Work Group Meeting: June 26, 2007
Summary Notes

Whole Effluent Toxicity (WET)

Summary of General Discussion:

The group suggested:

- The use IC25 instead of NOECs to smooth out the data. This is especially good for sublethal endpoints.
- That what constitutes a permit violation should be defined as well as what will lead to enforcement. Does a failure constitute a violation? California uses a median of 2 months of data, and North Carolina uses the median of a longer time period.
- For sublethal effects, use a 99% confidence interval (CI) instead of the 95% CI in order to tighten up the data and be more confident that false positives are less likely to occur. M. Hess expressed some concern that if we use the 99% CI and other conditions, such as persistent lethality, are we really using the same level of protection? A SWQM team member pointed out that we must ensure we are using an appropriate sample size (the number of organisms exposed) with the 99% CI.
- Find a way to address sublethal toxicity due to TDS. A possible approach would be with species substitution. This was a major concern from Tom (who?)
  - Lynda Clayton's response was that EPA does not think this is justification for toxicity in a chronic test. The effluent must be treated to a level that is non-toxic.
  - The group would like to see data that shows correlations between instream effects and sublethal toxicity.

The group questioned:

- Is there actual data on the correlation of sublethal toxicity to instream impacts?
- What are other states doing? Peggy Glass is looking into this and plans on providing information to TCEQ.
- What if there is cessation of lethality during a TRE? EPA has indicated that the cessation of lethality clause will go away.
- Three Steps (from Randy Palachek):
  - Reasonable Potential (RP)
  - Up front WET limits
  - Sublethal WET limits
  - Does TCEQ think their current process is inadequate?
- Is there an approach to address an inverted dose response? Invalidation of the test
- Is it really an option to "continue the current process"? Not really.

Questions for General Discussion:

1. should lethal and sublethal be assessed the same way?

Sara Burgin: Is there data showing a correlation between a failure of a sublethal test and toxicity in streams (receiving water impacts)? It is very difficult to identify a sublethal toxicant.

Myron Hess: the test for sublethality is all we have to protect against nonlethal effects. M. Hess suggested that there needs to be a change from where we are now.

Randy Palachek: Why is EPA so adamant that TCEQ change its procedures?

Lynda Clayton: EPA considers lethal and sublethal to be toxicity. They do not think TCEQ has correctly implemented the WET program.

Ted Long: Has there been any consideration that sublethal failures may be due to the incoming waters? For example, TDS in source water?

Faust Parker: For TDS- request a different species, different test length.

David Rutledge from TXU: TDS in once-thru cooling water, Phil Jennings understands the problem.

San Marcos River Foundation: Let's remember the objective is to make sure our streams can support life. The fathead minnow is a very hardy species.

Lynda Clayton: Yes, but water fleas are more sensitive which is why we test both.

Peggy Glass: Questioned whether sublethal represents real toxicity. There is difficulty in identifying toxicant using sublethal, and there is a high rate of false positives. She suggested that the use of the 99% CI is a good thing to help with this.

Pat Radloff: Concerning the inherent variability of WET testing. Are there any studies that the group can look at? If so, let's look at the studies rather than just grabbing a number, especially for sublethal. If CI is used there needs to be a rational for changing it.

Myron Hess: All factors interact together – 2 failures out of 3 tests, the C.I., and compounding factors can all reduce protection.

Lauren Kalichek: C.I. is for assessing one particular test.

Myron Hess: If you say you are 99% confident, why use 2 out of 3?

Sara Burgin: Magnitude and persistence are both important.

Lynda Clayton: If TCEQ used 99% CI, we would be more confident that the result is an actual failure.

Myron Hess: We may miss some real failures using a higher C.I.

** Group: Assess sublethal differently than lethal.

2. How should we address RP for lethal toxicity?

Sara Burgin: Does TCEQ see a need to change our procedures for lethal?

Peggy Glass: What TX currently does is done in other states and is recommended by other groups like the National Water Quality Assessment Program (NAWQA).

Tom and Karalyn Hardaway: There's a history of sublethal TREs being difficult. RP-WET does not serve the permit process. Our TRE process has worked. The problem gets identified and fixed (for a specific chemical).

