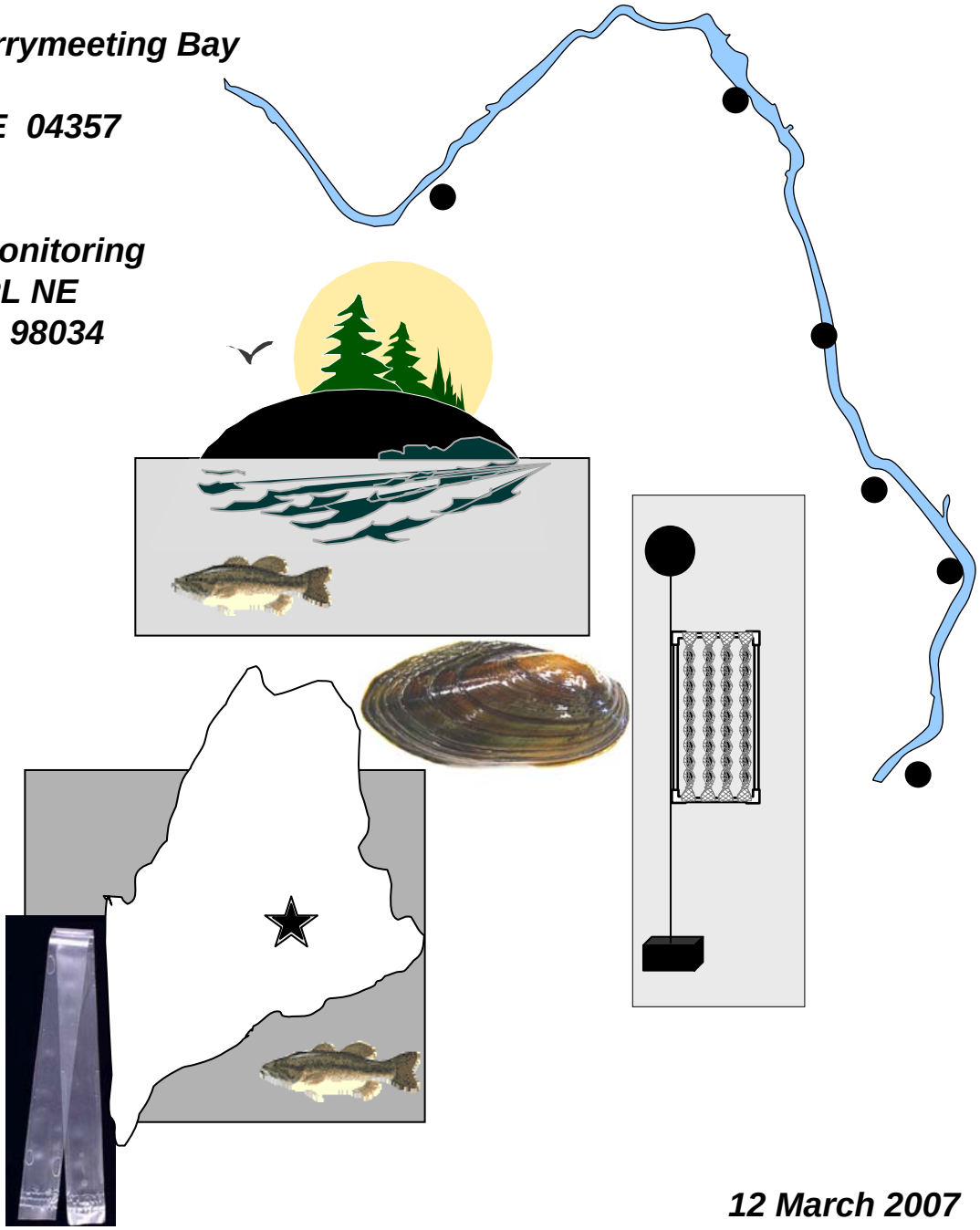
FINAL REPORT

Critique of Caged Mussel Monitoring on the Kennebec River: Environmental Significance of Exposure and Effects Measurements in 2000, 2003, 2005, and 2006

Submitted to:
Ed Friedman
Friends of Merrymeeting Bay
PO Box 233
Richmond, ME 04357

Prepared by:
Applied Biomonitoring
11648 - 72nd PL NE
Kirkland, WA 98034



12 March 2007

Critique of Caged Mussel Monitoring on the Kennebec River: Environmental Significance of Exposure and Effects Measurements in 2000, 2003, 2005, and 2006

Applied Biomonitoring, Kirkland, WA
12 March 2007

ABSTRACT

Applied Biomonitoring was contracted by Friends of Merrymeeting Bay to compare the biomarker results between 2003 and 2005 on the Kennebec River and to address inconsistencies between results from those two years of monitoring. To adequately address all of the potential problems associated with the data and their interpretation, it was necessary to review the results of all caged mussel studies conducted on the Kennebec River in 2000, 2003, 2005, and 2006. Applied Biomonitoring conducted the 2003 caged mussel study but the biomarker results were analyzed and reported by Environment Canada. Applied Biomonitoring conducted the original caged mussel monitoring in 2000 but DEP insisted on deploying mussels only at stations 13 miles above the mill and 11 miles below because that is where DEP routinely catches fish as part of human health and environmental effects monitoring. Applied Biomonitoring did not participate in the 2005 or 2006 caged mussel studies and those results are more difficult to interpret. The 2006 results may not be usable. A number of problems were identified with experimental design, the methods used to conduct the tests, and data interpretation, particularly in 2005 and 2006. This does not mean however that all future caged mussel monitoring should be terminated. The generic methods are sound but there were clear indications that inexperienced workers and those without the appropriate background and training were unable to conduct the tests satisfactorily or interpret the data in the most scientifically defensible manner. Some of these problems were identified by reviews conducted by a Dioxin Peer Review Panel (contracted by DEP) and Patrick Gwinn of AMEC (a consultant contracted by the SAPPI mill). The most surprising thing however, is that DEP seemed to pay little attention to any of the comments and continues to produce reports that have little substance, poorly explained, and poorly referenced.

Problems: 1) Stations too far apart for above-below testing; 2) Inappropriate comparisons; 3) Lack of an integrated program and inadequate peer review; 4) Insufficient replication for biomarker; 5) Lack of oversight for biomarker work; 6) Adversarial relationship between DEP, mills, and environmental interest groups; 7) DEP reports are poorly documented, scientifically unsound, and largely not credible.

1. INTRODUCTION

It now appears as if dioxins and furans are not the chemicals responsible for adverse effects in fish downstream of pulp and paper mill effluents around the world. It also appears as though many of these effects are decreasing due to mill process changes. These two factors are making it more and more difficult to quantify any effects and highlights the need for an ecological risk assessment-based approach. This is another problem associated with what DEP has done. It is this mixing of scientific and regulatory agendas that has caused much of the problem with characterizing exposure and effects on the Kennebec River. The net result is that they have provided a regulatory approach that was doomed to failure at the outset due to a scientifically indefensible position. The above/below design is not appropriate on the Kennebec River because of reference and treatment sites being separated by 24 miles and interceded by an impoundment with completely different conditions. It has been suggested that a true reference station in the field is virtually impossible (Landis 2002). It is not clear why the mills did not object to this approach from the beginning unless their complaints have not been

acknowledged or unless that have been satisfied with maintaining the status quo. Even if statistically significant differences were found when comparing above/below sites associated with the SAPPI mill on the Kennebec River, these results would not be scientifically justifiable nor would they withstand the rigorous review of litigation for the reasons mentioned previously. Results from the 2003 caged mussel study, the last study conducted by Applied Biomonitoring on the Kennebec River clearly show differences in exposure and effects (dioxins and furans, mussel growth and vitellin production) associated with distance from the mill (maybe not so clearly and maybe not implicating the mill but nonetheless differences that deserved future investigation). Results that suggest some relationship with the mill or other sources. Finally, a close inspection of the physical, chemical and biological changes within the impoundment, it just doesn't make sense (Common sense in Port Valdez) to compare sites for any ecological conclusions. The situation is different for human health because effects on fish and accumulated chemicals are important in any human health assessment at any site. This is another problem with the DEP approach; mixing objectives of ecological and human health. Fish data are important for human health at any sites because of fish consumption. This is not the case when assessing potential exposure and effects associated with mill effluents. DEP will ***NEVER*** complete a successful test on the Kennebec River using fish, mussels, SPMDs, or any other monitoring methodology using above/below sites as the yardstick for comparison.

Most monitoring programs have emphasized either ***exposure-based*** or ***effects-based*** approaches using indigenous populations of fish and most of these programs are established to measure regulatory compliance rather than characterizing and understanding processes for a more meaningful monitoring program. The ecological risk assessment (ERA) framework can be used to characterize exposure and effects to the extent necessary for understanding the subtle effects of mill effluents and establishing links with chemical exposure. The above/below test is an impediment to understanding these processes. Furthermore, the emphasis on wild fish in many monitoring programs has contributed to the uncertainty in these assessments and limited the ability to identify the chemicals associated with measured effects. In addition to an ***ERA-based*** approach, we suggest integrating field experiments with caged bivalves or any other field experiments into regular monitoring programs to facilitate characterizations of exposure and effects and to help establish cause-and-effect relationships. DEP has accomplished this with their biomonitoring program but has not integrated this approach into monitoring for the SAPPI mill on the Kennebec River. Caging bivalves or fish along suspected chemical gradients helps to identify sources of accumulated chemicals and facilitates the measurement of effects by controlling the location and duration of exposure as well as the exposure and genetic history of the test animals.

The caged bivalve approach in particular can be used to characterize exposure and effects over space and time under environmentally realistic conditions in the field and has a number of advantages over using wild fish. An integrated approach using fish, bivalves and rock baskets in laboratory and field experiments, with supporting water, sediment, and tissue chemistry data, would enhance the assessments and help identify the chemicals associated with adverse effects. Yes, there is even an approach for deploying sediment bags to assess accumulated chemicals. Since it now appears that dioxins and furans are not responsible for the effects measured in fish, these exposure-based approaches are now somehwat problematic from a regulatory perspective but not necessarily from a scientific perspective. It is important to highlight the difference between ***exposure-based*** and ***effects-based*** approaches. Indicators of exposure simply indicate an exposure to chemicals of concern (just like bioaccumulation in caged mussels) and not necessarily any adverse effects. Indicators of effects (such as caged mussel growth) suggest adverse environmental impacts. This distinction is even more important for fish and bivalve biomarkers because some have been linked to effects on the

population while others have not. This is a source of continuing controversy in the literature and is relevant to the biomarkers measured on the Kennebec River. Nevertheless, a number of pulp mill representatives have supported the biomarker approach (Borton & Firth 1997). Evidence has been provided that vitellin is an indicator of adverse environmental impacts but data interpretation is unclear. We are unaware of similar evidence being provided for lipid peroxidase for example. All of these problems in data interpretation emphasize the need for an ***ERA-based*** approach in characterizing exposure and effects, weight of evidence, and multiple lines of evidence. Most of the work conducted on the Kennebec River,is an example of an exposure-based approach and compared the accumulation of dioxins and furans in caged bivalves, fish, and semipermeable membrane devices (SPMDs). More recently, after terminating the dioxin monitoring program, DEP has switched to more of an effects-based monitoring program by measuring effects endpoints in captured fish and caged bivalves. Most of the work in Canada is an example of an effects-based study that emphasized measuring effects on fish. The level of uncertainty in both approaches could have been reduced by using an ERA-based approach. This paper focuses on the need for more integrated exposure and effects studies on pulp and paper mill effluents, and routine monitoring with equal emphasis on characterizing exposure and effects in the lab and in the field.

There is a growing need for more integrated approaches to monitor pulp and paper mill effluents in Maine and around the world. This need has been hampered by the fact that even after decades of monitoring the causative agent for the most significant effects have not yet been clearly identified. The State of Maine has tried over 78 different tests and none have been completely successful. "Since the development of the Above/Below test began in 1997, over 78 tests have been conducted for different dioxins, species, tissue types, and other surrogates in an attempt to develop a test powerful enough to accurately measure any differences above and below a mill." (Mower 2001). None have been completely successful. Part of the problem, at least on the Kennebec River is attempting to detect statistically significant differences that are not environmentally significant when the above/below stations are located 24 miles apart. There are serious problems associated with any test comparing sites that are that far apart. While there are no perfect monitoring tools, bivalves satisfy many of the criteria used to select appropriate biomonitors and standardized caging protocols have been developed that facilitate in-situ characterizations of chemical exposure and associated biological effects. No test is powerful enough to overcome the liabilities of an inappropriate experimental design. The required power of this phantom test is never defined. In other words, how will DEP ever know if it has achieved the necessary power? The implication is that if they get the same answer three years in a row that the results will be meaningful. This could happen (or not) just by chance alone. The experimental design is flawed and not scientifically defensible.

Some of the difficulties associated with establishing cause-and-effect relationships come from uncertainties associated with making extrapolations across fish species used in field and laboratory assessments and variations in the response pattern associated with location and duration of exposure Van Der Kraak et al 1998). Many of these problems could be reduced by using the same species of bivalves in national monitoring and assessment programs. These uncertainties demonstrate the importance of characterizing exposure and effects. These are the major analysis elements of ERA. The United States Environmental Protection Agency (USEPA) (USEPA 1998) ERA paradigm provides a focus to environmental monitoring and assessment because it includes a characterization of exposure and effects. We have previously described how caged bivalves could be used to characterize exposure and effects associated with pulp and paper mill effluents (Salazar & Salazar 1997) and provided an exposure-dose-response triad framework to show how caged bivalves could be used to support an integrated ecological risk assessment-based monitoring strategy (Salazar & Salazar 1998

2004). Others have outlined an ERA-based field and laboratory approach to assess endocrine disruption and suggested that the ERA framework is sufficiently robust to accommodate the specific characteristics of endocrine-disrupting chemicals (Kendall et al 1998). The ERA paradigm may be most appropriate for assessing ecological risks associated with pulp and paper mill effluents and various applications of this approach have been outlined (Bright et al 2003, Rodgers & Thomas 2004).

Experimental field approaches are needed to reduce uncertainties associated with fish surveys such as location and duration of exposure, comparisons with different species, and exposure to historical rather than current discharges. Establishing a relationship between exposure, dose, and response in the field would help predict effects and establish causality.

Canada's Aquatic Effects Technology Evaluation (AETE) Program for mining (AETE 1997), recommends the following important questions serve as a framework for monitoring and assessment: 1) Are contaminants entering system? 2) Are contaminants bioavailable? 3) Is there a measurable response? 4) Are contaminants causing the response? Exposure-based monitoring successfully addresses questions 1 and 2. Effects-based monitoring only addresses question 3. Neither approach alone completely addresses question 4. In addition to identifying key questions that need answering in the context of characterizing exposure and characterizing effects, these questions highlight the need to integrate the correct elements. In other words, just because a monitoring and assessment program is integrated does not mean that it includes all the elements necessary to answer the most important questions (Borgmann et al 2001).

1.1 Applied Biomonitoring – Experience and Expertise

Applied Biomonitoring was first contacted by Ed Friedman of Merrymeeting Bay in 1999 because of our extensive experience and expertise in conducting caged mussel studies and more specifically, freshwater studies conducted in collaboration with Environment Canada scientists in Montreal on the St. Lawrence River with Elliptio complanata. One of the reasons for contact was dissatisfaction with respect to the record of DEP conducting over 78 tests on the Kennebec River and having none completely successful. In retrospect, it now appears that DEP never really wanted to conduct these tests and agreed to do them because of the persistence of Ed Friedman. Because of the rush to get a contract in place before the work could begin, there was no agreed upon work plan before we arrived in Maine to conduct the test. To our surprise DEP insisted on placing all of the cages at only two stations outside of the impoundment where fish had been routinely collected for several years. The gradient design that we had intended was never utilized.

Applied Biomonitoring scientists have written the first and only consensus-based protocols on caged bivalve monitoring and accomplished this through the American Society for Testing and Materials (ASTM) through the peer-review process (ASTM 2001). Subsequently, these generic methods were also adopted by Environment Canada as an alternative to monitoring fish as part of the Environmental Effects Monitoring (EEM) for pulp and paper mills in Canada (Walker et al 2002). Similarly, a chapter on these methods was included in Standard Methods for the Examination of Water and Wastewater (Salazar & Salazar 2005a) and in a book on Techniques in Aquatic Toxicology (Salazar & Salazar 2005b). These methods are now well established in many countries and have been developed further into a conceptual framework for freshwater bivalve ecotoxicology (Salazar & Salazar 2006).

1.2 State of Maine Approach

According to DEP Reports (Mower 2001, 2002), the state of Maine has adopted the most stringent environmental regulations for dioxins in the US, with the primary objective of their dioxin-furan monitoring program to assess potential ecological and human health effects by measuring chemical exposure in fish tissues (Mower 2001, 2002). These regulations were established to regulate the discharge of dioxins from pulp and paper mills. A secondary objective of dioxin monitoring in Maine is to document the status and trends of dioxin-furan exposures, evaluate progress in reducing environmental concentrations by compliance with existing regulations, and the need for even more stringent regulations. The third, and most specific objective is to determine if kraft pulp mills are currently discharging dioxins or furans into the rivers of Maine. In practice, environmental exposures of dioxins and furans estimated by measuring concentrations in fish tissues or some surrogate, cannot be higher downstream of a pulp mill discharge than upstream. This is commonly referred to as the "above-below" test. Their monitoring program is based on resident fish, and they rely completely on the ability to detect concentrations of dioxins and furans in fish tissues at 1 part per trillion or less. Since their program only measures tissue chemistry, it could be referred to as exposure-based monitoring. Since it turns out that dioxins and furans are probably not responsible for effects on fish measured in other studies, all of the chemical monitoring conducted by DEP has little value today except that concentrations of dioxins associated with mill discharges have apparently decreased. An ERA-based approach would have provided much more useful information with respect to the status and trends of environments associated with mill discharges.

