Fraser River Action Plan

Evaluation of the PEEP Index and Recommended Toxicity Tests for the Fraser River Basin

DOE FRAP 1993-09

Environment Canada Environnement Canada

EVALUATION OF THE PEEP INDEX AND RECOMMENDED TOXICITY TESTS FOR THE FRASER RIVER BASIN

DOE FRAP 1993-09

Prepared for:

Environment Canada
Environmental Protection
Fraser Pollution Abatement Office
224 West Esplanade
North Vancouver, B.C.
V7M 3H7

Prepared by:

Environmental Management Associates

and

Hydroqual Laboratories Ltd.
Calgary, Alberta

M a y 1 9 9 3

ACKNOWLEDGEMENT

The Fraser Pollution Abatement Office recognizes the efforts of the Centre Saint-Laurent in Montreal, especially Dr. Norm Bermingham, in the development of the PEEP Index.

This consultant's report was funded by Environment Canada under the Fraser River Action Plan through its Fraser Pollution Abatement Office. Environment Canada is not responsible for the content of this report but has made it available for public distribution.

Any comments regarding this report should be forwarded to:

Fraser Pollution Abatement Office
Environment Canada
224 West Esplanade
North Vancouver, B.C.
V7M 3H7

EXECUTIVE SUMMARY

The Fraser Pollution Abatement Office (FPAO) has a mandate to significantly reduce discharges of environmentally disruptive effluents to the Fraser River Basin. To accomplish this task, the FPAO needs a reliable toxicity testing program to rank and compare waste discharges throughout the basin and to monitor the progress of pollution abatement measures. One approach being considered for these purposes is the Potential Ecotoxic Effects Probe (PEEP) which was developed for testing and ranking effluents in the St. Lawrence River Basin.

PEEP uses results from four small-volume bioassays, selected to incorporate a range of trophic levels and a variety of acute and chronic endpoints. Three of the tests are repeated before and after the sample is subjected to a 5-d stimulated aerobic biodegradation test intended to render a measure of persistence of the effluent's toxicity. Threshold effects levels from each test endpoint, before and after biodegradation, are then entered into the PEEP formula which collapses them into a single unitless number, taking into account the strength and breadth of the toxic response and the volume of discharge.

The applicability of the PEEP index approach to the Fraser River basin was evaluated using results of a pilot wastewater characterization study undertaken in April and May 1992. Three representative effluent types were sampled: (1) primary treated domestic sewage from greater Vancouver (Annacis Island Wastewater Treatment Plant); (2) final effluent from a bleached Kraft pulp mill (Northwood Pulp and Timber Ltd., Prince George); and (3) urban runoff from a Vancouver storm sewer. A flow-weighted, composite sample from each site was tested for acute toxicity with rainbow trout, *Daphnia* and the bacterial luminescence (Microtox) test by Environment Canada (North Vancouver). Analex Inc. (Laval, Québec) also tested the effluents and computed the PEEP index for each.

Although the idea of a single index to facilitate comparisons among effluents is attractive, the PEEP index was judged unsuitable for the Fraser River Basin in its present form. It was felt that it under-emphasizes acute toxicity while favouring a screening test of genotoxicity (SOS-Chromotest). In addition, the biodegradation test is nonstandard and untested, and there are problems with the structure of the mathematical formula used to derive the index.

A battery of bioassays is recommended for use in the Fraser River Basin, comprising these tests: (1) bacterial luminescence (Microtox); (2) algal growth, (3) survival and reproduction of *Ceriodaphnia* and (4) acute mortality of rainbow trout. The battery satisfies the key requirements that a diversity of taxonomic groups and trophic levels and a range of acute and chronic endpoints, should be represented. All four tests are widely accepted and are supported by Environment Canada protocols. The test battery is identical to that used for the PEEP index, except that the SOS-Chromotest is replaced with the acute trout test.

A simplified version of the PEEP index was derived that adapts it to the test battery recommended for the Fraser River. The new index includes only average toxicity and discharge, and the mean is weighted toward acute responses on the ground that these are more serious than chronic toxicity. Several other alternatives are possible, including using a dilution ratio instead of discharge, and a graphical approach that plots average toxicity against effluent volume. A quick, cheap, screening survey of all the effluents, followed by detailed assessment of the most toxic, is another approach to the problem which deserves consideration.

TABLE OF CONTENTS

TABLE OF CONTENTS CONCLUDED

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Under the auspices of the Fraser River Action Plan, the Fraser River Pollution Abatement Office (FPAO) is charged with identifying sources of pollution, characterizing wastewater discharges, and finding ways to reduce pollution entering the Fraser River drainage. The FPAO has a goal of reducing by 50% discharges of "environmentally disruptive" substances to the Fraser River Basin by 1997. To accomplish this task, and to keep interested parties and the general public apprised of progress toward reducing pollutant loads, the FPAO has need of a reliable toxicity testing program to compare waste discharges and to assign priorities for abatement. The program must be quantitative, scientifically defensible, directly relevant to the Fraser River ecosystem, and applicable to widely different effluent types. Toxicity testing would be used both to rank effluents with respect to their environmentally disruptive character, and to monitor the progress of pollution abatement actions.

As a first step in the development of a toxicity testing framework for the Basin, the FPAO has undertaken a pilot wastewater characterization study, using three effluent sources that are both common in the Basin and representative of the range of major effluent types (Thomas and Dwernychuk 1992). Single flow-weighted, composite samples were taken from:

(1) municipal wastewater (primary treatment effluent)
(2) pulp mill effluent (bleached Kraft, secondary treatment)
(3) urban runoff.

All three samples were subjected to chemical analysis to identify their major components, and toxicological testing, involving standard acute bioassays (trout, *Daphnia*, and bacterial bioluminescence) and a novel index called the Potential Ecotoxic Effects Probe (PEEP), developed by the Centre Saint-Laurent in Montréal. PEEP is itself a combination of four toxicity tests incorporating both acute and chronic effects, and lethal, sublethal and genotoxic endpoints (Costan et al. 1992). PEEP is a framework for effluent assessment used in the St. Lawrence Action Plan. It provides an index of toxic potential by integrating results from a set of bioassays, considering the strength and specificity of the toxicity and the volume of the effluent. The PEEP index appears to have worked well in assessing effluents on the St. Lawrence River (Costan et al. 1992), so its applicability to the Fraser River Basin warrants consideration.

The present study had two interrelated objectives: first, to evaluate the PEEP index for environmental relevance and applicability to the Fraser River Basin, and second, to recommend a test or array of tests best suited to assaying toxicity of wastewater discharges entering the Fraser River Basin. The evaluation draws on results of the pilot wastewater characterization study for chemical profiles of three typical effluents, and for comparisons of toxicity determined by PEEP and by standard acute bioassays.

To be useful, the recommended toxicity test program must be applicable to all kinds of effluent discharges present in the Basin, and hence must be flexible and robust. The program must also provide results that are quantitative and reproducible, so that it can be used to track changes in toxicity of effluents and monitor the success of abatement techniques. Therefore, a set of rigorous guidelines for the implementation of the recommended tests is appended to this report, to help standardize methods of sample collection and storage, and to facilitate quality control. The report also includes a brief discussion of alternative ideas and approaches to assessing the environmentally disruptive effects of effluents in the Fraser River Basin.

2.0 EVALUATION OF PEEP

2.1 DEFINITION OF PEEP INDEX

The Potential Ecotoxic Effects Probe (PEEP) was developed to help assess and compare the toxic potential of industrial effluents discharged to the St. Lawrence River, Québec, under a river basin management program similar to the Fraser River Action Plan. The index is based on results of four small-volume screening bioassays, selected to incorporate (1) acute and chronic toxicity (2) genetic toxicity (3) a range of trophic levels. The tests are:

(1) Algal growth test with *Selenastrum capricornutum*
(2) SOS-Chromotest with *Escherichia coli*
(3) Microtox bacterial luminescence assay
(4) *Ceriodaphnia dubia* survival and reproduction assay

The effluent is first tested with all four assays. It is then subjected to a biodegradation procedure designed to simulate aerobic biological treatment of the wastewater (Costan et al. 1992). A 1-L effluent sample is first enriched with (1) a concentrated nutrient solution containing $MgSO_4$, $CaCl_2$, and $FeCl_2$, and (2) a phosphate buffer solution, pH 7.2, which also contains NH_4Cl. The sample is inoculated with a 5% commercial mixture of microbial seed and incubated for 5 d, in darkness at room temperature, with constant gentle aeration. After incubation, the effluent sample is re-tested with three of the four bioassays, excluding the expensive *Ceriodaphnia* test. The incubation procedure is intended to render a measure of persistence of effluent toxicity in the face of microbial degradation, although volatile constituents will also be lost.

Results of all tests are reported as threshold effect concentrations (TEC) instead of the more familiar median lethal concentration (LC50) or median effective concentration (EC50). The TEC is the geometric mean of the lowest concentration in the effluent dilution series at which adverse effects on the test organisms were observed (Lowest Observed Effects Concentration, or LOEC) and the highest effluent concentration that produced no effect (NOEC). For example, if the NOEC = 12.5% effluent and the LOEC = 25%, then TEC = $(25 \times 12.5)^{1/2}$ = 17.7%. As the name implies, the TEC is taken to estimate the effluent concentration at which toxicity begins.

To render toxicity data into numbers that can be added together in a formula, the TEC are translated into toxic units (TU), calculated as TU = 100/TEC (Sprague and Ramsey 1965). For example, an effluent that has a TEC for *Ceriodaphnia* reproduction of 25% contains 4 chronic TU; a TEC of 5% in the *Ceriodaphnia* survival test represents 20 acute TU. The PEEP is calculated by adding the TU of all endpoints, before and after biodegradation, according to the following formula:

$$P = \log_{10}\left[1 + n\left(\frac{\sum_{i=1}^{N} T_i}{N}\right)Q\right]$$

where:

P = PEEP value
n = number of bioassay endpoints exhibiting toxic responses
N = maximum number of obtainable toxic responses (endpoints)
T_i = Toxicity, in TU, according to a given bioassay endpoint, before or after biodegradation
Q = effluent flow in m^3/h.

The four tests listed earlier produce a total of six endpoints (the *Ceriodaphnia* and SOS-Chromotests each have double endpoints), and without *Ceriodaphnia* there are four endpoints after biodegradation. The value of *N* is therefore 10. The PEEP formula sums the TU from each endpoint and divides by *N* to derive an "average toxicity" of the effluent. Average toxicity is then multiplied by *n* to account for the breadth of toxic response, and by discharge to compute the total hourly number of toxic units entering the river. The logarithmic transformation of the result (adding 1 merely ensures that the log is computable) scales results to normally fall between 0 and 10 (Costan et al. 1992). Thus, PEEP collapses results of seven toxicity assays into a single number.

2.2 THE PILOT WASTEWATER CHARACTERIZATION STUDY

2.2.1 Effluent Sources and Methods

A pilot study to characterize a range of wastewater effluents in the Fraser River Basin was conducted in April and May, 1992. Three effluents were selected as representative of the range of common wastewater types entering the basin:

(1) primary treated domestic sewage from greater Vancouver (Annacis Island Wastewater Treatment Plant);
(2) final treated effluent from a bleached Kraft pulp mill (Northwood Pulp and Timber Ltd., Prince George); and
(3) urban runoff from a Vancouver storm sewer.

Annacis Island, the second largest of four wastewater treatment plants in the Greater Vancouver Regional District (GVRD), discharges primary treated municipal sewage into the main arm (Annieville Channel) of the Fraser River between North Delta and New Westminster. The Northwood Pulp and Timber Ltd. mill produces chlorine-bleached, Kraft pulp and discharges wastewater to the Fraser River about 10 km upstream from Prince George. Street runoff was sampled from a GVRD storm sewer located in the upper basin of Still Creek, near the eastern boundary of Vancouver. The sewer receives runoff from both commercial and residential areas, and eventually drains into the North Arm of the Fraser River.