Peggy Glass: A three year compliance period is not that helpful. You may not have a toxicant to identify. A permit limit is not that helpful. It just gives permittees a black eye. It does not help the environment.

Pat Radloff: Easily identified toxics are given reporting requirements or numeric limits (70-84% have a report clause in the permit, values greater than or equal to 85% have a limit placed in the permit).

Jim Davenport: EPA sees less toxicity when permittees have a WET limit.

<u>Break out group results - RP - additional comments aside from individual group notes:</u>

Concerns were expressed about anti-backsliding with regards to removing a WET limit. There would need to be a process stipulated in the IP, as well as language in the permit.

Need to resolve the issue of what percentage of permits would get up front WET limits based on RP. The group suggested 30% with best professional judgment (BPJ), EPA's percentage, or 80% without BPJ. L'Oreal Stepney indicated that TCEQ will verify these numbers and then double check EPA's numbers.

<u>Breakout Group Discussion Questions:</u>

Reasonable Potential:

1. Do you think Reasonable Potential (RP) should be determined by using EPA's model, which uses the worst-test result of the past 5 years?

a) Why or why not?

There is a general agreement that RP shouldn't be determined by EPA's method because it is

- Unreasonable
- not accurate
- statistically flawed
- biased

It is important to have multiple criteria.

General discussion

- Variability of effluent should be considered
- Recent failures should be weighted more then later failures
- There is concern that RP can be triggered if no failure has occurred
- TRE can't be seen if no toxicity is present
- Can a three-year history be used instead?
- The variability of the test is too much to apply statistics.
- DMR studies: Too much variability
- Will RP from EPA be used?
- What about issues over the 5 year history? Can fewer years be used?
- What about seasonal toxicity? Should a five year history be used for seasonality, or more?
- Should be careful about CI for failures, especially small sample sets.
- Different statistics should be used, for example, for example, binomial statistics instead.
- Dealing with mitigating circumstances should be written into the procedure.
- RP calculation looks at the worst case scenario.
    - It does not consider persistence and frequency and it should
    - A single data point is unreasonable.
    - Mitigating circumstances should be considered in the calculation.
    - Ionic imbalances should be addressed.
    - Already identified constituents should be excluded.
- There should be a different RP calculation for both lethal and sublethal.
- Should another period of record be used?
    - If using the EPA calculation; two years
    - If using another RP; a longer data set, such as, five years or the entire history
- Chronic sub-lethal TREs are hard to do. They are time consuming and expensive.
- EPA's model makes it difficult to NOT come up with an RP. It is a statistical calculation.
- TCEQ should consider other calculations, especially when considering something with high variability. It is estimated that, strictly according to EPA's proposed predictive reasonable potential calculation, that 80% of permits would be determined to have RP and with 1 failure alone a permittee could be assessed RP.
- There was general agreement that other models should be considered.
- At the end of three years, if there is no toxicity and the permittee has a WET limit, what happens?

- EPA's model is hard to understand. There was general agreement within the group that a decision cannot be made without understanding the model.
- The group was concerned that it is purely a number driven thing and that a discharger have enough data points. Some permits only test once a year.
- There could be a provision to do quarterly testing, and a minimum of 10 data points would be required.
- The RP's purpose is to identify a problem in order to protect the environment. If there is one spurious event over a five year period, is it really protecting anything?
- The group suggested that TCEQ address site-specific issues and use best professional judgment (BPJ).
- If there are few data points, there is a need for TCEQ staff to use BPJ.
- The group suggested that they would like a process to invalidate test results e.g. if one can identify source of toxicity and eliminate it.
- There should also be a process to validate a failure.
    - If sublethal effects can be found in controls, there is a dose response. There needs to be a way to ensure proper and consistent dose response.
- Intermittent streams get 100% critical dilution. Would these streams automatically get an RP because they are intermittent?
- The group suggested that the 90% rule be implemented and that it apply to intermittent streams as well.
- There was concern about the magnitude of the failures, that is, whether they are minor or major failures?

b) If not, what would you propose instead to determine RP?

More data points need to be considered. The statistical analysis needs to be looked at and mitigating circumstances.