Over recent years, the concentrations of dioxins and furans in fish tissues have declined (Mower 2001, 2002), and the Department of Environmental Protection (DEP) has expressed concern regarding the ability to detect statistically significant differences in dioxins and furans in fish collected from locations above and below the pulp mill discharge. Many consultant, academic and public environmental groups have expressed concerns whether the observed differences in above-below comparisons are real or associated with the many uncertainties attributable to monitoring mobile fish, including uncertainty associated with mobility, accumulation from other sources, accumulation from previous mill discharges sequestered in sediments, and the reluctance by DEP to collect fish near the mill discharge. In this pilot study, mussel tissues and lipids from the SPMDs were assessed as potential surrogates for dioxin monitoring in fish. The intent was to eliminate concerns with monitoring fish by using a surrogate, such as caged mussels, that could be deployed closer to the mill discharge, i.e., areas where fish are not currently collected.

In the first caged mussel test, ten cages of freshwater mussels (*Elliptio complanata*) were transplanted to each of two stations in the Kennebec River where fish are collected as part of the dioxin monitoring program: an upstream station 13 miles from the mill and a downstream station 11 miles from the mill (24 miles apart). The presence of fish was the primary criterion for station selection, rather than testing the ability of the caged mussels to identify a gradient in chemical concentration in the vicinity of the mill, because DEP wanted a direct comparison of dioxin accumulation between fish and caged mussels. It is virtually impossible to adequately characterize exposure with an above-below comparison. Furthermore, effects in fish were not measured at all the same sites.

Mean concentrations of total dioxins-furans in mussels increased from below detection at the beginning of the test to 4.33 and 4.67 ng/kg-ww at the upstream and downstream stations, respectively, after the 53-day deployment. Total dioxin and furan concentrations in caged mussel tissues were higher downstream than upstream on both a lipid-normalized and a

non-lipid normalized basis, although the differences were not statistically significant. Total dioxins and furans in SPMDs were higher upstream and downstream on both a lipid-normalized and a non-lipid-normalized basis; these differences were not statistically significant. The concentration of total dioxins-furans in fish tissues was significantly higher downstream (4.19 ng/kg-ww) than upstream (2.76 ng/kg-ww) of the mill. However, the lipid-normalized concentrations of total dioxins-furans in fish tissues were higher upstream than downstream, and there was no significant difference between upstream and downstream. These data reinforce the significance of the important questions mentioned earlier regarding where fish exposure to dioxins and furans occurred, whether fish accumulated dioxins and furans from sediment or food that was contaminated from previous, rather than recent mill discharges, or how long ago exposure and accumulation occurred. There was less uncertainty in the mussel data when compared to the SPMD and fish data. For mussels, 38% of the values exceeded the detection limit, compared to approximately 20% for fish and 6% for the SPMDs (Applied Biomonitoring 2002, Salazar & Salazar 2003). Mussels also had fewer non-detects and fewer values that were between non-detect and the detection limit.

There are too many uncertainties in the results from accumulation of dioxins and furans in caged mussels, SPMDs, and fish tissues to unconditionally accept the results and make important decisions regarding the utility of these three methods. Important questions regarding the fish data remain unanswered. Given the large percentage of non-detects, their ability to move and either avoid exposure or accumulate dioxins and furans through other exposure pathways, and the inability to distinguish between current and previous discharges, using the fish data may be problematic since the regulations are based on current and not previous mill discharges. Although the fish appeared to be the most suitable monitoring tool based on the ability to detect statistically significant differences between upstream and downstream concentrations of total dioxins and furans, the question is “why did this difference exist?” Why was the concentration of total dioxins and furans so low in the upstream fish? A statistically significant difference could not be found between the up- and downstream mussels because the mussels accumulated more dioxins and furans from the upstream station relative to the fish. Furthermore, no attempt has been made to collect fish in the impoundment where the mill discharge is located or to measure effects in fish. The regulations and the compliance monitoring is completely exposure-based.

The caged mussel and SPMD data further suggest that using these upstream and downstream locations is inappropriate because the upstream station appears to be contaminated by a source upstream of the mill. The downstream station was too far away to know whether the fish accumulated dioxins and furans associated with current mill discharges, from previous discharges from the mill (i.e., sediment-bound dioxins and furans), or via the food chain. While the experimental design in the pilot study may have been appropriate for comparing dioxin and furan exposures in fish with those in caged mussels and SPMDs, it was not appropriate for addressing the upstream-downstream issues concerning these potential surrogates. Caged mussels and SPMDs should have been placed as close to the mill discharge as possible for a more accurate evaluation of their ability to detect upstream-downstream differences and in a gradient design to determine if dioxins are furans are currently being discharged by the mill.

1.3 Advantages of Era-based Monitoring

In the context of understanding the fate and effects of pulp and paper mill effluents, one major advantage of ERA is that it provides a focus for environmental monitoring and assessment. It serves as a reminder that there should be equal emphasis on charactering exposure and effects and that emphasizing or eliminating one or the other may limit the ability to correctly interpret

the data.

1.4 Advantages of Caged Bivalves

Bivalves have a number of advantages over other species such as fish for characterizing exposure and effects in routine monitoring programs throughout the world (Widdows & Donkin 1992, Phillips 1980, Phillips & Rainbow 1993). 1) They are dominant members of many benthic communities in marine and freshwater environments, which minimizes problems associated with comparing results from different species. 2) They are sedentary, and therefore more appropriate than mobile species as indicators of both exposure and effects. 3) They are relatively tolerant but not insensitive to a wide variety of environmental conditions and chemicals. 4) Most are suspension feeders that pump large volumes of water as they feed, concentrating and integrating chemicals found in water into their tissues. These elevated concentrations in tissues make it easier to measure the chemicals of concern. 5) The measurement of chemicals in tissues also has a greater toxicological significance than measuring chemicals in water or sediment. 6) Compared to fish, bivalves have a limited capacity to metabolize most organic chemicals and therefore more accurately reflect environmental exposure. This limited metabolic capacity is particularly important for using caged bivalves to characterize chemical exposure from pulp and paper mill effluents where the chemicals causing effects in fish appear to be organic. It may also help explain why early efforts to identify chemical tracers of mill effluents were largely unsuccessful. 7) Bivalve populations are relatively stable and can be sufficiently large for repetitive sampling. 8) They can be easily transplanted and maintained in cages for extended periods of time even in areas where they might not be naturally found. Furthermore, caging facilitates measuring both exposure and effects endpoints and almost any clinical measurement such as biomarkers. It is more difficult to conduct field studies with caged fish because caging could affect pathways of chemical exposure and the health of the test animals. 9) Because many marine bivalves are commercially and recreationally important and many freshwater bivalves are imperiled, they are both ecologically relevant with respect to natural resources. Interestingly, bivalves are widely used as sentinel organisms for chemical exposure in marine environments throughout the world, but this approach has not been as well developed for freshwater environments. The use of bivalves as sentinel organisms for biological effects is relatively recent, with more emphasis on marine than freshwater bivalves.

1.5 Characterizing Exposure and Effects over Space and Time

In the context of understanding the fate and effects of pulp and paper mill effluents, one major advantage of the caged bivalve methodology is that it provides a characterization of exposure and effects over space and time and under environmentally realistic conditions. Deploying caged bivalves along suspected chemical gradients helps to identify the source of accumulated chemicals and facilitates the measurement of effects. The location and duration of exposure are controlled as well as the exposure and genetic history of the test animals. This approach has a number of advantages over using wild fish, and can also be used to establish cause-and-effect relationships.

It is important to note that collecting wild fish is not an experiment and does not facilitate hypothesis testing through experimentation. Without an ERA-based approach many will remain skeptical about the fish results. To understand the context of effects studies with fish, experiments must be conducted under environmentally realistic test conditions that simulate the real world. While the State of Maine has emphasized monitoring exposure and Canada has emphasized monitoring effects, the ERA-based approach suggests that it is necessary to

include both in any meaningful monitoring program.

1.6 Source Identification, Predicting Effects, and Establishing Causality

Concentrations of potentially toxic chemicals in pulp and paper mill effluents are decreasing due to better management practices, and therefore, it is extremely difficult to measure these chemicals in the receiving waters. It is also difficult to characterize exposure and effects in some resident organisms, such as fish, because of their mobility. These difficulties hinder the establishment of cause-effect, field-based relationships. Caged bivalves are a potentially powerful monitoring tool for pulp mill effluents because they are sedentary, can be placed in areas of concern, and are able to concentrate and integrate chemicals from water and sediment in their tissues. Bivalves can be used to quantify exposure and effects over space and time (Salazar & Salazar 1997, Widdows & Donkin 1992, Phillips 1980, Phillips & Rainbow 1993, Phillips & Segar 1986). Chemicals in bivalve tissues provide a direct link between chemical exposure and associated biological effects. This relationship provides a way to compare the results of bioassays and population or community responses in the field. The caged bivalve approach is accepted world-wide as demonstrated by its long-term use in Finland and Canada, and more recent use in Argentina, Australia, Brazil, France, Germany, Hong Kong, Japan, New Zealand, Russia, and Sweden (Applied Biomonitoring 2002, Servos et al 1996, Stuthridge et al 2003, ASTM 2001). Environment Canada has recently adopted caged bivalve monitoring as an alternative to the required adult fish survey in their EEM program for pulp and paper mills in Canada. Standardized protocols have been developed through a consensus-based process in an international standards organization (ASTM 2001), and the approach is consistent with the ERA process of characterizing exposure through bioaccumulation and characterizing effects through growth and other endpoints. Although caging freshwater mussels may be restrict their ability to bury in sediment, bioaccumulation and growth are generally not affected because these are filter feeding organisms that generally utilize overlying water for food. Furthermore, we have developed caging methods where freshwater mussels are not placed in compartmentalized cages and are free to bury themselves in contaminated sediment (Salazar et al 2002 2003). Field bioassays with caged bivalves have several advantages over assessments with wild fish, with the most important probably being a defined and controlled exposure period. In addition, studies with caged bivalves offer more experimental control, although not as much as in standard laboratory bioassays.

Important links can be established between exposure and effects by using cultured or wild mussels, transplanting them along suspected chemical gradients, and analyzing their tissues for groups of chemicals suspected of inducing effects in fish. For example, in the caged mussel study at the Port Alice Pulp and Paper Mill we established a significant relationship between campesterol in mussel tissues and mussel growth rates (Applied Biomonitoring 2000, Salazar & Salazar 1999). While this does not establish causality, it is a working hypothesis that could be used for other chemicals as well. Conversely, in a series of caged mussel studies in Canada as part of the EEM program, no chemical measurements were made and no information was gained regarding possible relationships between exposure, dose, and response (Martel et al 2003). There is no question that controlled laboratory tests within the internal waste streams of the mill could be used to establish links between exposure, dose, and response as suggested recently (Hewitt et al 2003). The conceptual approach of first confirming that effects in the field have occurred and then moving into the lab to confirm causality could be viewed as an ERA-based approach. However, characterizing exposure an effects as part of an iterative process in the laboratory and the field would provide additional insights into causal relationships. Furthermore, using the caged bivalve methodology provides a practical approach for gathering the information necessary to establish those relationships under environmentally realistic

conditions with field experimentation.

Field studies with caged bivalves are appropriate for both marine and freshwater ecosystems. Pulp and paper mill effluents are discharged to both freshwater and marine environments and can be assessed with both freshwater and marine bivalve species. As a whole, freshwater mussels may be as important a resource as fish because many have an imperiled status. There are, however, many freshwater bivalve species that are not of imperiled or endangered status, and these are the species used in the work cited herein and recommended for future studies. Marine bivalves, which include mussels, oysters, and clams, are also an economically important resource, but few marine species have the imperiled or endangered status that many freshwater species have. Both freshwater and marine bivalves are recommended for environmental monitoring to better understand the consequences of chemicals in the environment and protect those species that are imperiled. All studies should be conducted with abundant species, and collections should limit harvesting effects on native populations.

1.7 Establishing Links Between Effects in Bivalves and Fish

We have been working with Environment Canada scientists at the St. Lawrence Center in Montreal over the last five years to develop a suite of biomarkers for marine and freshwater bivalves that have been tested upstream and downstream of a municipal effluent and other sites that could be applied to pulp and paper mill effluents. These include an assay for immunocompetence (Blaise et al 1999, 2002), several biomarkers including cytochromeP450, DNA damage (Gagne et al 2002), a vitellin assay that was linked to possible endocrine disruption and concentrations of coprostanol in caged mussel tissues (Gagne et al 2001a,b,c) and experimentally-induced sex reversal in mussels caged downstream of a municipal effluent for a period of 1 year (Blaise et al 2003). During this process we have also measured mussel growth to help calibrate the sensitivity of the various biochemical responses. Furthermore, we have developed a benthic cage that facilitates holding mussels in bottom sediment for a period of 1 year to measure all of those responses (Blaise et al 2003, Salazar et al 2003). Collectively, these studies demonstrate that most effects endpoints commonly measured in fish can also be measured in caged bivalves. Perhaps more importantly, similar effects on hepatic vitellin and reproductive function were demonstrated in spottail shiners at several of the same sites downstream of the same municipal effluent by other Environment Canada scientists at the St. Lawrence Center (Aravindakshan et al 2004). Caged freshwater mussels (other species of *Elliptio*; *E. buckleyi* and *E. icterina* also showed significant endocrine and reproductive effects downstream of a pulp and paper mill in Florida that were similar to those reported for largemouth bass (Kernaghan et al 2004).

1.8 Integrated Monitoring Strategy

An integrated ERA-based monitoring strategy is suggested to reduce uncertainties in current environmental assessment approaches and to establish causality (Salazar & Salazar 2004). There are three basic components to ecological risk assessment: problem formulation, analysis, and risk characterization. Most assessments of pulp and paper mill effluents have not fully utilized the ERA approach in all aspects of their monitoring, but have used only selected elements in the lab and the field. In the examples provided earlier, the State of Maine emphasizes exposure characterization and Environment Canada emphasizes effects characterization. These are their basic conceptual models. However, if the conceptual model developed during the problem formulation is flawed or deficient, the resulting components that follow will also have increased uncertainty. Therefore, results of the analysis phase, which is supposed to include an integration of characterizing exposure and characterizing effects, may

be biased because the analysis is either exposure-based on effects-based. Similarly, there can be no reliable characterization of risks in the third phase, because the appropriate elements have not been included previously. It may be misleading or inappropriate to suggest that either exposure-based, stressor-based, or effects-based monitoring will successfully reduce uncertainty and answer critical questions that remain with respect to effects on wild fish. An integrated ERA-based approach is the best way to accomplish those tasks.

2.0 DIOXIN PEER REVIEW PANEL REPORT: A Summary & Critique

The review panel was asked to review, comment and make recommendations on the following relating to the Dioxin Above/Below (A/B) test: 1). ***Sensitivity***; 2) ***Specificity***; 3) ***Variability***; and 4) ***Overall Validity of A/B Test***. Their report is structured according to these 4 elements (Adams et al 2004). The review panel considered monitoring data from prior years as well as 2003 data. These data included dioxin congener concentrations from small mouth bass (SMB and) white suckers (WHS) from the sites shown in Figure 1 and caged mussels from studies conducted in 2000 and in 2003 and semi-permeable membrane devices (SPMD) deployed in 2003. We have outlined differences in approach for 2000, 2003, 2005, and 2006 in Table 1 and differences in results for 2000, 2003, 2005, and 2006 in Table 2. As mentioned previously, the 2006 mussel growth rate results appear spurious and will not be discussed in detail here. It is important to keep in mind however that the review panel completed their report before the 2005 and 2006 caged mussel studies were completed.

2.1 Sensitivity of the Above/Below Test

The sensitivity of the Above/Below test depends strongly on the precision and accuracy of the analytical data. The review panel discusses identifies improvements in the fish tissue chemistry data and concludes that the 2003 chemistry data are generally more precise than data produced prior to this time. However, they also point out that there continues to be a problem with the t=0 mussel chemistry data, where the two laboratories involved had poor agreement in an intercalibration exercise. Concerns were also raised with respect to the accuracy and precision of the mussel percent lipid data. Some of these observations are now moot since dioxins and furans and no longer measured. Although we included percent lipids in the 2003 report, they were not included in the DEP 2005 Monitoring Report (Maine Department of Environmental Protection 2006, Maine Department of Environmental Protection Surface Water Ambient Toxics Monitoring Program 2006). In fact, there is little documentation or citations to explain the results of the 2005 caged mussel monitoring study. In fact, it is referred to as the vitellin study and yet the vitellin results in 2005 were much different than in 2003. It appears as though the biomarker results from the 2005 caged mussel study were simply copied and pasted into the report without any clarification, explanation, supporting information, or citations. Expectations are that the 2006 report will be more of the same.

2.2 Variability

The 2003 data indicate that there are important sources of variability in the data that can interfere with an A/B test. Many of these issues are related to the use of wild fish as indicator species. However, there is a value in using fish in the A/B test because they are caught and consumed by the human population. The peer review panel also fails to make the distinction between fish data for purposes of human health versus fish data for ecological effects and potential effects associated with mill discharges. The panel does identify that surrogates can be an important complement to fish and should be part of an A/B test because they can provide valuable information when fish species provide more variable results.

