23219-4074

Frederic P. Andes, Esq.
Erika K. Powers, Esq.
Barnes & Thornburg
One North Wacker Drive
Suite 4400
Chicago, IL 60606-2809

John C. Hall, Esq.
Hall & Associates
1101 15th Street, N.W.
Suite 203
Washington, DC 20005

F. Paul Calamita, Esq.
Christopher D. Pomeroy, Esq.
AquaLaw PLC
801 E. Main Street, Suite 1002
Richmond, VA 23219

Roberta L. Larson, Esq.
Somach, Simmons & Dunn
Hall of Justice Building
813 Sixth Street, Third Floor
Sacramento, CA 95814

David E. Evans, Esq.
Stewart T. Leeth, Esq.
Richard H. Sedgley, Esq.
McGuireWoods LLP
901 East Cary Street
Richmond, VA 23219

Alexandra D. Dunn, Esq.
General Counsel
Association of Metropolitan Sewerage Agencies
1816 Jefferson Place, N.W.
Washington, DC 20036-2505

David S. Gualtieri
Counsel for Respondents

United States
Environmental Protection
Agency

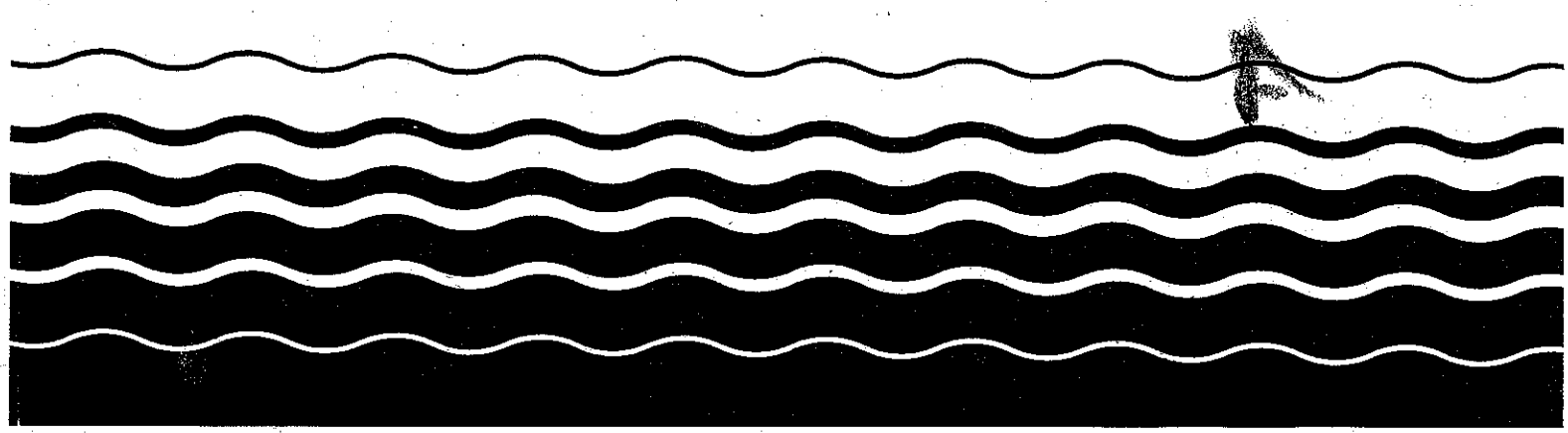
Office Of Water
(EN-336)

EPA/505/2-90-001
PB91-127415
March 1991

EPA

Technical Support Document For Water Quality-based Toxics Control

EXHIBIT 1



4. EXPOSURE AND WASTELOAD ALLOCATION

4.1 INTRODUCTION

At this point in the toxics control process, a water quality problem has been identified. Screening analyses may have been done to assess the extent of toxicity, or a wasteload allocation (WLA) based on an existing total maximum daily load (TMDL) may already have been established. A TMDL is the sum of the individual WLAs for point sources and load allocations (LAs) for nonpoint sources of pollution and natural background sources, tributaries, or adjacent segments. WLAs represent that portion of a TMDL that is established to limit the amount of pollutants from existing and future point sources so that surface water quality is protected at all flow conditions.

The TMDL process uses water quality analyses to predict water quality conditions and pollutant concentrations. Limits on wastewater pollutant loads are set and nonpoint source allocations are established so that predicted receiving water concentrations do not exceed water quality criteria. TMDLs and WLAs/LAs should be established at levels necessary to attain and maintain the applicable narrative and numerical water quality standards, with seasonal variations and a margin of safety that takes into account any lack of knowledge concerning the relationship between point and nonpoint source loadings and water quality. Determination of WLAs/LAs and TMDLs should take into account critical conditions for stream flow, loading, and water quality parameters. Conditions that will protect the receiving water have been determined from State numeric or narrative water quality criteria.

This chapter is divided into sections that explain the steps that precede establishment of a WLA and then the methods and tools (models) that can be used to determine the WLA. Section 4.2 briefly discusses TMDLs and how they relate to waters identified as requiring a water quality-based approach for toxics control. The section also discusses different WLA schemes. Sections 4.3 and 4.4 discuss mixing zones, areas described as allocated impact zones where acute and chronic water quality criteria may be exceeded. Section 4.3 provides background information on mixing zones and discusses EPA's mixing zone policy and how this policy affects the allowable toxic load that can be discharged from a point source. State mixing zone dimensions and the determination of mixing zone boundaries are also discussed.

Section 4.4 discusses mixing zone analyses for situations in which the discharge does not mix completely with the receiving water within a short distance. Included in Section 4.4 are discussions of outfall designs that maximize initial dilution in the mixing zone, critical design periods for mixing zone analyses, and methods to analyze and model near-field and far-field mixing.

Section 4.5 discusses the calculations of the WLA and LA and the types of EPA-recommended mathematical models available to determine WLAs in completely mixed situations for both aquatic life and human health. The WLA models listed in Section 4.5 can be used to predict ambient concentrations and to calculate the effluent quality required to meet the criteria and protect designated and existing uses of the receiving water. The data requirements of each of these models are also described so that the effluent characterization procedures described in Chapter 3 can be designed to support the specific types of WLA modeling selected by the regulator. Section 4.6 discusses human health considerations and how to determine WLAs for human health toxicants.

EPA is currently working on methods to develop sediment criteria. Once developed, point source discharges could be further limited to prevent accumulation of pollutants in the bed sediment; such accumulation impairs beneficial uses. Although the criteria are not yet available for this document, they will be addressed in future documents. In the meantime, some of the models discussed in Section 4.5 are capable of simulating interactions between the water column and sediment and between toxic transport and transformation in the sediment. EPA is encouraging the States to consider the role of sediments in WLA.

4.2 TOTAL MAXIMUM DAILY LOADS AND WASTELOAD ALLOCATIONS

4.2.1 *Total Maximum Daily Loads*

The Federal Clean Water Act (CWA), under Section 303(d), requires the establishment of TMDLs for "water quality limited" stream segments. In such segments, water quality does not meet applicable water quality standards and/or is not expected to meet applicable water quality standards even after the application of the technology-based effluent limitations. A TMDL includes a determination of the amount of a pollutant, or property of a pollutant, from point, nonpoint, and natural background sources, including a margin of safety, that may be discharged to a water quality-limited waterbody. Any loading above this loading capacity risks violating water quality standards. TMDLs can be expressed in terms of chemical mass per unit of time, by toxicity, or by other appropriate measures. Permits should be issued based on TMDLs where available.

The establishment of a TMDL for a particular waterbody is dependent on the location of point sources, available dilution, water quality standards, nonpoint source contributions, background conditions, and instream pollutant reactions and effluent toxicity. All of these factors can affect the allowable mass of the pollutant in the waterbody. Thus, two issues must be determined in conjunction with the establishment of the TMDL: (1) the definition of upstream and downstream boundaries of the waterbody for which the TMDL is being determined, and (2) the definition of critical conditions. For the following discussion, the waterbody boundaries are delineated as the portion of the waterbody be-

organic chemicals and transformation products of those chemicals. The reaction and transfer processes included are hydrolysis, oxidation, photolysis, biodegradation, volatilization, and sorption. Sorption is modeled as a first-order kinetic process in which a desorption rate and an equilibrium partition coefficient for each of the three solid types must be specified. Resuspension and settling of silts and clays (cohesive solids) are defined in terms of shear stress at the sediment-water interface. For sands, the system's capacity to transport sand at a particular flow is calculated and resuspension or settling is defined by the difference between the sand in suspension and the calculated capacity. Sediment exchanges with surficial benthic sediments are modeled as sorption/desorption and deposition/scour. Underlying sediment and pore water are not modeled.

- SARAH2 [60] is a steady-state, near-field model for calculating acceptable concentrations of hazardous organic chemicals discharged to land disposal or wastewater treatment facilities. Acceptable leachate or treated industrial waste discharge constituent concentrations are estimated by a "back calculation" procedure starting from chemical safety criteria in surface water, drinking water, or fish. For steady or batch waste streams, SARAH2 considers the following concentration reductions: dilution and loss during treatment, initial Gaussian mixing at the edge of a stream, lateral and longitudinal diffusion in the mixing zone, sorption, volatilization, hydrolysis, and bioaccumulation in fish. The user must specify appropriate concentrations for protection of the aquatic community and of humans exposed through consumption of fish and water. The benthic community is not presently considered. Treatment loss is handled empirically. SARAH2 contains data sets for three disposal-watershed scenarios that can be easily modified and employed. The model is designed for screening analysis and contains numerous assumptions that should be verified before the model is used in actual cases.

- MINTEQA2 is an equilibrium metals speciation model for dilute aqueous systems [61]. It does not have any transport and transformation processes and must be run with one of the above models. It can be used to calculate the mass distribution at equilibrium among dissolved, absorbed, and solid phases and the species distribution within each phase. MINTEQA2 contains a chemical component data set for major ions commonly found in aqueous systems (e.g., Ca, Fe, and S), trace metals/metalloids of pollution interest (e.g., Cd, Cr, Ni, Pb, and Zn), and organic ligands of significant affinity for metal complexation. The model can be used to calculate the concentrations of adsorbed metals via any of seven different adsorption algorithms.

- FGETS is a toxicokinetic model that simulates the bioaccumulation of nonpolar organic chemicals by fish from both water and food [62]. Both of these routes of exchange are modeled as diffusion processes that depend upon physicochemical properties of the pollutant and morphological/physiological characteristics of the fish. FGETS contains a moderately sized data base of allometric relationships for gill morphology with which it can simulate the direct gill/water exchange of organic chemicals for essentially any fish species, assuming certain default values. FGETS also contains a limited data base of physiological/morphological relationships that are used to set parameters for food exchange. In addition to simulating bioaccumulation of organic toxicants, FGETS can calculate time to death from chemicals whose mode of action is narcosis. This calculation is based on the existence of a single, lethal, internal chemical activity for such chemicals. The concentrations of toxic chemical to which the food chain is exposed may be specified by the user or may be taken directly from the values calculated by the exposure concentration model WASP4. Thus FGETS may be executed as a separate model or as a postprocessor to WASP4.

- FCM2 is a generalized model of the uptake and elimination of toxic chemicals by aquatic organisms [63]. It generates a mass balance calculation in which the rates of uptake and elimination are related to the bioenergetic parameters of the species. A linear food chain or a food web may be specified. Fish tissue concentrations are calculated as a function of time and age for each species included. Exposure to the toxic chemical in food is based on a consumption rate and predator-prey relationships that are specified as a function of age. Exposure to the toxic chemical in water is functionally related to the respiration rate. Steady-state concentrations also may be calculated. The concentrations of the toxic chemical to which the food chain is exposed may be specified by the user or may be taken directly from the values calculated by the exposure concentration model WASP4. Thus FCM2 may be executed as a separate model or as a postprocessor to WASP4. Migratory species, as well as nonmigratory species, may be considered. Separate nonmigratory food chains may be specified, and the migratory species is exposed sequentially to each food chain based on its seasonal movements.

4.5.5 Effluent Toxicity Modeling

To apply the steady-state, continuous simulation, or probabilistic methods to effluent toxicity modeling, the percent effluent measurements should be converted to toxic units (TUs). As discussed in Chapters 1, 2, and 3, it is necessary to convert toxicity to units that can be directly related to mass. When comparing toxicity among chemicals, the relationship between toxicity and concentration is inverse; chemicals that have toxic effects at low concentrations have a greater "toxicity" than chemicals that have toxic effects at higher concentrations. The modeling of toxic effluents is based on mass balance principles; therefore, toxicity needs to be in units that increase when the percent of the effluent of the receiving stream increases. Thus, a TU is the reciprocal of the dilution that produces the test endpoint, i.e., acute toxicity endpoint (ATE) or chronic toxicity endpoint (CTE). An acute toxic unit (TU_a) is the reciprocal of an ATE. A chronic toxic unit (TU_c) is the reciprocal of a CTE. The TMDL must ensure that the CMC and the CCC are met in the receiving water at the desired duration and frequency. The CMC for toxicity is recommended as 0.3 TU_a. This is a value that should prevent lethality unless the duration of exposure exceeds 1 hour.