A full description of sampling techniques is contained in a previous data report (Thomas and Dwernychuk 1992). Briefly, each of the three sites was sampled at two-hour or three-hour intervals for 18 or 24 hours, and discrete samples were combined into a single, flow-proportionate composite for each site. Chemical profiles of each effluent were produced by analysing the composite samples for physical parameters (pH, suspended solids), nutrients, oil and grease, biochemical oxygen demand (BOD), metals, and organics.

Split samples from each effluent source were also sent to two Environment Canada biological laboratories for toxicity testing. Portions of a well-mixed composite sample were sent to Environment Canada's laboratory in North Vancouver for acute toxicity tests and to Analex Inc., Laval, Quebéc, for assays required by the PEEP index. Environment Canada tested each effluent with the following assays, following standard protocols:

(1) Rainbow trout 96-h LC50 (acute)
(2) *Daphnia magna*, 48-h LC50 (acute)
(3) Bacterial luminescence (Microtox) assay (5 and 15 min).

2.2.2 Results

Results of toxicity tests on the three effluents are summarized in Table 1. Results for chronic tests are given as TEC and LOEC, the endpoints reported by Analex. Chemical profiles of the effluents may be found in Thomas and Dwernychuk (1992). In summary, sewage effluent from Annacis Island was acutely toxic to trout and *Daphnia*, and chronically toxic to *Ceriodaphnia*, probably because of ammonia and metals. Pulp mill effluent showed weak acute toxicity presumably from resin acids, and severe suppression of algal growth from unknown sources. Storm sewer runoff slightly inhibited reproduction in *Ceriodaphnia*, for which metals are again seen as probably responsible. All three effluents exhibited genotoxicity to bacteria, but that in pulpmill effluent was especially severe; a variety of compounds or elements could be causing the genotoxicity.

TABLE 1
RESULTS OF TOXICITY ASSAYS ON THREE WASTEWATERS,
IN PERCENT OF FULL STRENGTH EFFLUENT

TEST	LABORATORY	ENDPOINT	ANNACIS ISLAND	NORTHWOOD PULP	GVRD STORM SEWER
Trout	EC	LC50	54.6	>100[2]	>100
Daphnia	EC	LC50	100	>100[3]	>100[3]
Ceriodaphnia					
Survival	Analex[4]	LOEC	>100[5]	>100	>100
Reproduction		LOEC TEC	25 17.5	25 17.5	100 71.4
Algal Growth	Analex[4]	LOEC	>100[5]	0.33	>100
		TEC	>100	0.22	>100
SOS Chromotest	Analex[4]				
-S9		LOEC	10	1.56	10
		TEC	4.4	1.11	4.4
+S9		LOEC	10	3.1	>50
		TEC	4.4	2.2	>50
Microtox	EC	EC50	7	>100	>100
	Analex[4]	LOEC	25	50	>50
		TEC	17.5	35.7	>50

1 EC, Environment Canada, Vancouver, B.C.; Analex, Analex Inc., Laval, Quebec
2 20% mortality in 100% effluent.
3 Zero mortality in 100% effluent.
4 LOEC and TEC back-calculated from toxic units reported by Analex (1992).
5 No adverse observed effect in 100% effluent. Therefore LOEC and TEC given as >100%.

2.3 COMPARISON OF TOXICITY DATA

The problem of assessing the applicability of PEEP to the Fraser River Basin can be decomposed into two questions. First, are the toxicity assays that subtend the index appropriate? Second, is the formula for combining test results appropriate? The questions are not entirely independent because the formula may work better for some test batteries than for others. The chronic test on *Ceriodaphnia* is treated differently than other tests, for instance, and tests with double endpoints have a greater influence than single-endpoint tests. Notwithstanding, the PEEP index itself is simply a way of reducing test results to a single number, and should be applicable to any battery of tests from which effects thresholds can be derived. It follows, therefore, that the PEEP formula can be applied to results of the acute toxicity tests done by Environment Canada for the pilot wastewater study (Table 1).

This provides an opportunity to test the accuracy of PEEP at predicting toxicity as measured by assays other than those intended to be used in its computation. PEEP values for the three wastewaters (Analex 1992) are listed in Table 2 alongside index values calculated by substituting acute toxicity results with rainbow trout, *Daphnia* and Microtox into the formula. For the latter calculations $N = 3$, and there are no data after biodegradation but the formulae were otherwise applied identically. The numeric values of the PEEP index and the acute toxicity index are on different scales, but they should ideally produce identical relative scores and ranks for the three effluents.

TABLE 2

COMPARISON OF PEEP INDICES FOR THREE WASTEWATERS WITH AN INDEX BASED ON ACUTE TOXICITY

EFFLUENT SOURCE	VOLUME (m^3/h)	PEEP		ACUTE TOXICITY*	
		INDEX	RANK	INDEX	RANK
Annacis Island	17328	6.0	2	5.7	1
Northwood Pulp	5204	6.8	1	3.3	2
GVRD Storm Sewer	396	3.3	3	0	3

* based on trout, *Daphnia*, and Microtox, as assayed by Environment Canada.

In reality, the PEEP index disagrees with the acute toxicity index both in the ranking of the effluents and in their relative strengths. PEEP ranks Northwood Pulp first (i.e., most toxic) because of the strong responses in the algal growth test and the SOS-chromotest (Table 1) coupled with a substantial effluent volume. It ranks Annacis Island effluent a close second. In contrast, on the basis of acute toxicity Annacis Island would be ranked first, and Northwood Pulp would fall in far second place. Both indices rank storm sewer water third (Table 1). It cannot be claimed that the acute toxicity index is the more accurate; rather, neither index is sufficiently comprehensive.

The difference in ranks arises because the various effluents differ both in the strength and the nature of their toxicities: Northwood Pulp effluent is chronically toxic, while Annacis Island toxicity is largely acute. The PEEP index weighs heavily against acute toxicity in favour of chronic toxicity and genotoxicity assays, and therefore underrates acutely toxic effluents. The only PEEP test with lethality as an endpoint is survival of *Ceriodaphnia*, which is not sensitive to ammonia, and the only truly acute test is Microtox, which, as a screening assay, is not entirely analogous to tests with *Daphnia* or rainbow trout.

The omission of acute toxicity assays is not an oversight. In the development of the index for the St. Lawrence Action Plan, Costan et al. (1992, p. 19), posited that chronic effects are more important in the St. Lawrence River and therefore should be heavily favoured. That decision, however, does not appear to apply to the Fraser River Basin. Chronic toxicity is certainly the more widespread problem, especially now that most industrial and municipal effluents undergo some kind of treatment to remove gross organic contaminants, BOD, and other conspicuous toxicants. Absence of severe acute lethality is now a regulatory requirement for industries such as pulp and paper. Acutely toxic effects may be rapidly attenuated by dilution in the river, and many acutely toxic constituents (e.g., ammonia) are rapidly degraded or volatized and therefore do not persist in the environment. Hence various forms of chronic toxicity are by far the more prevalent problem in wastewater discharges of all kinds. However, as the results of the pilot study demonstrate, acute toxicity of effluents has not been eliminated in the Fraser River Basin, and where it does occur, acute toxicity should weigh heavily in any estimation of environmentally disruptive effects.

There are numerous grounds for this criterion:

1. Acute toxicity represents death of organisms and therefore is a more severe form of toxicity than chronic sublethal effects, which lead only to impaired physiological function.

2. Acute toxicity represents a more serious and immediate threat to populations because individuals and their potential fecundity are lost.

3. From the above, it follows that acute toxicity has a greater potential to impair ecosystem function through loss of all or part of a taxon, guild or trophic level.

4. Acute toxicity will often mask chronic effects, especially in complex mixtures. In the *Ceriodaphnia* assay, for example, it is sometimes observed that one component of a mixture causes acute lethality and another apparently reduces reproductive success, but the threshold for the latter effect is above the threshold where adult mortality begins, confounding the assay. Hence, the strength of chronic toxicity cannot be fully known until acute toxicity is removed.

5. Acute toxicity of an effluent is usually more easily cured, and therefore is more efficient to attack, than chronic toxicity.

6. Acute toxicity is easily understood by the general public.

Therefore, in a comparison of two toxic effluents like Annacis Island and Northwood Pulp, of which one exerts acute toxicity (i.e., lethality) and the other chronic, all other things being equal, the acutely toxic effluent always should be ranked higher in terms of potential for environmental degradation and priority for treatment, just the reverse of the priority given by PEEP. Certainly in the Fraser River Action Plan, effluents causing mortality to trout must not be ignored if the program is to gain public acceptance. On the other hand, as regulations become more restrictive and wastewater treatment improves, acute toxicity is likely to become less and less common; chronic toxicity, in various manifestations, will be the continuing problem. To accommodate both these needs, the ideal toxicity test battery would include several chronic endpoints to only one or a few acute endpoints, but would assign greater weight to acute toxicity results.

There is a second problem with the weights given to different tests in PEEP. Two of the tests have double endpoints: the *Ceriodaphnia* survival and reproduction assay and the SOS-Chromotest. For reasons of cost, the *Ceriodaphnia* test is excluded from the assessments done after biodegradation (Analex 1992). While the cost issue cannot be ignored, especially since the *Ceriodaphnia* assay is one of the most expensive tests, its exclusion in PEEP creates a bias away from cladocerans and in favour of other assays. This is the more regrettable because the *Ceriodaphnia* assay is the only test of lethality, and indeed the only animal test, included in PEEP.

The problem with SOS-Chromotest is more serious. The appropriateness of this test in the battery is considered in detail later (see Section 2.4.4). The issue here is the way in which results with and without rat liver enzymes are considered. Some organic compounds, notably many PAHs, must be metabolised before they become genotoxic. To detect indirect mutagens, the SOS-Cromotest can be run with addition of rat liver extract (S9) in which enzymes have been induced by exposure to Arocolor, a potent PCB. Running the test without S9 detects direct-

acting mutagens (that may or may not be rendered inactive when S9 is added). But the addition of S9 is only a means of distinguishing between two modes of mutagenic action. Thus, results with S9 and without S9 are not two separate endpoints like survival and fecundity in the *Ceriodaphnia* test.

The appropriate manner to compare effluents with the SOS-Chromotest would be to take the lower of the two TECs in each test and discard the other. Yet the PEEP index adds both results into the total as if they were two endpoints and thus creates a bias in favour of genotoxic effluents. This bias is worsened because the SOS-Chromotest is one of three tests run after biodegradation, while *Ceriodaphnia* is deleted. Hence, of the ten endpoints that measure toxicity in PEEP, four are from SOS-Chromotest. The index is heavily weighted toward potential genotoxicity.

The interim conclusion from this analysis is that the PEEP index, as presently formulated, should not be used to rank effluents in the Fraser River Basin. The index is designed for assessment of chronic toxicity and genotoxicity and does not adequately rank acutely toxic effluents. Both the battery of tests used for PEEP and the formula for calculation of the index need to be carefully evaluated if this approach is to be modified for application to the Fraser River Basin. Since the PEEP formula, or a derivative, can be applied to any set of comparable endpoints, the selection of appropriate tests for the Fraser River Basin is the next task. Then the PEEP formula will be examined in more detail (Section 3).

2.4 SELECTION OF TOXICITY TESTS FOR THE FRASER RIVER

Selection of a toxicity test or tests for the Fraser River must take into account the goals of the FPAO, the intended uses of the toxicity testing (ranking effluents for abatement actions and monitoring changes in effluent toxicity) and environmental conditions in the Fraser River. Ordinary application of the toxicity test or tests would be complementary to chemical analysis of the effluent under consideration. Thus simultaneous chemical data would be available to assist interpretation of the toxicity test results, and to verify that changes in toxicity correlate with changes in levels of purported toxic components.