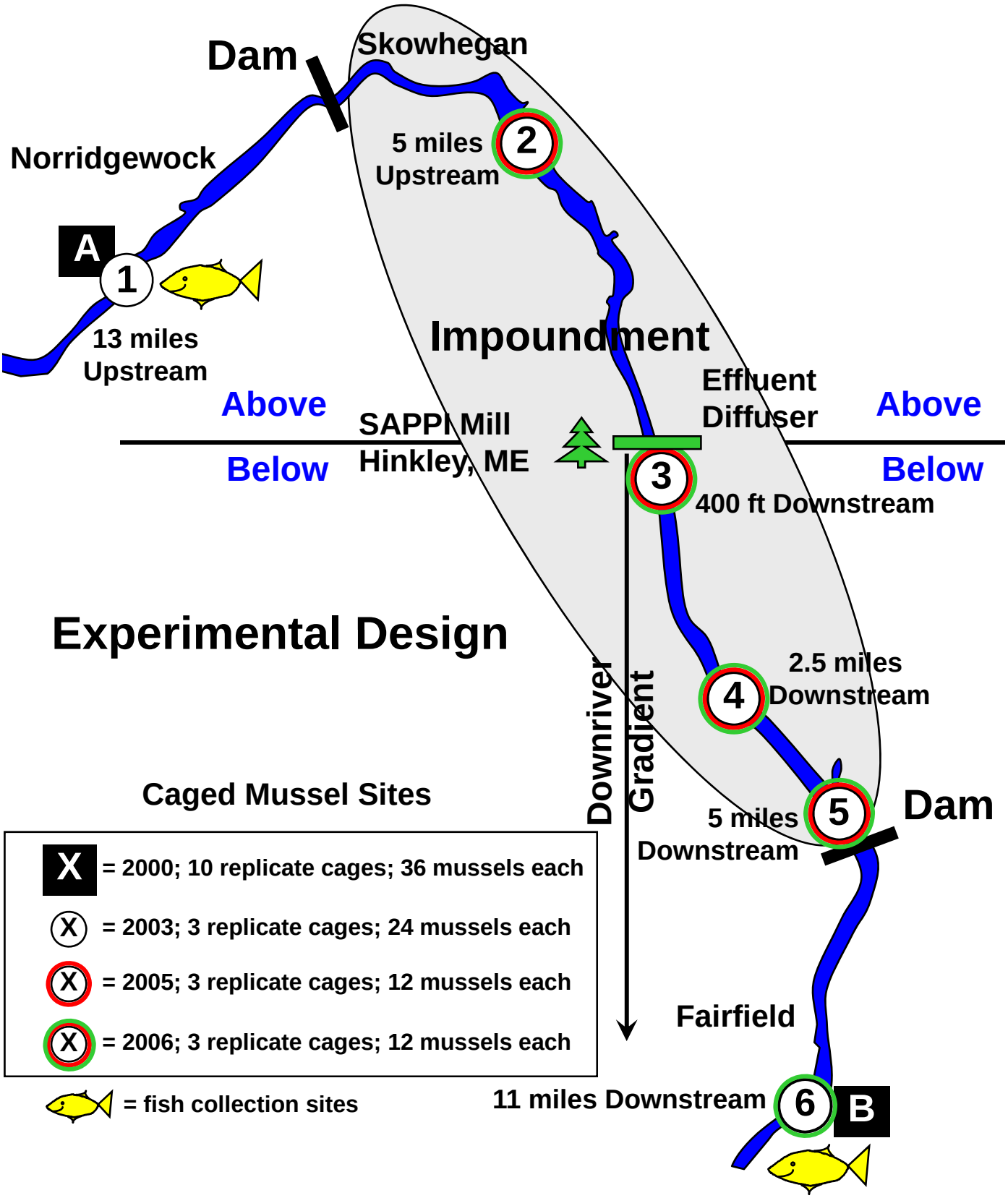


Figure 1. Station locations on the Kennebec River – 2000, 2003, 2005, 2006

Table 1: Kennebec River - Approach

Differences: 2000 vs 2003 vs 2005 vs 2006

	2000	**2003**	**2005**	**2006**
Size (length; mm) range @ T_o	58.0-67.2	62.4-66.9	58.1-65.2	65.7-75.6
Overall range (mm)	9.2	4.5	7.1	9.9
Total # mussels used	900	504	180	216
# mussels T_0 Chemistry	180	72	36	36
# mussels deployed	720	432	144	144
# mussels/station: growth	360	72	36	36
# mussels/station: biomarkers	na	12	12	12
Sampling design	Above/below	Gradient	Gradient	Gradient
Number of stations	2	6	4	5
Number of mussels/cage	36	24	12	12
Exposure duration	53-d	66-d	71-d	69-d
Exposure period	8/3 - 9/26	7/28 – 10/7	7/28 – 10/7	7/21 – 9/28

Table 2: Kennebec River – Results

Differences: 2000 vs 2003 vs 2005

	2000	**2003**	**2005**
% Survival	99.7%	98.8%	100%
Mean Temperature	**20.7°C**	**21.7°C**	**21.8°C**
Max Temperature	**28.0°**	**25.8°**	**26.5°**
Min Temperature	**16.2°C**	**15.5°C**	**15.9°C**
% Δ weight (g)	3.3%	6%	12.5%
% Δ length (mm)	0.4%	1.1%	2.1%
% Δ tissue wt. (g)	**15.5%**	**25.3%**	**4%**

Variability in Percent Lipid Values:

The peer review panel suggests that options for surrogates include caged mussels and semi-permeable membrane devices (SPMDs). In contrast to fish, caged mussels provide a more uniform population and therefore a large population of individuals is not required for analysis. Each of the 3 replicates at a site consists of a pooled sample of 20 individuals for chemical analysis, ensuring less site-specific variability in data and strengthening the confidence in A/B test results. Additionally, caged mussels can be deployed closer to putative sources and reflect primarily water-borne exposures, not bottom sediment exposures or historical contamination. Should mussels be used as a surrogate in the DMP, the following is recommended: 1) 4 arrays per site should be used rather than 3; 2) only one above and below site approximately equidistant from each mill should be used; and 3) for statistical purposes, each array should consist of 5 pools of approximately 12 mussels each instead of 3 pools of 20 mussels each. They do not understand the importance of pooling more mussels for growth, chemistry and biochemistry. A number of studies have shown that for chemical analysis the optimum number of replicates for minimizing costs and maximizing discriminating power is between 20 and 25 mussels (Baez & Bect 1989, Besada et al 2002, Daskalakis 1996, Gordon et al 1980, Mitra & Choudhury 1993, Wright et al 1985).

It is important to note that the peer review panel clarifies that mussels do not have a lower bioaccumulation potential simply because they are at a lower trophic level. This is one of the myths perpetuated by DEP in previous reports (Mower 2000, Mower 2001) in attempting to justify their initial decision to proceed with fish and SPMDs and monitoring tools. The panel also reports that mussels lack cytochrome P450-based oxidative detoxifying enzymes that vertebrates employ to detoxify and excrete xenobiotics. While this is not exactly true (mussels have a much more primitive system), this is one of the reasons why mussels are commonly used in Mussel Watch Programs around the world. This fact takes on even greater importance when considering that the responsible chemical could be a PAH-type compound associated with natural wood products and not directly related to mill processing. The features of the 2003 mussel study design indicate that the mussels were deployed for a sufficient duration to reach a steady state with respect to ambient dioxin levels in the water and will reflect the current levels of bioavailable dioxins in the water column in all forms (dissolved and particulate).
The peer review panel also dispels the assertion by DEP (Mower 2000, 2001) that SPMDs are a more promising monitoring tool. The panel concludes that the limitations of SPMDs outweigh their advantages. For example, SPMDs are not a realistic surrogate because they are not representative of the food chain, their sensitivity of chemical detection is low, they are not specific to dioxin compounds released from paper mills, and they have several technical difficulties. On the other hand, the biological tests also have their own unique sets of advantages and limitations but these generally complement each other which provides a much more comprehensive basis for assessing A/B differences. The other big advantage is that fish and mussels are appropriate for characterizing exposure and effects in a more natural way than the artificial SPMDs. SPMDs are not living organisms.

2.3 Overall Validity

The review panel's recommendation is that a preponderance of evidence (POE) approach be considered using analysis results from multiple species in an A/B protocol. The panel properly identifies that aquatic environments are variable and complex ecosystems that are controlled and regulated by a variety of physicochemical and biological processes. In addition, aquatic organisms are subjected to a variety of natural and anthropogenic stressors, both of which vary spatially and temporally. High variability of environmental factors combined with synergistic and

cumulative interactions of these factors can complicate the interpretation and evaluation of environmental data. Because of this high variability and complex interactions among ecosystem components, it is often necessary to apply multiple endpoints or measures when assessing the effects of environmental factors on biological components of aquatic systems. Use of a variety of multiple endpoints which reflect different sensitivities, specificities, and response time scales to environmental factors are needed in environmental monitoring and assessment programs (Adams et al. 2002). In assessing the effects of environmental factors on aquatic organisms, individual variables or endpoints are generally inadequate indicators of effects and the exclusive use of only one endpoint may lead to invalid conclusions regarding the true status of a system (Capuzzo 1985). Multiple endpoints or measures that are complementary to each other relative to their respective advantages and limitations are needed in the design of environmental assessment programs (Adams et al. 2003).

As suggested previously, Borgmann and his colleagues (Borgmann 2000, Borgmann et al. 2000) have concluded that traditional approaches such as the sediment quality triad successfully address questions 1 and 3 (associated with exposure) but do not directly address questions 3 and 4 (associated with effects). A growing number of scientists agree with that assessment and is one of the reasons why we have developed the exposure-dose-response triad; i.e., to place more emphasis on bioaccumulation and directly address questions 2 and 4.

Borgmann et al. (2001) suggest that: 1) bioaccumulation is the most direct and reliable method of estimating bioavailable chemicals; 2) comparison with critical body residues provides a means for identifying the cause of toxicity; and 3) this approach has less uncertainty than relying on chemical concentrations in water or sediment.

The sediment quality triad is generally considered an effects-based approach because two of the three basic elements quantify effects. It generally consists of sediment chemistry, a biological response measured in the laboratory and field studies that generally utilize benthic community structure. By contrast, we have referred to the exposure dose response triad as an EcoRisk-based approach because it places equal emphasis on characterizing exposure and effects as does the ecological risk assessment paradigm. In the context of a sediment assessment, it would consist of sediment chemistry (could include water as well), tissue chemistry (dose) and effects measures such as bioassays or benthic community structure. Rather than two effects measures, it includes one element of external exposure (sediment chemistry), one element of internal exposure (tissue chemistry) and one element of response (biological effects). In the context of weight of evidence or multiple lines of evidence, it seems as though the exposure dose response triad generally includes more different kinds of information that is more consistent with ecological risk assessment. In a larger context, a recent SETAC Pellston Workshop concluded as part of their re-evaluation of water quality criteria that a more holistic approach was necessary and that this should include harmonization of water, sediment, and tissue quality guidelines (Reiley et al. 2003).

This is the problem with the suggested POE approach by the peer review panel. All lines of evidence are not created equal. More to the point, the above/below approach with fish is an observational approach while the caged bivalve methodology is an experimental approach. To clarify the difference in this application we need to make the distinction between POE (WOE) and multiple lines of evidence. By advocating two observational methods (WHS and SMB, and caged mussels the panel has arbitrarily selected two approaches that are not scientifically defensible and only one that is. It would have been more appropriate to select one or two observational approaches if DEP is insistent on continuing the ill-advised above/below test with indigenous fish ***AND*** two experimental approaches such as caged mussels and caged fish for

example. In other words, there are more than the four A/B tests evaluated by the panel. Other experimental approaches include fish embryos (Bailey et al. 2006, Larsson & Forlin 2002, and caged fish (van den Heuvel 2005). A summary of Bailey et al. 2006 is provided in the Appendix. Although there could be problems with some caged fish tests because of the need to swim or move about, the main advantage of embryo tests or eel tests is that they do not have these same limitations. There are also approaches that used sediment bags that assess uptake of chemicals by sediments placed along suspected chemical gradients like the caged mussels (Nix & Merry 1990, Nix & Daykin 1992, Nix et al 1993). There are also the observational biomonitoring methods developed by DEP and the experimental field approaches such as rock baskets. It is not clear why these methods have not been used on the Kennebec River in conjunction with the caged mussel studies. All of these approaches would be valuable complements to help characterize and understand physical, chemical, and biological processes associated with the mill effluent.

2.4 Application to the Dioxin A/B Tests

Four separate A/B test are possible: SMB, WHS, mussels and SPMDs. Each test has its limitations and advantages relative to application to the A/B test. Most of the possible combinations of tests complement one another (ie, limitations of one test are supported by the advantages of another). With respect to independent applicability, not each test is equal relative to the four criteria of sensitivity, specificity, variability, and overall validity (the four criteria we are to address). If each test was equally sensitive, specific, variable, etc. then perhaps an independent applicability approach could be applied. However, if each test was similar relative to all of these criteria, then perhaps only one test would really be necessary. Each test system is associated with diverse advantages and limitations. As an example of using independent applicability for assessing the A/B test assume, for example, that the WHS, mussel and SPMD tests all had to demonstrate that there was no A/B difference in dioxin levels in order for the mills to continue operating. One of the main problems with WHS is that they may reflect historical rather that recent discharges. With respect to variability, SMB data shows very high variation in percent lipid (and suspected unrealistically low body levels) that may skew the lipid-normalized dioxin to higher levels than are actually present in the environment. Because each test has at least one major limitation, each test alone would not likely be defensible as a valid A/B test under the scrutiny of a legal challenge. When multiple tests are evaluated together and a POE approach (eg, 2/3 criterion) is used, the conclusions are more scientifically defensible. This is a good point but one must keep the problems associated with any reference site in mind. The reference site is really an impossible dream, particularly in terms of effects monitoring (Landis 2000). While the AMEC review (Gwinn 2007) raises questions about comparing results relative to T_0 values, comparing results with reference sites is also fraught with difficulty. This is one of the reasons we have promoted the gradient design (ASTM 2001).

3.0 AMEC's REVIEW: A Summary & Critique

A critical review of the analytical results from the 2005 Caged Mussel Study was provided by AMEC in January 2007 and was also used as part of the current review for Friends of Merrymeeting Bay (Gwinn 2007). We must acknowledge here that we agree with many of the criticisms raised by the AMEC review and must assume some responsibility for the some of the ambiguous interpretations in the 2000 and 2003 Applied Biomonitoring reports, particularly those that included the vitellin assay. With that caveat it should also be made clear that the vitellin assay and other biomarker tests were performed independently by Environment Canada and were not the prime focus of the caged mussel study and Applied Biomonitoring only participated in the 2000 and 2003 caged mussel studies. Furthermore, it was our

understanding that the 2003 vitellin component was only a range-finding test and that future studies would include more replication. In addition, while the caged bivalve protocol has undergone extensive peer review during the development of an ASTM Standard guide, the vitellin assay as performed by Environment Canada has not undergone a similar review and is subject to considerable variability in results and data interpretation based on the number of males, females, and indeterminate sex in each replicate. For those reasons and others mentioned previously with respect to the widespread use of mussels, problems with the test identified in the AMEC review are not sufficient to discard the test. In other words, many of the problems identified in the AMEC memo and the current critique is that the problems were largely associated with the way the test was conducted and the interpretation of the data rather than some inherent flaw in the methodology. There are few inherent flaws in the caged bivalve methodology since it has undergone intense scrutiny through the peer review process. This does not mean however that investigators without extensive experience and expertise can conduct a successful test. The methods are simple but not necessarily easy and the slightest deviation from established protocols. For example, if the mussels are not allowed to equilibrate for a sufficient period of time and incorporate the water that was in the shells at either the beginning or end of the test the growth rates will be uncomparable from year to year. Similarly, even though the protocols clearly state that enough room should be allowed in each chamber and that the plastic cable ties not so tight as to prevent the mussels from opening or they will be in poor health. It seems most likely that one or both of these occurred in the 2006 caged mussel experiment because the results were so different from previous years. It is also possible that the BOT and EOT measurements were not properly paired but if this were the case one would expect more random variability in positive and negative growth rates.

Again, it is important to remember that Applied Biomonitoring only participated in the 2000 and the 2003 studies. While the 2005 study was conducted primarily by FOMB personnel the results appear largely comparable with the 2000 and the 2003 results. The 2006 study was conducted by DEP personnel. Applied Biomonitoring was not consulted on the experimental design or the conduct of the test. This is not to say that Applied Biomonitoring scientists are the only ones in the universe capable of conducting this test. There are scientists around the world currently using this or some closely related approach. However, it takes dedication and persistence and DEP has not demonstrated either with respect to the caged bivalve approach. They have been resistant from the beginning and have never truly embraced the methodology. They also do not seem to understand the importance of either calibrating bioaccumulation or biomarkers with the growth results as a form of ground-truthing. In fact, it was only through the course of re-evaluating the growth results that we identified problems similar to those identified in the AMEC review.

It goes well beyond that however, using only 4 animals per cage or a total of 12 per station does not appear to be adequate. Our Canadian colleagues maintain that this is all that is necessary to detect statistically significant differences among sites. However, this hypothesis has yet to be tested on the Kennebec River. This also points out another problem with respect to lack of oversight on behalf of FOMB and DEP in 2005 and 2006 (when Applied Biomonitoring was not involved in the studies). In 2005 Environment Canada was sent all 12 animals per cage with the assumption that all would be used in the vitellin measurements. It was only discovered after the fact that Environment Canada had only used four animals/cage. Given the variability in mussel growth rates, bioaccumulation and other biomarkers in other studies this does not seem like enough. Most studies on bioaccumulation suggest between 20 and 25 animals per sample to reduce the variability while keeping costs at a minimum (Baez & Bect 1989, Besada et al 2002, Daskalakis 1996, Gordon et al 1980, Mitra & Choudhury 1993, Wright et al 1985). It was only recently that DEP clarified the contract to include vitellin measurements on all the samples sent.

It was also requested that Environment Canada provide some estimate of tissue weights since they were not measured in 2005 and were probably not going to be measured in 2006. In 2000 and 2003, we measured tissue weights and shell weights and considered them to be an integral part of the POE, MLO approach. WE attempted to estimate tissue weights based on regressions developed from previous studies but the results appear quite different than expected or measured in 2000 or 2003. It is not clear if this is a problem with the regressions (which are not very accurate) or the measurements made in 2005 and 2006, particularly in 2006

It is important however to make a clear distinction between the basic caged mussel bioassay (which generally includes bioaccumulation and growth) and the biomarker assays. This distinction is important to a reasonable interpretation of the AMEC criticisms. The first issue has to do with the use of a "reference station." We generally do not like to use that term because we feel that it is virtually impossible to obtain a true "reference site" in the field. (Landis 2000) Therefore, we have promoted the use of the gradient design instead of comparisons with reference sites. While we agree in concept with the points raised by AMEC with respect to beginning-of-test comparisons we have typically used this approach as a benchmark in addition to results from the gradient design and regression analysis. We still believe that this is a useful piece of information and should be used in a risk assessment-based approach that includes mutliple lines of evidence and weight of evidence. It should be made clear that Environment Canada scientists are the ones who introduced the term "reference" and used it in their reports. The other problem is that the Environment Canada reports were very limited in scope and we apparently never completely understood the limitations of the test and the results. As an example, it was not clear to us that biomarker results could not be compared from year to year. We are not sure why this is so different from mussel growth rates which can be compared from year to year and to establish long term trends. We only learned this through discussions with Francois Gagne of Environment Canada in December of this year as part of this review. Another reason for these apparent misunderstandings is associated with the fact that while we conducted five different caged mussels studies with Environment Canada on the St. Lawrence River, Applied Biomonitoring was not part of the contractual mechanism with Environment Canada and was not responsible for the contents of its reports. In retrospect, our verification of their results was less than it should have been. Similarly, there was nobody within DEP or Friends of Merrymeeting Bay who reviewed these results in great detail. Finally it is also important to note that we were not involved in the 2005 Caged Mussel Study. This study was conducted independently by Friends of Merrymeeting Bay and DEP. Similarly, we were not involved in the conduct of the 2006 caged mussel study.