The CCC for toxicity measured with chronic tests is recommended as the following:

$$CCC = 1.0\ TU_c.$$

The first step in the TMDL process is to calculate the allowable acute effluent toxicity that meets the CMC in the receiving water at the duration and frequency discussed in Appendix D.

The next step in the TMDL process is to calculate the allowable chronic effluent toxicity that meets the CCC in the receiving water at the duration and frequency discussed in Appendix D. To compare the allowable acute toxicity value to the allowable chronic toxicity value, the numbers must be converted to the same units as follows:

$$TU_a = (ACR)(TU_c)$$

where the acute-to-chronic ratio (ACR) is determined from tests on the effluent. It is important that the ACR used for TMDL purposes be based on actual data and not be assumed to be 10 or 20, as in the screening procedure (Chapter 3). The value of this ratio will influence whether the acute or chronic TMDL is more stringent and is used to calculate the permit limit using the methods described in Chapter 5.

At the present time, the fate of effluent toxicity in a receiving water is not fully understood. Even if a decay rate for toxicity can be measured on a given day in a site-specific situation, there is no way as yet to know how this rate is affected by temperature, pH, or other environmental conditions. There is also no way to know how this rate may change when new treatment is installed. Instream measurements of toxicity should be made at least once per season to identify any time-varying trends in site-specific fate processes. These monitored decay rates can then be used in steady-state or continuous simulation fate and transport models to predict receiving water toxicity, assuming that the rates will not change with future treatment.

Without specific information concerning the persistence of toxicity, it is recommended that effluent toxicity be limited to dilution estimates and that toxicity be assumed to be additive and conservative. Toxicity is expected to be additive even when the toxicity of one effluent affects selected biota while the toxicity of a downstream discharge affects different biota. For rivers and run-of-river reservoirs with a detention time of less than 20 days, the following dilution equation should be used, assuming completely mixed conditions:

$$C = \frac{C_s Q_s + C_e Q_e}{Q_e + Q_s}$$

where

- C = downstream concentration (TU_c or TU_a)
- C_s = upstream concentration (TU_c or TU_a)
- Q_s = upstream flow (cfs)
- C_e = effluent concentration (TU_c or TU_a) and
- Q_e = effluent flow (cfs).

For multiple dischargers, this equation must be applied sequentially to find the concentration as a function of distance downstream. The equation can be used for a steady-state analysis if Q_s is set equal to the design flow, Q_e is set equal to the historical plant flow, and C_e is calculated to meet the CMC and CCC. This equation can also be used with the continuous simulation, log-normal probabilistic, or Monte Carlo methods. For these dynamic analyses, a series of C_e, Q_e, C_s, and Q_s values would be used.

If instream toxicity measurements are available and a first-order decay rate for toxicity can be estimated, the following equation should be used:

$$C = C_o e^{-K(x/u)}$$

where

- C = downstream concentration (TU_c or TU_a)
- C_o = concentration after the point source discharge has mixed completely with the river (TU_c or TU_a)
- x = distance downstream of complete mix point
- u = velocity of river
- K = measured decay rate.

Additional statistical approaches are available that might provide better statistical fits to the available data. However, these models are somewhat more limited than the example provided above.

The same equations used for toxicity analyses in **rivers** can also be used in steady-state, continuous simulation, or probabilistic analysis of long, narrow, shallow impoundments with high inflow velocities. **Wider, deeper lakes** require more complicated analyses since prolonged detention times (>20 days) and stratification exert a significant impact on water quality. The prolonged detention times make it essential that receiving water measurements of toxicity be available to estimate decay factors. These measurements should be made at least once per season to identify any time-varying trends in toxicity fate processes. Steady-state or continuous simulation fate and transport models for lakes can then be run with monitored decay rates for toxicity. A simple steady-state analysis can be performed using the following equations [64]:

$$T_w = V/Q$$

$$C = C_{in}/(1+T_w K)$$

where

- T_w = mean hydraulic residence time
- V = lake volume at design conditions
- Q = mean total inflow rate at design conditions
- C = steady-state lake concentration (TU_c or TU_a)
- C_{in} = steady-state inflow concentration (TU_c or TU_a)
- K = first-order decay rate.

If effluent is discharged into a stratified lake and mixes only with the hypolimnion or epilimnion, the volume of the layer should be used only to calculate mean hydraulic residence time (T_w). The mean total inflow rate (Q) and the inflow concentration (C_{in}) should be calculated as the sum of all sources to the lake, including point source, nonpoint source, and tributary inputs.

Dilution calculations for effluent toxicity discharges to an estuary are complicated by the oscillatory motion of the tides and possible stratification of the estuary. The prolonged detention times make it essential that field measurements of toxicity be available to estimate decay factors. These measurements should be made at least once per season to identify any time-varying trends in toxicity rate processes. Steady-state or continuous simulation fate and transport models for estuaries can then be run with monitored decay rates for toxicity. A simple steady-state analysis can be performed using the following equations for each nonconservative pollutant entering from the river at the head of an estuary [64]:

$$C_i = C_{i-1} \frac{(f_i)}{(f_{i-1})} B_i$$

where

$$B_i = \frac{r_i}{1 - (1-r_i)e^{-kt}}$$

r_i = exchange ratio for segment i as defined by modified tidal prism method
t = flushing time
f_i = fraction of freshwater in segment i
C_i = nonconservative pollutant concentration in segment i (TU_a or TU_c)
k = decay rate of pollutant.

The following equations should be used for each nonconservative pollutant entering along the side of an estuary:

For segments downstream of outfall:

$$C_i = C_o \prod_{i=1}^{n} \frac{f_i}{f_o} \left[\frac{r_i}{1 - (1-r_i)e^{-kt}}\right]$$

For segments upstream of outfall:

$$C_i = C_o \prod_{i=1}^{n} \frac{S_i}{S_o} \left[\frac{r_i}{1 - (1-r_i)e^{-kt}}\right]$$

where

C_i = nonconservative pollutant mean concentration in segment i (TU_c or TU_a)
C_o = nonconservative pollutant mean concentration in segment of discharge
r_i = exchange ratio for segment i as defined by the modified tidal prism method
n = number of segment away from outfall
f_i = fraction of freshwater in segment i
f_o = fraction of freshwater in segment with discharge
S_i = salinity in segment i
S_o = salinity in segment of discharge
k = decay rate
t = flushing time.

The details of how to calculate exchange ratios and flushing times for estuaries are included in Part 2 of EPA's water quality assessment manual [64]. This manual also describes how to perform these calculations for stratified estuaries using a two-dimensional box model analysis.

4.6 HUMAN HEALTH

4.6.1 Human Health Considerations

Human exposure to pollutants should be evaluated as completely as available information will allow. Exposure information is used in calculating the human health reference ambient concentration (RAC) from the formulas in Chapter 2, Water Quality Standards. This information should be used to estimate exposures due to fish consumption and drinking water ingestion, background concentrations, and other exposure routes, such as recreational, occupational, drinking water, dietary (other than fish), and inhalation. Factors in the formulas for which information is not available can be omitted from the calculation. If States choose, bioaccumulation factors also can be modified.

4.6.2 Determining the TMDL Based on Human Health Toxicants

TMDLs are typically necessary only where mixing is allowed. Mixing zones are used at the discretion of the States. If a State does not allow a mixing zone or the assumption of complete mixing, then the RAC is applied at the end of pipe and no TMDL determination is typically necessary.

With persistent or bioconcentratable pollutants, special mixing zone considerations apply. Bioconcentratable pollutant criteria exceedances within the mixing zone can potentially result in tissue contamination of organisms directly or indirectly through contamination of bed sediments with subsequent incorporation into the food chain. For discharge situations with incomplete mixing (e.g., large rivers, lakes, estuaries, oceans), States need to carefully consider whether mixing zones for persistent or bioconcentratable pollutants are appropriate. Where a mixing zone is allowed, one TMDL should be calculated to achieve the RAC or criterion selected above [65]. Because most human health criteria are chronic only, a TMDL to protect against acute effects will usually not be needed, although EPA's Office of Drinking Water does have acute criteria for some pollutants.

For the purpose of the following discussion, use of simple, steady-state dilution models is assumed. However, these models may be inappropriate for certain situations where sediments serve as a sink for bioconcentratable pollutants and where additional factors need to be considered. Dynamic models, where available, are useful tools for accounting for an array of variables that may have an impact on the fate of bioconcentratable pollutants in the food chain. These models may be used by States for surface waters in appropriate instances.

In simple situations, the TMDL is determined from the RAC and the design flow of the receiving water. In more complicated situations, e.g., where mixing is not rapid or where lakes or estuaries are involved, a spatial averaging scale must be chosen. Selection of the spatial scale must be consistent with reasonable assumptions about the behavior of aquatic organisms and the target human population.

In some cases, it may be necessary to apply the chronic human health criterion within a mixing zone if it is reasonable to assume that the bioconcentrating aquatic organisms have little mobility, thus spending most of their time within the mixing zone; and the target human population consistently consumes fish from the mixing zone (over a 70-year lifetime, for carcinogenic risks).

The procedure for developing TMDLs/WLAs generally requires determining values for the following parameters, based upon water quality considerations: (1) the duration of the averaging period applicable to the WLA; (2) design considerations, e.g., flow; (3) the discharge (WLA) concentration that will result in meeting the ambient water quality criterion during the design condition; and (4) the allowable probability (or frequency) of the discharge's exceeding the WLA, averaged over the appropriate

Appendix E

Lognormal Distribution and Permit Limit Derivations

LOGNORMAL DISTRIBUTION AND PERMIT LIMIT DERIVATIONS

Introduction

This appendix provides supporting information for the statistical methodology used in permit limit calculations. The methodology described in this appendix applies to many types of data including data that are used to develop both technology-based and water quality-based permit limits. The appendix is divided into two sections. The first section gives an overview of permit limits: the derivation of water quality-based limits and the consistency among different permit limits. The second section describes the statistical methodology for the normal distribution, the lognormal distribution, the delta-lognormal distribution, methods of checking distributional assumptions, and correlation. This section also provides guidance on the application of each distribution to permit limits. Tables E-1, E-2, and E-3 at the end of the appendix summarize the permit limit calculations. This appendix describes the statistical methodology for three distributions that are often used in determining permit limits. Other distributions can be used, and this topic is discussed in the subsection, Other Distributions.

Section 1: Overview of Permit Limits

Two types of permit limits are contained in the effluent guidelines regulations: daily maximum limits and monthly average limits. The daily maximum permit limit is the maximum allowable value for any daily sample. The daily maximum limits are usually based on the 99th percentile of the distribution of daily measurements. The monthly average permit limit is the maximum allowable value for the average of all daily samples obtained during 1 month. Monthly average limits are in most cases based on the 95th percentile of the distribution of averages of daily values.

The following two subsections discuss the derivation of water quality-based limits and the consistency among different permit limits.

Derivation of Water Quality-based Limits

Water quality-based limits are derived from the required treatment system performance necessary to comply with the wasteload allocation (WLA). Technology-based effluent limits are derived from treatment system performance. The mathematical expressions for water quality-based limits are the same as those for technology-based effluent limits; the major difference is that the means and standard deviations in those expressions are derived from the WLA. This topic is discussed in Chapter 5.

Consistency Among Different Permit Limits

The current Technical Support Document for Water Quality-based Toxics Control (TSD) procedures provide consistency among different permit limits. The stringency of permit limits is independent of monitoring frequency and is determined entirely by the WLA and permit limit derivation procedures. The daily maximum limit is constant regardless of monitoring frequency. The numerical value of the monthly average limit decreases as monitoring frequency increases only because averages become less variable as the number of values included in the average increases. For example, an average based on 10 samples is less variable than an average based on 4 samples. This phenomenon makes monthly average permit limits based on 10 samples **appear** to be more stringent than the monthly limit based on 4 samples. A permittee performing according to the WLA specifications will in fact be equally capable of meeting either of these monthly average limits when taking the corresponding number of samples. The stringency of the TSD procedures, accordingly, is **constant** across monitoring frequencies.

Section 2: Statistical Methodology

The statistical procedures that are used in permit limit development involve fitting distributions to effluent data. The estimated upper percentiles of the distributions form the basis of the limits. This section describes the statistical methodology applied to permit limits in the following subsections: the normal distribution, the lognormal distribution, the delta-lognormal distribution, methods of checking distributional assumptions, and correlation. Before discussing these topics several definitions are made for notation, assumptions, coefficients of variation, and variability factors.

Notation

In the calculations in this appendix, natural logarithms (i.e., logarithms to the base e), denoted by ln(x), are used. The calculations can be modified to use logarithms to the base 10 by replacing $\log_{10}(x)$ for ln(x) in the formulas.