2.4.1 Single Test versus Battery of Tests

The first matter to resolve is the question of whether a single test, or a test battery, would be most effective. A single test has several attractions. It is simple to apply and straight-forward to compare and there is a significant cost savings compared with a multi-test battery, especially where many effluents are to be compared. However, there are a number of arguments against a single test.

Results of the pilot study (Table 1) demonstrate that effluents discharging to the Fraser River have widely different toxicity profiles. None of the six single assays used in the pilot study

successfully diagnosed the strength of toxicity of all kinds in all three effluents, although the *Ceriodaphnia* test came closest. To highlight the most striking example, nothing in the acute toxicity data with trout, *Daphnia*, or Microtox suggested the powerful inhibition of algal growth by pulpmill effluent. Conversely, none of the chronic tests suggested that Annacis Island effluent would be acutely toxic to trout. The toxicity program selected for the Fraser River must be both flexible and robust enough that it can rank effluents of dramatically different character on a linear scale, and it is evident from the pilot study that no single test can accomplish this task.

The pilot study results accord with broader toxicological research. There is a substantial body of evidence now that sensitivity to toxicants varies greatly among species, especially taxonomically dissimilar species, to the extent that effects of pollutants on aquatic ecosystems are difficult to extrapolate from tests with any single organism (Cairns 1983, 1986). Multi-species testing, using organisms from very different trophic levels and taxonomic groups, is now routinely recommended, especially for complex mixtures like municipal and industrial effluents, wherein different species may be affected by entirely different groups of toxicants (van Straalen and Denneman 1991, USEPA 1991). Multi-species tests also produce a better picture of ecosystem-level effects. The conclusion here is that the toxicity testing framework for the Fraser River Basin should be based on a battery of tests.

2.4.2 Selection Criteria

From the considerations discussed above, the testing program for the Fraser River should consist of a battery of tests encompassing a broad range of (1) taxonomic groups (2) trophic levels and (3) test endpoints. Acute lethality and several chronic endpoints should be included. The ideal battery would include representatives of bacteria (decomposers), plants (primary producers), invertebrates (primary consumers) and fish (secondary consumers). For reasons of economy and practicality, the recommended toxicity test program is limited to at most five bioassays, all to be done in the laboratory.

A comprehensive review of freshwater toxicity tests for use in Canada has recently been completed by Keddy et al. (1992). They reviewed 123 aquatic bioassays, and ranked each first according to three essential criteria, which they considered absolutely necessary, and then against 12 secondary criteria describing desirable attributes in an assay. For each of the tests that met all of the essential criteria and most of the desirable criteria, they then reviewed other information on the test, including trophic level represented, reproducibility, field validation and applicability to the Canadian environment. The work of Keddy et al. lays most of the groundwork for selection of tests suitable for the Fraser River, as their criteria match those required by the FPAO. Therefore, candidate tests were selected from the list provided by Keddy et al. (1992), and then evaluated again for their suitability specifically for the Fraser River Basin.

The essential criteria of Keddy et al. are:

(1) a readily available printed method;
(2) specified reference toxicant;
(3) acceptability criteria.

Acceptable printed methods are those produced by provincial or state, national and international standards organizations, (including Environment Canada, provincial Ministries of Environment, Organization for Economic Cooperation and Development (OECD), United States Environmental Protection Agency (USEPA) and the American Society for Testing and Materials (ASTM)) as well as test protocols published in the scientific literature. Reference toxicants are chemicals of known toxicity to the test organism that are used to assess the reproducibility of an assay. To meet this criterion, a specific chemical had to be named as the reference toxicant, and the expected toxic concentration (e.g., LC50) under the described test conditions had to be given. Acceptability criteria are measurable aspects of the health of control organisms which, if exceeded, invalidate a test. For example, in the trout acute assay, results are considered invalid if more than 10% of the control fish die (Environment Canada 1990a).

The 12 desirable criteria are more technical and mostly pertain to laboratory details of execution of each test. These criteria cover: (1) test organism (must be identified to species), (2) endpoints, (3) organism selection (based on age, weight, size, etc.), (4) number or organisms and replicates, (5) observation frequency, (6) volume of test solution, (7) volume of test vessel, (8) collection, storage, and preparation of test solutions, (9) culture and handling of organisms, (10) environmental conditions during the test, (11) dilution water, nutrient solutions, and (12) statistical analysis. Those tests that survived the selection process of Keddy et al. (1992) were then scrutinized to decide their ecological relevance to the Fraser River ecosystem.

2.4.3 Recommended Toxicity Assays for the Fraser River

The recommended test battery for the Fraser River Basin comprises the following tests:

1. Bacterial luminescence with *Photobacterium phosphoreum* (Environment Canada 1991)

2. Algal growth with *Selenastrum capricornutum* (Environment Canada 1992a)

3. *Ceriodaphnia dubia* survival and reproduction (Environment Canada 1992c)

4. Rainbow trout (*Onchorynchus mykiss*) acute mortality (Environment Canada 1990a)

These four tests satisfy the requirements that the battery should contain representatives of a diversity of taxonomic groups from the four major aquatic trophic levels, and should include a range of acute and chronic endpoints (Table 3). All four tests scored highly in the assessment of Keddy et al. (1992) and all are protocols presently maintained by Environment Canada. The test battery is identical with the battery composing the PEEP index except that the SOS-Chromotest has been discarded in favour of trout. The rationale for each of the tests is briefly given next.

TABLE 3

TROPHIC LEVELS AND ENDPOINTS INCLUDED IN THE RECOMMENDED TEST BATTERY

TEST	ENDPOINT	TROPHIC LEVEL
Trout survival	acute	Predators (secondary consumers)
Ceriodaphnia - adult survival	acute	Grazers (primary consumers)
- reproduction	chronic	
Algal growth	chronic	Primary producers
Bacterial luminescence	chronic	Decomposers

1. Bacterial Luminescence Test

Bacteria are essential in all freshwater ecosystems for the decomposition of organic matter, natural and anthropogenic, and for mineralization of nutrients. In lakes and slow-moving rivers, suspended bacteria, the so-called nanoplankton, are an important food source for micro-plankton and thus form the base, along with phytoplankton, of the pelagic food-chain. Bacteria are equally important to benthic organisms which either consume them directly or in combination with detrital particles.

The bacterial luminescence test, first marketed under the Microtox trademark in 1978, is variously considered an acute sublethal test or a chronic test. The distinction is imperfect because the endpoint measured, natural light production, could be caused by cell death or just interference with normal metabolism. The test has the advantage of a well standardized protocol that has seen

wide use and is thus very familiar. It is quick (15 minutes) and requires only minute samples.

Bacterial luminescence has been widely compared with other tests (see review in Keddy et al. 1992) and while results are often contradictory, depending on the organisms, conditions and test substance, the general trend is that the bacterial luminescence assay is as sensitive as acute lethality tests with trout or *Daphnia* for organic compounds, but markedly less so for inorganics (Ankley et al. 1990). The test appears to have low sensitivity to metals (Miller et al. 1985). However, the test is consistently more sensitive than tests of microbial processes such as respiration, glucose mineralization or oxygen consumption. Reproducibility of the test is quite good, with coefficient's of variation for light inhibition of <20% (Keddy et al. 1992).

The ecological relevance of the bacterial luminescence test has been questioned on the ground that it uses a species of marine bacterium to test freshwater toxicity. The criticism is valid, but the other attractions of the test, notably its speed and simplicity and the good correlations with other assays, outweigh the disadvantages of a non-indigenous organism. *P. phosphoreum* requires salt to grow and produce light, so tests must be carried out in a 2% NaCl solution instead of pure water. Salts can modify the availability and activity of toxicants, which may explain the low sensitivity of this test to metals. A modification using sucrose instead of salt for osmotic adjustment has been proposed but has not yet been adopted.

Several other microbially based assays exist. The Toxi-chromotest uses a genetically engineered strain of *Escherichia coli*, in which cellular damage is linked to induction of the enzyme beta-galactosidase. The test uses a simple colorimetric endpoint based on mineralization of galactose. The few extant comparisons indicate the test may be considerably less sensitive than algae or *Daphnia* (Keddy et al. 1992). In addition the test protocol lacks acceptability criteria or adequate reference toxicant information. Further, *E. coli*, while not a marine bacterium, is not a truly freshwater species either, as its normal habitat is the digestive tracts of warm-blooded animals, so the problem of a non-indigenous species remains. Other bacterial tests, with *Pseudomonas putida* or *Spirillum volutans* do use freshwater organisms, but the protocols are too incomplete to be alternatives to the bacterial luminescence assay. Therefore, the bioluminescence assay with *P. phosphoreum* is recommended for the Fraser River program until a suitable replacement assay based on a freshwater microbe is available.

2. Algal Growth Test

Algae, and to a much lesser extent rooted plants, are the base of the autochthonous food-chain in rivers, and therefore of fundamental significance to the functioning of the entire ecosystem. In the slower reaches of the Fraser River mainstem, algal production in the water column and on submersed surfaces is probably substantial. In the headwaters and in faster tributaries, phytoplankton are replaced by benthic algae that perform the same function (e.g., in the

Thompson River; Bothwell et al. 1992). *Selenastrum capricornutum* is a minute, single-celled green alga that has a long history of use in toxicological testing.

The algal growth test is included in the battery because it is a well documented, widely used, and extremely sensitive test that provides an indication of chronic effects on a critical ecosystem process, primary production. The test is simple and the endpoint unambiguous. Sensitivity of the test has been repeatedly examined (reviewed in Keddy et al. 1992) and is usually comparable with tests on trout, *Daphnia* and *P. phosphoreum*. For a number of substances especially critical to plant growth, such as herbicides, the algal growth test is far more sensitive than the others. Reproducibility of the test is excellent, and the test protocol meets all the critical and desired criteria of Keddy et al. (1992).

An interesting feature of the algal growth test is the possibility of a positive response, i.e., increased growth from a non-toxic effluent that provides nitrogen and phosphorus. Robust cells with a ready supply of essential nutrients are more able to combat toxicity, so in complex effluents that contain both nutrients and toxicants, the growth inhibition will be mitigated. In strongly toxic effluents, however, the effect of nutrient enrichment is completely overwhelmed, leading to a marked inhibition of growth as was observed for the pulpmill effluent in the pilot wastewater study (Section 2.2.2).

Selenastrum capricornutum is indigenous to North America, but not Canada. Again, however, the widespread use of the test, its complete and standardized protocol, and the wealth of information accumulated on its sensitivity and behaviour weigh heavily in its favour. *S. capricornutum* is evidently as sensitive or more so than other algal species, and is otherwise a typical component of the phytoplankton of fresh waters. Hence, results of the algal growth test with *S. capricornutum* should be fully applicable to the Fraser River Basin.

There are few present alternatives to the algal growth test. A toxicity assay based on plant growth is essential to a comprehensive test battery. Many other algal species have been used as toxicity test species, notably *Chlorella vulgaris*, *Scenedesmus* spp. and a wide variety of diatoms. While some of these are undoubtedly native to Canada, and to the Fraser River basin, there would be no real gain in information from preferring a test with another species over that with *S. capricornutum*, and the test species and protocol would be less fully known. The only test involving higher plants is based on duckweed (ASTM 1991) but the species used, *Lemna gibba*, is not native to Canada and its sensitivity to toxicants is less completely known. Keddy et al. (1992) have recommended development of a bioassay using native duckweed, *L. minor*. When such a test is sufficiently developed, its inclusion along with, or in place of, the algal growth test should be considered.

3. *Ceriodaphnia* Survival and Reproduction

Cladocerans are small, free-swimming crustaceans that compose a major part of the zooplankton in most lakes, ponds and slow-flowing rivers, including the Fraser River (Northcote et al. 1976). They can at times be very numerous and exert a powerful influence over the standing crop and species composition of the phytoplankton. In addition, cladocerans are food species for many pelagic fishes, and thus form an integral link between primary producers and higher trophic levels (Wetzel 1975).