Maybe: Different types of statistics should be looked at, since sublethality is so variable. We need a better understanding of the EPA model.

Pro: if toxicity appears without an external explanation, you technically have a reasonable potential for toxicity.
Con: Number of data points used.

Is this protective of the environment? Do permit limits on spurious results protect anything?

Keep the component of best professional judgment, such as number of tests, dilution magnitude and the number of failures.

**There needs to be a clear process to invalidate bad test results, and a process to validate a sublethal failure. Why? Because these are delicate organisms; some sublethal effects are not wholly uncommon in nature. How? Need to check lab sheets, control charts, scientific analysis, and ensure proper dose response curve.

**The 90% critical dilution exception must apply to protect dischargers into small/intermittent streams.

Use frequency and magnitude, need enough data points to do statistical analysis

Option for RP: if a permittee fails the first test, retest to determine if the toxicity is persistent and increase the frequency. If toxicity is persistent then enter a TRE. Sublethal is more of a concern.

Lial Tischler's binomial approach was mentioned as a viable alternative. Other factors should include frequency and magnitude. There needs to be adequate data on which to base the evaluation. WET limits shouldn't be considered for a new permittee. P. Glass existing permit compliance protocols now are like a RP, if a permittee fails the first test then they test again to determine the presence of persistent toxicity. If toxicity is shown, then the permittee should conduct a TRE.

c) Are there additional factors you would like considered in assessing RP?

The group indicated that they were unclear on the number of sublethal tests that would be needed, the number of failures, and the magnitude of those failures.

They were also unclear whether there would be a difference in stream values and source water issues. For example, Receiving water vs. Source water, could a separate control be used for both.

The group indicated that for sublethal effects TCEQ needed to consider ion imbalance. Use only valid tests. If a previous TRE was successful in identifying a toxicant, use only test results since then as part of the RP analysis. If EPA is only using a single point, then use only the previous two years of data. If using the statistical approach, use a lot more data.

Take source water issues into account, some tests will need multiple controls

The group suggested that 50% effluent could be a trigger for permit limits for sublethal? A greater percentage would be too random.

d) Should different factors contributing to the assessment of RP be weighted differently?

Treat lethal and sublethal effects differently. One group needed more information.

2. Should RP lead directly to WET limits (lethal and/or sublethal) without a TRE being performed? If not, what would you propose as an alternative?

No; sublethal should go to a TRE; lethal should maintain the status quo.

A qualified no from M. Hess; the major issue is that an upfront WET limit does not take enough information into account and a permittee will be doing the TRE anyway. A WET limit may appear more protective but is it justified? TREs will narrow the field of possibilities for sources of toxicity. A three year compliance period may not be enough time when the permittee does not have a known contaminant. Permittees believe EPA thinks they will not try to fix the WET problems without WET limits; however, the permittees disagree with this assumption. Others in the group wondered what is wrong with the current process. EPA wants proactive action, they don't want to wait for a failure but address the potential during the permitting process.

There could be a TRE placed in the permit if the RP is determined during the permit review. A TRE results in a better product because you can potentially identify the problem. However, there should be some flexibility for TCEQ.

A WET limit does not do anything to provide incentives to identify the problems.

Upfront WET limits do not help find the issue. A TRE is the tool to find and fix the problem. WET limits are different than chemical specific limits

- One group discussed wanting the opportunity to do a TRE before getting a WET limit. A compliance limit would work well with a known contaminant. A WET limit with a three year compliance period may not provide enough time to determine toxicity.
- EPA believes permittees won't take the initiative to take TRE steps unless there is a WET limit. This is not true in many cases.
- If a permittee has toxicity issues that aren't getting resolved then a TRE may be appropriate.
- EPA is trying to be proactive with RPs, by using them before a failure, not reactive.
- With RPs, a discharger doesn't know what to look for. It would be the same as doing a TRE with the "joy" of enforcement.
- The benefit of an RP is that it may speed a facility taking action.
- A WET limit is more proactive but not necessarily justified.

3. What would be appropriate justification for removing effective WET limits?

If a permittee can identify and remove toxicity; or if there was monitoring data for 3-4 years without problems.