Assumptions

The distribution fitting methods assume that the daily measurements are independent, uncorrelated observations.

The fundamental assumptions underlying the discussion on calculating limits are:

- Daily pollutant measurements are approximately lognormally distributed for values above the detection limit
- Maximum n-day monthly averages for $n \leq 10$ are approximately lognormally distributed above the detection limit
- Maximum n-day monthly averages for $n > 10$ are normally distributed.

Recommendation of the use of the lognormal distribution for daily pollutant measurements is based on practical rather than theoretical consideration. Usually environmental data sets possess the basic lognormal characteristics of positive values and positive skewness. In addition, the lognormal distribution is flexible enough to model a range of nearly symmetric data. Furthermore, in comparison to other positive valued, positively skewed distributions that could be used to model environmental data, the lognormal is relatively easy to use.

When lognormal data are log transformed, the properties of the normal distribution apply to the transformed data. The section on statistical methodology describes the properties of the normal distribution and its relationship to the lognormal distribution. The delta-lognormal distribution is a generalization of the lognormal distribution and may be used to model data that are a mixture of non-detect measurements with measurements that are lognormally distributed. In delta-lognormal procedures, nondetect values are weighted in proportion to their occurrence in the data.

In determining permit limits based on averages (e.g., monthly average permit limits), a distribution should be used that approximates the distribution of an average of pollutant measurements. The lognormal distribution can be used for approximating the distribution of averages for small sample sizes where the individual measurements are approximately lognormally distributed. For larger sample sizes, a powerful statistical result, called the Central Limit Theorem, provides theoretical support for determining limits based on averages of individual measurements. According to the Central Limit Theorem, when the sample size n is large enough, the average of the n sample values will be approximately normally distributed regardless of the distribution of the individual measurements. The section on statistical methodology provides procedures and guidance for calculating averages for both small and large samples sizes where the individual measurements are lognormally distributed.

shape of the observed data is the key factor in evaluating a distributional model. For environmental data lognormal distribution is usually appropriate. The critical question in a given situation is how well a ticular distribution models the shape of observed data. Although the lognormal does not provide an exact n all cases, it usually provides an appropriate and functional fit to observed environmental data. Graphical plays and goodness-of-fit tests, as described in the subsection, Other Distributions, may be used as a guide in ying assumptions and selecting a distribution.

efficients of Variation

The coefficient of variation (denoted by "CV") is the ratio of the standard deviation to the mean. Thus, the CV a dimensionless measure of the relative variability of a distribution. Estimates of the CV can be used when the actual CV cannot be calculated or if the available data sets for calculating the CV are small. In such cases, ifferent values for the CV should be used in the permit calculations to assess the effect of the CV on the final ermit limit. Typical values of the CV for effluent data usually range from 0.2 to 1.2. The CV is a measure of the relative variation in observed data. In many cases, changes in the CV will have little impact on the final ermit limit. In assessing the sensitivity of the permit limit to the CV, the calculations may include CV = 0.6 a conservative estimate (assumes relatively high variability). If the final permit values vary greatly with ifferent CV values either of two approaches may be used. The first approach is to use a conservative estimate of the CV that assumes relatively high variability (e.g., CV = 0.6) in the final permit limit. The second approach is to collect additional data to obtain a more definitive value for the CV.

Variability Factors

An important component of the process used by the Environmental Protection Agency (EPA) for developing technology-based limits are variability factors. The variability factor is the ratio of a large concentration level of a pollutant to the average level determined from that particular plant. The ratio expresses the relationship between the average treatment performance level and large values that would be expected to occur only on rare occasions in a well-designed and operated treatment system. Such factors are useful in situations where little data are available to characterize the long-term performance of a plant.

In cases where only a small number of observations are available from a plant, EPA has been reluctant to estimate a variability factor. In the Organic Chemicals, Plastics, and Synthetic Fibers (OCPSF) rulemaking [1], a minimum of seven daily observations from a plant, with at least three of the seven above the detection limit, was established for calculation of a plant level priority pollutant variability factor. However, EPA has not established a minimum number of observations required for calculating variability factors for all pollutants in all industries.

The calculations for variability factors for the daily maximum and the monthly average are included in the discussion of the different distributions below.

Normal Distribution

The normal distribution plays a central role in the methods described in this appendix. In most cases, the normal distribution is not an appropriate model for individual pollutant measurements; however, the normal distribution is related to the lognormal distribution that is used to establish many permit limits. In most cases, the simple logarithmic transformation of effluent and water quality data results in data distributions that are normally distributed. Such data are referred to as being lognormally distributed. When lognormal data are log transformed, the properties of the normal distribution apply to the transformed data. Since the normal and lognormal distributions are related in a straightforward manner, the methods of analysis for normal and lognormal data also are easily related. The normal distribution is described below and is followed by a discussion of the lognormal distribution and its relationship to the normal distribution.

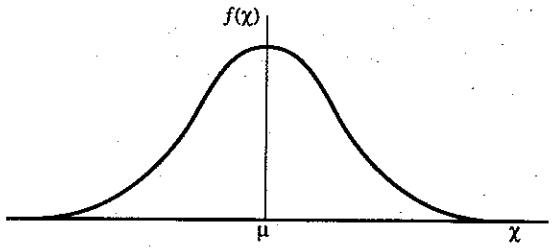


Figure E-1. Normal Probability Distribution

The normal probability distribution is encountered in a number of applications. The bell-shaped curve of the normal distribution is shown above in Figure E-1. Excellent introductions and reviews of the normal distribution are found in numerous statistical, engineering, and scientific texts, as for example in Reference 2. Only a brief review is given here.

A sample of independent observations, denoted by $x_1, x_2,...,x_k$, from a normally distributed population can be used to estimate the mean, μ, and variance, σ^2, according to the formulas below:

$$\begin{aligned}
\hat{\mu} &= \text{estimated mean} \\
&= \Sigma[x_i] / k,\ 1 \le i \le k \\
\hat{\sigma}^2 &= \text{estimated variance} \\
&= \Sigma[(x_i - \hat{\mu})^2] / (k - 1),\ 1 \le i \le k \\
\hat{\sigma} &= \text{estimated standard deviation} \\
&= (\hat{\sigma}^2)^{1/2} \\
\hat{cv} &= \text{estimated coefficient of variation} \\
&= \hat{\sigma} / \hat{\mu}
\end{aligned}$$

The characteristics of the normal distribution are the range is defined for positive and negative values, and the frequency curve is bell-shaped and symmetric about the mean. In most cases, the normal distribution is not an appropriate model for the distribution of individual pollutant measurements. Environmental data rarely are symmetric, which is a fundamental property of the normal distribution. In addition, the normal distribution is defined over a range that includes negative values while pollutant measurements are restricted to nonnegative values. Thus, fitting a data set to a normal distribution allows for the possibility, however small, of observing negative values. The lognormal distribution, or any positive valued distribution, is not defined for negative values and thus avoids assigning any probability to negative values.

Daily Maximum Permit Limits Based on the Normal Distribution

For data sets which have the characteristics of the normal distribution, the daily maximum permit limits can be calculated. The upper percentile daily maximum permit limits for the normal distribution are calculated using the quantity z_p, the standardized Z-score for the pth percentile of the standardized normal distribution (i.e., normal distribution with mean = 0, and variance = 1). For example, the Z-score for the 95th percentile is 1.645. Z-scores are listed in tables for the normal distribution (in most statistical textbooks and references). The pth percentile daily maximum limit is estimated by:

$$\begin{aligned}
\hat{X}_{.p} &= \text{pth percentile daily maximum limit} \\
&= \hat{\mu} + z_p \hat{\sigma}.
\end{aligned}$$

For example:

$\hat{X}_{.95}$ = 95th percentile daily maximum limit
= $\hat{\mu} + 1.645\ \hat{\sigma}$
$\hat{X}_{.99}$ = 99th percentile daily maximum limit
= $\hat{\mu} + 2.326\ \hat{\sigma}$

Note:

$Z_{95} = 1.645$
$Z_{99} = 2.326$

The daily variability factors (denoted by VF_1) are estimated by:

Daily maximum 95th percentile $VF_1 = \hat{X}_{.95} / \hat{\mu}$

Daily maximum 99th percentile $VF_1 = \hat{X}_{.99} / \hat{\mu}$

Monthly Average Permit Limits Based on the Normal Distribution

The normal distribution can be used to model the averages of the individual measurements for a wide range of circumstances. Although the normal distribution usually is not an appropriate model for individual pollutant measurements, the averages of those individual measurements can often be modeled by the normal distribution. This subsection explains the theory behind using the normal distribution for averages and provides the general formulas.

A powerful statistical result, called the Central Limit Theorem, provides theoretical support for determining limits based on averages of individual measurements. According to the Central Limit Theorem, when the sample size n is large enough, the average of the n sample values will be approximately normally distributed regardless of the distribution of the individual measurements. In determining permit limits, the calculations incorporate the number of samples that will be required for monitoring purposes during the specified time period (usually a month). For the purposes of permit writing, monitoring sample sizes greater than 10 are recommended to be sufficiently "large enough" to assume the sample average is approximately normally distributed. The above formulas can be modified for finding the estimated mean and variance for the average from a sample of size n (e.g., for 14-day monthly average, n = 14 samples during the month for monitoring purposes). The parameters μ_n and σ_n^2 denote the mean and variance, respectively, of the distribution of the average of n values. The estimates of the n-day average and the variance of the n-day average are denoted by $\hat{\mu}_n$ and $\hat{\sigma}_n^2$, respectively.

$\hat{\mu}$ = estimated mean of distribution of X
$\hat{\sigma}^2$ = estimated variance of distribution of X
$\hat{\mu}_n$ = mean of distribution of the n-day monthly average
= $\hat{\mu}$
$\hat{\sigma}_n^2$ = variance of distribution of the n-day monthly average
= $\hat{\sigma}^2 / n$
$\hat{\sigma}_n$ = standard deviation
= $(\hat{\sigma}_n^2)^{1/2}$
$\hat{cv}_n$ = coefficient of variation
= $\hat{\sigma}_n / \hat{\mu}_n$.

The upper percentile limits are:

$$\hat{X}_{.p} = \text{pth percentile n-day monthly average limit}$$
$$= \hat{\mu}_n + z_p \hat{\sigma}_n$$

where z_p is the pth percentage point of the standard normal distribution.

For example:

$$\hat{X}_{.95} = \text{95th percentile n-day monthly average limit}$$
$$= \hat{\mu}_n + 1.645\ \hat{\sigma}_n$$
$$\hat{X}_{.99} = \text{99th percentile n-day monthly average limit}$$
$$= \hat{\mu}_n + 2.326\ \hat{\sigma}_n$$

Note:

$$Z_{95} = 1.645$$
$$Z_{99} = 2.326.$$

The monthly average variability factors (denoted by VF_n) are estimated by:

Monthly average 95th percentile $VF_n = \hat{X}_{.95} / \hat{\mu}$

Monthly average 99th percentile $VF_n = \hat{X}_{.99} / \hat{\mu}$

The above discussion of the normal distribution can be modified for data from the lognormal distribution. The next subsection explains the modifications.

Lognormal Distribution

Experience has shown that daily pollutant discharges are generally lognormally distributed. The distributional fit of the data varies somewhat from application to application, but not enough to alter the conclusion that effluent pollutant discharges are generally lognormally distributed. Ambient water quality data also are often lognormally distributed. Figure E-2 displays the positively skewed shape of the lognormal distribution.

The distribution fitting methods assume that the daily measurements are independent, uncorrelated observations. Although, in general, this assumption is not satisfied exactly, the lognormal distribution has been used in the effluent guidelines program primarily because it consistently provides a reasonably good fit to observed effluent data distributions. Figure E-3 shows the lognormal distribution applied to data used in the development of the OCPSF effluent guidelines regulation [1].