The chronic survival and reproduction test with *Ceriodaphnia dubia* fulfils all the criteria set out by Keddy et al. and it is attractive for its sensitivity and relative rapidity for a chronic test. A special advantage of the *Ceriodaphnia* test is that it simultaneously produces estimates of adult survival (chronic lethality) and fecundity (chronic sublethal toxicity) in a single test. It thus produces twice as much quantitative information about toxicant strength as do most other tests. The *Ceriodaphnia* test is also one of the few well-established protocols to use reproductive success as an endpoint. (There is also an acute test with *Ceriodaphnia* which measures mortality only).

The chronic *Ceriodaphnia* test generally is at least as sensitive as other commonly used assays (Ankley et al. 1990), and the reproduction endpoint is often extremely sensitive. This is particularly true of metals, but is also observed for many organic substances. The greater sensitivity of the reproduction endpoint was demonstrated in the pilot wastewater study, in which all three effluents reduced fecundity to some degree, two of them strongly, yet adult survival was never affected (Table 1). *Ceriodaphnia* reproduction was also one of only two endpoints out of six to respond to all three effluents. Because of its sensitivity and double endpoint, the *Ceriodaphnia* test is seeing increasing use in water and wastewater monitoring.

Reproducibility of tests with *C. dubia* falls within acceptable limits, although the reproduction endpoint tends to be more variable and there is sometimes difficulty quantifying reproductive inhibition if lethality to adults is high. The test is also one of only a few for which extensive validation studies have been completed. Results are far from invariant, but a correspondence between toxicity to the daphnid and in-stream biological effects is commonly reported (Keddy et al. 1992). For instance, toxicity to *C. dubia* and changes to the benthic invertebrate communities in 43 streams receiving sewage treatment plant effluent were highly correlated (Eagleson et al. 1990). The *Ceriodaphnia* test was one of two assays used by the USEPA to assess the utility of whole-effluent tests in the Complex Effluent Toxicity Testing Program (USEPA 1991). The test has already been used successfully to assay pulp mill effluents discharged to the Fraser River (Beak 1990).

There are a few drawbacks to the assay. Crustaceans as a group require calcium for their exoskeletons, and are therefore more abundant in hard waters. While *C. dubia* is common in

surface waters of varying calcium content, there may be problems with testing or maintaining cultures in extremely soft water. The *Ceriodaphnia* assay is also the only one for which cost is a primary consideration. A single assay with *C. dubia* costs the equivalent of 3-4 trout assays or 6-7 algal growth tests (based on average industry prices). While the cost of the assay is offset by the double endpoint, it may not be economical to include this test where funds are limited. Maintenance of *C. dubia* cultures requires considerable care, and many laboratories do not yet support colonies for routine analysis.

The best alternative to the cladoceran survival and reproduction test with *C. dubia* is the acute lethality test with *Daphnia* spp. (Environment Canada 1990b). *Daphnia magna* has been long used for toxicity testing, and protocols have now been developed for the more widespread *D. pulex* (Environment Canada 1990b). Toxicity tests with *Daphnia* meet all of the critical and desirable criteria of Keddy et al. (1992) and the test sees wide regulatory use, which requires that the protocol be firmly established and accepted. However, the standard test with *Daphnia* is a test of acute lethality, and thus does not fill the same niche in the test battery as does *Ceriodaphnia*; the acute *Daphnia* test also lacks the second endpoint of the chronic test.

Chronic reproduction tests for *Daphnia* have been developed (OECD 1991, Biesinger et al. 1987) but the long duration of these tests (3 weeks, versus about 7 days for *C. dubia*) makes them unattractive for routine use. Tests with other invertebrates, particularly benthic insects, are either too new to be broadly applied, or restricted in application to the kinds of waterbodies where the test species are normally found. Hence, none of them are suitable for a river basin like the Fraser, which traverses nine biogeoclimatic zones and many different riverine and lake habitats (Birtwell et al. 1988). If the *Ceriodaphnia* test is judged unsuitable because of cost or other factors, the acute test with *Daphnia* would be the best replacement.

4. Rainbow Trout Acute Mortality

The 96-hr acute mortality test with rainbow trout is one of the longest standing and certainly the most widely used toxicity test in the country. It is included in the test battery for several reasons. First, to be complete and balanced the test battery requires an acute lethality assay, and a test using fish, to represent secondary consumers and vertebrates. Second, the rainbow trout test has been used so long and so widely that there is a huge compendium of data on its sensitivity to various toxicants and mixtures (Environment Canada 1990a); the test thus provides a baseline for comparison with results of other tests. Third, protection of salmonids is a key issue in the Fraser River basin, which supports all five species of anadromous salmon and provides irreplaceable spawning and rearing habitat (Birtwell et al. 1988). Fourth, the rainbow trout test is familiar and accepted by government, industry and the general public, an important consideration in the sphere of the FPAO.

Rainbow trout are indigenous to the Fraser River system, along with three other trout species (Scott and Crossman 1973). Toxicity test results are thus directly applicable to species in the river. The rainbow trout has become the worldwide standard cool-water fish for studying and monitoring freshwater pollution and for research in aquatic toxicology (Keddy et al. 1992). Trouts have long been known to be among the most sensitive fishes; and results of the acute lethality test tend to be highly correlated with acute tests with *Daphnia magna* and *Photobacterium phosphoreum* (Doherty 1983). Of course, given the usual interspecific variation, trout are more sensitive to some toxicants and less sensitive to others. As the pilot wastewater study showed (Table 1), trout are quite sensitive to ammonia, more so than the other species tested, and without that species in the battery an important source of fish toxicity would have been missed.

The trout test protocol is firmly standardized and reproducibility is good, usually with endpoint C.V.'s of 20% or less (Keddy et al. 1992). The Environment Canada protocol requires 40 L of effluent, and more if the test is to be repeated, a much larger volume than required by the other recommended tests.

The principal arguments against the test is that it is potentially irrelevant if waste treatment removes acute toxicity from effluents, and earlier life stages are more sensitive than fingerlings. The trout test could be replaced in the bioassay battery if acute toxicity were seldom observed, but results of the pilot study indicate that is not yet a safe assumption. An early life stage test with rainbow trout has been developed by Environment Canada (1993). The shortest option for that test is a 7-day test on newly fertilized embryos. When local laboratories have had a chance to become practised with this assay it might be effectively added to the battery in replacement of the present trout assay. Other early life-stage tests with alevins or fry require far too long (30-80 d) for routine application (APHA 1988, Environment Canada 1992c). Another alternative to the trout test is the survival and growth test with laval fathead minnows (Environment Canada 1992b). This test uses a sensitive early life stage and is at least as responsive to most toxicants as juvenile rainbow trout (ammonia is one exception) (Keddy et al. 1992). The test requires only small sample volumes and is reasonably rapid (7 days). The greatest attraction of this test is that it measures both mortality of larvae and growth (weight gain) of survivors and thus incorporates a lethal and a sub-lethal endpoint into the same test.

Unfortunately, fathead minnows are not native to the west coast (Scott and Crossman 1973), which creates an impediment to implementation of the test in British Columbia. Facilities wishing to maintain stock of fathead minnows in B.C. must obtain approval from the Transplant Committee of the Department of Fisheries and Oceans, who have prohibited introduction of this species into waters of B.C. Until this hurdle can be overcome, the 7-day embryo test is probably a better alternative to the fathead minnow test. Other measures of biological response in fish, such as mixed-function oxidase induction, while useful, are not straightforward to interpret and are not strictly toxicity tests.

2.4.4 Genotoxicity

The test battery recommended here differs from that used in PEEP by the exclusion of the SOS-Chromotest for genotoxicity. The SOS system is an inducible enzyme system that repairs lesions in bacterial DNA. The Chromotest uses a genetically engineered strain of *Escherichia coli* in which an unrelated enzyme, beta-galactosidase, has been linked to the genes producing SOS proteins. Hence, any time the cell suffers genetic damage that induces the SOS repair mechanism, beta-galactosidase will also be produced, which can be determined colorimetrically.

The test has several attractions. It is simple and fast, requiring only a few hours, and large numbers of samples can be processed by using 72-well microplates. Results are easily quantified. A parallel measurement of cellular metabolism, based on the activity of alkaline phosphatase and another colour reaction, allows genotoxic effects to be separated from cell mortality. The test thus provides two endpoints with each assay. (Note that these are not the same as genotoxicity with and without S9, considered endpoints in PEEP; see Section 2.3.)

The SOS-Chromotest does not measure mutagenic effects or carcinogenicity, but rather genotoxicity. Induction of the SOS system means only that the cell has suffered a DNA lesion and that repairs are being attempted. Most often, repairs will be successful or the cell will die; in a few instances errors during DNA repair will cause a mutation, i.e., an alteration to the genetic structure of the cell that may be heritable. In this respect the SOS-Chromotest is similar to the Mutatox test (also called the Ames test), which uses a strain of histidine-dependent *Salmonella* sp. Mutations that replace the ability to synthesize histidine result in growth of mutated colonies on histidine-free medium.

Neither the Ames test nor the SOS-Chromotest can predict carcinogenicity. They are intended as preliminary screening assays to flag compounds that have genotoxic potential, which can then be subjected to further testing. There are many layers of defenses at the cellular and tissue levels that act to prevent mutations from being transmitted or expressed. Unlike the Ames test, in which positive responses (growth on histidine-free medium) indicate that a heritable mutation has occurred, the SOS-Chromotest indicates only that the cell has incurred genetic damage, which may not have lead to a mutation at all. It is therefore an unwarranted extrapolation to construe a positive response in this test as an indication that an effluent contains carcinogens. Given the highly tentative value of results, and public unease about environmental carcinogens, we do not recommend the SOS-Chromotest for inclusion in the testing battery for the Fraser River Basin. However, research linking genotoxicity to incidence of tumours or other health effects on fish or other organisms would be useful.

2.4.5 Source of Dilution Water

Dilution water from two sources can be used when testing a waste effluent for toxicity. Ordinarily, standard tap water from the laboratory is used, but sometimes dilution water is drawn from the river, just above the effluent outfall, to better represent the environmental conditions at that point in the river. While there are good arguments in favour of using river dilution water, we do not recommend its use for initial testing of effluents in the Fraser River.

The rationale for using river water is that test results translate more easily to site-specific effects in nature. Water quality in the river, including hardness, salinity, and concentrations of suspended solids and dissolved organic matter can modify toxicity in complicated ways; the best way to account for these factors is to include them in the test. In particular, using river water allows for any background toxicity from degraded environmental conditions and residual effects from effluents upstream.

However, river dilution water creates a number of practical problems with test implementation and data analysis. First, sample size becomes extremely large, especially for large-volume tests like the trout assay. More critically, to ensure a valid test, controls must be run in duplicate, once with river water and once with laboratory water. Statistical interpretation of results becomes delicate if there is significant toxicity in the river water controls. In particular, if background toxicity is high, there is the real possibility of encountering a reverse dose-response curve, wherein toxicity increases as the effluent is successively diluted, because river water is more toxic than the effluent. What statement can be made about the effluent's toxicity in that situation? As much to the point, using river water in a program to rank wastewaters imposes a non-standard criterion, and more effectively measures cumulative effects from all sources rather than just the wastewater in question. Tests with river water dilution are very useful in more intense local studies to unravel effects of several overlapping effluent plumes or to evaluate site-specific impacts (Beak 1990) but we do not believe they are best for the initial survey of wastewater toxicity.

3.0 DEVELOPMENT OF A MODIFIED TOXICITY INDEX

3.1 OBJECTIVE

The FPAO intends to use toxicity tests to rank wastewater effluents in the Basin according to their potential for environmental disruption, and to monitor improvements in wastewater treatment. If a battery of tests is used, producing five or more estimates of toxicity (endpoints) for each effluent, then some means of collapsing these statistics into a single number for comparison would be useful. Note that the choice of toxicity tests is independent of whether results are to be combined into an index. Decisions on clean-up priorities can be made from the results of the toxicity assays alone; an index is merely a compact way of summarizing the data.