Would WET limit removal constitute backsliding? How would the antibacksliding clause be avoided? The WET limit removal process would need to be clearly documented, even in the permit. Can permittees see the actual RP analyzer model/spreadsheet?

WET limits should be able to be removed if a particular chemical is identified and can be controlled; also if a BMP may be used then a WET limit should be removed. What does compliance history mean? Currently, after one year of quarterly testing, the permittee may reduce the testing frequency to semi-annual or annual depending on the species.

Possible alternatives: The WET limit may be removed if

- the permittee can show that no reasonable potential exists any longer or
- after one year of monthly testing (maybe even 2-3 years) depending on the discharger.

There should be different standards for release from a WET limit for those who were assessed a WET limit upfront just from a RP determination as opposed to releasing a permittee from a WET limit post TRE.

The group suggested that

- three years worth of positive results (no lethal or sublethal) or the permittee has removed what is causing the toxicity (e.g. no longer using a metal that caused the toxicity).
- If the permittee fixes the problem, then they get a chemical specific limit. Monitor for 3 years and no toxicity issues.
- There was some skepticism expressed on whether or not the WET limit would be removed even though there was the "promise" that the limits will be removed after five years of no reported significant toxicity. If sufficient clean data was reported, the group suggested that the WET limit should be removed. The wet limit should also be removed:
    - if a toxicant is actually identified,
    - If the chemical has been identified and is controlled.
    - If there is a history of compliance
    - There are no failures based on the frequency of testing.
- 1 year of quarterly testing can move to annual with no failures.
- Some in the group were not totally comfortable with the above. They felt it depended on the type of discharger. Others suggested one year of quarterly testing if a typical discharger. If a variable discharger, more years worth of data would be needed.

Sublethal Toxicity Issues

1. Do you think sublethal TREs are ever appropriate?

a) Why or why not?

Yes, if repeated issues are real at the 99% CI or if the critical dilution is 5% with some conditions or procedures or evidence of evaluation.

Yes, if you have a validated failure of an appropriate magnitude (with an appropriate dose response and the test is repeatable). The group questioned how many sublethal TREs have been (1) started and (2) successful in identifying the problem?

Yes, for persistent sublethal toxicity with a large enough signal.

Yes, only if you have repeated real issues or a low critical dilution and if you realize that they are different than lethal TRE's. Low critical dilution- higher dilutions are harder to find problems with than low dilutions. The group asked for a definition of what is a failure? Failures trigger seven monthly retest samples.

Yes, from a permit protective basis.

Based on a dose response and the test's repeatable which results in a validated failure, then Yes. Sublethal TREs are also appropriate assuming the magnitude is at a sufficient percent NOEC. If not, the group proposed just to increase the frequency of monitoring. The group questioned how many sublethal TREs have been successful at identifying the toxicant? Is sublethal toxicity at high magnitude easier to ID?

b) If not, what would you propose instead to address sublethal toxicity?

Increase monitoring frequency.

2. TRE triggers

a) How many sublethal failures should a facility have before a TRE is triggered?

For a TRE to be triggered

- there should be a failure rate of 50% of a minimum of 8 samples at the 99% CI.
- use 2 failures out of 3 attempts if the failures have been validated.
- use BPJ if results are ambiguous and then go to 3 of 4 or something more appropriate.
- The group suggested
    - an initial failure rate of 2 or more out of 3 failures
    - 50% with a minimum size of 8 at 99% CI

Peggy Glass: There needs to be a way out of a TRE if the data shows that toxicity has disappeared; such as using a non-punitive escape clause similar to what is currently in the Implementation Procedures for lethal TREs.

b) Over what time period?

The group suggested

- Over a two year period or until 4 failures occurred.
- Monthly
- 4 failures
- Consecutive failures are an issue that should be addressed.

c) What NOEC would be appropriate to use to trigger a TRE?

The group comments included:

- An appropriate NOEC to trigger a sublethal TRE would be 70%.
- It depends.
- NOEC = 50%? or IC25
- There needs to be a non-punitive escape clause.
- A TRE should be triggered by a 70% NOEC
- Use an IC25 or an NOEC = to 50%, the use of an IC25 is recommended because it smoothes out the curve of the data.

d) Could anything else trigger a TRE?