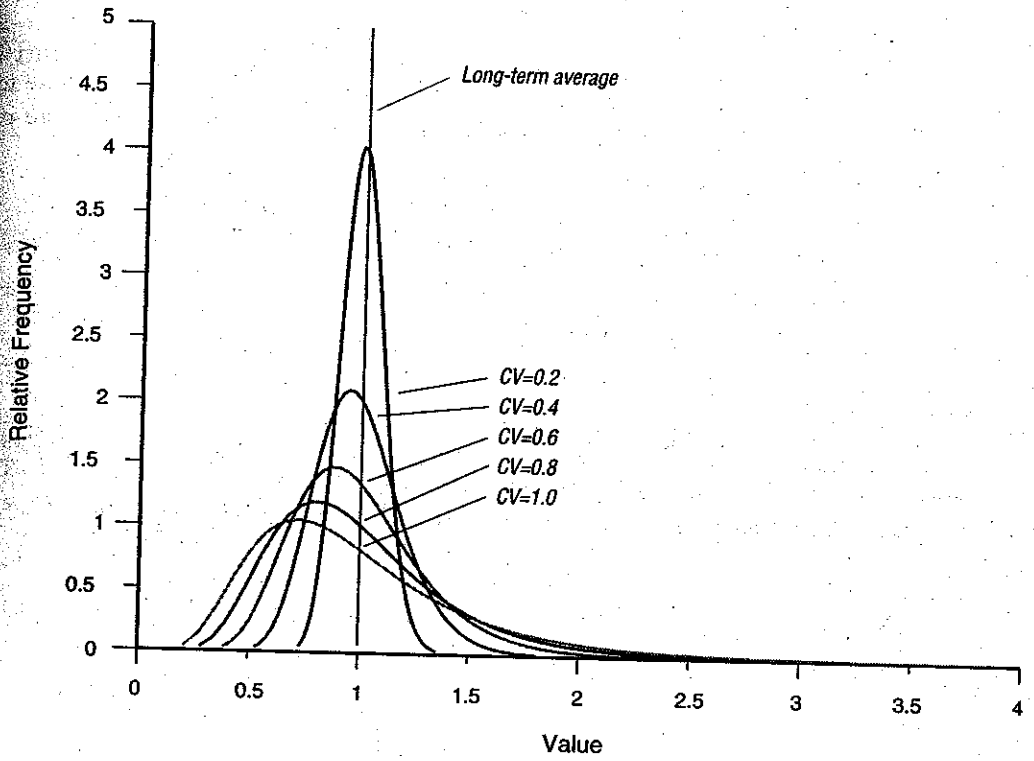


Figure E-2. Examples of Lognormal Densities

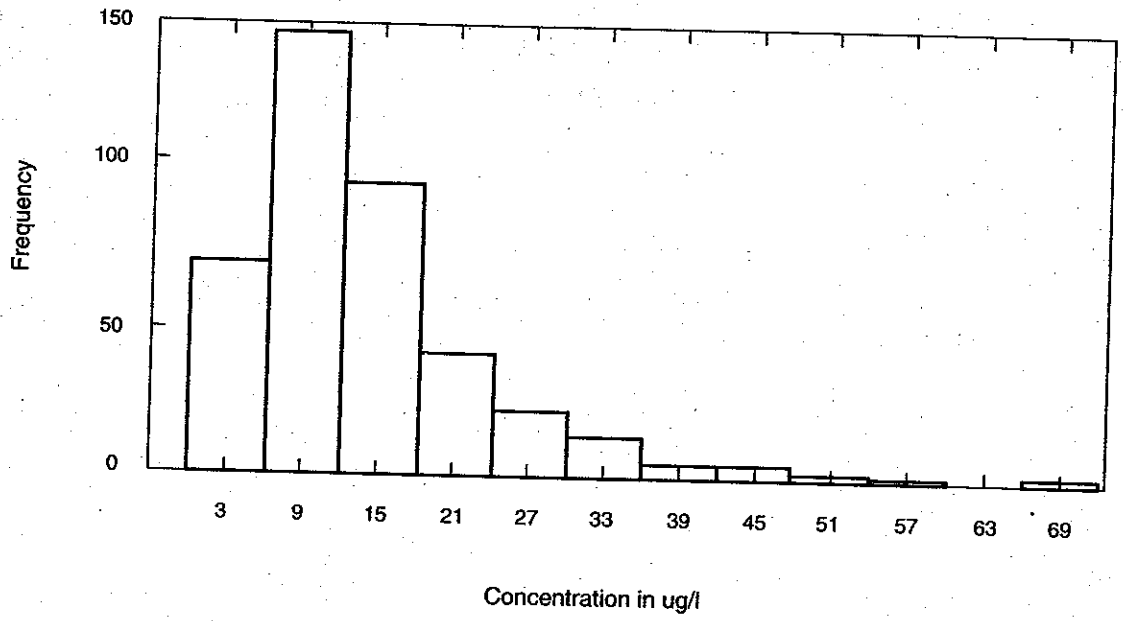


Figure E-3. BOD Frequency Distribution - Plant C

The logarithmic transformation of the random variable X, $Y = \ln(X)$ results in a random variable Y that is normally distributed. Therefore, the analysis procedures for analyzing lognormal data are similar to those for the normal distribution. The mean and variance from the normal distribution of the random variable Y are σ_y, and $\hat{\sigma}_y^2$ respectively. These parameters can be estimated by:

$$\hat{\mu}_y = \Sigma(y_i) / k$$

and

$$\hat{\sigma}_y^2 = \Sigma[(y_i - \hat{\mu})^2] / (k - 1), \text{ respectively}$$

where

$$y_i = \ln(x_i) \text{ for } i=1,2,...k.$$

When data are lognormally distributed, these values from the normal distribution can then be used to calculate the mean, variance, and coefficient of variation for the random variable X that is lognormally distributed. The mean, variance, and coefficient of variation of the random variable X may be estimated by $\hat{E}(X)$, $\hat{V}(X)$, and $\hat{cv}(X)$, respectively.

$$\hat{E}(X) = \text{daily average} = \exp(\hat{\mu}_y + \hat{\sigma}_y^2 / 2)$$

$$\hat{V}(X) = \text{variance} = \exp(2\hat{\mu}_y + \hat{\sigma}_y^2)[\exp(\hat{\sigma}_y^2) - 1]$$

$$\hat{cv}(X) = \text{coefficient of variation} = [\exp(\hat{\sigma}_y^2) - 1]^{1/2}.$$

Daily Maximum Permit Limits Based on the Lognormal Distribution

The upper percentile limits for the random variable X (which is lognormally distributed) are:

$$\hat{X}_{.p} = \text{pth percentile daily maximum limit} = \exp[\hat{\mu}_y + z_p \hat{\sigma}_y^2]$$

where z_p is the pth percentage point of the standard normal distribution.

For example:

$$\hat{X}_{.95} = \text{95th percentile daily maximum limit} = \exp[\hat{\mu}_y + 1.645\,\hat{\sigma}_y]$$

$$\hat{X}_{.99} = \text{99th percentile daily maximum limit} = \exp[\hat{\mu}_y + 2.326\,\hat{\sigma}_y].$$

Note:

$$Z_{95} = 1.645$$
$$Z_{99} = 2.326.$$

The daily maximum variability factors (denoted by VF_1) are estimated by:

Daily maximum 95th percentile $VF_1 = \hat{X}_{.95} / \hat{E}(X)$

Daily maximum 99th percentile $VF_1 = \hat{X}_{.99} / \hat{E}(X)$.

Monthly Average Permit Limits Based on the Lognormal Distribution

This subsection contains the formulas required to approximate the distribution of the average of a small number of lognormally distributed values with another lognormal distribution. Although, the Central Limit Theorem holds that the average of a sample of independent measurements is normally distributed provided that the number of measurements, n, is sufficiently large, the minimum value for n required in specific cases may vary considerably. In cases where the individual values are lognormally distributed, the minimum required for the average to be normally distributed may be quite large. As a consequence, the distribution of the average of a small number of lognormally distributed values may be better approximated by another, related lognormal distribution [3]. For sample sizes larger than 10 when the data are lognormally distributed, it is recommended that the calculations given in Table E-3 should be used. For the purposes of permit writing, monitoring sample sizes of 10 or less are recommended to be "small enough" to assume the sample average is approximately lognormally distributed. The mean, variance, and coefficient of variation of the distribution of the average of n daily values are $\hat{\mu}_n$, $\hat{\sigma}_n^2$, and $\hat{cv}$, estimated by:

$$\hat{\sigma}_n^2 = \text{variance} = \ln\{\hat{V}(X) / [n[\hat{E}(X)]^2] + 1\}$$

$$\hat{\mu}_n = \text{n-day monthly average} = \ln(\hat{E}(X)) - 0.5\,\hat{\sigma}_n^2$$

$$\hat{\sigma}_n = \text{standard deviation} = (\hat{\sigma}_n^2)^{1/2}$$

$$\hat{cv}_n = \text{coefficient of variation} = [\exp(\hat{\sigma}_n^2) - 1]^{1/2}$$

where

$$\hat{E}(X) = \exp(\hat{\mu}_y + \hat{\sigma}_y^2 / 2)$$

$$\hat{V}(X) = \exp(2\,\hat{\mu}_y + \hat{\sigma}_y^2)\,[\exp(\hat{\sigma}_y^2) - 1].$$

The upper percentile limits of the maximum n-day monthly average are:

$$\hat{X}_{.p} = \text{pth percentile n-day monthly average limit} = \exp[\hat{\mu}_n + z_p\,\hat{\sigma}_n]$$

where z_p is the pth percentage point of the standard normal distribution.

For example:

$$\hat{X}_{.95} = \text{95th percentile n-day monthly average limit} = \exp[\hat{\mu}_n + 1.645\,\hat{\sigma}_n]$$

$$\hat{X}_{.99} = \text{99th percentile n-day monthly average limit} = \exp[\hat{\mu}_n + 2.326\,\hat{\sigma}_n]$$

Note:

$$Z_{95} = 1.645$$
$$Z_{99} = 2.326.$$

The variability factors are:

Monthly average 95th percentile $VF_n = \hat{X}_{.95} / \hat{\mu}_n$
Monthly average 99th percentile $VF_n = \hat{X}_{.99} / \hat{\mu}_n$.

Delta-Lognormal Distribution

The delta-lognormal distribution is a generalization of the lognormal distribution. The delta-lognormal distribution may be used when the data contain a mixture of nondetect values and values above the detection limit and can be used to model nondetects in water quality-based limits. In delta-lognormal procedures, nondetect values are weighted in proportion to their occurrence in the data. The values above the detection limit are assumed to be lognormally distributed values. The delta-lognormal distribution can be used in setting daily maximum limits and for setting limits on monthly averages with the recommended number of monitoring samples being 10 or less.

The delta-lognormal distribution models data as the combination of two distributions: the lognormal distribution and a distribution with discrete probability of obtaining observations at or below the detection limit. The lognormal distribution models the observations above the detection limit. The nondetect values are modeled by the distribution with discrete probability of obtaining observations at or below the detection limit. The organic priority pollutant data set shown in Figure E-4 contains a number of observations that were reported as "nondetect." These detection limit measurements are observations that are censored at the detection limit and are represented by the left-most bar in the histogram. Data sets of this form are fairly typical of organic chemicals in wastewater. The delta-lognormal distribution often provides an appropriate and computationally convenient model for analyzing such data.

The estimation procedure for the delta-lognormal distribution assumes that a certain proportion, δ, of values are at the detection limit, which is denoted by D. (The estimation procedure when D = 0 is detailed in Reference 4.) These values set to D are observations that can only by quantified as nondetect (ND) at some minimum level. This minimum level is the detection limit as established by the laboratory performing the chemical analysis.

Let $x_1, x_2, \ldots, x_r, x_{r+1}, \ldots, x_k$ denote a random sample of size k, with r observations recorded as nondetects, and k-r observations greater than the detection limit. The k-r positive observations are assumed to follow a lognormal distribution. The entire data set is assumed to follow the delta-lognormal distribution with censoring point equal to the detection limit D. Let $\hat{\mu}_y$ and $\hat{\sigma}_y^2$ be the sample mean and variance of the distribution of the logarithmic transformation $Y = \ln(X)$ of the observations greater than the detection limit. Let $\hat{\delta}$ be the sample proportion of nondetects. Then the estimates of the mean and variance of the delta-lognormal distribution are estimated by:

$$\hat{E}(X^*) = \text{daily average}$$
$$= \hat{\delta} D + (1 - \hat{\delta}) \exp(\hat{\mu}_y + 0.5 \hat{\sigma}_y^2)$$

$$\hat{V}(X^*) = \text{variance}$$
$$= (1 - \hat{\delta}) \exp(2\hat{\mu}_y + \hat{\sigma}_y^2) [\exp(\hat{\sigma}_y^2) - (1 - \hat{\delta})] + \hat{\delta} (1 - \hat{\delta}) D [D - 2 \exp(\hat{\mu}_y + 0.5 \hat{\sigma}_y^2)]$$

$$\hat{cv}(X^*) = \text{coefficient of variation}$$
$$= [\hat{V}(X^*)]^{1/2} / \hat{E}(X^*)$$