Given that we did not conduct the caged mussel studies in 2005 and 2006 we should not be held completely responsible for the way the tests were conducted, the way the results were analyzed or the data interpretation. Therefore, many of the questions included in the AMEC review are directed at DEP. 1) The clarity of the reference station has already been addressed. In summary, the Reference Station was selected and compared by Environment Canada. 2) Has DEP verified the gonado-somatic index (GSI) that underlie the statistical analysis about this metric? We believe that the short answer is no and we have identified some of the same problems as AMEC. 3) Has DEP verified the Lipid Peroxidation (LPO) calculations for either the digestive gland or the gill. We believe that the short answer is no and we have identified some of the same problems as AMEC. 4) Has DEP evaluated the Vitellogenin-like compound data with respect to other stations on the Kennebec River? We believe that the short answer is no and we have identified some of the same problems as AMEC. More specifically, we agree that there are no significant differences in vitellin response among stations on the Kennebec River in 2005. Furthermore, we agree that the results of the 2003 caged mussel study are subject to various interpretations. The 2004 Applied Biomonitoring report suggests that the

protein-normalized ALP measurements were “not as responsive” as the other vitellin measurements and that is why they were not presented. If the protein-normalized data are the most meaningful, as suggested in other studies, but not explained in the Environment Canada report, we too would have concluded that the mill was not having an effect on mussel vitellin responses. In the end, the AMEC memo concludes that “…one could argue strongly that these data provide ample evidence against continued caged mussel research and sampling.” We would like to make the distinction between additional vitellin sampling as previously conducted and any caged mussel studies whatsoever.

We believe that in-situ testing along a suspected chemical gradient is one of the few ways to characterize and understand processes on the Kennebec River associated with the mill discharges. There are other methods that could be used but the caged bivalve methodology is one of the few that have undergone rigorous peer review and become an ASTM Standard Guide (ASTM 2001). We have suggested on numerous occasions that there is a disconnect between mill monitoring and other monitoring methods within DEP. For example, when we first suggested caged bivalve monitoring in 1999, we used the analogy of the rock basket monitoring currently being used by DEP to establish links with these other monitoring programs. Another recently developed method is the use of In situ hatchbox studies (fish embryos) to evaluate water quality effects Bailey et al. (2006). A summary of this approach is included as an appendix at the end of this report. In other words, the use of in-situ testing along suspected chemical gradients is a well-established protocol with various methodologies available to use in testing. The DEP program would also be helped by more physical-chemical monitoring of the Kennebec River and a routine peer review process to make sure the program stays on track and modifies its approach as more data become available. One approach that should be carefully scrutinized is the continued sampling of fish above and below the mill on the Kennebec River when those sites are separated by 11 miles. With this experimental design as a basic element of DEP monitoring, it is not surprising that they have tried over 78 tests and none have been completely successful. The problem lies not with these basic methodologies that have been standardized over time but the way these methods have been used by DEP. Similarly, it may eventually be shown that the vitellin assay can be a useful monitoring tool with the appropriate application and data interpretation. Now however, this should be left to an outside and unbiased peer review committee to decide before moving forward with additional vitellin measurements.

4.0 ENVIRONMENT CANADA’S VITELLIN ASSAY

Applied Biomonitoring assisted Environment Canada scientists with caged mussel studies on the St. Lawrence River over a 5-year period between 1999 and 2003. Their work on endocrine disruption in bivalves and supporting studies have recently been summarized (Gagne and Blaise 2003). This is important because both Applied Biomonitoring and Environment Canada have been criticized recently for both the caged mussel methodology and the vitellin assay. This series of papers have resulted in a number of publications in their peer-reviewed literature that gives credibility to the work. Furthermore, in addition to Environment Canada St. Lawrence Center being recognized as a leader in the vitellin assay, the fact that they contracted Applied Biomonitoring on the other side of the continent is a testament to Applied Biomonitoring scientists being a leader in the field as well. While the series of papers demonstrates the credibility of both methods, this does not mean that the systems are perfect. We are continually improving the caged mussel methodology with each test. Similarly, the vitellin assay improves with each test as well. While the caged bivalve methodology has been standardized through the consensus-based process at ASTM, the vitellin assay has not undergone a similar review. Surprisingly, the vitellin results were not included as part of the peer review panel report. This

indicates another problem with DEP in that they directed the panel and asked the questions they wanted answered.

It should also be mentioned that Environment Canada scientists are not the only ones developing the vitellin assay. A number of papers have been written on the development and use of the vitellin assay. Problems that have occurred in the implementation of the test do not merit discarding the test completely but fixing the problem. While one of the strongest reasons for using this test is the pairing of the vitellin assay in fish and bivalves. This will help explain some of the fish results. Alternatively, as long as DEP persists in monitoring fish above and below the mill on the Kennebec with sampling stations located 24 miles apart, the results will never be explained and the results never scientifically justifiable.

Perhaps the most compelling evidence for endocrine disruption when we experimentally induced sex reversal in *Elliptio complanata* in a 1-year benthic cage experiment (Salazar et al 2003). Preliminary monitoring of natural Elliptio populations showed a higher percentage of females downstream when compared to the upstream location. The field experiment was designed to determine if this form of endocrine disruption could be experimentally induced and the test was successful in doing just that. Therefore, the weight of evidence and multiple lines of evidence has demonstrated that some component of the City of Montreal effluent is capable of endocrine disruption. It is not clear that the components of the pulp mills on the Kennebec River or elsewhere are capable of the same effects. Results from monitoring fish in Canada and elsewhere suggest that dioxins and furans are not the component of altering sex steroid ratios in fish and causing adverse effects. Although the chemical responsible has not been identified, and apparently occurs in very low concentrations, the effects persist. It is important to note however, that even if effects were measured above and below the mill on the Kennebec River outside of the impoundment, any result would be scientifically unjustifiable because of the large distance between the sampling sites of 24 miles. One of the apparent reasons why DEP continues to monitor fish above and below the mill is for human health reasons. This dichotomy between environmental effects and human health effects has clouded the justification for this fish monitoring for some time. The problem is exacerbated by the poor quality of the dioxin monitoring reports (and the surface water monitoring reports). The biomonitoring reports are quite different for example. There appears to be little or no internal or peer review of the dioxin monitoring reports.

The problem of the 24-mile distance between the fish sampling stations above and below the mill on the Kennebec River is exacerbated by the use of only 12 vitellin samples per station for the caged mussels in any attempt to compare these results with the fish results. Combining the problem of the 24-mile distance with only 12 samples per site increases variability and the validity of any comparison.

5.0 A COMPARISON OF ALL CAGED MUSSEL STUDIES

5.1 Experimental Design

Problems in experimental design are clearly identified in Figure 1 and in Tables 1 and 2. The figure shows the overall experimental designs of the caged mussel studies conducted on the Kennebec River in 2000, 2003, and 2005. The three most important geographical features are the location of the dam above Skowhegan, the dam at Shawmut, and the impoundment created by these two dams. With respect to monitoring potential discharges from the mill, all stations north of the mill are considered “above” and all stations south “below” although fish are routinely caught at only two locations outside of the impoundment. Caged mussels were deployed at six

stations on the Kennebec River in 2003. Differences in approach and results are identified in Tables 1 and 2, respectively. Differences in growth rates associated with distance from the mill are shown in Figures 2 and 3, respectively.

In 2000, there were only two stations, co-located with fish collections sites designated “A” and “B” in Figure 1. In 2003, Stations 1 and 6 were located as in the 2000 caged mussel pilot study and outside of the impoundment. Stations 2, 3, 4, and 5 were all located within the impoundment and utilized a gradient design to examine trends in various metrics with distance from the mill dischare. Station 2 was approximately 5 miles above the effluent diffuser within the impoundment. Station 1, a second upstream station, was approximately 13 miles above the effluent diffuser, at the same upstream location used in the 2000 study. Stations 3, 4, and 5 were located along a suspected chemical gradient downstream from the effluent diffuser at distances of 400 ft, 2.5 miles, and 5 miles. Station 6 was located approximately 11 miles from the diffuser, outside of the impoundment and below the Shawmut Dam. It was the same location used for the “below mill” station in the 2000 study. It is important to note how the number of mussels decreased substantially from 720 in 2000 to 432 in 2003 and only 144 in 2005 and 2006. It is also important to identify that only 12 mussels per cage were used for biomarker analyses in 2003 and it was considered part of a range-finding experiment. These data were never intended to be used for making regulatory decisions. As suggested previously, the intent was to increase that number in 2005 but that did not happen either. The utility of the 2006 biomarker results remain to be determined. If the mussels were not properly equilibrated for the WAWW growth measurements, this will not affect the biomarker results. If however, the mussels were stressed in the cages because the cable ties were placed to tightly around the shells, the biomarker results will be affected as well. Therefore, the biomarker results may provide insights into which problem caused the aberrant growth rate measurements. This is another reason for always measuring growth; to calibrate bioaccumulation and biochemical measurements. This should be part of the POE/MLO approach but neither results are discussed in the 2005 caged mussel results on the DEP website. In fact, it is referred to as the vitellin test. Measuring growth is an integral part of the test but DEP has never acknowledged that fact. Furthermore, when we initially suggested problems with the vitellin assay DEP became uncomfortable because we were questioning the validity of the results after first suggesting the method. We should not be held responsible for suggesting methods that appeared appropriate and then were not carried out as we had intended. DEP also suggested that the vitellin analyses continue on the 2006 samples to help establish links in biochemical effects in caged mussels and resident fish. Surprisingly, DEP did not deploy mussels at Station 1 where fish are collected as part of the A/B test. Given that fish are only collected at two sites on the Kennebec River it seems that mussels should have been caged at both locations to help identify these links, if any.

Due to budgetary constraints and other issues, the 2005 caged mussel study only included stations within the impoundment; i.e., Stations 2,3,4,5. A similar experimental design is being used in the 2006 caged mussel study which is currently underway. Also as a result of budgetary constraints Figure 1 shows that the number of replicates/site also decreased between 2000 and 2003 and again between 2003 and 2005. Fewer growth measurements were also made in 2003 and 2005 when compared to 2000 and more emphasis was placed on biomarkers.

This decision was made on the basis that the biomarkers were more specific to potential endocrine disruption and effects associated with current mill discharges. While this decision was entirely justified, it now appears that the biomarker results may be more variable than the growth results and with the low level of replication (only 12 mussels/station used for biomarker

analyses) not environmentally significant. While the growth results are less specific and known for relatively high variability, more work has been conducted on factors affecting growth in bivalves than with the relatively new biomarker and particularly vitellin analysis in freshwater mussels.

Using the caged bivalve model in combination with the gradient design model is a potentially powerful tool but all of the exposure and effects measurements need to be refined for monitoring pulp mill effluents on the Kennebec River. It should be recognized that monitoring programs around the world with much larger research and development programs have also been stymied in developing an environmentally realistic monitoring program for this application. Therefore, the DEP is pursuing the most appropriate course of action in their program but it will take time. It was probably unrealistic to expect a "silver bullet" indicator that would show conclusive results immediately. Nevertheless, the existing program would probably be enhanced by adding the more generic approach used for other applications in Maine rivers and streams that include experimental field approaches similar to the caged mussel approach (rock bags & rock baskets). The addition of more traditional metrics such as benthic community structure and productivity would also be helpful. Finally, to help eliminate the possibility that natural factors might be affecting the effects measurements, it would be useful to include more physical-chemical monitoring in addition to the temperature measurements made in 2000, 2003, and 2005. Collectively, this information integrated into a comprehensive monitoring program based on ecological risk assessment would provide the most conclusive results in confirming the relationship between chemical exposure and associated biological effects. This approach would also help identify the chemicals causing the observed effects.

5.2 Mussel Growth

Given the differences in personnel making the measurements, possible differences in natural and anthropogenic stressors between years, and differences in replication, the growth results are remarkably similar, except for 2006 which are much lower than in previous years. All three metrics compared here (Figure 2,3) also showed the same gradient of increasing growth with distance from the discharge. Assuming that there is no co-occuring natural or anthropogenic gradient that co-occurs with decreasing effluent exposure with distance from the discharge, the results from these two years of monitoring suggest that the discharge is adversely affecting mussel growth. It should also be mentioned that similar results were obtained in 2000 but due to the above-below experimental design and the large distance between test sites, these results are not environmentally meaningful (as were the fish results from those studies).

Increases in whole animal wet weight (WAWW) are often the most meaningful of all the mussel growth metrics because weight is more easily quantified than length. There is more error in length measurements because of the curvature of the shell. This is shown in Figure 2 where the WAWW growth rates among years is much more similar than Length growth rate among years. Whereas the slopes are virtually identical for WAWW with distance from the mill (Figures 2,3), there is much more of a disparity in the length data, particularly at Station 4. Interestingly, Station 4 is also the station in 2005 where 11 of the 12 test animals were females. This could have affected both the growth rate and biomarker results. It could also be a pseudo-correlation. The other interesting observation is that the WAWW growth rates are systematically higher in 2005 than in 2003 whereas some of the length metrics are higher and some are lower at each station when the years are compared. This is not surprising.

There is a similar increasing gradient of tissue weight with distance from the mill discharge but the slope is much less significant (Figures 2,3). The other possible explanation for the

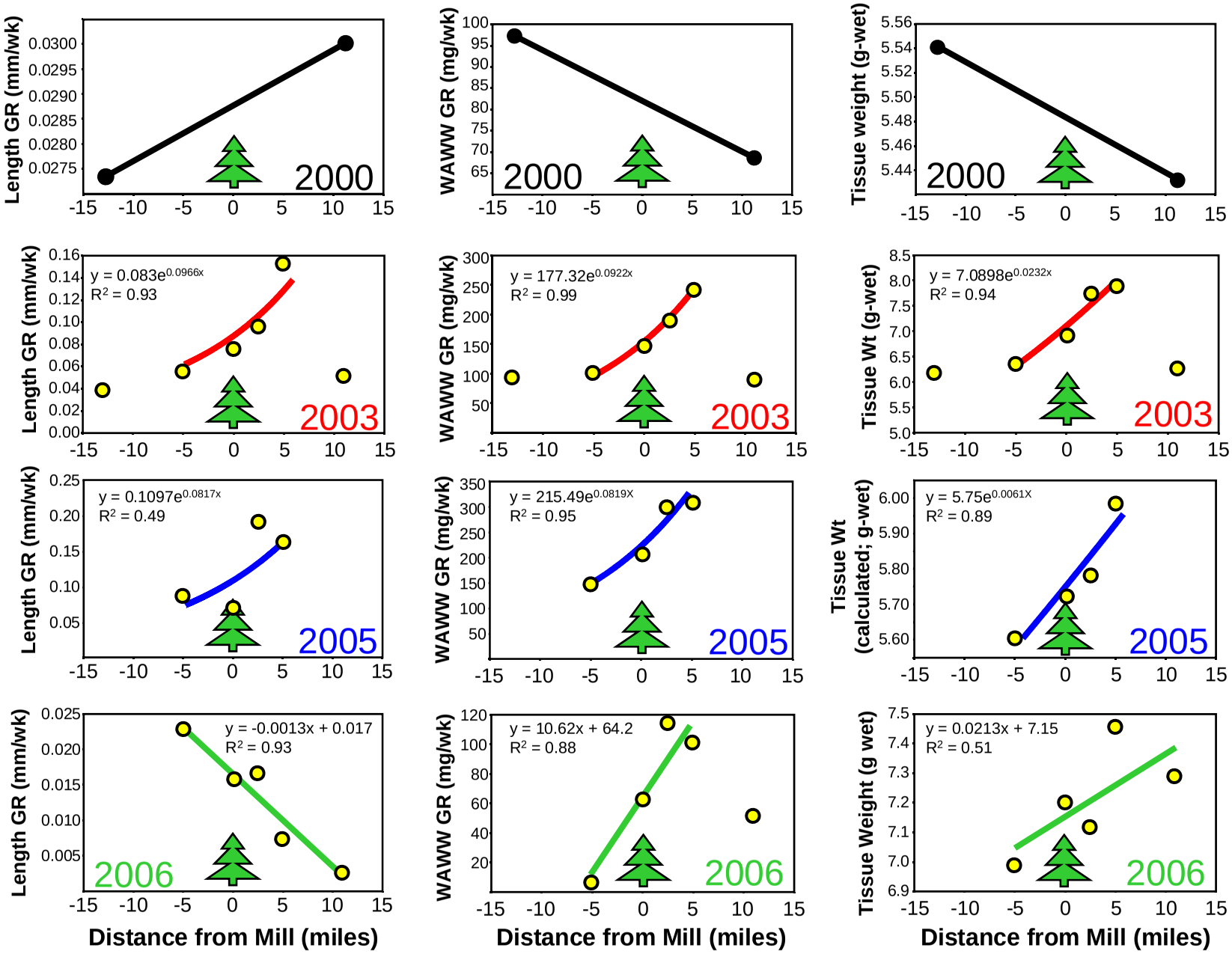


Figure 2. Relationship between various mussel growth metrics and distance from the mill (2000, 2003, 2005, 2006).

systematic differences between 2003 and 2005 is that the tissue weights were measured directly in 2003 while they were only estimated in 2005 and 2006. Furthermore, these tissue weights were estimated after the whole animals were frozen and then defrosted. This adds an additional source of error because of the frozen water inside the tissues. These tissues also degrade after freezing. This would also explain the systematic differences between years and the fact that the tissue weights were consistently higher (as were the WAWW measurements). Since the tissue weights were estimated by the differences between WAWW and shell weights, this would explain those differences as well. Again, the similarities between years was surprising and adds corroborative evidence of potential effects associated with the mill discharge.