The PEEP index is one means of summarizing results from several toxicity tests, but as the pilot wastewater study demonstrated, its original configuration is not appropriate for the Fraser River Basin (See Section 2.3). The germ of the idea may be sound, however. This section presents a detailed evaluation of the PEEP index, concentrating on the logical base and practical effect of each element of the formula. That analysis leads to conclusions as to whether the PEEP index is applicable with modifications to the FPAO program, or whether a new approach should be developed.

3.2 DETAILED EVALUATION

There are seven features of the PEEP formula that must be assessed to judge the utility of the index. Each of these features represents an assumption about what information should be included in the index and how it is to be manipulated. Briefly, the features are: toxic units, threshold effects concentrations, average toxicity, number of responses, persistence of toxicity, effluent discharge, and logarithmic transformation. These are examined in turn below.

1. Use of Toxic Units

The PEEP index first translates toxicity test results into toxic units (TU). This is a sensible and well-established practice (Brown 1968, Esvelt et al. 1973, USEPA 1991). Toxic units reverse the confusing inverse relationship between toxicity and bioassay endpoints (as the LCSO increases, toxicity decreases) and normalizes results from different tests. Further, TU allow toxicity results to be manipulated mathematically like other dissolved constituents. For example, a 50% solution of an effluent containing 100 mg/L of calcium and 20 TU of toxicity will contain 50 mg/L of Ca and 10 TU. This eases comparisons of toxicity with toxicant concentrations and makes dose-response relationships clearer.

2. Use of TEC as Test Endpoint

PEEP calculates TU based on the threshold effects concentration (TEC). Other endpoints, in particular the LC50 or EC50, are more standard. Notwithstanding, the TEC is a good choice because it is the only endpoint that can be calculated under almost all circumstances. In marginally toxic effluents the EC50 is often not achieved (for example, trout lethality, Northwood Pulp), and the Inhibition Concentration (the concentration of effluent causing a specified reduction in the biological function of interest) depends upon a solid dose-response curve. The TEC requires only an LOEC to be calculable. This is an important feature in an index for wide application because it ensures that most results can be applied.

Use of TEC does create a problem when other assays, such as the rainbow trout test, are substituted into the test battery. LOEC are normally calculated as a statistic based on mean results from several replicates; the standard five-dilution trout test does not include replication. There are several solutions to this problem. The whole test could be run in triplicate, although that would be expensive and inefficient; an equivalent measure, such as the LC20, could be substituted; the lowest concentration in which fish died could be taken as the LOEC for the formula irrespective of statistics; or, given that many effluents will not be acutely toxic, a screening assay with full-strength effluent could be run first, followed by more detailed, replicated testing if needed.

3. Averaging Toxic Responses

PEEP calculates the "average toxicity" of an effluent as the simple arithmetic mean of all test endpoints, including those that showed no response. As pointed out earlier (Section 2.3), this produces a heavy bias toward the SOS-Chromotest, but that is as much a reflection of the choice of assays as of the formula. Using the simple average makes the implicit assumption that all toxic responses are equally significant. A better assumption might be to weight the average toward acute toxicity, on the ground that it signifies a more urgent problem. With the recommended battery, for example, an average calculated as

$$\text{Mean TU} = [(10 * TU_a) + (1 * TU_c)]/N$$

would ensure that effluents lethal to trout would be ranked higher than non-lethal effluents with somewhat greater chronic toxicity. Of course other weightings are possible.

Another alternative to the simple average is to replace the TU with a score from 1 to 10 indicating the strength and significance of the toxic response. This would allow expert judgement, which cannot realistically be excluded from interpretation of test results, to be incorporated into the index. The fixed scale would also compensate for tests that derive high scores by virtue of an extremely powerful response in one test (example: algal growth inhibition

by Northwood Pulp effluent). For uniformity, all tests in the battery should be included in the index calculation. Nevertheless, occasionally a test will fail or give ambiguous results. Taking the average ensures that index values are still comparable, but expert judgement is called for again to ensure the index is not biased by the missing test.

4. Weighing for Number of Test Responses

The PEEP index includes a multiplier for the number of tests or endpoints exhibiting a toxic response. Inclusion of this factor reflects the originators' assumption that moderate toxicity to several trophic levels is more serious than strong toxicity to one trophic level (Costan et al. 1992). That assumption is reasonable, although the term in the equation may not be necessary. If routine testing uses standard dilutions series, then the number of TU that can be contributed by any one test is limited. Consequently mean toxicity values beyond a certain point must include toxicity from more than one assay.

5. Including Persistence of Toxicity

PEEP includes toxicity tests performed before and after a test of biodegradation, to allow for persistence of toxicity in the formula. The factor is included by the multiplier n, which is the sum of all responding tests before and after biodegradation. The potential for this feature to bias results when different test batteries are used before and after degradation has been discussed in Section 2.3. Even when that problem is corrected, the persistence measurement is one of the weakest parts of the index.

The biodegradation procedure involves massive nutrient additions, plus pH stabilization and seeding with microbial decomposers. It therefore entails potentially dramatic alterations in the chemistry and toxicological properties of the effluent irrespective of any biological degradation. Solubility, bioactivity and chemical form of metals and organic acids could be drastically altered by the addition of a pH 7.2 buffer, and there is the potential for unpredictable complexations and reactions between elements of the effluent and chemical additives. Tests such as algal growth will be influenced by the added nutrients. The procedure uses an unpublished, in-house formula that has apparently not been tested anywhere outside the originating laboratory. Test conditions, in particular aeration rate, are incompletely specified and a correlation with persistence in the field has not been established.

The proposed method does not allow for the possibility that the undiluted effluent may be toxic or bacteriostatic to the decomposers themselves. Volatilization will be included along with biodegradation because of the bubbling aeration. The procedure, based on 1-L samples, is unworkable for assays like the trout test that require large volumes of effluent. Finally, the effect of biodegradation, which in some instances leads to increased toxicity, is included as an integral part of the formula; it is not possible to separate the effect of biodegradation from raw-effluent

toxicity (the way flow can be separated by deleting the flow multiplier, Q) because both are included in the multiplier n.

The conclusion is that the persistence measurement in the PEEP index is inappropriate for use in the Fraser River Basin. If a measurement of biodegradability is required, it should be based on a standard procedure such as BOD_5 or aeration of unaugmented samples at a fixed, measured, rate. Again, judgement would be called for in the interpretation of results, and biodegradation is probably best deleted as a factor in the toxicity index.

6. Including Effluent Volume

The PEEP formula multiplies mean toxicity by the hourly volume of effluent, to compute the total TU entering the river. This calculation makes sense because between equally toxic effluents the one with the larger volume is certainly the more environmentally disruptive. But discharge volume is relevant only when compared against the dilution potential of the river at the point of discharge, which varies immensely along the length of the Fraser River and its tributaries. Hence, a ratio of effluent volume to river flow would be more accurate, but the question then arises as to which volume estimate to use, given the variability of river and effluent discharges. Rather than include effluent volume exactly in the formula, it might be better to use a scale of 1-10 again based on the average dilution ratio, but incorporating such factors as annual flow variability, seasonality of effluent discharge, and rapidity of mixing. Alternatively, the index could be reported as two numbers, one expressing the simple toxicity of the effluent and the other incorporating flow.

7. Use of the Logarithmic Transformation

The final PEEP calculation takes a base-10 log of the mean toxicity times flow. This step is done for convenience (to reduce the scale to 0-10). Logarithmic transformation disturbs the linearity of the toxicity response and distorts differences between index values (untransformed values of 100, 1000 and 10000 each differ by 1 after log-transformation). This use of logs is not recommended in a toxicity index for the Fraser River.

3.3 CONCLUSIONS AND MODIFIED INDEX

Section 2.3 demonstrated that the test battery used in PEEP is unbalanced and incomplete, while the above text argues that there are additional considerations related to the formula used as well. If the problematic elements of PEEP are removed, then the remainder forms a simpler index that should be effective for the Fraser River Basin. The following assumes that a test battery like that recommended in Section 2.4 is used.

The first, preferred, option is essentially a collapsed PEEP index:

$$P_1 = [(5 * \Sigma TU_a) + (\Sigma TU_c)/(5*N_1 + N_2)] * Q$$

where TU_a and TU_c are toxic units for acute and chronic endpoints, respectively, N_1 and N_2 are the number of acute and chronic endpoints, and all other symbols are as previously defined. This formula takes the mean toxicity, weighted toward acute lethality, and multiplies it by hourly discharge. Toxic units would be based on TEC, but the biodegradation test, weighting for number of endpoints responding, and log transformation are deleted. The values before and after multiplying by Q would be reported. The weighting factor for acute toxicity is necessarily arbitrary, and reflects the seriousness with which acute toxicity is viewed. Here, a value of five has been used, based on the observation that thresholds for chronic effects commonly begin at toxicant concentrations one-fifth of the acute toxicity threshold (USEPA 1991). Higher weightings could probably be justified.

In the application of the recommended test battery to this index, endpoints would be labelled acute or chronic as follows: rainbow trout mortality, acute; *P. phosphoreum* light emission, acute; algal growth, chronic; *Ceriodaphnia* mortality and reproduction, both chronic. The battery therefore includes two acute endpoints to three chronic endpoints. Mortality of *Ceriodaphnia* is considered a chronic endpoint because the test extends over a significant part of the life cycle of the organism.

Table 4 compares P_1 index values for the three sampled wastewaters with values of the PEEP index. (See Appendix 2 for calculations). PEEP values have also been given without logarithmic transformation, to ease comparisons with P_1. Despite the weighting toward acute toxicity in the new index, the rank order of the three wastewaters is unchanged from that given by PEEP, although the difference between Annacis Island and Northwood Pulp is reduced. The extraordinarily strong response in the algal growth test inflates the apparent toxicity of Northwood's effluent, regardless of which index is used.

Several variants of the P_1 formula are possible. The flow variable Q could be replaced with a dilution factor based on effluent volume and river discharge. The seven-day low flow with a ten-year return period (7Q10) is commonly used for comparisons of this kind. As an alternative to the weighting factor toward acute toxicity, two separate indices could be calculated, one for acute toxicity alone, and the other for chronic toxicity.

TABLE 4

COMPARISON OF A PEEP INDEX FOR THREE WASTEWATERS WITH A MODIFIED INDEX, P_1

EFFLUENT SOURCE	VOLUME (m^3/h)	PEEP		P_1	
		INDEX	ANTILOG	MEAN TOXICITY (%)	INDEX
Annacis Island	17328	6.0	1 000 000	1.64	28465
Northwood Pulp	5204	6.8	6 310 000	36.1	187888
GVRD Storm Serwer	396	3.3	1995	0.11	43

The second option incorporates expert opinion more fundamentally:

$$P_2 = [(TU_1 + TU_2 + ...TU_5)/N] * F_q$$

where TU_i is a score for each test endpoint, scaled from 1 to 10, and F_q is a similarly scaled factor for flow based on the dilution ratio, mixing zone length and sensitivity of the region. This index could not be calculated by the technical personnel doing the toxicity testing. P_1 is more mechanical to calculate, but it still requires judgement to interpret. P_2 might be a better index for a second round of testing where site-specific factors are to be considered more explicitly. Other formulae incorporating aspects of both P_1 and P_2 are also possible. For example, the 7Q10 flow might serve for F_q. Nevertheless, the interpretation of toxicity test results will always demand professional judgement.