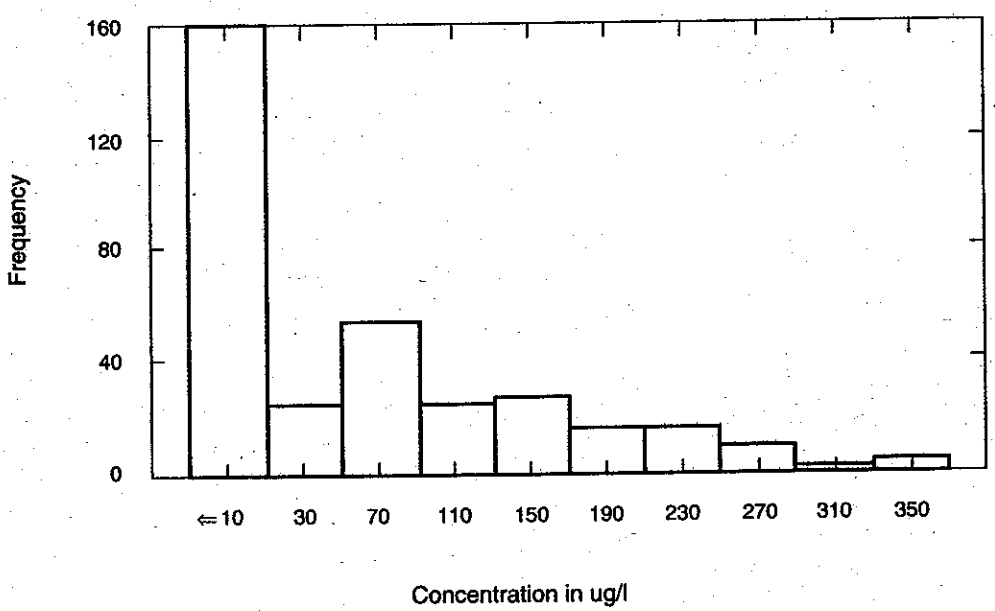


Figure E-4. Organic Priority Pollutant Frequency Distribution—Plant B

where

k = number of samples
D = detection limit
r = number of nondetect values in sample
$k-r$ = number of values greater than the detection limit
$y_i = \ln(x_i)$ $\quad r+1 \le i \le k, \quad r < k$
$\hat{\mu}_y = \Sigma(y_i) / (k - r)$ $\quad r+1 \le i \le k, \quad r < k$
$\hat{\sigma}_y^2 = \Sigma(y_i - \hat{\mu}_y)^2 / (k - r - 1)$ $\quad r+1 \le i \le k, \quad r < k$
$\hat{\delta} = r / k$

Daily Maximum Permit Limits Based on the Delta-Lognormal Distribution

The 95th and 99th upper percentile limits for the random variable X (which is delta-lognormally distributed) are given by the following formulas:

The estimated 95th percentile daily maximum limit is:

$$\hat{X}_{.95^*} = \begin{cases} D & \hat{\delta} \ge 0.95 \\ \max [D, \exp(\hat{\mu}_y + z^* \hat{\sigma}_y^2] & \hat{\delta} < 0.95 \end{cases}$$

where

$$z^* = \Phi^{-1}[(0.99 - \hat{\delta}) / (1 - \hat{\delta})].$$

The estimated 99th percentile daily maximum limit is:

$$\hat{X}_{.99^*} = \begin{cases} D & \hat{\delta} \ge 0.99 \\ \max [D, \exp(\hat{\mu}_y + z^* \hat{\sigma}_y)] & \hat{\delta} < 0.99 \end{cases}$$

where

$$z^* = \Phi^{-1}[(0.99 - \hat{\delta}) / (1 - \hat{\delta})].$$

$\Phi^{-1}[\,]$ is the mathematical notation for Z-scores. For example, when $\hat{\delta} = 0$, then the corresponding value is $\Phi^{-1}[.99] = Z_{99} = 2.326$. Values of $\Phi^{-1}[\,]$ are available from tables of the normal distribution (available in most statistical textbooks and references).

The variability factors (denoted by VF) are estimated by:

Daily maximum 95th percentile VF = $\hat{X}_{.95} / \hat{E}(X)$
Daily maximum 99th percentile VF = $\hat{X}_{.99} / \hat{E}(X)$.

Delta-Lognormal Distribution of Averages

The derivation of the formulas for the averages computationally is difficult and beyond the scope of this appendix. However, the formulas for n-day averages are included in Table E-2. The derivation of 4-day monthly averages using the delta-lognormal distribution is available in Appendix VII-F of the Development Document for the OCPSF regulation [1]. For the purpose of permit writing, it is recommended that data sets of greater than 10 samples be assumed to fit the normal distribution and the averages be calculated using the formulas given in Table E-3.

cking Distributional Assumptions

methods of checking distributional assumptions are goodness-of-fit and probability plots. When checking ibutional assumptions, the sample size must be large enough. Small sample sizes may lead to erroneous usions.

ness-of-Fit Tests

me cases, statistical goodness-of-fit tests may indicate that a particular distribution provides a reasonable o a data set of pollutant measurements. Such cases should be evaluated carefully to verify that the frequency re for the data also show the shape characteristic of the distribution.

obability Plots

e of probability plots is one method of determining whether a normal distribution is appropriate for odeling a population using only a limited set of measurements. The set of measurements should have at least observations [5]. Consider an independent sample of size k, labeled $x_1,x_2,...,x_k$. Let $u_1,u_2,...,u_k$ be the dered sample of x-values in ascending order in which $u_1 \leq u_2 \leq,...,\leq u_k$. Now for each u_i, find z_i from the normal ble (in any statistical reference or textbook) such that $P[Z \leq z_i] = i/(k+1)$ and plot each pair (z_i,u_i) on linear aph paper (or use a computer graphics software package). If the data are from a normal distribution, they will ll approximately along a straight line.

his same method can be adapted to check the assumption of lognormality. Log-probability plots are similar to probability plots used for the normal distribution. To construct a log-probability plot, set $y_i = \ln(x_i)$ for =1,2,...,k and then prepare a probability plot for the y_i, first by ordering the data as described in the previous ection. If the data are from a lognormal distribution, they will fall approximately along a straight line, as illustrated by Figure E-5.

Other Distributions

If the probability plots or the log-probability plots show serious deviation from straight lines, other distributions should be considered. Nonparametric methods, which do not require the assumption that the data follow a particular distributional form, are often useful for this type of data. Further details are available in many statistical references (e.g., Reference 6).

Correlation

Up to this point, we have assumed that all the observed pollutant levels are independent, i.e., uncorrelated with one another. This subsection is not intended to address correlation between observed pollutant levels and plant operating factors that influence and control treatment performance.

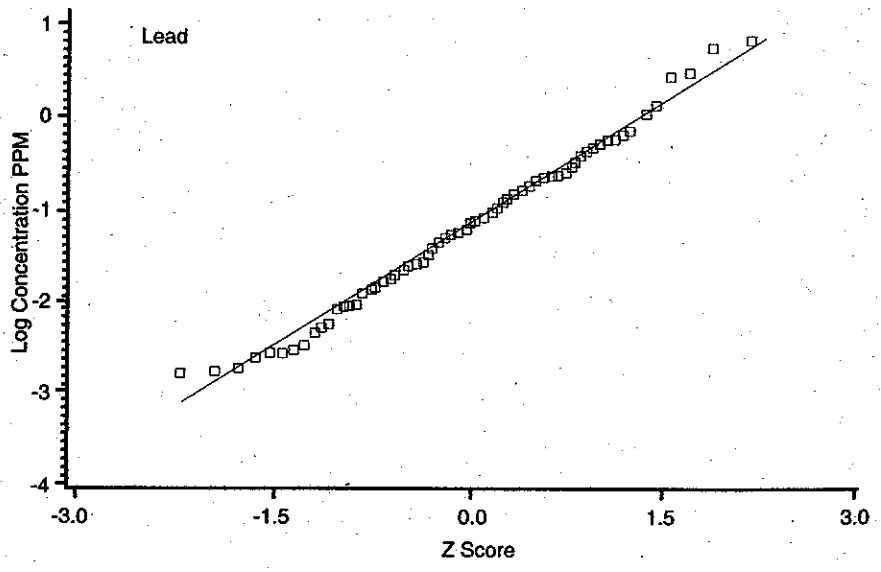


Figure E-5. Example of a Log-Probability Plot with a Normal Distribution

In the case of the monthly average limit derivation, the assumption that observed pollutant levels are independent can be quite important. If the effluent levels are correlated, the actual monthly average limit can be substantially higher than that derived from the analysis based on the independence assumption. However, correlation has essentially no effect on the calculated daily permit limits. This sub-section provides guidance on determining when levels may be correlated, and adjusting the sample size.

A major factor that determines whether effluent levels are highly correlated is the retention time of the wastewater treatment system. If the retention time is large relative to the time between effluent samples, then those samples will tend to be correlated with each other in most cases. In municipal systems, for example, the retention time is frequently a matter of days, and sampling is often conducted on a daily basis. The effluent levels, consequently, may be substantially correlated. However, in many industrial systems, for instance a physical/chemical treatment system for electroplating wastewaters, the treatment system retention time is relatively short 4 to 8 hours. Daily effluent levels from these kinds of systems are generally uncorrelated, i.e., statistically independent. These general patterns are the same irrespective of the kind of pollutant in question. Significant correlation between observed pollutant levels, when present, should be factored into monthly average permit limits.

Several different methods can be used to account for correlation in determining limits. One general approach involves time series modeling. Another possible approach is to use a direct computation of the covariance among the observed data to adjust the variance of the average used in determining the limit. Help in adjusting the sample size for correlation is available from the OW Statistics Section (phone number [202] 382-5397).

Table E-1. Daily Maximum Permit Limit Calculations

The daily maximum permit limit is usually the 99th upper percentile value of the pollutant distribution. In certain cases the 95th percentile value may be allowable. The following gives the formulas:

WITH ALL MEASUREMENTS > DETECTION LIMIT (based on lognormal distribution)

$\hat{X}_{.95}$ = 95th percentile daily maximum limit

$= \exp[\hat{\mu}_y + 1.645\ \hat{\sigma}_y]$

$\hat{X}_{.99}$ = 99th percentile daily maximum limit

$= \exp[\hat{\mu}_y + 2.326\ \hat{\sigma}_y]$

where

x_i = daily pollutant measurement i

y_i = $\ln(x_i)$

k = sample size of data set

$\hat{\mu}_y = \Sigma(y_i) / k \qquad 1 \le i \le k$

$\hat{\sigma}_y^2 = \Sigma[(y_i - \hat{\mu}_y)^2] / (k - 1) \quad 1 \le i \le k$

$\hat{E}(X) = \exp(\hat{\mu}_y + 0.5\ \hat{\sigma}_y^2)$

$\hat{V}(X) = \exp(2\ \hat{\mu}_y + \hat{\sigma}_y^2)\ [\exp(\hat{\sigma}_y^2) - 1]$

$\hat{cv}(X) = [\exp(\hat{\sigma}_y^2) - 1]^{1/2}$

Table E-1. Daily Maximum Permit Limit Calculations (continued)

WITH SOME MEASUREMENTS < DETECTION LIMIT (based on delta-lognormal distribution)

$\hat{X}_{.95}$ = 95th percentile daily maximum limit

$$\hat{X}_{.95} = \begin{cases} D & \hat{\delta} \geq 0.95 \\ \max [D, \exp(\hat{\mu}_y + z^* \hat{\sigma}_y)] & \hat{\delta} < 0.95 \end{cases}$$

with $z^* = \Phi^{-1}[(0.95 - \hat{\delta}) / (1 - \hat{\delta})]$

$\hat{X}_{.99}$ = 99th percentile daily maximum limit

$$\hat{X}_{.99} = \begin{cases} D & \hat{\delta} \geq 0.99 \\ \max [D, \exp(\hat{\mu}_y + z^* \hat{\sigma}_y)] & \hat{\delta} < 0.99 \end{cases}$$

with $z^* = \Phi^{-1}[(0.99 - \hat{\delta}) / (1 - \hat{\delta})]$

where

x_i	=	daily pollutant measurement i	
k	=	sample size of data set	
D	=	detection limit (as established by the laboratory)	
r	=	number of nondetects	($x_1, x_2, \ldots, x_r$ are ≤ D)
k - r	=	number of detects	($x_{r+1}, x_{r+2}, \ldots, x_k$ are > D)
y_i	=	$\ln(x_i)$	for $r+1 \leq i \leq k$
$\hat{\delta}$	=	r / k	
$\hat{\mu}_y$	=	$\Sigma(y_i) / (k - r)$	$r+1 \leq i \leq k$ (exclude values ≤ D from sum)
$\hat{\sigma}_y^2$	=	$\Sigma[(y_i - \hat{\mu}_y)^2] / (k - r - 1)$	$r+1 \leq i \leq k$

$$\hat{E}(X^*) = \hat{\delta} D + (1 - \hat{\delta}) \exp(\hat{\mu}_y + 0.5 \hat{\sigma}_y^2)$$

$$\hat{V}(X^*) = (1 - \hat{\delta}) \exp(2 \hat{\mu}_y + \hat{\sigma}_y^2) [\exp(\hat{\sigma}_y^2) - (1 - \hat{\delta})] + \hat{\delta} (1 - \hat{\delta}) D [D - 2 \exp(\hat{\mu}_y + 0.5 \hat{\sigma}_y^2)]$$

Table E-2. Monthly Average Permit Limit Calculations for Ten Samples or Less

The monthly average permit limit is usually based on the estimates of the 95th percentile of the distribution of the average of the daily effluent values. For sample sizes less than or equal to 10, the data are assumed to be lognormally distributed (or delta-lognormally distributed if the data includes nondetects).