The differences in the environmental significance of the mussel growth parameters between 2003 and 2005 is more easily visualized with the regression analyses shown in Figure 3. Although the regressions include only four data points, the regression coefficients (R^2)values range between 0.93 and 0.99 in 2003. In 2005 however, these same regression coefficients range between 0.49 and 0.95. Since it is generally the most accurate measurement it is not surprising the the highest value (0.95) occurs for WAWW growth rate. It is also not surprising that the lowest value occurs for length (0.49) since these measurements are generally associated with the greatest error (ASTM 2001). Given that the tissue weights were estimated and not measured directly, it is somewhat surprising that the tissue weight values gave a coefficient of 0.77. This could be attributed, at least in part to the fact that they were estimated from WAWW metrics.

5.3 Biomarker Results

As suggested previously, the biomarker results are not as easily interpreted as the growth results. However, they are potentially more useful because of their specificity an in this case, the vitellin metric can be used as an indicator of endocrine disruption. In Figure 4, The Gonad Index is actually more like a growth metric and it is not surprising that the regression with distance from the mill is similar to that of the growth metrics. However, given the reduced level of replication and variability in these measurements, the environmental significance is unclear. The Gill Index also shows a similar increase with distance from the mill discharge but this metric is not often used as an effects indicator. Nevertheless, taken together these two metrics seem to corroborate the growth rate results. The digestive gland ratio shows no apparent gradient with distance from the mill discharge.

The original data for lipid peroxidation in gills and digestive glands show almost opposite results although the gradients are less obvious and not necessarily intuitive. These results are further complicated by the fact that the originator of these data appeared to reverse the tabulation so that we are not sure which is showing the apparent increasing gradient and the other an apparent increasing gradient. The results are better explained, if they are in the correct order. The digestive gland response is associated with the discharge due to the dramatic increase between Stations 2 and 3 and then decreases with distance downstream (Figure 4). The gill response could be associated with another downstream stressor as we have shown previously with other dioxin and furan measurements apparently associated with a downstream source. These data required further analysis and consultation with the developer of the metrics.

We have subsequently confirmed through discussions with Environment Canada scientists (also identified in the AMEC review) that the data were reversed. This is yet another example of the lack of oversight and peer review provided by DEP. Remember that Applied Biomonitoring did not participate in the 2005 caged mussels study. Interestingly, with the data corrected the LPO

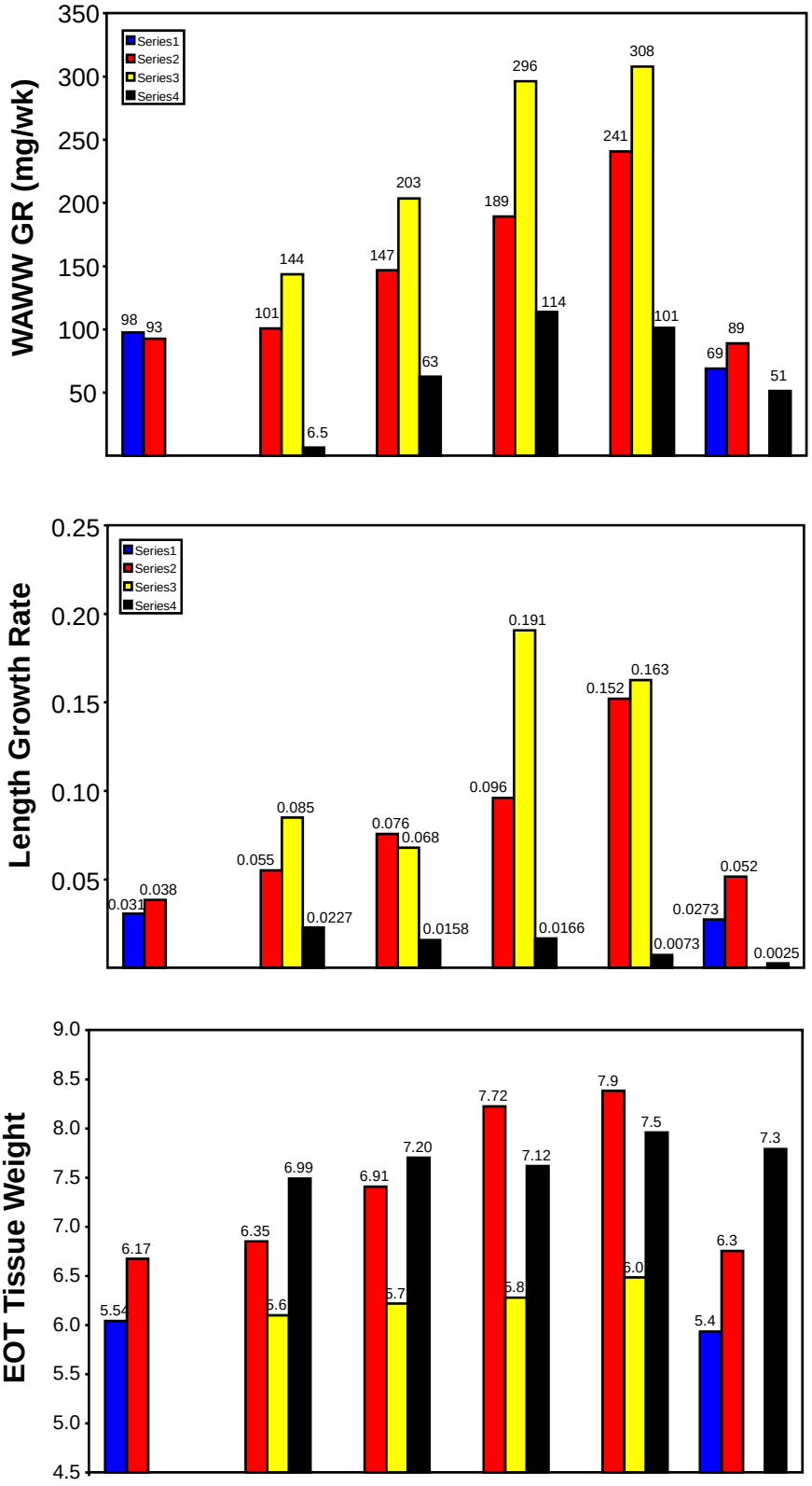


Figure 3. Differences in mussel growth metrics among sites and years in 2000, 2003, 2005, and 2006. Note the extremely low values in 2006 for WAWW growth rate and length growth rate and the extremely high values for EOT tissue weights (which were estimated in 2005 and 2006 based on regressions).

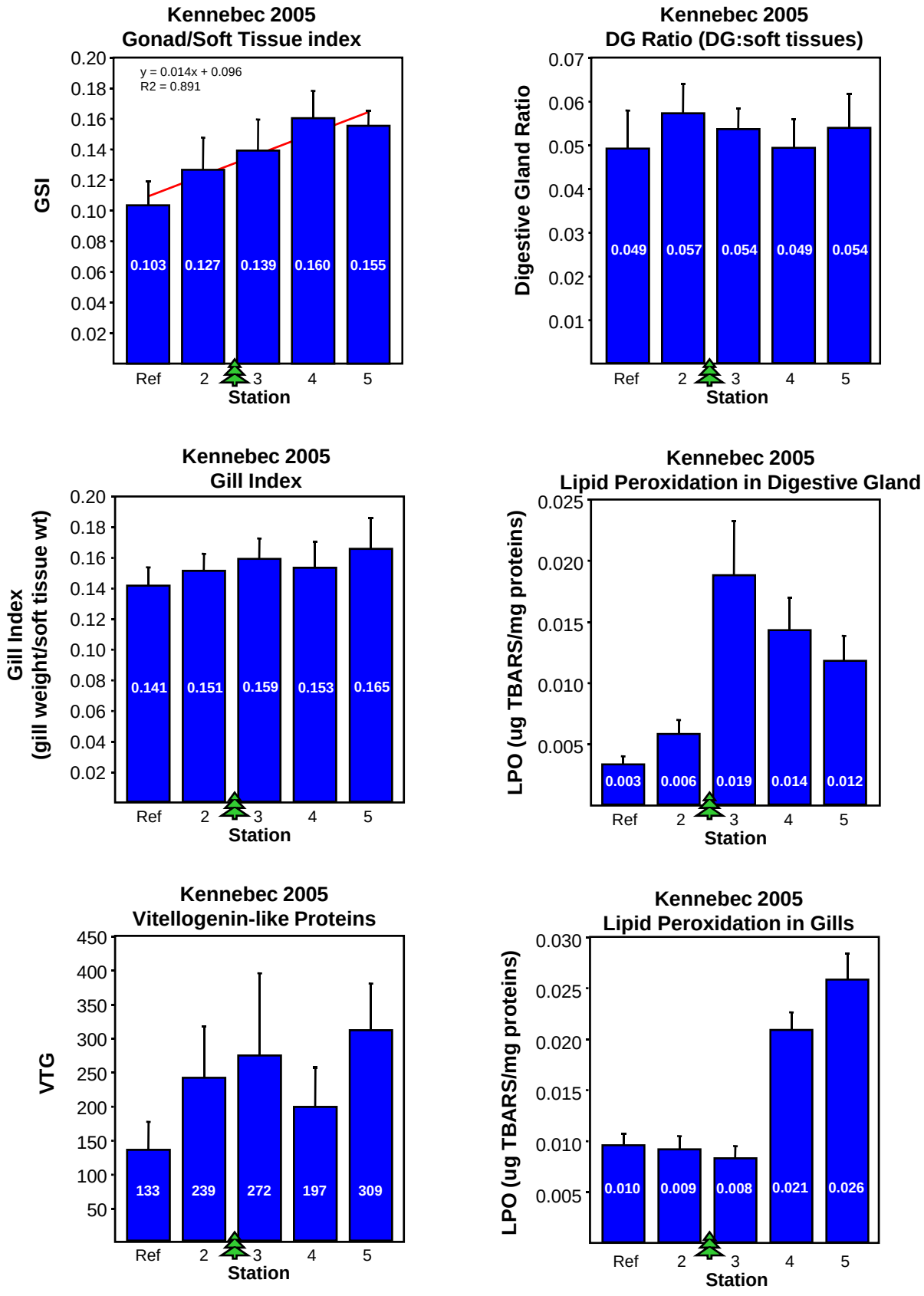


Figure 4. Biomarker results for 2005.

results in gills and digestive gland are quite similar. They both show an elevation but the former suggests the primary effect between Stations 3 and 4 while the latter between stations 2 and 3. The other difference is that the former shows no decreasing gradient with distance from the mill while the latter suggests a clearly decreasing gradient. These results remain complicated at best but could be resolved with additional physical chemical and biological monitoring.

Unfortunately, the most potentially useful biomarker the vitellin response, is also the most complicated. Two reasons for this are the limited level of replication. We have already discussed the consequences of limited replication. Another reason is the variability in the measurements associated with variability in sex ratios at each site. As an example in Table 3, T_0 had only 2/12 females whereas Station 4 had 11/12 females. Since the females are generally considered more responsive in vitellin assays, this could explain some of the differences in results. In addition, looking at the results for Station 3, the first two values range from about 45 to 655. This variability is alarming and raises serious questions about the distinguishing power and the environmental significance of the results. As suggested previously, the vitellin biomarker has not been used for that long and more validation may be necessary for applications on the Kennebec River.

5.4 Vitellin Results

Increased vitellin production is an indication of potential endocrine disruption and reproductive effects in bivalves and is comparable to vitellogenin in fish. It is expected that the females will have some increase in vitellin because they are getting ready for the next spawning cycle. However, excessive vitellin production in the females and the males, is an indication of adverse effects. The three graphs in Figure 5 show that the results are entirely dependent on how the data are normalized. Using the straight vitellin results (ug ALP/mL), there is an apparent increasing gradient with increasing distance from the mill for both 2003 and 2005. Furthermore, vitellin at all four sites within the impoundment is higher than at T_0 for both years as well.

Using vitellin normalized to tissue mass (ug ALP/tissue mass) however, the results are very different. There is not much of a gradient in the 2003 data while the 2005 data show an apparently increasing gradient with distance from the mill. However, vitellin production normalized to tissue mass is higher by about a factor of 5 at T_0. These data suggest that vitellin production is being supressed by the mill discharge. This is counter to previous results.

The 2003 data using vitellin normalized to protein (ug ALP/mg protein), which is the preferred metric, is difficult to interpret. The 2003 data show a slight gradient approaching the mill and vitellin production is higher than T_0 at all test sites at T_{end} except Station 6 (Figure 6). The protein normalized female vitellin data show an increasing gradient toward the mill and then a decreasing gradient downstream. The most dramatic difference occurs with the protein nomalized male vitellin data where there is a sharp increase between Stations 2 and 3 and then a general decreasing gradient with distance downstream. These data are remarkably similar to the lipid peroxidation data in the digestive gland for 2005. In 2005, all the sites had higher vitellin production than T_0 and there is an apparent increasing gradient with distance from the mill. Surprisingly, vitellin production at Station 4, with 11/12 females and lower growth rates than expected also had lower vitellin production than expected if the gradient were consistent.

The ALP assay is an indirect method to determine the relative levels of vitellin in biological tissues. The ELISA was not performed because the available kits are for fish vitellogenin and the appropriate antibodies do not cross-react well with bivalves. The ALP assay, because it is indirect, is validated with gel electrophoresis where vitellogenin-like protein bands are quantified

Table 3: Kennebec 2005 – Female Vitellin (12 mussels/station)

	Reference (T_0)	Station 2	Station 3	Station 4	Station 5
	214.38	240.26	655.41	305.54	407.83
	114.64	178.80	44.79	312.53	365.92
		581.79	211.39	208.57	424.51
		205.11	157.60	218.75	303.19
		131.43	532.42	173.83	314.68
			326.68	121.80	380.03
			87.16	217.18	361.36
			272.70	63.14	
			604.99	391.35	
				185.75	
				112.23	
Mean	164.5	267.5	321.5	210.1	365.4
Count	2	5	9	11	7
95% CI	97.7	157.9	147.8	57.1	33.1
SD	70.5	180.1	226.2	96.7	44.7

Comparison between 2003 & 2005 T_0 Vitellin

T_0	2003	2005
Females	3	2
Males	9	2
Indeterminate	0	8

Another reason not to compare with T_0

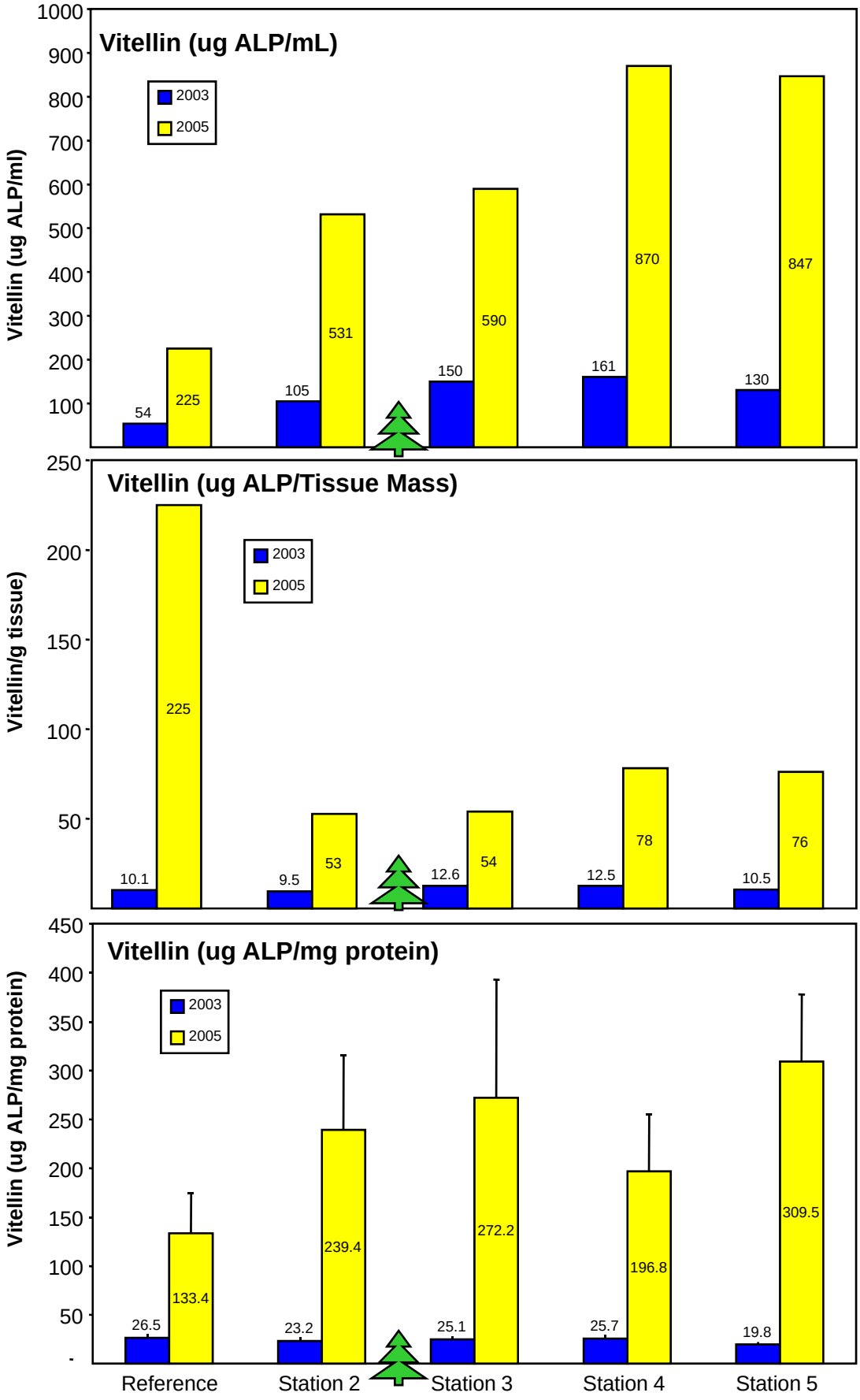


Figure 5. Comparing Vitellin results in 2003 and 2005.