4.0 ALTERNATIVES

Some alternative approaches to the problem of ranking effluents should be considered. The ultimate objective of the FPAO is to reduce the influx of environmentally disruptive effluents to the Fraser River. The first step toward achieving that objective is to rank the effluents for priority of remediation based on toxicological response. But there are > 300 licensed discharges entering the Fraser River, and innumerable smaller, unlicensed sources such as storm sewers. The toxicity testing program presented in the previous sections assumes an equal effort for every discharge, regardless of its size or known potential to contain deleterious substances. Further, extensive chemical testing would be conducted in parallel with the toxicity testing. Given that a few discharges will almost certainly contribute most of the toxic chemical entering the river, and that many discharges will be essentially harmless, this approach may not be the most efficient use of resources.

A better alternative might be to use a two-phase approach in which all discharges would be quickly screened for potential toxicity, but more extensive second-phase testing would be reserved for discharges that showed significant results in the first phase. The key to success of this approach is to establish reliable decision criteria to separate innocuous discharges from potentially harmful ones based on results of preliminary tests. Multi-species bioassay batteries can be used in a modified form to do this screening. In screening mode, the tests are run identically as in ordinary dilution-series tests except that the organisms are exposed to undiluted effluent only, and results are scored as toxic or not (i.e., pass/fail) according as a biological response occurred. Much of the expense in conducting full-dilution bioassays required to accurately determine the strength of toxicity by statistical endpoints (LC50, TEC, etc) is thus avoided. An effluent would be considered potentially harmful if any of the test endpoints, chronic or acute, showed a significant response compared with controls.

The battery of tests recommended previously could equally well serve for screening; the only differences is that initially they would only be tested with 100% effluent. The tests would otherwise follow established protocols. Tests should not only render a qualitative result (*viz.*, toxic or nontoxic); quantitative information on the number of mortalities, or reproduction or growth rates should also be reported to ensure the tests are as informative as possible. While both would score as toxic, an effluent that killed all test fish in one day is probably a more serious problem than one that kills three fish in four days.

The advantage of the screening assays is that a greater number of effluents can be tested at a substantially lower cost and with a smaller sample volume; for the four recommended tests, screening assays require 20 L, compared with 45 L for full-dilution tests (see Appendix Table 1). The disadvantage is that the first phase testing will only reveal which effluents have potential to cause adverse effects; there is no information upon which to rank the toxic effluents further, beyond what can be gleaned from observations on the speed or strength of response in full-

strength effluent. The second phase of analysis would thus be conducted on the effluents that exhibited toxic responses in phase one. Assessments in this phase would follow the full protocols for multiple-dilution testing and chemical characterization. Toxicity endpoints and chemical data would then be used to rank each effluent, using a toxicity index such as described in the previous sections.

If the two-phase approach is adopted, strict attention must be paid to quality control, particularly in phase one. When conducting single pass/fail tests, there is always a risk of encountering false negative results, i.e., concluding that the effluent is nontoxic when in fact it is toxic. This possibility can be minimized by insisting on strict quality control. Fortunately, by the nature of bioassays and biological systems generally, false positive responses, i.e., classifying a harmless effluent as toxic, are far more common than false negatives. False positive results would create a minor inefficiency because the miscategorized effluent would then be retested needlessly in the next phase, but they err on the side of extra protection for the river.

A second alternative pertains to presentation of results. The potential for ecological damage from an effluent depends upon its toxicity and flow. If these are plotted orthogonally as in Figure 1, the relative ranking of each effluent can be seen at a glance. The graph could be subdivided into any number of zones by isopleths defining priorities for further work. In Figure 1, four zones have been used. Effluents in zone A are of low toxicity and low volume, and hence of low priority. Zones B and C define areas of successively higher discharge volumes or toxicity, and are of intermediate priority. The isopleths are angled, under the assumption that high-volume, low-toxicity effluents are equally important as low-volume, high-toxicity effluents. The most serious effluents lie in Zone D, where toxicity and flow are both high. As an example, Figure 1 suggests where the three effluents in the pilot wastewater study might lie. This graphical approach may be useful to help visualize the differences among effluents, or for presentation to the public.

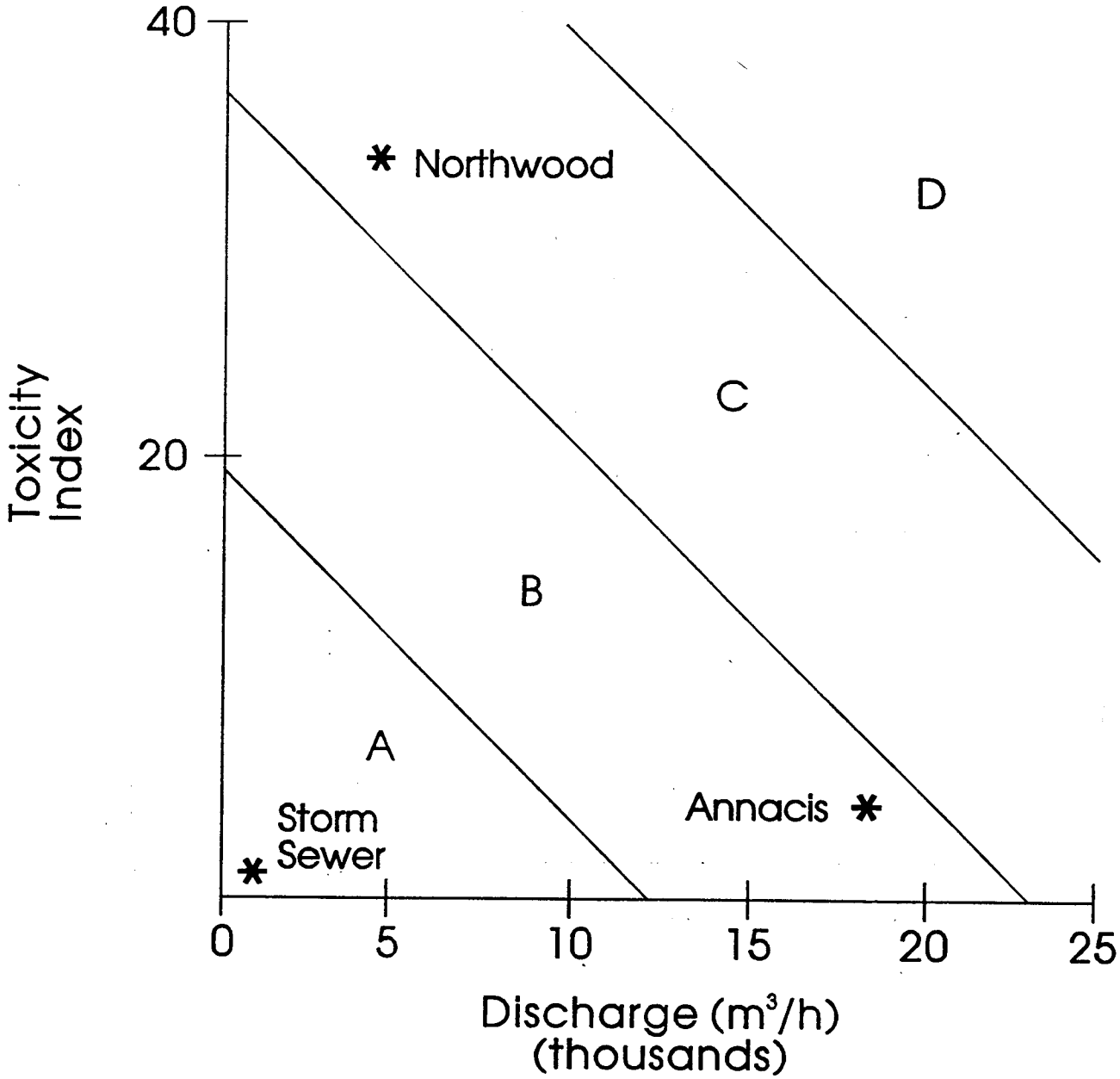

Figure 1. Theoretical toxicity-discharge graph for comparing wastewater effluents.

5.0 QUALITY ASSURANCE

The assurance of quality data is important for the implementation of the FPAO testing program. Guidelines are required to cover all aspects of sample collection, transport, storage and testing along with the analysis and reporting of the results. Adherance to guidelines makes it easier to interpret the test data and provides a basis for dealing with outliers. In this chapter, important factors relating to quality assurance in toxicity testing are described. Specific quality assurance requirements for the FPAO program are provided in the Instruction Manual in Appendix I.

5.1 STANDARD OPERATING PROCEDURES

Different laboratories will be conducting the tests for the FPAO; to ensure comparable and reliable results, each facility must have defined standard operating procedures (SOPs) for conducting tests, managing data, and performing other laboratory functions such as maintaining cultures and equipment. Basic toxicity test procedures are well defined in published protocols. However, there is some room for interpretation in the implementation of each protocol. These kinds of details must be documented as part of a laboratory's standard operating procudures.

5.2 QUALITY ASSURANCE PROGRAM

As well as standard operating procedures, testing facilities should have a Quality Assurance Program that outlines their quality assurance and quality control practices. Quality assurance encompasses all aspects of laboratory operations which may directly or indirectly involve a specific test. This can include everything from the collection of the sample, transport, storage, and testing to data recording, analysis, interpretation and reporting. Direct factors are the procedures and conditions under which the test is conducted. Indirect factors are those that may not directly effect a test but which could have some influence on the test results. Indirect factors include the care and maintenance of the test organisms, environmental conditions in the test facility, dilution water quality and staff training.

Quality control is an integral component of quality assurance. Quality control procedures are used to measure and assess data quality. Reference toxicants and the use of warning charts, along with participation in inter-laboratory test programs are quality control procedures. Control charts can be used to establish data quality objectives and ensure that the test is in control and the organisms are responding in a consistent and reproducible fashion.

5.3 CAEAL ACCREDITATION PROGRAM

Quality assurance guidelines for the toxicity testing of aqueous samples as part of the FPAO program recommended in Appendix I are based on the proposed requirements for accreditation under the Canadian Association of Environmental and Analytical Laboratories (CAEAL 1992).

The CAEAL Biological Testing Laboratory Accreditation Program will be implemented in 1993. The program encompasses the stringent requirements established by the USEPA (USEPA 1990) for laboratories performing aquatic toxicity tests. The CAEAL program is also the first national program to set industry-wide standards for toxicity testing. It is supported by Environment Canada and covers: routine testing; environmental monitoring or biological testing; and Good Laboratory Practices or GLP programs.

Routine tests are biological test methods developed and supported by Environment Canada (see references). Each procedure specifies the quality assurance and control requirements for conducting the test along with reporting information. GLP is a specific program for product registration. Environmental monitoring programs are more routine and include testing of effluents and wastewaters for licence compliance. The requirements for both programs are well defined although the GLP requirements are more rigorous.

Toxicity testing laboratories will be able to obtain certification for all Environment Canada protocols which are included in the FPAO program. (The SOS-Chromotest is not included, but we do not recommend this test.) Accreditation will include submission of specific documentation on the operation of the facility, test methods, and personnel, participation in inter-laboratory performance evaluations, and site evaluations. An integral component is the maintenance of control or warning charts for accredited tests using one or more recommended reference toxicants (Appendix Table 3).

The Quality Assurance Guidelines we recommend for the FPAO test program cover sampling, testing, and reporting (Appendix I). Sample procedures and reporting of test data are not explicitly covered in the CAEAL accreditation program, which is primarily concerned with testing of samples. Sampling and reporting requirements are based on those recommended in the Environment Canada test protocols, with additional details added. Many of the procedures require documentation to ensure in retrospect that the sample was collected and handled according to instructions and the test was conducted and reported correctly.

5.4 LABORATORY SELECTION

A reliable toxicity testing laboratory must be selected before sampling begins. Selection of testing facilities is based on a number of simple criteria. It is important to determine if a facility can run a particular test in a competent fashion. This can be determined from the answers to the following questions:

1. Can they run the Environment Canada protocol specified in the FPAO test program?

2. Do they have a written, standard operating procedure (SOP) for this test detailing the test procedure followed by the testing facility?