ALL MEASUREMENTS > DETECTION LIMIT (based on lognormal distribution)

$\hat{X}_{.95}$ = 95th percentile n-day monthly average limit
= $\exp[\hat{\mu}_n + 1.645\ \hat{\sigma}_n]$

$\hat{X}_{.99}$ = 99th percentile n-day monthly average limit
= $\exp[\hat{\mu}_n + 2.326\ \hat{\sigma}_n]$

where

x_i = daily pollutant measurement i

$y_i = \ln(x_i)$

k = sample size of data set

$\hat{\mu}_y = \Sigma(y_i) / k \qquad 1 \le i \le k$

$\hat{\sigma}_y^2 = \Sigma[(y_i - \hat{\mu}_y)^2] / (k - 1) \qquad 1 \le i \le k$

$\hat{E}(X) = \exp(\hat{\mu}_y + 0.5\ \hat{\sigma}_y^2)$

$\hat{V}(X) = \exp(2\ \hat{\mu}_y + \hat{\sigma}_y^2)[\exp(\hat{\sigma}_y^2) - 1]$

$\hat{\sigma}_n^2 = \ln\{\hat{V}(X) / (n[\hat{E}(X)]^2) + 1\}$

$\hat{\mu}_n = \ln(\hat{E}(X)) - 0.5\ \hat{\sigma}_n^2$

$\hat{cv}_n = [\exp(\hat{\sigma}_n^2) - 1]^{1/2}$

Table E-2. Monthly Average Permit Limit Calculations for Ten Samples or Less (continued)

SOME MEASUREMENTS < DETECTION LIMIT (based on delta-lognormal distribution)

$\hat{X}_{.95}$ = 95th percentile n-day monthly average limit

$$\hat{X}_{.95} = \begin{cases} D & \hat{\delta} \geq 0.95 \\ \max [D, \exp(\hat{\mu}_n + z^* \hat{\sigma}_n)] & \hat{\delta} < 0.95 \end{cases}$$

with $z^* = \Phi^{-1}[(0.95 - \hat{\delta}) / (1 - \hat{\delta})]$.

$\hat{X}_{.99}$ = 99th percentile n-day monthly average limit

$$\hat{X}_{.99} = \begin{cases} D & \hat{\delta} \geq 0.99 \\ \max [D, \exp(\hat{\mu}_n + z^* \hat{\sigma}_n)] & \hat{\delta} < 0.99 \end{cases}$$

with $z^* = \Phi^{-1}[(0.99 - \hat{\delta}) / (1 - \hat{\delta})]$

where

x_i	=	daily pollutant measurement i	
k	=	sample size of data set	
D	=	detection limit (as established by the laboratory)	
r	=	number of nondetects	($x_1, x_2, \ldots, x_r$ are $\leq D$)
k - r	=	number of detects	($x_{r+1}, x_{r+2}, \ldots, x_k$ are $> D$)
y_i	=	$\ln(x_i)$	for $r+1 \leq i \leq k$
$\hat{\delta}$	=	r / k	
$\hat{\mu}_y$	=	$\Sigma(y_i) / (k - r)$	$r+1 \leq i \leq k$ (exclude values $\leq D$ from sum)
$\hat{\sigma}_y^2$	=	$\Sigma[(y_i - \hat{\mu}_y)^2] / (k - r - 1)$	$r+1 \leq i \leq k$

$$\hat{E}(X^*) = \hat{\delta}\, D + (1 - \hat{\delta}) \exp(\hat{\mu}_y + 0.5\hat{\sigma}_y^2)$$

$$\hat{V}(X^*) = (1 - \hat{\delta}) \exp(2\hat{\mu}_y + \hat{\sigma}_y^2) [\exp(\hat{\sigma}_y^2) - (1 - \hat{\delta})] + \hat{\delta}(1 - \hat{\delta}) D[D - 2\exp(\hat{\mu}_y + 0.5\hat{\sigma}_y^2)]$$

$$\hat{\sigma}_n^2 = \ln\{(1 - \hat{\delta}^n)[1 + A + B + C]\}$$

with

$$A = \hat{V}(X^*) / [n(\hat{E}(X^*) - \hat{\delta}^n D)^2]$$

$$B = -[\hat{\delta}^n D^2 (1 - \hat{\delta}^n)] / (\hat{E}(X^*) - \hat{\delta}^n D)^2$$

$$C = (2\hat{\delta}^n D) / (\hat{E}(X^*) - \hat{\delta}^n D)$$

$$\hat{\mu}_n = \ln[(\hat{E}(X^*) - \hat{\delta}^n D) / (1 - \hat{\delta}^n)] - 0.5\hat{\sigma}_n^2$$

Table E-3. Monthly Average Permit Limit Calculations for More Than Ten Samples

The monthly average permit limit usually is based on the estimates of the 95th percentile of the distribution of the average of the daily effluent values. These daily values are assumed to be lognormally distributed. For sample sizes larger than 10, the averages (represented by the random variable X_n) are assumed to be normally distributed.

$\hat{X}_{.95}$ = 95th percentile n-day monthly average limit

$= \hat{E}(X_n) + 1.645\,[\hat{V}(X_n)]^{1/2}$

$\hat{X}_{.99}$ = 99th percentile n-day monthly average limit

$= \hat{E}(X_n) + 2.326\,[\hat{V}(X_n)]^{1/2}$

where

x_i = daily pollutant measurement i

y_i = $\ln(x_i)$

k = sample size of data set

$\hat{\mu}_y = \Sigma(y_i) / k, \qquad 1 \le i \le k$

$\hat{\sigma}_y^2 = \Sigma[(y_i - \hat{\mu}_y)^2] / (k - 1) \qquad 1 \le i \le k$

$\hat{E}(X) = \exp(\hat{\mu}_y + 0.5\,\hat{\sigma}_y^2)$

$\hat{V}(X) = \exp(2\,\hat{\mu}_y + \hat{\sigma}_y^2)\,[\exp(\hat{\sigma}_y^2) - 1]$

$\hat{E}(X_n) = E(X)$

$\hat{V}(X_n) = \hat{V}(X) / n$

$c\hat{v}(X_n) = \hat{V}(X_n)^{1/2} / (X_n)$

APPENDIX E
REFERENCES

1. U.S. EPA. 1987a. Development Document for Effluent Limitations Guidelines and Standards for the Organic Chemicals, Plastics and Synthetic Fibers. EPA 440/1-87/009.

2. Mendenhall, W., R. Scheaffer, and D. Wackerly. 1981. Mathematical Statistics with Applications. 2d ed. Wadsworth, MA.

3. Barakat, R. 1976. Sums of Independent Lognormally Distributed Random Variables. *J. Optical Soc. Am.* 66:211-16.

4. Aitchison, J., and J. Brown. 1963. The Lognormal Distribution. Cambridge University Press.

5. Johnson, R., and D. Wichern. 1982. Applied Multivariate Statistical Analysis. New Jersey: Prentice-Hall.

6. Hollander, M., and D. Wolfe. 1973. Nonparametric Statistical Methods. New York: Wiley.

ADDITIONAL REFERENCES

Kahn, H. 1989. Memorandum: Response to Memorandum from Dr. Don Mount of December 22, 1988. August 30, 1989. U.S. EPA, Washington, DC, to J. Taft, U.S. EPA, Permits Division, Washington, DC.

Kahn, H., and M. Rubin. 1989. Use of Statistical Methods in Industrial Water Pollution Control Regulations in the United States. *Environmental Monitoring and Assessment* 12:129-48.

ORAL ARGUMENT OCCURRED ON OCTOBER 15, 2004

UNITED STATES COURT OF APPEALS
FOR THE DISTRICT OF COLUMBIA CIRCUIT

EDISON ELECTRIC INSTITUTE, *et al.*,	)	
	)	
Petitioners,	)	
v.	)	Docket No. 96-1062
	)	(and consolidated cases)
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY, *et al.*,	)	
	)	
Respondents.	)	
	)	

DECLARATION OF JOHN F. FOX, PH.D. ON BEHALF OF RESPONDENT UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Pursuant to 28 U.S.C. § 1746, I, John F. Fox, affirm and state as follows:

1. I am a Mathematical Statistician in the Office of Research and Development in the United States Environmental Protection Agency ("EPA").

2. I was a member of the team that developed EPA's Whole Effluent Toxicity ("WET") test methods and the Interlaboratory Study used to validate the test methods. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms, EPA-821-R-02-013 (4th ed. Oct. 2002) ("Methods Manual"), J.A. 1554-1813; Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Whole Effluent Toxicity Test Methods, Vols. 1 & 2, EPA 821-B-01-004 (Sept. 2001) ("Interlaboratory Study"), J.A. 1123-1336.

3. I obtained my Bachelor of Science in Biology from Johns Hopkins University in 1967 and my Ph.D. in Biology from the University of Chicago in 1974. I obtained my Master of Science in Statistics from Florida State University in 1994.

4. I have been a statistician for EPA since June, 1995. Before that I was an assistant professor of biometrics at the University of Alaska, Fairbanks, where I taught graduate courses in statistical research design, data analysis and statistics, and ecological topics, and where I advised faculty and graduate students on experimental and observational study designs and statistical data analysis.

EXHIBIT 2

5. I have served as a biostatistics reviewer for several academic journals, including *Ecology*, *The American Naturalist*, and *The Canadian Journal of Forest Research*. I have authored and co-authored 23 publications on ecology and biostatistics in peer-reviewed journals and monographs. I am a member of the American Statistical Association.

6. I coauthored a peer-reviewed paper on variability of WET methods for chronic toxicity (2003, Env. Toxicol. & Chem. 22:2323-2328) and the article *Whole Effluent Toxicity* for the Encyclopedia of Environmetrics (vol. 4, 2002, ed. A. H. El-Shaarawi & W.W. Piegorsch). I coauthored EPA's guidance document on WET variability, Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System, EPA 833-R-00-003 (June 2000) ("Variability Guidance Document"), for which I led the work on data analysis and data quality and wrote most of two chapters and three appendices. I also to the draft National Whole Effluent Toxicity (WET) Implementation Guidance Under the NPDES Program (DRAFT) (EPA 2004) ("2004 Draft Guidance Document").

7. I have reviewed the Court's decision, *Edison Electric Institute v. EPA*, 391 F.3d 1267 (D.C. Cir. 2004), Petitioners' Petition for Panel Rehearing, and relevant materials in the administrative record, including the Interlaboratory Study, the Variability Guidance Document, and EPA's *Technical Support Document for Water Quality-Based Toxics Control* (March 1991) ("TSD"), J.A. at 359-473. I also reviewed the affidavit of Robert F. Rockwell ("Rockwell Aff.") and the 2004 Draft Guidance Document.

8. After reviewing these materials and based on my personal knowledge and experience regarding EPA's promulgation of the WET test methods, it is my understanding and opinion that: (a) WET test method variability, including coefficients of variation ("Cvs"), should not be quantified based on WET test results arrived at by hypothesis testing, *i.e.*, No Observed Effect Concentration ("NOEC") data; (b) Table 2 on page 25 of Petitioners' Reply Brief ("Reply Brief Table") contains at least one major methodological error in that Petitioners calculated CVs from NOEC data; (c) the Court did not commit a statistical error by calculating the CVs as it did, 391 F.3d at 1270 n.4; (d) the data in question cannot fit the delta lognormal distribution; and (e) Petitioners based their calculations on a dataset containing many "high" values and a few relatively "small" values but did not follow the standard practice for interlaboratory studies of rejecting data outliers. The Petitioners then converted the dataset to TUs, which, in this case, produced an inflated CV.

9. I reconstructed the Court's calculations in footnote 4, which appear to be based on the NOEC data in the seventh column from the left of Table 9.8, "Results for *Ceriodaphnia* chronic test performed on reference toxicant samples," Interlaboratory Study at 81-82, J.A at 1224-25. A reference toxicant is a known toxic chemical that is routinely tested to evaluate the consistency and precision of toxicity tests. *Id.* at xx, J.A. at 1142. Petitioners based the third

row of their Reply Brief Table on the same dataset. The test results reported on this table are based on two statistical techniques: hypothesis testing and point estimation. Hypothesis test results are reported as NOEC values, and point estimation results are reported as LC_{50} and IC_{25} for organism survival and organism reproduction, respectively.

10. The Court's first calculation yielded a CV of 0.43. Based on my reconstruction of the Court's calculations, it appears that the Court used the 36 numerical values for NOEC reported in Table 9.8 of the Interlaboratory Study. The sample standard deviation and sample mean (average) of these NOECs are 34.46 and 79.86, respectively. The CV (the sample standard deviation divided by the sample mean) of this dataset is 0.43. While EPA advises against using NOEC data for this type of calculation, see, *infra*, ¶¶ 14 & 23, the Court's first calculation is mathematically correct.