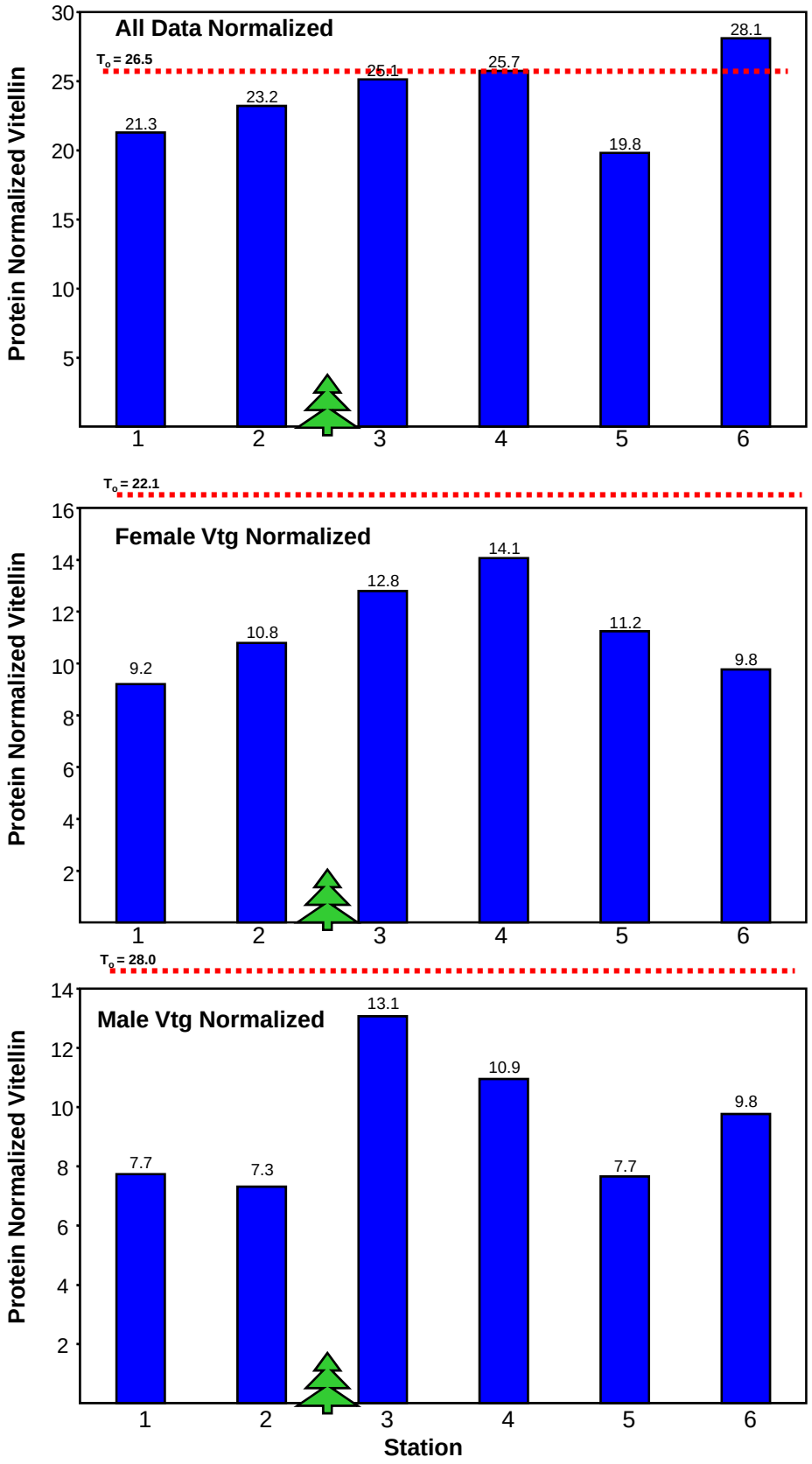


Figure 6. Protein-normalized vitellin results: 2003.

by densitometric analysis. In this study, there was a significant relationship between the ALP and gel electrophoretic assay indicating that ALP could be interpreted as vitellin-like proteins.

The 2005 data are about an order of magnitude higher at all stations within the impoundment (2,3,4,5) than in 2003. This is alarming in that it suggests that the effect on endocrine disruption have increased dramatically between 2003 and 2005. Similarly, the BOT samples were also about 5x higher in 2005 than in 2003. While it is possible than there could have been dramatic increases at all sites between years it seems more likely that these differences were attributable to insufficient replication (only 12 mussels/site in both years), variable number of females at each site (e.g. only 2/12 at BOT compared to 11/12 at station 4 in 2005. This would seem to make it inappropriate to compare BOT with sites where the number of females is significantly different. It also seems odd that while there were no indeterminate sex categories in 2003 that there were 8/12 at BOT in 2005. This suggests that either there were significant changes at the collection site or that there was a problem in the methodology. More work will be necessary to confirm these results. The other problem associated with the limited replication and perhaps other factors is the extreme variability in the data. This also suggests a problem with the test. At station 3 for example the first two data points for females are 655.4 and 44.8. This is over an order of magnitude difference and attributes to the huge standard deviation of 226 with a mean value of only 272.

It is also possible that there have been changes at the mill that might have affected the results and these could have either increased or decreased the potential exposure to caged mussels. This is another reason why it is always good to measure some chemical that can be used as a tracer or indicator of the relative discharge, regardless of whether or not it is the chemical causing the measured effects. Dioxins and furans have been used as tracers and indicators of potential effects but the concentrations are so low now that they are difficult and costly to measure and are probably not the primary chemical of concern associated with ongoing effects in Maine and around the world. Environment Canada has been working on tracers for a long time but they have had difficulties as well. The switch by DEP from exposure-based to effects-based monitoring and from above-below to the gradient design is a good decision based on current funding constraints. Nevertheless, it would be good to have some chemical indicator of exposure (a more traditional indicator) given the differences in the biomarker results between years and the need to explore factors affecting growth rates or any other indicator system. It is also good to use a more traditional indicator of effects like mussel growth rates to validate the biomarker results. So far it looks like the growth rate data for the two years are very similar and this suggests that the exposures were also similar. If we had matching tissue chemistry data we could have confirmed this.

To further examine the environmental significance of protein-normalized vitellin we re-examined the 2003 data which we believe is the most reliable since we conducted the tests. Interestingly, using all data, the protein-normalized vitellin shows almost no difference among sites. The female normalized vitellin shows an increasing gradient with distance approaching the mill and slightly beyond (Stations 3 and 4). The male-normalized vitellin shows the most direct response to what appears to be an effect associated with the mill. The response at Stations 1 and 2 above the mill (Station 2 within the impoundment) are relatively flat but there is a sharp increase at Station 3 nearest the mill discharge and then an apparent decreasing gradient with distance from the mill. Surprisingly, these values are all significantly different than BOT values by about a factor of two.

6.0 OTHER DATA, GRADIENT DESIGNS, AND STATUS AND TRENDS

While most of the freshwater caged mussel studies conducted by Environment Canada scientists from the St. Lawrence Center have suggested estrogenic effects in the freshwater mussels *Elliptio complanata* and *Dreissena polymorpha* (feminization; Blaise et al 2003, Gagne et al 2001a,b,c, 2002a, Quinn et al 2004) associated with the City of Montreal municipal effluent. Their work in the Saguenay Fjord with a marine species (*Mya arenaria*) suggests androgenic effects (masculinization; Blaise et al 1999 2002, Gagne et al 2002b 2003). Similarly, most of the work conducted by other scientists studying pulp and paper mill effect also suggest androgenic effects (masculinization) of both fish and bivalves (Aravindakshan 2004, Kernaghan et al 2004). One fish study demonstrates how measurements conducted over space and time can establish the status and trends of environmental effects associated with pulp and paper mill effluents. It also shows how a fortuitous opportunity created by a planned mill shutdown confirmed that effects were associated with a mill discharge and then later how effects seem to be decreasing as processing became more clean (Larsson & Forlin 2002). Initial studies on potential effects were reported previously (Larsson et al 1987 2000).

All of these studies were conducted over a 4-year period and the results are shown in Figure 7 (Larsson & Forlin 2002). The most compelling evidence of mill effects occurred when the mill shutdown occurred and effects were not observed (1999). There is also the suggestion that there were improvements at the mill because the closest site to the south did not show statistically significant differences in 2000. On the contrary, not only were there statistically significant differences in 1997 and 1998. We interpolated the data from 1997 and 1998 and provided regressions for the three closest stations in 1997 ($R^2 = 0.93$) and 1998 ($R^2 = 1.00$). These eelpout broods and resulting embryonic sex ratios have provided important information about potential effects associated with the mill, cause and effects due to the fortuitous mill shutdown, and the status and trends of exposure and effects associated with the mill.

Another important distinction here is the difference between observational and experimental approaches. Rather than wait for some fortuitous occurrence such as a mill shutdown, the caged bivalve methodology (or any other experimental approach) is that the cages can be placed along these suspected chemical gradients over space and time to assess exposure and effects. It is also more conducive to hypothesis testing than the observational approaches. As suggested previously, if DEP really wants to concentrate on fish, there are several other experimental approaches such as the fish embryo test (Bailey et al. 2006).

7.0 EFFECTS OF TEMPERATURE & OTHER PHYSICAL-CHEMICAL VARIABLES

We found statistically significant differences in water temperature between Stations 1 and 6 in the 2000 and 2003 caged mussel studies (Figures 8, 9). There were also significant differences in temperature at the outfall plume (Figure 9). These data provide additional evidence that Station 1 is not an appropriate reference site, particularly for effects measurements (Landis 2002). Figure 9 also shows that temperatures at Station 6 were more similar to the outfall plume temperatures than the reference site. This is probably due, at least in part to the shallow water there and higher temperatures compared to the impoundment. There is a need for more physical-chemical monitoring to better characterize and understand processes associated with the mill effluent.

Although other important physical, chemical, and biological chemicals were not measured during most of the caged mussel studies, other studies have shown the change in physical, chemical, and biological variables with distance downstream from various outfalls (Hynes 1960;

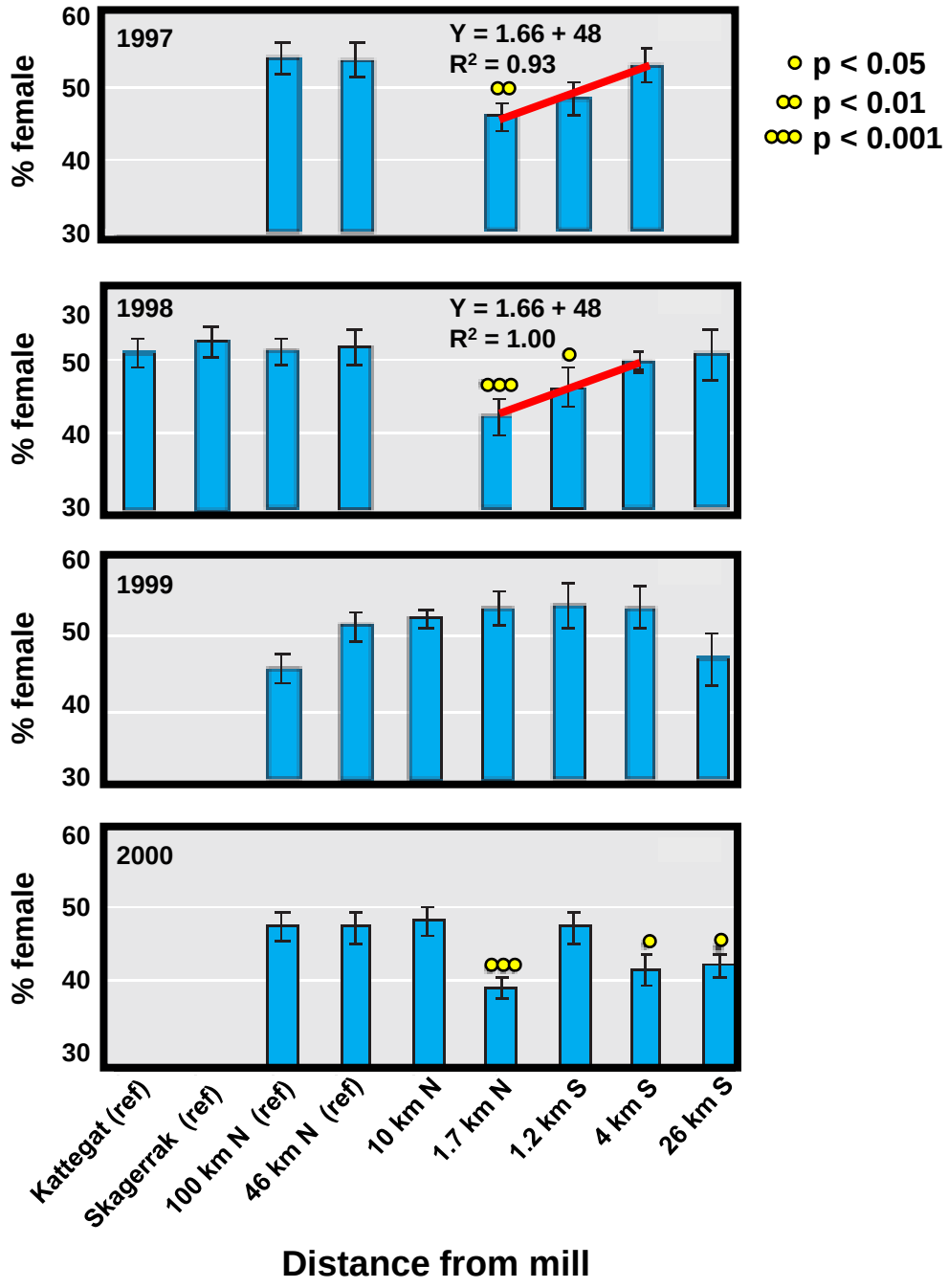


Figure 7. Differences in embryonic sex ratios of eelpout broods (mean ± among sites and years in 1997, 1998, 1999, and 2000 sampled at reference sites (ref) or various distances north (N) or south (S) of a pulp mill on the Swedish Baltic coast. Redrawn from Larsson & Forlin 2002 and including estimated regressions from their graphs.

○ = statistically significant difference from the pooled reference sites within the same year.

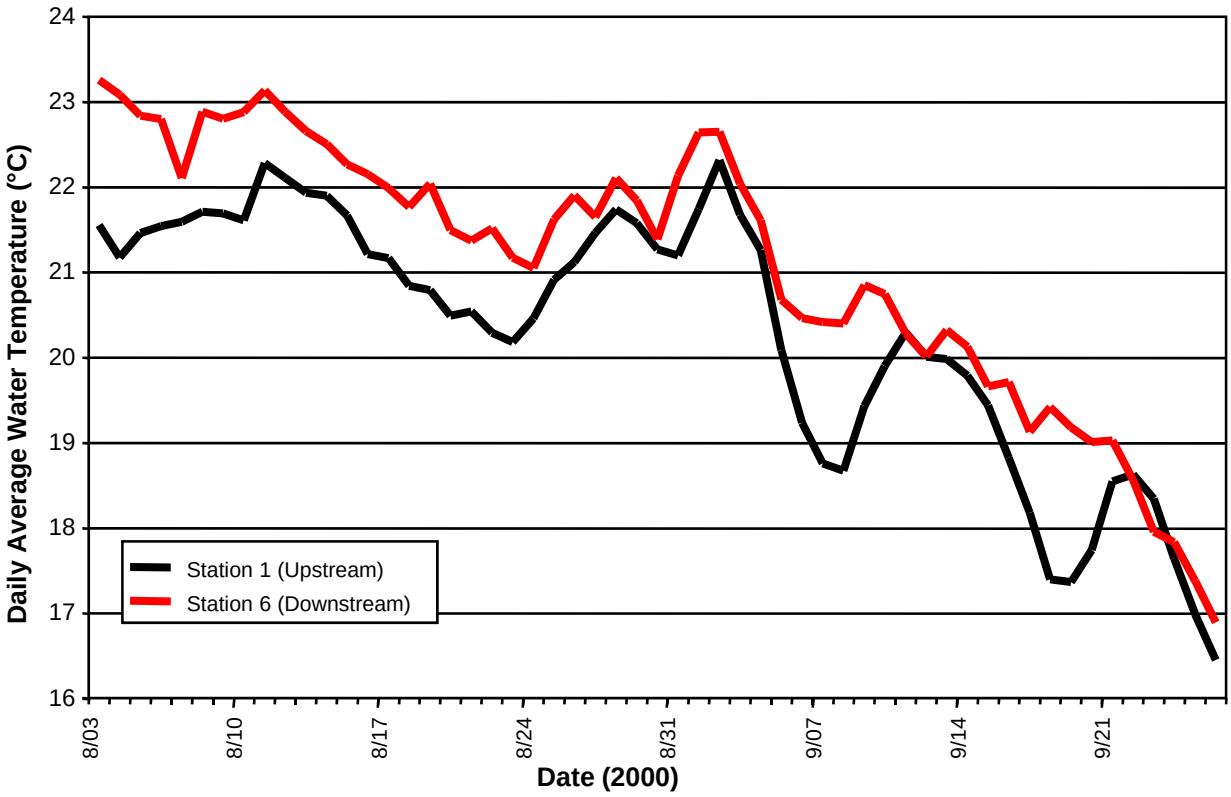


Figure 8. Daily average water temperatures at Stations 1 and 6 (2000).