Yes, see a) and b) above.

General Discussion

- Currently for lethality, a TRE is initiated when there is a failure and the permittee fails a retest. This assumes that the tests are valid with a reasonable dose response curve.
- Louisiana uses 2 of 3 validated, repeatable, test failures before initiating a sublethal TRE. Initiating a sublethal TRE also depends on the dilution at which the failure occurred.
- Is failing 2 of 3 tests a valid enough approach to indicate a legitimate failure? Perhaps it would be better to use 2 of 3 failures with BPJ to ensure a real sublethal toxicity problem.

3. If a TRE is performed and a toxicant is not identified, what should happen next?

The group suggested:

- Use the same procedure as now
- If a TRE is initiated but toxicity never reappears after one year (cessation of lethality) then no enforcement and no WET limit, the regulatory agency should exercise BPJ.
- WET limit or continued high frequency testing; possibly extend the TRE indefinitely?
- Consider on a case-by-case basis.
- TRE for life or have a WET limit or chemical specific limit.
- Ion imbalance addressed in TRE.
- Continuing to fail with a WET limit in place?
- WET limits,
- continued high frequency testing,
- extend the TRE indefinitely,
- be aware that it can be harder for cities than for industries to identify a toxicant, sort of like looking for a needle in a haystack (from permittees experience).

- Pat Radloff expressed concern with letting unidentified toxicity "go." Perhaps the regulatory agency should exercise BPJ on a case-by-case basis.
- 12 months – no permit repercussions if cessation of toxicity, lethal and sublethal.
- A facility with reasonable variability, constant stream variability

Break out group results - sublethal - additional comments aside from individual group notes:

Bruce Huther questioned what if the toxicant is there, but the lab can't identify it? What if the toxicant disappears? Will there be a cessation of lethality?

Changes in site-specific uses and criteria in classified and unclassified waters – Charles Bayer

Mr. Bayer presented four handouts. Two concerned potential changes to Appendix A, which lists uses and criteria for classified segments, and two for potential changes to Appendix D, which are smaller unclassified streams. He also presented information on a linear regression model for Black Cypress Bayou that explains approximately 95% of the observed variation in 24-hr average dissolved oxygen concentrations in the bayou. Black Cypress Bayou is a least-impacted ecoregion reference stream that experiences naturally occurring low dissolved oxygen concentrations during low stream flow periods in the summer. Despite the periods of low dissolved oxygen concentrations, the bayou maintains excellent fish community characteristics. The linear regression model also explains the well documented low dissolved oxygen concentrations in other East Texas low-gradient streams and bayous. Mr. Bayer also discussed TDS changes in some reservoirs that are dependent on the amount of flow coming into the reservoir and the water level in the reservoir.

Next meeting and topics

August 7 meeting has been moved to September 6. The topics to cover in this meeting are:

1. Additional questions for 307.8 and .9 and standards attainment and assessment
2. Nutrient criteria implementation and assessment
3. Follow up on recreational uses and criteria

The next workgroup would be scheduled late October or early November and the topics to cover are:

1. WET follow-up
2. Potential changes to 307.2 and 307.5

Another meeting could be held in December to cover the IPs.

Suggestions for the next work group meeting:

- Questions be more specific for a more focused discussion.
- More time was needed to work through the questions.
- The background for the WET was well put together.
- For the breakout groups a staff person available to answer questions would be helpful to keep the group on track.
- May be possible to better divide out the people based on the groups they are representing. Some groups were tweaked with individuals moving to another group.
- Would like the handouts at least a week ahead and with links to information that would be helpful for those with little background. The agenda needs to be available even earlier.
- e-mail notification when things are placed on the webpage.
- Provide more technical information with criteria numbers.
- Small groups worked very well.
- Topics are a little too technical. Provide more information.
- NOEC explanation was needed.
- Separate upfront WET and TRE based WET limits
- Post reference material on the web (i.e. 2004 EPA guidance document)
- Post the agenda as far in advance as possible.