11. The Court's second calculation yielded a CV of 1.47. Based on my calculations, it appears that the Court used the same dataset as in the first calculation, but converted the WET test results, *i.e.*, the NOECs, to TUs before calculating the CV. When the NOEC values are transformed to TUs, (*i.e.*, TU = 100 divided by the NOEC), the standard deviation and mean (average) of these TUs are 3.842 and 2.611, respectively, and the resulting CV is 1.47.

12. The Court's second calculation is mathematically correct. It also shows that calculating a CV based on NOEC data that have been converted to TUs, to evaluate interlaboratory variability, produces, in this case, a CV more than three times larger than that produced when using the percentages.

13. I also attempted to reconstruct the Petitioners' calculations in the third row of the Reply Brief Table, where they reported a CV of 179.3% (equivalent to 1.793). I first converted the NOEC data reported by each laboratory into a TU. (TU = 100 ÷ NOEC). I then applied the "delta-lognormal model" from page E-10 in Appendix E of the TSD. *See* Reh'g Pet. at 8. In doing so, I treated each TU that equaled "1" as if it were a "non-detect" value. This produced an estimated standard deviation of 4.869 and an estimated mean of 2.715. The resulting CV is 1.793.

14. Use of NOECs to quantify method variability leads to distorted and unreliable results, because the hypothesis test that produces a NOEC value considers only a finite set of discrete (percentage concentration) values, *e.g.*, 6.25, 12.5, 25, 50 or 100. The NOEC result is the highest concentration at which we cannot reject the hypothesis that the effluent is not toxic (the "null hypothesis"). The set of possible results is limited to the actual concentrations tested. It is not possible to have a result that falls between the test concentrations. Thus, NOEC data are not continuous. Widening or narrowing the range of concentrations and changing the ratios between concentrations used in the test will alter the CV of NOECs. Accordingly, CVs based on NOEC values are likely to overstate or understate (*i.e.*, distort)

the resulting characterization of variability. EPA has cautioned against calculating CVs for NOEC data and instead advises that point estimate data, such as IC_{25} and LC_{50}, be used when calculating CVs. TSD at 5-6 & 11-12; Variability Guidance at 3-2, 6-1 & 6-4; Interlaboratory Study at 4 & 66; Methods Manual at 15-16 & 39, J.A. at 383-84, 389-90, 837, 859, 862, 1147, 1209, 1583-84 & 1607.

15. Use of TUs based on these data produces an inflated estimate of the CV. The CV of the TUs will tend to be smaller or greater than the CV of the NOECs, depending on where most of the values lie within the dataset. This particular dataset for *Ceriodaphnia* reproduction consists of many "high" values equaling 100 and a few relatively "small" values, *e.g.*, 6.25. The CV of all of the values, when they are converted to TUs, *i.e.*, 100 ÷ NOEC, will be greater than the CV of the NOEC data, because the inverses of a few small values will inflate the standard deviation disproportionately more than they increase the mean.

16. Thus, the presence of a few small percentages in this dataset caused the CV for TUs to be larger than the CV for corresponding percentages, *i.e.*, the NOEC or IC_{25} data. When evaluating a test method's variability, it is standard practice to apply statistical methods to identify and reject outliers. Had the Petitioners analyzed the data in a manner consistent with their assumptions, they would have identified and excluded the outliers, and – notwithstanding the inappropriateness of CV calculations based on NOEC data – at least two of the lowest datapoints would have been excluded from the CV calculations, which would have dramatically reduced the CV.

17. Petitioners claim that the Court incorrectly assumed that the data were normally distributed. Reh'g Pet. at 7-8. Based on my review of the Court's opinion and my reconstruction of the Court's calculations, there is no indication that the Court assumed normality to compute the CV for these data in order to illustrate that converting NOEC test results to TUs for CV analysis had a distorting effect.

18. It is not correct that it is "a fundamental rule of statistical analysis that the distribution of a data set dictates which statistics should be performed on the data." *See* Reh'g Pet. at 3, Rockwell Aff. at ¶ 6. The sample mean, standard deviation, and CV, all of which are statistics that describe the data, may be calculated legitimately for any data set, regardless of the data's distribution. It is a fundamental, elementary theorem in mathematical statistics that the sample mean and sample variance produce unbiased estimates of the population mean and population variance for any and all distributions that possess a finite mean and finite variance.

19. The *Ceriodaphnia* dataset at issue here does not fit a "delta-lognormal" distribution. NOEC data are limited to a finite set of discrete values, *see supra*, ¶ 14, and cannot be assumed to fit any distribution that describes a continuous random variable, such as the normal,

lognormal or delta-lognormal distributions. NOEC should be modeled using a discrete distribution, because NOEC can take only the discrete values of the measured test concentrations.

20. Therefore, statistical methods that assume normality or lognormality, including the delta-lognormal method, do not apply to the NOEC data at issue. Thus, the procedures in Appendix E of the TSD do not apply.

21. EPA's recommendation to use TUs to characterize variability in specific statistical procedures does not apply to determining test method variability. Rather, EPA's recommendation to use TUs applies to effluent variability determinations for specific permitting purposes. For example, effluent variability is relevant when determining the need for and when deriving a permit limit. The data used to characterize effluent variability are effluent monitoring data collected *over a period of time* from *one discharge point*. It is reasonable to assume, unless the data show otherwise, that treated effluent data are lognormally distributed. TSD at 93-94, 95 ("From the vast amount of data that EPA has examined, it is reasonable to assume, (unless the data show otherwise) that treated effluent data follow a lognormal distribution."), Variability Guidance at 6-1, J.A. at 445-46 & 859; TSD E-2, E-3, E-6, E-7, Exh. 1.

22. Point estimate data, *i.e.*, IC_{25} and LC_{50}, underlying *test method variability* determinations, on the other hand, are based on test results collected from multiple laboratories that tested *samples having the same reference toxicant concentration*, and the normal distribution can be assumed to provide a reasonable approximation for such data, although normality is not essential to characterizing test method variability (*see* ¶¶ 17 & 18, *supra*).

23. I am not aware of any materials in the administrative record in which EPA asserts that TU is an appropriate basis for evaluating WET test method variability. In its WET methods documents and related reports, EPA quantifies WET test method variability based on the CV for WET test results that are point estimates, *i.e.*, IC_{25} and LC_{50}. For WET test results that are based on hypothesis testing, *e.g.*, NOEC, EPA evaluates test method variability based on ranges. TSD at 12; Variability Guidance at 3-2, 6-1 & 6-4; Interlaboratory Study at 4, 66; Methods Manual at 15-16 & 39, J.A. at 390, 837, 859, 862, 1147, 1209, 1583-84 & 1607.

24. WET test results are expressed in percentages. The percentage reflects the concentration of the test sample that produces a specific result. For example, the IC_{25} is the percentage concentration at which there is a 25 percent reduction in test organism growth or reproduction. EPA Br. at 17; Methods Manual at 37, J.A. at 1605. Percentages are used to calculate WET test method variability and to compare their variability to that of chemical-specific test methods in all previously-published, peer-reviewed reviews and studies that I am aware of. *See, e.g.*, Warren-Hicks, W., B.R. Parkhurst, D. Moore, and S. Teed, 1999,

Water Environment Research Foundation, Whole effluent toxicity testing methods: Accounting for Variance. Project 95-PQL-1 (ISBN 1-893664-01-5). Moreover, none of the studies I have reviewed contain any discussion of using TUs rather than WET test results themselves to compare WET test variability to that of chemical-specific test methods. Accordingly, EPA's use of percentages to calculate WET test method variability, *see* Memorandum from M. Kelly, EPA Engineering and Analysis Division (October 16, 2002), J.A. 1814-23, was consistent with the practice established in the relevant scientific literature.

25. Whole effluent toxicity is a method-defined analyte, and WET test results represent the *effect* of pollutants on aquatic life. 40 C.F.R. § 122.2. In this way, the concept of "mass" does not apply to whole effluent toxicity in the same way that it applies to the measurement of chemical analytes. Thus, the mere inversion and multiplication by 100 of WET test results, to generate a TU value, is not an expression of the "mass" of toxicity in a sample. Petitioners refer to the use of TU values to model effluent toxicity in the development of regulatory waste load allocations. Reh'g Pet. at 6-7 (citing TSD at 85). The development of regulatory waste load allocations relies on mass balance principles and requires toxicity to be in units that increase when the percent of effluent in the receiving stream increases in order to calculate necessary dilutions. *See* TSD at 85. The evaluation of WET test method variability, by contrast, does not employ mass balance principles, such that the "mass" of toxicity is both irrelevant and indeterminable. Thus, WET test results themselves, and the related percentage scale, are appropriate for comparisons between WET test methods and test methods used to determine the concentrations of specific chemical analytes. In addition, as discussed with respect to the NOEC values in this particular dataset, which consists of data from multiple laboratories measuring samples having the same reference toxicant concentration, the TU scale will tend to distort measures of variability, like CV, especially if the NOEC values contain a few extreme values (*see* ¶¶ 15&16, *supra*).

26. Petitioners' conclusions about the arithmetic averaging of WET test results in relation to physical mixing, Reh'g Pet. at 6, n.6, Rockwell Aff. at ¶ 14, ignore at least two important principles. First, because whole effluent toxicity is the aggregate toxic effect of an effluent, 40 C.F.R. § 122.2, it cannot be assumed that mixing two samples that produced different WET test results will produce a composite sample with a WET test result that is the arithmetic average of the two. For example, the pollutants present in one sample may amplify the toxic effects of the pollutants in another sample when they are combined. Second, toxicity typically follows a sigmoidal concentration-response curve, *see* Methods Manual at 50, J.A. 1618, and not necessarily a straight line. This means that the measured toxic effect of a physical mixture of two equal volumes of different test samples will not necessarily fall at a midpoint between the two concentration-response curves of those two test samples, regardless of whether NOEC, IC_{25}, LC_{50}, or TU values are used. Thus, rudimentary arithmetic averaging, *per se*, is not analytically useful to the assessment of WET test method variability.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct to the best of my knowledge.

Executed this 15th day of March, 2005:

John F. Fox

United States Environmental Protection Agency | Office of Wastewater Management (4203M) | EPA 832-B-04-003 (Month/Year) http://www.epa.gov/npdes/permitbasics

National Whole Effluent Toxicity (WET) Implementation Guidance Under the NPDES Program

DRAFT

November 2004

(Released on December 28, 2004)

EXHIBIT 3

Glossary[4]

Acute-to-Chronic Ratio (ACR) is the ratio of the acute toxicity of an effluent or a toxic to its chronic toxicity. It is used as a factor for estimating chronic toxicity on the basis of acute toxicity data, or for estimating acute toxicity on the basis of chronic toxicity data.

Acute Toxicity Test is a test to determine the concentration of effluent or ambient waters that causes an adverse effect (usually death) on a group of test organisms during a short-term exposure (e.g., 24, 48, or 96 hours). Acute toxicity is measured using statistical procedures (e.g., point estimate techniques or a hypothesis test).

Ambient Toxicity is measured by a toxicity test performed using solely receiving water.

Average Monthly Limit (AML) is the highest allowable value for the average of daily discharges obtained over a calendar month.

Chronic Toxicity Test is a short-term test, usually 96 hours or longer in duration, in which sublethal effects (e.g., significantly reduced growth or reproduction) are usually measured in addition to lethality. Chronic toxicity is defined as $TU_c = 100/NOEC$ or $TU_c = 100/EC_p$ or IC_p.

ChV is the chronic value, the geometric mean of the NOEC and LOEC.

Coefficient of Variation (CV) is a standard statistical measure of the relative variation in a distribution or set of data, defined as the standard deviation divided by the mean. It is also called the relative standard deviation (RSD). The CV can be used as a measure of within-laboratory and between-laboratory precision or as a measure of precision among replicates for each treatment concentration.

Criterion Continuous Concentration (CCC) is the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed indefinitely without causing unacceptable effects such as the exceedence of a chronic water quality criterion.

Criterion Maximum Concentration (CMC) is the highest in-stream concentration of a toxic or an effluent to which organisms can be exposed for a brief period of time without causing an acute effect such as the exceedence of an acute water quality criterion.

Discharge Monitoring Report (DMR) is EPA's standardized reporting form for the reporting of self-monitoring results by permittees. DMRs must be used by "NPDES-approved States," as well as by EPA. The EPA standardized forms may be modified to substitute the State agency's name, address, logo, and other similar information, as appropriate, in place of EPA's.

Effect Concentration is a point estimate of the toxic concentration that would cause an observable adverse effect (e.g., death, immobilization, or serious incapacitation) in a given percent of the test organisms. EC_{25} is a point estimate of the toxic concentration that would cause an observable adverse effect in 25 percent of the test organisms.

Effluent Flow (Qe) is the flow (in cubic feet per second or million gallons per day) of a wastewater discharge from a NPDES-regulated facility expressed in standard NPDES formulas by permit writers as "Qe" to calculate water quality based effluent limits.