3. Do they have a control chart for this test with one of the specified reference toxicants?

4. Have they run a reference toxicant for this test in the last month?

5. Do they have the preparation time to implement the test?

The first question is to determine if the laboratory is familiar with a specific protocol. Most biological tests require a certain amount of time to implement. If a facility is not familiar with a test protocol and they are not currently running the test then it is questionable whether they can provide reliable test data.

The test facility must have an internal standard operating procedure for all the tests they offer (they should also have SOPs for all aspects of the operation). This is a fundamental component of a quality assurance plan. SOPs are detailed sets of instructions for performing a specific tasks. For biological tests, they are based on standard protocols and further refined during implementation of the test. The existence of a written SOP for a test indicates that the test facility has implemented the test and that they have some experience conducting the test.

Control charts are an integral component of a quality assurance program. They provide a means to determine if the test organisms are responding in a reproducible fashion based on historical data. The control chart incorporates many direct and indirect factors that can influence a test. It also provides some assurance about the facility's ability to conduct the test. Test facilities should run reference toxicants on a regular basis; if they have not run a reference toxicant in the last month, then one should be requested prior to submission of any samples. The reference toxicant result should then be compared against their control chart for that test. If the test is under control, then the facility has demonstrated it's ability to run the test and provide reliable data.

Although not desirable, it may be necessary to send samples to different test facilities for different tests. When a new laboratory is to be used, it is recommended that FPAO personnel visit the facility and conduct an informal audit. Guidelines are available for auditing toxicity testing facilities (CAEAL 1992; USEPA 1990).

6.0 RECOMMENDATIONS

1. We recommend that PEEP as presently formulated not be used in the Fraser River Basin. The PEEP test battery is too heavily weighted toward chronic and genetic toxicity, the procedure for measuring persistence of toxicity is of uncertain relevance and the mathematical formula needs to be refined.

2. The following four tests are recommended for the FPAO:

    1. Bacterial luminescence with *Photobacterium phosphoreum* (Environment Canada 1991)

    2. Algal growth with *Selenastrum capricornutum* (Environment Canada 1992a)

    3. *Ceriodaphnia dubia* survival and reproduction (Environment Canada 1992b)

    4. Rainbow trout (*Onchorynchus mykiss*) acute mortality (Environment Canada 1990a)

3. We recommend that the SOS-Chromotest not be used for effluent assessment because it is only an indicator of potential mutagenesis, not a true toxicity assay, and results suggesting potential carcinogenicity are too easily misconstrued. We also recommend that the FPAO consider inclusion of the survival and growth test with fathead minnow larvae, and the seven-day embryo mortality test with rainbow trout as alternatives to the standard acute trout test.

4. We suggest a simplified index of toxicity, calculated as the mean TEC in each of the assays, weighted toward acute lethality, multiplied by flow. Alternatively, the bioassay results can be scored on a fixed scale according to the strength of ecological effects they imply, and the average multiplied by a factor for dilution ratios. Circumspect expert judgement must be applied to interpretation of either index.

7.0 REFERENCES

American Society for Testing and Materials (ASTM). 1988. Standard guide for conducting early life-stage toxicity tests with fishes. 1991 Annual Book of Standards, Volume 11.04 E 1241-88: 857-882.

American Society for Testing and Materials (ASTM) 1991. Standard guide for conducting static toxicity tests with *Lemna gibba* G3. 1991 Annual Book of ASTM Standards. Volume 11.04. E 1415-91: 1137-11446.

Analex Inc., and Environment Canada (St. Lawrence Centre) 1992. Ecotoxicological evaluation of effluents from a pilot project on the Fraser River (British Columbia). 22 pp.

Ankley, G.T., Peterson, G.S., Amato, J.R. and Jenson, J.J. 1990. Evaluation of sucrose as an alternative to sodium chloride in the Microtox assay.: comparison to fish and cladoceran tests with freshwater effluents. Environmental Toxicology and Chemistry 9: 1305-1310.

Beak Consultants Ltd. 1990. Effluent toxicity of Fraser River Pulp Mills. Prepared for Environmental Conservation and Protection, Pacific Region, Environment Canada. Vancouver, British Columbia. 6 pp.

Biesenger, K.E., Williams, L.R. and van der Shalie, W.H. 1987. User's Guide: Procedures for conducting *Daphnia magna* bioassays. EPA/600/8-87/011. 57 pp. (Cited in Keddy et al. 1992).

Birtwell, I.K., Levings, C.D., Macdonald, J.S. and Rogers, I.H. 1988. A review of fish habitat issues in the Fraser River system. Water Pollution Research Journal of Canada. 23: 1-29.

Bothwell, M.L., Derksen, G., Nordin, R.N. and Culp, J.M. 1992. Nutrient and grazer control of algal biomass in the Thompson River, British Columbia: a case history of water quality management. *In*: Robarts, R.D. and Bothwell, M.L. (Eds.). Aquatic ecosystems in semi-arid regions: implications for resource management. NHRI Symposium Series 7, Environment Canada, Saskatoon. pp. 253-265.

Brown, V.M. 1968. The calculation of the acute toxicity of mixtures of poisons to rainbow trout. Water Research 2: 723-733.

Cairns, J. Jr. 1983. Are single species toxicity tests alone adequate for estimating environmental hazard? Hydrobiologia 100: 47-57.

Cairns, J. Jr. 1986. The myth of the most sensitive species. Bioscience 36: 670-672.

Canadian Association for Environmental Analytical Laboratories (CAEAL) Inc. 1992. National Biological Testing: Laboratory Accreditation Program, Program Description Manual. Ottawa, Ontario..

Costan, G., Bermingham, N., Blaise, C. and Ferard, J.F. 1992. Potential Ecotoxic Effects Probe (PEEP): a novel index to assess and compare the toxic potential of industrial effluents. Unpublished manuscript. Centre Saint-Laurent, Environment Canada, Montréal, Québec.

Doherty, F.G. 1983. Interspecies correlations of acute aquatic median lethal concentrations for four standard testing species. Environment Science and Technology 17: 661-665.

Eagleson, K.W., Lenat, D.L., Ausley, L.W. and Winborne, F.B. 1990. Comparison of measured instream biological responses with responses predicted using the *Ceriodaphnia dubia* chronic toxicity test. Environmental Toxicology and Chemistry 9: 1019-1028.

Environment Canada. 1990a. Biological Test Method: Acute Lethality Test Using Rainbow Trout. Report EPS 1/RM/9. 51 pp.

Environment Canada. 1990b. Biological Test Method: Acute Lethality Test Using *Daphnia* spp. Report EPS 1/RM/11.

Environment Canada. 1990c. Biological Test Method: Guidance Document on Control of Toxicity Test Precision using Reference Toxicants. EPS Report 1/RM/12.

Environment Canada. 1990d. Biological Test Method: Reference Method for Determining the Acute Lethality of Effluents to Rainbow Trout. Reference Method EPS 1/RM/13.

Environment Canada. 1990e. Biological Test Method: Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*. Reference Method EPS 1/RM/14.

Environment Canada. 1991. Toxicity testing using luminescent Bacteria (*Photobacterium phosphoreum*). Report EPS 1/SRM/1. 64 pp.

Environment Canada. 1992a. Biological Test Method: Growth Inhibition Test Using the Freshwater Alga *Selenastrum capricornutum*. Report EPS (in press).

Environment Canada. 1992b. Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows. Report EPS 1/RM/22.

Environment Canada. 1992c. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*. Report EPS 1/RM/21. 73 pp.

Environment Canada. 1993. Biological Test Method: Early Life-Stage Tests Using Salmonid Fish. Report EPS (in press).

Esvelt, L.A., Kaufman, W.J. and Sellock, R.E. 1973. Toxicity assessment of treated municipal wastewaters. Journal of the Water Pollution Control Federation 45: 1558-1564.

Keddy, C., J.C. Greene and M.A. Bonnell. 1992. A review of whole organism bioassays for assessing the quality of soil, freshwater sediment and freshwater in Canada. CCME Subcommittee on Environmental Quality Criteria for Contaminated Sites, Canadian Council of Ministers of Environment (CCME). Ottawa, Ontario.

Miller, W.E., Peterson, S.A., Greene, J.C. and Callahan, C.A. 1985. Comparative toxicology of laboratory organisms for assessing hazardous waste sites. Journal of Environmental Quality 14: 569-574.

Northcote, T.G., N.T. Johnston, and K. Tsumura. 1976. Benthic, epibenthic and drift fauna of the lower Fraser River. Westwater Research Centre Technical Report No. 11, University of British Columbia, Vancouver, B.C.

Organization for Economic Cooperation and Development (OECD). 1991. Guideline Annex 2, Test Guideline - Prolonged Toxicity Study with *Daphnia magna*: effects on reproduction, Proposal for update of 202 , Part II. OECD Guidelines (May) Annex 2: 14 p. (Cited in Keddy et al. 1992).

Scott, W.B. and Crossman, E.J. 1973. Freshwater Fishes of Canada. Fisheries Research Board of Canada, Bulletin 184. Ottawa, Ontario. 966 pp.

Sprague, J.B. and Ramsey, B.A. 1965. Lethal levels of mixed copper-zinc solutions for juvenile salmon. Journal of the Fisheries Research Board of Canada. 22: 425-432.

Thomas, G.P. and Dwernychuk, L.W. 1992. Characterization of selected wastewaters. Final Report (Volume I). Prepared for Environment Canada by Hatfield Consultants, Vancouver, B.C. 40 pp. + 7 app.

United States Environmental Protection Agency (USEPA). 1990. Manual for the Evaluation of Laboratories Performing Aquatic Toxicity Tests. D.J. Klemm, L.B. Lobring and W.H. Horning. EPA/600/4-90/031.

United States Environmental Protection Agency (USEPA). 1991. Technical support document for water quality-based toxics control. Office of Water, Washington, D.C.

van Straalen, N.M. and Denneman, C.A.J. 1991. Ecotoxicological evaluation of soil quality criteria. Ecotoxicology and Environmental Safety 18: 241-251.

Wetzel, R.L. 1975. Limnology. W. B. Saunders Co., Philadelphia, PA. 743 pp.

APPENDIX I

TOXICITY TESTING OF AQUEOUS SAMPLES

INSTRUCTION MANUAL

CONTENTS:

The intent of this document is to provide concise instructions for the collection and testing of aqueous samples for the FPAO. There are three sections dealing with the following topics:

- sample collection
- sample transport and storage
- testing and reporting

Each step involved in the collection and toxicity testing of an aqueous sample is covered in detail. Checklists and sample forms for required documentation are appended.

The instructions encompass quality assurance practices required within the accreditation program of the Canadian Association for Environmental Analytical Laboratories (CAEAL) Inc. along with those specified in the test protocols published and supported by Environment Canada (see references). These quality assurance practices are paramount for the collection of reliable and meaningful data.

SAMPLE COLLECTION

This first component involves the following tasks:

- coordination of sampling and testing
- collection of the sample
- transport of the sample to the test facility
- storage during transport or prior to testing

Before collecting the samples, first contact the test facility and give them notice of your intent to collect a sample. One to two weeks is usually sufficient advance notice. Coordinating sampling with the test facility will ensure that they will be ready to process the sample in an efficient manner. This is particularly important for the more complicated chronic tests. If notice is given verbally, follow it up immediately with a written communication of your intent to collect a sample.

The samples must be collected in new plastic containers (polyethylene or polypropylene). Containers must not be recycled or reused. Glass and teflon-lined containers can also be used but these are often too impractical and expensive for larger volumes. The test facility may provide sample containers. We recommend 5-gallon (20-L) polyethylene pails with lids. These pails are inexpensive, easy to fill, stackable, they can be labelled directly, and the lids form watertight seals. Rectangular, polyethylene gasoline cans (Jerry-cans) also make convenient containers. The type of container and its composition (polyethylene, polypropylene or teflon) should be recorded. If a plastic liner is used, its composition should be recorded as well.