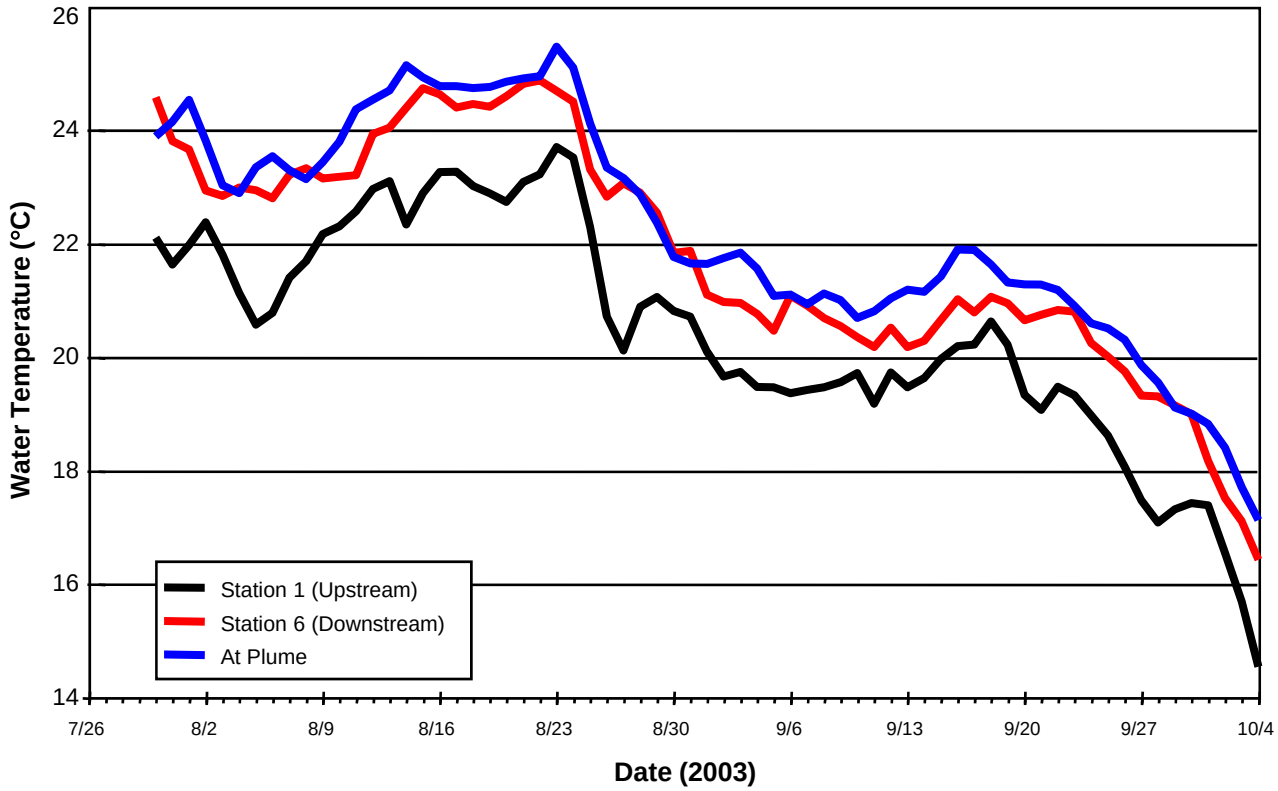


Figure 9. Daily average water temperatures at Stations 1, 6, and the Plume (2003).

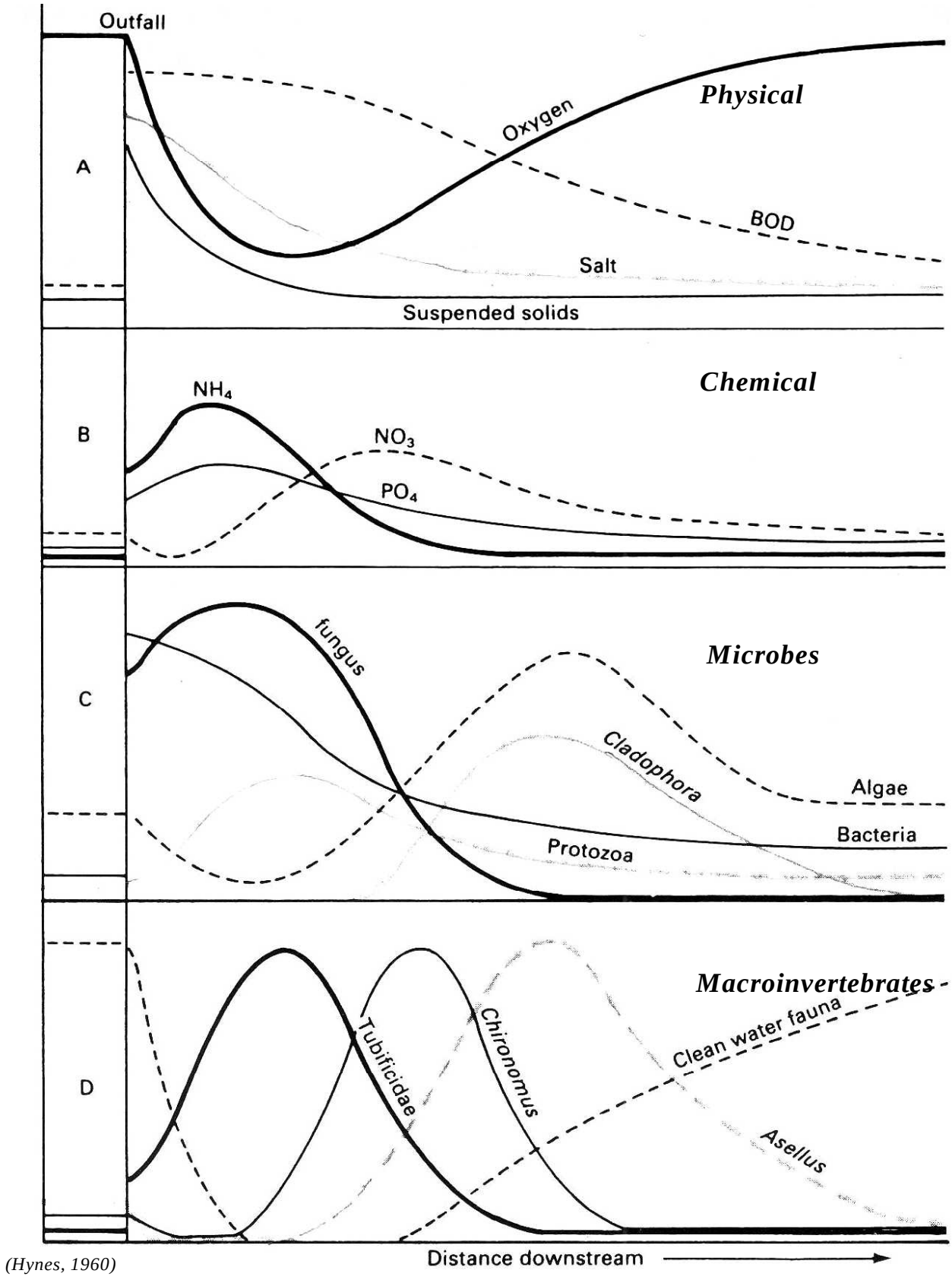


Figure 10. Graphic above copied from “Biology of Freshwater Pollution”, Chapter 3 (Organic Pollution), pg. 76. The cross-reference is from: Hynes, H.B.N. (1960). *The biology of polluted waters.* Liverpool University Press, Liverpool.



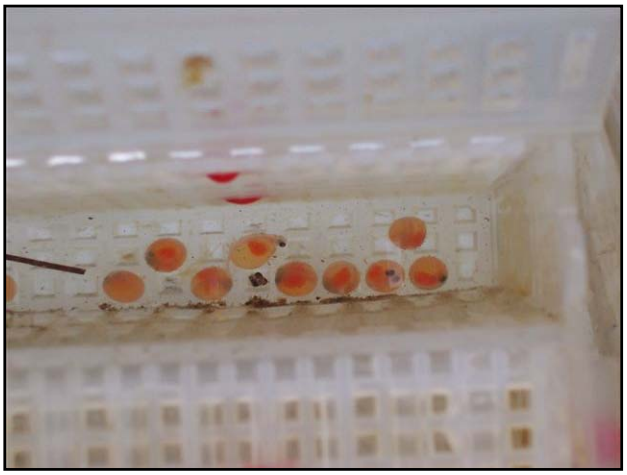



Figure 11. A summary of the in-situ fish embryo embryo test developed by Bailey & Elpchick (2006). More detail on this method is provided in the Appendix.

Immature mayfly nymph

Freshwater Unionids

Rock-filled baskets

Rock-filled riffle bag for sampling streams less than 10 cm deep

Remote-retrievable rock-filled cone and retrieval funnel forsampling non-wadeable rivers

Figure 12. Rock baskets and other observational and experimental monitoring methods developd by DEP but not used for monitoring on the Kennebec River in conjunction with caged mussel studies. (From, State of Maine, Biomonitoring Retrospective: Fifteen Year Summary for Maine Rivers and Streams, 1999)

Figure 10). These factors should be kept in mind when attempting to interpret the subtle effects of potentially toxic chemicals discharged from the mill. There are complex interactions between many of these physical-chemical parameters and mussel responses that are not clearly understood. This is another reason for measuring those chemicals and other variables as part of the caged mussel studies.

It is not clear why DEP has not utilized other experimental approaches in their assessments of the Kennebec River. As discussed previously, these include fish embryos (Bailey et al. 2006, Larsson & Forlin 2002), and caged fish (van den Heuvel 2005). A summary of Bailey et al. 2006 is provided in Figure 11 and in the Appendix. Although there could be problems with some caged fish tests because of the need to swim or move about, the main advantage of embryo tests or eel tests is that they do not have these same limitations. There are also approaches that used sediment bags that assess uptake of chemicals by sediments placed along suspected chemical gradients like the caged mussels (Nix & Merry 1990, Nix & Daykin 1992, Nix et al 1993). There are also the observational biomonitoring methods developed by DEP and the experimental field approaches such as rock baskets shown in Figure 12. These biomonitoring approaches of DEP in particular have a history of assessing the importance of physical chemical parameters in affecting benthic community structure. All of these approaches would be valuable complements to help characterize and understand physical, chemical, and biological processes associated with the mill effluent.

8.0 CONCLUSIONS

- The 2000 mussel growth data suggest mill effects but due to distances involved the results are not environmentally significant.
- Mussel growth data in 2003 and 2005 suggest possible adverse effects associated with the mill but it should be acknowledged that the measured differences could be associated with other unmeasured factors.
- Biomarker data presented for 2003 suggested possible adverse effects associated with the mill effluent but interpretation was hampered by inconsistent comparisons endpoints with varying environmental significance, and lack of transparency in the methods and results.
- Biomarker data presented for 2005 suffered from some of the same inconsistencies, less-than-meaningful endpoints, and lack of transparency in the methods and results.
- There also seemed to be some inconsistencies between the data reports and reports issued by DEP and that the data, analysis, and interpretation were not adequately reviewed by DEP.
- The adversarial relationship between DEP, industry and environmental interest groups has hampered the scientific aspects of monitoring and assessment.
- There is a need for routine peer review outside of existing advisory boards to identify the best approaches to combine the needs of compliance and research monitoring and develop a more transparent process.
- The real problem lies not with the caged bivalve methodology, but the way the tests have been conducted, data analyzed, and results interpreted by DEP.

9.0 RECOMMENDATIONS

The approach being utilized by DEP with respect to effects-based monitoring appears justified based on funding constraints. Nevertheless, there should be documentation specifically stating the reasons for this approach and that the preferred approach would be risk assessment-based monitoring that includes both exposure and effects. There needs to be some type of tracer

confirming that differential exposure has occurred along the gradient regardless of whether the causative agent has been identified. The gradient design paired with mussel growth measurements has been successful but a greater effort is needed to corroborate these measurements with other endpoints such as those used in other DEP monitoring programs. The biomarker endpoints need to be more carefully evaluated with performance criteria with respect to the level of replication and corroborative physical-chemical measurements. Extensive peer review of biomarker monitoring and other aspects of the program are needed.

10.0 REFERENCES

Adams, S. M. 2003. Establishing Causality Between Environmental Stressors and Effects on Aquatic Ecosystems. Human Ecol. Risk Assess. 9(1):17-35.

Adams, S. M., W. Halteman, C. Mattingly, and D. Page. 2004. Maine Department of Environmental Protection. Dioxin Monitoring Program. Review Panel Report. January 28, 2004.

Adams, S. M., W. R. Hill, M. J. Peterson, M. G. Ryon, J. G. Smith, and A. J. Stewart. 2002. Assessing recovery from disturbance in a stream ecosystem: application of multiple chemical and biological endpoints. Ecological Applications 12:1510-1527.

Applied Biomonitoring. 2000. Final Report. Caged Mussel Pilot Study: Port Alice Mill, EEM Program. Regional Manuscript Report: MS 00-01. 85 pp. plus appendices.

Applied Biomonitoring. 2002. Kennebec River Caged Mussel Pilot Study. Submitted to Maine Department of Environmental Protection and Friends of Merrymeeting Bay. 1 May 2002. Contract No. MDEP 2000, MDEP Agreement No. 800389. 73 Pp.

Aquatic Effects Technology Evaluation (AETE). 1997. Aquatic Effects Technology Evaluation Program, Background Document. Canada Centre for Mineral and Energy Technology, Natural Resources Canada.

Aravindakshan, J., V. Paquet, M. Gregory, J. Dufresne, M. Fournier, D. J. Marcogliese, and D. G. Cyr. 2004. Consequences of xenoestrogen exposure on male reproductive function in spottail shiners (*Notropis hudsonius*). Toxicol. Sci. 78:156-165.

ASTM. 2001. E-2122. Standard Guide for Conducting In-Situ Field Bioassays With Marine, Estuarine and Freshwater Bivalves. 2001 Annual Book of ASTM Standards. Conshohocken, PA. American Society for Testing and Materials (ASTM). pp. 1546-1575.

Baez, B. P. F. and M. S. G. Bect. 1989. DDT in *Mytilus edulis*: statistical considerations and inherent variability. Mar. Pollut. Bull. 20(10):496-499.

Besada, V., J. Fumega, and A. Vaamonde. 2002. Temporal trends of Cd, Cu, Hg, Pb and Zn in mussel (Mytilus galloprovincialis) from the Spanish North-Atlantic coast 1991-1999. Science of the Total Environment 288(ER3):239-253.

Blaise, C., F. Gagne, and J. Pellerin. 2003. Bivalve population status and biomarker responses in *Mya arenaria* clams (Saguenay Fjord, Quebec, Canada). Fresenius Environ. Bull. 12(9):956-960.

Blaise, C., F. Gagne, J. Pellerin, P.-D. Hansen, and S. Trottier. 2002. Molluscan shellfish

biomarker study of the Qubec, Canada, Saguenay Fjord with the soft-shell clam, *Mya arenaria*.

Blaise, C., F. Gagne, M. Salazar, S. Salazar, S. Trottier, and P.-D. Hansen. 2003. Experimentally-induced feminisation of freshwater mussels after long-term exposure to a municipal effluent. Fresenius Env. Bull. 12(8):865-870.

Blaise, C., F. Gagné, J. Pellerin, and P.-D. Hansen. 1999. Determination of vitellogenin-like properties in *Mya arenaria* hemolymph (Saguenay Fjord, Canada): a potential biomarker for endocrine disruption. Environ. Toxicol. 14(5):455-465.

Borgmann, U. 2000. Methods for assessing the toxicological significance of metals in aquatic ecosystems: bio-accumulation-toxicity relationships, water concentrations and sediment spiking approaches. Aquatic Ecosystem Health and Management 3(3):277-290.

Borgmann, U. and W. P. Norwood. 2000. What is wrong with the sediment quality triad. In: 27th Annual Aquatic Toxicity Workshop. Hotel Newfoundland October 1-4, 2000, St. John's Newfoundland. Program and Abstracts.

Borgmann, U., W. P. Norwood, T. B. Reynoldson, and F. Rosa. 2001. Identifying cause in sedment assessments: bioavailability and the Sediment Quality Triad. Can. J. Fish. Aquat. Sci. 58(5):950-960.

Borton, D. L. and B. K. Firth. 1997. Biomarkers move beyond discharges to measure mill effluent effects. Pulp Paper 71(5):4 pp.

Bright, D. A., P. V. Hodson, K. J. Lehtinen, A. B. McKague, J. H. Rodgers Jr, and K. R. Solomon. 2003. Chapter 1.3. Evaluation of Ecological Risks Associated With the Use of Chlorine Dioxide for the Bleaching of Pulp -- Scientific Progress Since 1993. In : T. R. Stuthridge, M. R. van den Heuvel, N. A. Marvin, A. H. Slade, and J. Gifford, Environmental Impacts of Pulp and Paper Waste Streams. Proceedings of the 3rd International Conference on Environmental Fate and Effects of Pulp and Paper Mill Effluents. Rotorua, New Zealand, November 9-13, 1997. Pensacola, FL. SETAC Press. pp. 18-26.

Capuzzo, J. M. 1985. Biological effects of petroleum hydrocarbons on marine organisms: Integration of experimental results and predictions of impacts. Mar. Environ. Res. 17(2-4):272-276.

Daskalakis, K. D. 1996. Variability of metal concentrations in oyster tissue and implications to biomonitoring. Mar. Pollut. Bull. 32(11):794-801.

Gagne, F. and C. Blaise. 2003. Chapter 10. Biomarkers to Assess Endocrine Disruption of Reproduction in Bivalves. In: C. Mothersill and B. Austin (Eds), In Vitro Methods in Aquatic Ecotoxicology. Springer Verlag. pp. 221-240.

Gagne, F., C. Blaise, I. Aoyama, R. Luo, C. Gagnon, Y. Couillard, and M. Salazar. 2002a. Biomarker study of a municipal effluent dispersion plume in two species of freshwater mussels. Environ. Toxicol. 17(149-159).

Gagne, F., C. Blaise, B. Lachance, G. I. Sunahara, and H. Sabik. 2001a. Evidence of coprostanol estrogenicity to the freshwater mussel *Elliptio complanata*. Environ. Pollut. 15:97-106.

Gagne, F., C. Blaise, J. Pellerin, and S. GauthierClerc. 2002b. Alteration of the biochemical properties of female gonads and vitellins in the clam *Mya arenaria* at contaminated sites in the Saguenay Fjord. Marine Environmental Research 53(ER3):295-310.

Gagne, F., C. Blaise, J. Pellerin, E. Pelletier, M. Doubille, S. Gauthier-Clerc, and L. Viglino. 2003. Sex alteration in soft-shell clams (*Mya arenaria*) in intertidal zone of the Saint Lawrence River (Qubec, Canada). Compar. Biochem. Phisol. 134:189-198.

Gagne, F., C. Blaise, M. Salazar, S. Salazar, and P. D. Hansen. 2001b. Evaluation of estrogenic effects of municipal effluents to the freshwater mussel *Elliptio complanata*. Comp. Biochem. Physiol. C 128: 213-225.