[4] These terms may have other meanings in other EPA programs or documents, including other programs under the CWA.

2 SETTING APPROPRIATE WATER QUALITY GOALS FOR WET

State water quality standards (WQS) provide the foundation for water quality-based pollution control programs. The purpose of WET limits and monitoring requirements is to ensure compliance with State numeric or narrative water quality criteria established to protect the designated uses of the water body. This chapter provides a brief description of the WQS that are the basis for WET water quality-based effluent limits (WQBELs) in NPDES permits.

2.1 EPA'S RECOMMENDED WATER QUALITY CRITERIA FOR WET

Water quality *standards* are provisions of State (or, in certain instances, Federal) law which define the water quality goals of a water body, or portion thereof, by designating the use or uses to be made of the water body and by setting criteria necessary to protect the uses. States adopt water quality standards to protect public health or welfare, enhance the quality of water, and serve the purposes of the Clean Water Act. Such standards serve the dual purposes of establishing the water quality goals for a specific water body and serve as the regulatory basis for the establishment of water quality-based treatment controls and strategies beyond the technology-based levels of treatment required by sections 301(b) and 306 of the CWA (40 CFR 131.2).

Water quality *criteria* are elements of State WQS, expressed as constituent concentrations, levels, or narrative statements representing a quality of water that supports a particular use. When criteria are met, water quality will generally protect the designated use (40 CFR 131.3). While states have adopted a variety of criteria expressed as constituent concentration levels (or *numeric* criteria) for various pollutants for the protection of aquatic life, all states have adopted criteria expressed as narrative statements (or *narrative* criteria). These narrative criteria, often referred to as "free-from" (or, in the case of toxicity, "no toxics in toxic amounts") criteria, are an effective tool for controlling the discharge of pollutants where numeric criteria are not available. Narrative criteria are a basis for establishing WET controls in the NPDES permitting regulations at 40 CFR 122.44(d)(1).

Section 304(a) criteria are developed by EPA under authority of section 304(a) of the Act based on the latest scientific information on the relationship that a constituent concentration has on a particular aquatic species and/or human health. This information is issued periodically to the States as guidance for use in developing criteria. In adopting criteria to protect their designated uses, States may establish criteria based on (1) section 304(a) guidance; (2) section 304(a) guidance modified to reflect site-specific conditions; or (3) other scientifically defensible methods.

Although EPA has not published numeric water quality criteria under section 304(a) for whole effluent toxicity, EPA has provided general guidance on appropriate WET limits. The TSD (USEPA 1991a) recommends 0.3 acute toxic units (TU_a) as an acute criterion and 1.0 chronic toxic units (TU_c) as a chronic criterion, for most water bodies. While effective for CWA purposes only in the Great Lakes Basin, the final Water Quality Guidance for the Great Lakes System in 40 CFR 132 (also known as the Great Lakes Initiative) also establishes 0.3 TU_a and 1.0 TU_c, either as numeric criteria or as equivalent numeric interpretations of narrative criteria, for receiving waters of the Great Lakes [40 CFR 132, Appendix F, Items A.1. and A.2].

Federal regulations at 40 CFR 122.44(d)(1)(i) establish different approaches for implementing a water quality criterion for toxicity in NPDES permits, depending on whether the criterion is expressed in a numeric or narrative form. If the State has not adopted a numeric criterion for WET, EPA expects the permitting authority to *interpret* the State narrative criterion so that the appropriate effluent limits, including any necessary toxicity numeric limits, can be established. States should identify the method they intend to use in regulating toxics based on narrative criteria and describe how their toxics control

EPA has developed the following three-step approach to using a steady-state model and facility-specific WET data to determine whether a discharge causes, has reasonable potential to cause, or contributes to an excursion above a numeric or narrative toxicity criterion. This approach assumes a lognormal distribution of effluent data. (The basis for assuming lognormality of WET data is included in EPA's TSD, Appendix E, pages E-6 through E-10 and Tables E-1 through E-3.)

Step 1. Determine whether a discharge *causes* an excursion above numeric or narrative water quality criteria for aquatic life (i.e., toxicity). When assessing the need for a WET limit using a steady-state model, the equation for determining the resultant RWC for WET after effluent discharge is:

$$RWC = \frac{QeCe}{Qe + (pmf)Qs}$$

The value used for *Ce* is the maximum observed TU value. Once *RWC* is calculated, this magnitude is compared to the toxicity criterion. If the calculated magnitude of the *RWC* is greater than the toxicity criterion (e.g., $RWC > 0.3$ TU_a for acute toxicity or $RWC > 1.0$ TU_c for chronic toxicity), the discharge causes an excursion above the toxicity criterion (see Box C in Figure 2), and a WET limit is necessary (see Box D in Figure 2).

Regulatory authorities should calculate the facility-specific CV using point estimate techniques to determine the need for and derive a permit limit, even if the permit compliance monitoring test results will be determined using hypothesis test procedures (USEPA 2000c, see sections 3.4.1 and 6.2). Point estimates make the best use of the WET test data for purposes of estimating the CV, LTA, and RP factor and calculating the permit limit.

Alternatively, if the calculated magnitude of *RWC* is less than or equal to the toxicity criterion (e.g., $RWC \leq 0.3$ TU_a or $RWC \leq 1.0$ TU_c), the discharge does not cause, but may still have reasonable potential to cause, an excursion above the toxicity criterion. The permitting authority then evaluates whether the discharge has reasonable potential to cause an excursion above the applicable toxicity criterion based on the sample size [i.e., the number of available WET data points (n), see Box E in Figure 2].

Step 2. Determine whether a discharge *has reasonable potential* to cause an excursion above numeric or narrative water quality criteria for aquatic life (i.e., toxicity). EPA's statistical approach for determining whether a discharge will have reasonable potential to cause an excursion above a water quality criterion is outlined below and described in Section 3.2.2 of the TSD. This approach accounts for effluent variability and uncertainty associated with small effluent data sets by calculating a projected effluent value associated with the upper 95-percent or 99-percent confidence bound representing the 95th or 99th percentile of the lognormal distribution. Censored data may be encountered in WET test results. These are data reported as "greater than" or "less than," for example "> 100% effluent" or "< 6.25% effluent." Such data would be expressed in toxic units for reasonable potential calculations ("< 1.0 TU_c" or "> 16.0 TU_c"). It is not appropriate to ignore such data or assign an arbitrary value such as detection limit or one-half the detection limit, and to then proceed with calculations as if such values were uncensored data. Two reasonable methods are available for accommodating censored data in calculations used to determine reasonable potential and permit limitations (see text box *Treatment of Censored Data Sets*).

Steps for determining RP using a reasonable potential multiplying factor (RPMF) are described below. Permitting authorities also can calculate RP multipliers directly using the equations in Appendix E

of EPA's TSD. The RPMF is a function of the number of WET tests (n); the TSD recommends use of a default value of CV = 0.6 if n is less than 10. The process is described below (*NOTE: Use point estimates rather than NOEC for these calculations*):

a. If 10 or more valid, facility-specific WET data points for the most sensitive species are available ($n \geq 10$), calculate the facility-specific variance of log(TU) or log(concentration) using the equations in EPA's TSD (Appendix E, page E-8, page E-15) (see Box F in Figure 2). [Do not calculate variance from the CV as shown in Box 5-2 (page 100).]

-or-

b. If fewer than 10 valid, facility-specific WET data points are available ($n < 10$), the permitting authority should use the recommended default CV of 0.6. This CV is an empirical estimate of effluent variability in relation to the mean developed using WET data presented in EPA's TSD, Appendix A (see Box I in Figure 2).

Treatment of Censored Data Points

1. The delta-lognormal method described in the TSD (USEPA 1991a) may be applied, provided there is only one censoring level, for example, 100% (1 TU_c). The delta-lognormal is expected to provide reasonable estimates for the mean and variance, even if the distribution is a mixture of a lognormal and a point mass at the censoring level. For the case of one or more censoring levels, one should consider the ROS method (below) or maximum likelihood methods (Cohen 1991).

2. The regression-on-order-statistics (ROS) methodology is expected to provide reasonable estimates for the mean, variance, and possibly for high percentiles (Helsel and Cohn 1988, Hirsch and Stedinger 1987, Kroll and Stedinger 1996, Shumway et al. 2002). These authors evaluated the method for estimating the 90th percentile with small (n = 10) samples. For higher percentiles, we are not aware of any similar evaluations, but one must expect greater variability than reported for the 90th percentile. The ROS method accommodates multiple detection limits. Be warned that the articles by Helsel and Cohn (1988) and Hirsch and Stedinger (1987) contained an inaccurate formula for Cj, and the articles did not address ties in uncensored observations or the occurrence of uncensored observations below the lowest detection limit. The formula for Cj is: Cj = Bj - (Aj-1 + Bj-1), where
   A0 = B0 = 0.

3. If the frequency of data at a detection limit is greater than expected from the censoring of a single parent distribution, a mixture model may be needed. The ROS method is not intended to deal with this situation. The delta-lognormal method (USEPA 1991a) and the approach described by Taylor et al. (2001) are two ways of handling this situation

Cohen, A.C. 1991. Truncated and Censored Samples. New York: Marcel Dekker.

Helsel, D.R., and T.A. Cohn. 1988. Estimation of Descriptive Statistics for Multiply Censored Water Quality Data. Water Resour. Res. 24:1997–2004.

Hirsch, R.M., and J.R. Stedinger. 1987. Plotting Positions for Historical Floods and Their Precision. Wat. Resour. Res. 23:715–727.

Kroll, C.N., and J.R. Stedinger. 1996. Estimation of Moments and Quantiles Using Censored Data. Wat. Resour. Res. 32:1005–1012.

Shumway, R.H., R.S. Azari, and M. Kayhanian. 2002. Statistical Approaches to Estimating Mean Water Quality Concentrations with Detection Limits. Environ. Sci. Technol. 36:3345–3353.

Taylor, D.J., L.L. Kupper, S.M. Rappaport, and R.H. Lyles. 2001. A Mixture Model for Occupational Exposure Mean Testing with a Limit of Detection. Biometrics 57:681–688.

APPENDIX A – EXAMPLE: PERMIT LIMIT DERIVATION PROCEDURES

This appendix presents an example of how to derive a permit limit.

The mean and standard deviation of log(TU) or log(chemical concentration are calculated for each pollutant using historical effluent data). Where historical data regarding effluent variability are insufficient (e.g., n < 10), the default CV should be 0.6 (see TSD, Appendix E, pg. E-3). In that case only, the variance of log(TU) or log(chemical concentration) is calculated from the CV using formulas in Box 5-2 of the TSD (page 100). Statistical derivation procedures for the AML for WET should assume that at least four samples (n) will be taken per month.

The WLA required to protect against both acute and chronic effects under critical conditions may be calculated using either steady-state or dynamic models. However, for derivation of the WLA, the equation is rearranged to solve for the effluent concentration (C_d), or WLA, necessary to achieve the appropriate applicable criterion. For compliance purposes, the water quality criterion for aquatic life (toxicity criterion) is set equal to C_r, where C_r is the applicable criterion:

$$WLA = C_d = [C_r(Q_d + \%Q_s)] - [(C_s)(\%Q_s)]/Q_d$$

where:

Q_d = waste discharge flow in cubic feet per second (cfs) or MGD

C_d = waste discharge pollutant concentration in TUs for WET (TU_a or TU_c)

Q_s = background in-stream flow in cfs or MGD above point of discharge

C_s = background in-stream pollutant concentration in TUs for WET (TU_a or TU_c); setting $C_s = 0$ is recommended for WET

$\%Q_s$ = percent of upstream flow allowed by mixing zone standard, if applicable

Q_r = resultant in-stream flow after discharge in cfs or MGD: $\%Q_s + Q_d$

C_r = applicable toxicity criterion = resultant in-stream pollutant concentration in TUs for WET (TU_a or TU_c), in the stream reach (after complete mixing)

In most cases, this steady-state model should be used to calculate the WLA (i.e., allowable effluent concentration) that will meet acute and chronic water quality criteria for the protection of aquatic life at the critical stream flow conditions, for example, 1Q10 and 7Q10 respectively (see TSD, Section 4.2, pg. 68). Ambient flow data from the U.S. Geological Survey are available on STORET.

When calculating the WLA, it should be noted that, if applicable State water quality standards and plans do not explicitly allow the application of mixing zones, the appropriate applicable criterion must be met at the end-of-pipe (i.e., applicable criterion = $C_r = C_d$ = WLA). Where mixing zones are allowed, appropriate State procedures should be applied.

If adequate receiving water flow and effluent concentration data are available to estimate frequency distributions, dynamic modeling techniques can be used to calculate allowable effluent loadings that will more precisely maintain water quality standards (see TSD, pg. 97). The steady-state mass balance equation, however, when coupled with the recommended conservative assumptions, should be adequately