The sample volumes for each test are specified in Table 1. A total of 44 L is required for the trout, *Ceriodaphnia*, bioluminescence and algal growth tests (a greater volume may be required for a salmonid larval growth test). We recommend that a total of 60 L be collected. The extra volume can be used for chemical testing or archived until the tests are completed.

CHECKLIST FOR SAMPLING AND TOXICITY TESTING OF AQUEOUS SAMPLES

ITEM	COMPLETED DATE/INITIAL
SAMPLING	
notify the testing facility of your intent to sample	
obtain containers and forms	
rinse containers three times with sample	
fill and seal containers, minimize headspace	
label containers and initial seals	
fill out forms	
transport containers to testing facility	
INFORMATION REPORTED BY THE TESTING FACILITY	
sample information (type of substance, location, sampler, date, how sampled, contact person)	
location of test facility (address and contacts)	
complete reference for the test method	
test conditions (dates started/finished, investigators, temperatures, light levels)	
test results (end points and method of calculation)	
reference toxicant data (endpoint and method of calculation)	
monitoring data (biological and chemical)	
observations during test and other comments	

TABLE 1

SAMPLE VOLUMES REQUIRED TO CONDUCT SELECTED BIOLOGICAL TESTS

	VOLUME (L)	
TEST	**FULL DILUTION**	**SCREENING**
rainbow trout lethality	**40**	**20**
bacterial luminescence inhibition	**0.1**	**0.01**
Ceriodaphnia survival and reproduction	**3**	**0.2**
algal growth stimulation/inhibition	**0.5**	**0.5**

The containers must be rinsed three times with the sample before filling. Each rinse should be with a volume not less than 5% of the total container volume. The rinse volumes can also be used to rinse the bottoms of the lids. Fill the containers to minimize the head space and then seal them firmly with lids. Wrap packing tape around the lid at the top of the pail and over any spouts in the lid. Initial the tape and label the pail and lid with a permanent black marker. The taping and initialling of the tape are simple things to do but they provide valuable information on sample integrity between the point and time of collection and receipt at the test facility. They provide a simple means to determine if the sample was opened or tampered with during transit.

Both the pail and lid must be labelled with the following information.

- type of substance (effluent, surface water, etc.)
- location of sample (company, municipality, etc.)
- name of sampler
- date sampled
- how sampled (grab or composite)
- contact person (name, address, and phone number)

If the sample is collected in more than one container then each container must be assigned one number out of the total number of containers (e.g., 1 of 3, 2 of 3, and 3 of 3).

SAMPLE TRANSPORT AND STORAGE

The containers must be shipped to the test facility accompanied by a test request and a chain-of-custody form. It is not necessary to ship samples packed in ice. A sample chain-of-custody form is appended. The chain-of-custody form is to insure that someone is always accountable for the sample and that custody is recorded in an orderly and clear fashion. If the sample changes carriers during transit, then the carrier must sign the chain-of-custody form accompanying the samples. The person receiving the sample at the test facility must also sign for the sample to complete the record. The temperature of the sample on receipt should be logged. Observations of sample conditions (e.g., ice present, container intact) should also be recorded. A copy of the completed form should then be returned to the sampler.

The sample must arrive at the test facility within 48 h of collection and chronic toxicity tests must be initiated within 72 h of collection. The trout and *Ceriodaphnia* protocols permit up to 5 days to elapse between the time of sampling and test initiation. For purposes of simplicity and uniformity, we recommend that all tests be initiated within 72 h of sample collection. The sample must be stored at 5°C ± 2°C.

TESTING AND REPORTING

Each test facility must have written standard operating procedures (SOPs) for conducting each test. SOPs must cover the method of data analysis and determination of endpoints. The standard operating procedures and records of performance in any inter-laboratory test programs must be available for inspection by the FPAO. The test facility must follow the appropriate Environment

CHAIN OF CUSTODY AND SAMPLE INFORMATION FORM

SAMPLE INFORMATION				
TYPE OF SUBSTANCE (effluent, surface water, etc.)				
LOCATION OF SAMPLE (site, address, location of outfall)				
SAMPLER:				
DATE SAMPLED:				
HOW SAMPLED (grab, composite, etc.):				
TYPE OF CONTAINER:				
CONTACT PERSON:				
ADDRESS:				
CHAIN OF CUSTODY				
DATE	TIME	CARRIER	SIGNATURE	PRINT NAME
TEST FACILITY				
ARE THE SEALS INITIALED?		TEMP?	CONDITION?	
INITIAL CHEMISTRY				
pH		DO	AMMONIUM	
COND		TEMP	FREE CHLORINE	
ODOUR		COLOUR		
TESTS REQUESTED:				

Canada test procedures. These procedures are published (see references) and available from the following Environment Canada offices:

EP Publications
Conservation and Protection
Ottawa, ON
K1A 0H3
(613) 953-5921

Communications Directorate
Conservation and Protection
224 West Esplanade
North Vancouver, BC
V7M 3H7
(604) 666-5900

The reporting requirements for each test are summarized in Table 2. This table was compiled from the test protocols. In summary, the following general information is required with each test report:

- sample information
- location of the test facility
- test method reference
- test conditions
- test result
- reference toxicant data
- monitoring data (biological and chemical)
- observations during test and other comments

This information is required to ensure data quality. Any additional information collected during the test must be kept on file by the testing facility for at least two years.

A control or warning chart for a reference toxicant must also be submitted with each test result. The reference-toxicant test result that corresponds with the sample must be clearly marked on this chart. Warning limits that are statistically defined from historical test data must also be marked on the control chart. Reference toxicants used in Environment Canada protocols are listed in Table 3.

REPORTING REQUIREMENTS FOR THE FPAO TEST PROGRAM

REPORTING REQUIREMENT	TROUT	BACTERIAL LUMINESCEN	CERIODAPHNIA	ALGAL GROWTH
SAMPLE INFORMATION				
type of sample	+	+	+	+
sample location	+	+	+	+
date & time of sampling	+	+	+	+
sampler	+	+	+	+
colour & odour	+	+	+	+
TEST CONDITIONS AND REPORTING				
method reference	+	+	+	+
test facility (name, location & contact person)	+	+	+	+
date test initiated	+	+	+	+
date test terminated	+	+	+	+
endpoints	LC50	IC50/IC20	SURVIVAL	IC50/NOEC/LOEC
			LC50/NOEC/LOEC	
			REPRODUCTION	
			IC50/IC25/NOEC/LOEC	
method of analysis	+	+	+	+
MONITORING DATA				
pH, conductance, dissolved oxygen	+		+	+
ammonium	+		+	
alkalinity, hardness, free chlorine			+	
mortality over time	+		+	
weight	+			
length	+			
young production over time			+	
immobility			+	
observations on behaviour	+		+	

TABLE 3

REFERENCE TOXICANTS RECOMMENDED FOR BIOLOGICAL TESTING

TEST	REFERENCE TOXICANTS			
	$ZnSO_4$	PHENOL	KCl	NaCl
rainbow trout lethality	+	+	+	
bacterial luminescence inhibition	+	+		
Ceriodaphnia survival and reproduction			+	+
algal growth stimulation/inhibition	+		+	+

REFERENCES

Canadian Association for Environmental Analytical Laboratories (CAEAL) Inc. 1992. National Biological Testing: Laboratory Accreditation Program, Program Description Manual. Ottawa, Ontario..

Environment Canada, 1990a. Biological Test Method: Acute Lethality Test Using Rainbow Trout. Report EPS 1/RM/9

Environment Canada, 1990b. Biological Test Method: Acute Lethality test Using *Daphnia* spp. Report EPS 1/RM/11.

Environment Canada, 1990c. Biological Test Method: Guidance Document on Control of Toxicity Test Precision using Reference Toxicants. EPS Report 1/RM/12.

Environment Canada, 1990d. Biological Test Method: Reference Method for Determining the Acute Lethality of Effluents to Rainbow Trout. Reference Method EPS 1/RM/13.

Environment Canada, 1990e. Biological Test Method: Reference Method for Determining Acute Lethality of Effluents to *Daphnia magna*. Reference Method EPS 1/RM/14.

Environment Canada, 1992a. Biological Test Method: Test of Reproduction and Survival Using the Cladoceran *Ceriodaphnia dubia*. Report EPS 1/RM/21.

Environment Canada, 1992b. Biological Test Method: Test of Larval Growth and Survival Using Fathead Minnows. Report EPS 1/RM/22.

Environment Canada, 1993a. Biological Test Method: Growth Inhibition Test Using the Freshwater Alga *Selenastrum capricornutum*. Report EPS (in press)

Environment Canada, 1993b. Biological Test Method: Toxicity Test Using Luminescent Bacteria (*Photobacterium phosphoreum*). Report EPS (in press)

Environment Canada, 1993c. Biological Test Method: Early Life-Stage Tests Using Salmonid Fish. Report EPS (in press)

APPENDIX II

EXAMPLE CALCULATIONS OF MODIFIED TOXICITY INDEX

APPENDIX II

CALCULATION OF MODIFIED PEEP INDEX, P1

SITE: Annacis Island

DISCHARGE (m3/h): 17316

TOXICITY DATA

TEST	END-POINT	TEC* (%)	TOXIC UNITS
Trout	acute	50.2	1.99
Ceriodaphnia			
-survival	acute	>100	0.00
-fecundity	chronic	17.5	5.71
Microtox**	chronic	17.5	5.71
algal growth	chronic	>100	0.00
	∑TUa =	1.99	
	∑TUc =	11.42	
	Na =	2	
	Nc =	3	

$P1 = [(5 * \Sigma TUa) + \Sigma TUc)/((5*Na)+Nc)] * Q$

MEAN TOXICITY= 9.95 + 11.42)/13
= 1.64

P1 = 1.64 * 17316
= 28465

REPORT MEAN TOXICITY AND P1

---

* TEC = Square root of (LOEC * NOEC)
if no toxicity is observed, TEC > 100 and
TU = 0

** Analex data used for calculation

CALCULATION OF MODIFIED PEEP INDEX, P1

SITE: Northwood

DISCHARGE (m3/h): 5204

TOXICITY DATA

TEST	END-POINT	TEC* (%)	TOXIC UNITS
Trout	acute	79.37	1.26
Ceriodaphnia			
-survival	acute	>100	0.00
-fecundity	chronic	17.5	5.71
Microtox**	chronic	35.7	2.80
algal growth	chronic	0.22	454.55
	_TUa =	1.26	
	_TUc =	463.06	
	Na =	2	
	N2c =	3	

$P1 = [(5 * \Sigma TUa) + \Sigma TUc)/((5*Na)+Nc)] * Q$

MEAN TOXICITY= 6.30 + 463.06)/13
= 36.10

P1 = 36.10 * 5204
= 187888

REPORT MEAN TOXICITY AND P1

---

* TEC = Square root of (LOEC * NOEC)
if no toxicity is observed, TEC > 100 and
TU = 0

** Analex data used for calculation

CALCULATION OF MODIFIED PEEP INDEX, P1

SITE: GVRD Storm Sewer

DISCHARGE (m3/h): 396

TOXICITY DATA

TEST	END-POINT	TEC* (%)	TOXIC UNITS
Trout	acute	>100	0.00
Ceriodaphnia			
-survival	acute	>100	0.00
-fecundity	chronic	71.4	1.40
Microtox**	chronic	>100	0.00
algal growth	chronic	>100	0.00
	_TUa =	0.00	
	_TUc =	1.40	
	N1 =	2	
	N2 =	3	

P1 = [(5 *Σ TUa) + ΣTUc)/((5*Na)+Nc)] * Q

MEAN TOXICITY= 0.00 + 1.40)/13
= 0.11

P1 = 0.11 * 396
= 43

REPORT MEAN TOXICITY AND P1

---

* TEC = Square root of (LOEC * NOEC)
if no toxicity is observed, TEC > 100 and
TU = 0

** Analex data used for calculation