Gagne, F., D. J. Marcogliese, C. Blaise, and A. D. Gendron. 2001c. Occurence of compounds estrogenic to freshwater mussels in surface waters in an urban area. Environ. Toxicol. 16(3):260-268.

Gagné, F., D. J. Marcogliese, C. Blaise, and A. D. Gendron. 2001. Occurrence of compounds estrogenic to freshwater mussels in surface waters in an urban area. Environ. Toxicol 16:206-268.

Gordon, M., G. A. Knauer, and J. H. Martin. 1980. *Mytilus californianus* as a bioindicator of trace metal pollution: variability and statistical considerations. Mar. Pollut. Bull. 11:195-198.

Hewitt, L. M., M. G. Dubé, J. M. Culp, D. L. MacLatchy, and K. R. Munkittrick. 2003. A Proposed Framework for Investigation of Cause for Environmental Effects Monitoring. Human Ecological Risk Assessment: Vol. 9, No. 1, Pp. 195-211 (2003) 9(1):195-211.

Hynes, H. B. 1960. The Biology of Polluted Waters. Liverpool, Liverpool University Press.

Kendall, R., Dickerson, R., Giesy, J., and Suk, W. 1998. Principles and Processes for Evaluating Endocrine Distruption in Wildlife. Pensacola, FL, SETAC Press.

Kernaghan, N. J., D. S. Ruessler, S. E. Holm, and T. S. Gross. 2004. An Evaluation of the Potential Effects of Paper Mill Effluents on Freshwater Mussels in Rice Creek, Florida. In : D. Borton, R. Fisher, T. Hall, and J. Thomas (Eds.), Fate and Effects of Pulp and Paper Mill Effluents. Proceedings of the 5th International Conference. Seattle, Washington, June 1-4, 2003. Lancaster, PA. DEStech Publications, Inc. pp. 455-463.

Landis, W. G. 2000. Eulogy for the reference site. In: SETAC 21st Annual Meeting. Environmental Sciences in the 21st Century: Paradigms, Opportunities, and Challenges. Nashville, TN. SETAC. pp. pp. 144-145.

Larsson, A., T. Andersson, L. Forlin, and J. Hardig. 1987. Physiological disturbances in fish exposed to bleached kraft mill effluents. Water Sci. Tech. 20:67-76.

Larsson, D. G. J. and L. Forlin. 2002. Male-biased sex ratios of fish embryos near a pulp mill: temporary recovery after a short-term shutdown. Environ. Health. Perspect. 110:739-742.

Larsson, D. G. J., H. Hällman, and L. Förlin. 2000. More male fish embryos near a pulp mill. Environ. Toxicol. Chem. 19(12):2911–2917.

Maine Department of Environmental Protection. 2006. Dioxin Monitoring Program. State of Maine. 2005. DEPLW0754-2006. Department of Environmental Protection. Augusta, ME, http://www.maine.gov/dep/blwq/docmonitoring/dioxin/.

Maine Department of Environmental Protection Surface Water Ambient Toxics Monitoring Program. 2006. 2005 Surface Water Ambient Toxics (SWAT) Monitoring Program Report. Department of Environmental Protection. Augusta, ME, http://www.maine.gov/dep/blwq/docmonitoring/swat/index.htm

Martel, P., T. Kovacs, R. Voss, and S. Megraw. 2003. Evaluation of caged freshwater mussels as an alternative method for environmental effects monitoring (EEM) studies. Environ. Pollut. 124:471-483.

Mitra, A. and A. Choudhury. 1993. Trace metals in macrobenthic molluscs of the Hooghly Estuary, India. Mar. Pollut. Bull. 26(9):521-522.

Mower, B. 2001. Dioxin Monitoring Program. State of Maine. 2000. DEPLW0472. Department of Environmental Protection. Augusta, ME.

Mower, B. 2002. Dioxin Monitoring Program. State of Maine. 2001. DEPLW0528. Department of Environmental Protection. Augusta, ME.

Nix, P. G. and M. M. Daykin. 1992. Fecal pollution events reconstructed and sources identified using a sediment bag grid. Bull. Aquacul. Assoc. Can. 92(3):14-16.

Nix, P. G., M. M. Daykin, and K. L. Vilkas. 1993. Sediment bags as an integrator of fecal contamination in aquatic systems. Water Res. 27(10):1569-1576.

Nix, P. G. and C. H. Merry. 1990. Use of sediment bags as a monitor of fecal pollution in streams. Bull. Env. Contam. Toxicol. 45:864-869.

Phillips, D. J. H. 1980. Quantitative Aquatic Biological Indicators - Their Use to Monitor Trace Metal and Organochlorine Pollution. London, Applied Science Publishers Ltd.

Phillips, D. J. H., and Rainbow, P. S. 1993. Biomonitoring of Trace Aquatic Contaminants. London, Elsevier Applied Science.

Phillips, D. J. H. and D. A. Segar. 1986. Use of bio-indicators in monitoring conservative contaminants: Programme design imperatives. Mar. Pollut. Bull. 17(1):10-17.

Quinn, B., F. Gagne, M. Costello, C. McKenzie, F. Wilson, and C. Mothersill. 2004. The endocrine disrupting effect of municipal effluent on the zebra mussel (*Dreissena polymorpha*). Aquatic Toxicol. 66(3):279-292.

Reiley, M. C., Stubblefield, W. A., Adams, W. J. , Di Toro, D. M., Hodson, P. V., Erickson, R. J., and Keating, F. J. 2003. Reevaluation of the State of the Science for Water-Quality Criteria Development. SETAC.

Rodgers, J. H. and J. F. Thomas. 2004. Evaluations of the Fate and Effects of Pulp and Paper Mill Effluents From a Watershed Multistressor Perspective: Progress to Date and Future Opportunities. In : D. Borton, T. Hall, R. Fisher, and J. Thomas (Eds.), Pulp and Paper Mill Environmental Fate & Effects.. Proceedings of the 5th International Conference. Seattle,

Washington, June 1-4, 2003. Lancaster. PA. DES*tech* Publications, Inc. pp. 135-146.

Salazar, M. H. and S. M. Salazar. 2005a. 8610 E. Field Test Procedures Using Freshwater and Marine Bivalves. In: L. S. Clesceri, A. D. Eaton, and E. W. Rice (Eds), Standard Methods for the Examination of Water and Wastewater. 21st Edition. American Public Health Association, American Water Works Association, Water Environment Federation.

Salazar, M. H. and S. M. Salazar. 1999. Campesterol in caged mussels as a tracer of pulp and paper mill effluents. In: R. Van Collie, R. Chasse, L. Hare, and others (Eds.), Proceedings of the 25th Annual Aquatic Toxicity Workshop. Canadian Technical Report of Fisheries and Aquatic Sciences N. 2260. Quebec City, Quebec. p. 49 (abstract only).

Salazar, M. H. and S. M. Salazar. 2005b. Chapter 6. Field Experiments With Caged Bivalves to Assess Chronic Exposure and Toxicity. In: G. K. Ostrander (Ed.), Techniques in Aquatic Toxicology, 2nd Edition. Boca Raton, FL. CRC Press. pp. 117-135.

Salazar, M. H. and S. M. Salazar. 2006. Chapter 9. Linking Bioaccumulation and Biological Effects to Chemicals in Water and Sediment: A Conceptual Framework for Freshwater Bivalve Ecotoxicology. In: J. Vanhassel and J. Farris (Eds.), Freshwater Bivalve Ecotoxicology. Boca Raton, FL. CRC Press/SETAC Press. pp. 215-255.

Salazar, M. H. and S. M. Salazar. 2003. Characterizing dioxin and furan exposure from pulp and paper mill effluents using caged mussels in the Kennebec River, Maine. Organohalogen Compounds 62:444-447.

Salazar, M. H. and S. M. Salazar. 2004. Using an Integrated Risk Assessment Strategy to Characterize Exposure and Effects in the Field With Caged Bivalves. In : D. Borton, T. Hall, R. Fisher, and J. Thomas (Eds.), Pulp and Paper Mill Environmental Fate & Effects. Proceedings of the 5th International Conference. Seattle, Washington, June 1-4, 2003. Lancaster. PA. DES*tech* Publications, Inc. pp. 439-454.

Salazar, M. H. and S. M. Salazar. 1998. Using Caged Bivalves As Part of an Exposure-Dose-Response Triad to Support an Integrated Risk Assessment Strategy. In: A. de Peyster and Day. K., Proceedings, Ecological Risk Assessment: A Meeting of Policy and Science, SETAC Special Publication. SETAC Press. pp. 167-192.

Salazar, M. H. and S. M. Salazar. 1997. Using caged bivalves to characterize exposure and effects associated with pulp and paper mill effluents. Wat. Sci. Tech. 35(2-3):213-220.

Salazar, M. H., S. M. Salazar, F. Gagne, C. Blaise, and S. Trottier. 2002. Developing a benthic cage for long-term in-situ tests with freshwater and marine bivalves. In: Proceedings of the 29th Annual Aquatic Toxicity Workshop. Canadian Technical Report of Fisheries and Aquatic Sciences 2438. Whistler, British Columbia. 62. pp. 34-42.

Salazar, M. H., S. M. Salazar, F. Gagne, C. Blaise, and S. Trottier. 2003. An in-situ benthic cage to characterize long-term organochlorine exposure and estrogenic effects. Organohalogen Compounds 62:440-443.

Servos, M. R., Munkittrick, K. R., Carey, J. H., and van der Kraak, G. J. 1996. Environmental Fate and Effects of Pulp and Paper Mill Effluents. Delray Beach, Florida, St. Lucie Press.

Stuthridge, T. R., van den Heuvel, M. R., Marvin, N. A., Slade, A. H., and Gifford, J. 2003.

Environmental Impacts of Pulp and Paper Waste Streams. Proceedings of the 3rd International Conference on Environmental Fate and Effects of Pulp and Paper Mill Effluents. Rotorua, New Zealand.

US EPA. 1998. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002F. April 1998. Final. Washington, DC, Risk Assessment Forum, US Environmental Protection Agency

van den Heuvel, M., M. Landman, and L. Tremblay. 2006. Responses of shortfin eel (*Anguilla australis*) exposed in situ to pulp and paper effluent. J. Toxicol. Environ. Health Part A 69(19):1763-1779.

Van Der Kraak, G., K. Munkittrick, M. E. McMaster, and D. L. MacLatchy. 1998. A Comparison of Bleached Karft Mill Effluent, ß-Estradiol, and ß-Sitosterol Effects on Reproductive Function in Fish. In: R. J. Kendall, R. L. Dickerson, J. P. Giesy, and W. P. Suk, Principles and Processes for Evaluating Endocrine Disruption in Wildlife. Pensacola, FL. SETAC Press. pp. 249-265.

Walker, S. L., K. Hedley, and E. Porter. 2002. Pulp and Paper Environmental Effects Monitoring in Canada: An Overview. Water Quality Research Journal of Canada 37(1):7-19.

Widdows, J. and P. Donkin. 1992. Mussels and Environmental Contaminants: Bioaccumulation and Physiological Aspects. In: E. Gosling (Ed.), The Mussel *Mytilus*: Ecology, Physiology, Genetics and Culture. Amsterdam. Elsevier Science Publishers. pp. 383-424.

Wright, D. A., J. A. Mihursky, and H. L. Phelps. 1985. Trace metals in Chesapeake Bay, USA, oysters: intra-sample variability and its implications for biomonitoring. Mar. Environ. Res. 16(3):181-198.

APPENDIX

Bailey, H. and Elphick, J. 2006. Use of In situ Hatchbox Studies to Evaluate Water Quality Effects. Presented at Pacific Northwest Society of Environmental Toxicology Annual Meeting, Port Townsend, Washington, April 2006.

Use of *In situ* Hatchbox Studies to Evaluate Water Quality Effects

Howard Bailey and James Elphick
Nautilus Environmental

Ben Chalmers, Gavin Dirom and Ivor McWilliams
NVI Mining Ltd, Myra Falls Operation

What is *In situ* testing?

- Controlled exposure at site of interest
- Bridges gap between laboratory exposures and bioassessment studies

Advantages

- Evaluate adverse effects under site-specific conditions
- Quantifiable response
- Known exposure
- Separate water quality from habitat effects
- Can be paired with bio-monitoring studies

Study Site

- Mine located in Strathcona Park, Vancouver Island, BC
- Discharges treated wastewater to Myra Creek
- Also has potential subsurface seeps from tailings deposits
- Has limited population of cutthroat trout

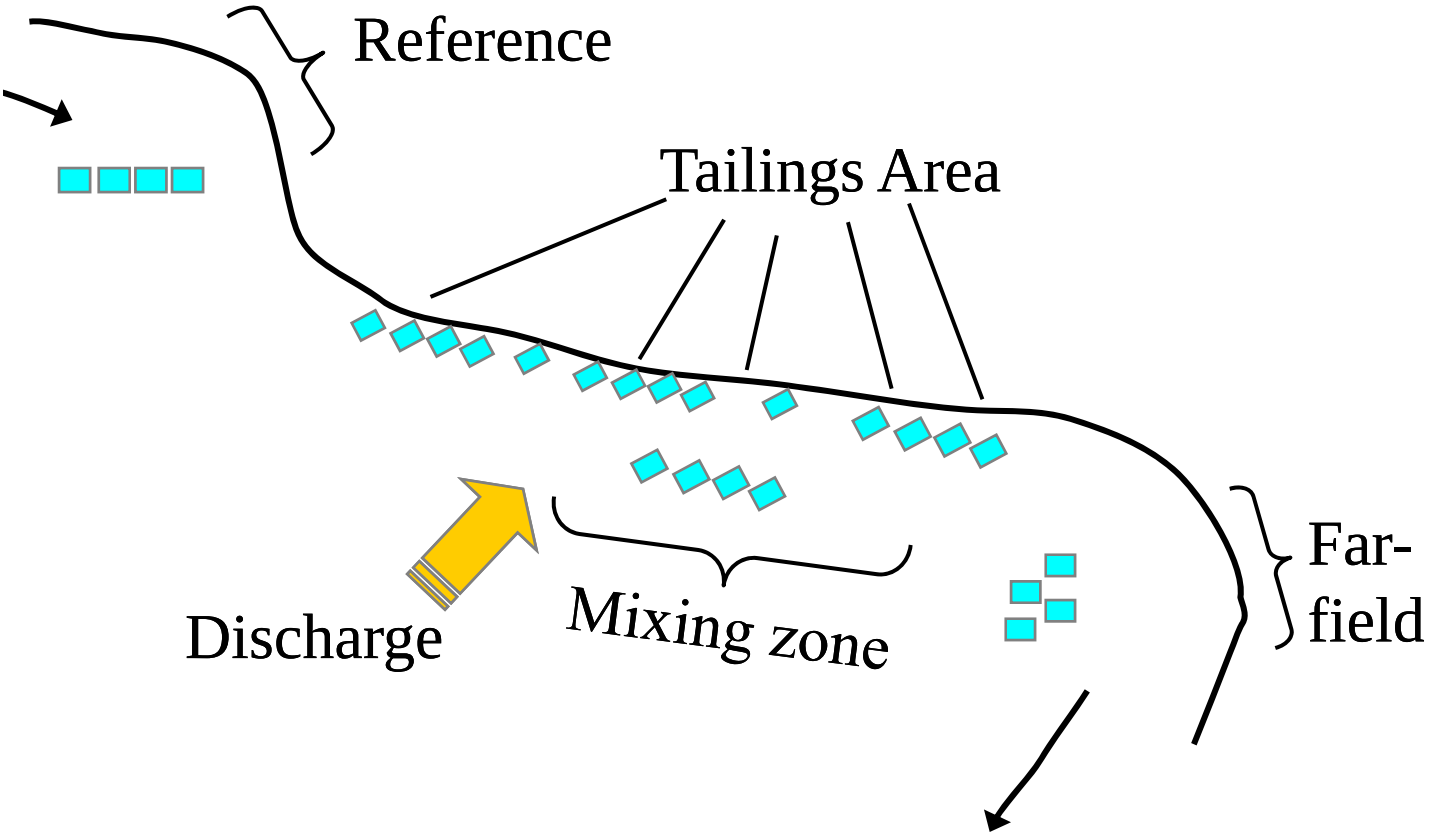
Why Use *In situ* Approach?

- Limited fish population
  - Adversely affected by destructive sampling
- Small creek
  - Fish population mobile; difficult to separate exposed from un-exposed groups
- Multiple contaminant sources (discharge points)
  - Can locate exposures in direct proximity to discharges

Study Design

- Primary
  - Upstream (reference); n=4
  - Downstream (near field); n=4
  - Downstream (end of mixing zone); n=4
- Secondary
  - Gradient design along foot of tailings area

Study Design



Study Design

- Cutthroat trout embryos- 25/box
- Boxes enclosed in “chicken basket”, surrounded by rocks (reduce impact of flow, sedimentation and light)
- Baskets placed in excavated area of stream bed; covered with rocks

Monitoring

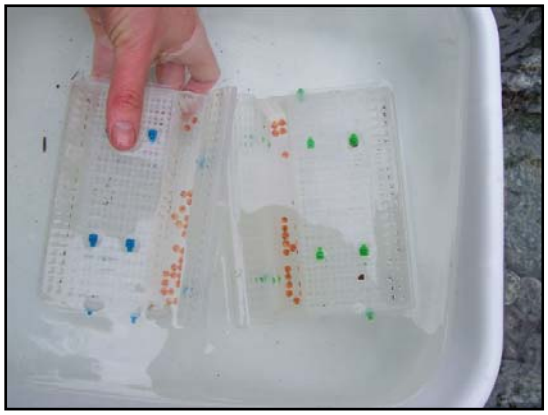
- Boxes checked at weekly intervals
  - Survival
  - Hatching success
  - Abnormalities
  - Growth
  - Swim-up
- Laboratory control
  - Monitor key developmental periods
  - Assure embryo quality

Exposure Locations

Reference Site

Engineered Channel

Hatch Box Placement



Results

- EFFECTS ON SURVIVAL
  - Exposure Site – No Detectable Effect
  - Ground Water Seep
  - Effect at discrete locations
    - 2 stations with 0% survival
- SUBLETHAL EFFECTS
  - Exposure Site – No Detectable Effect
  - Effluent Discharge Site – Detectable Effect

Key Findings

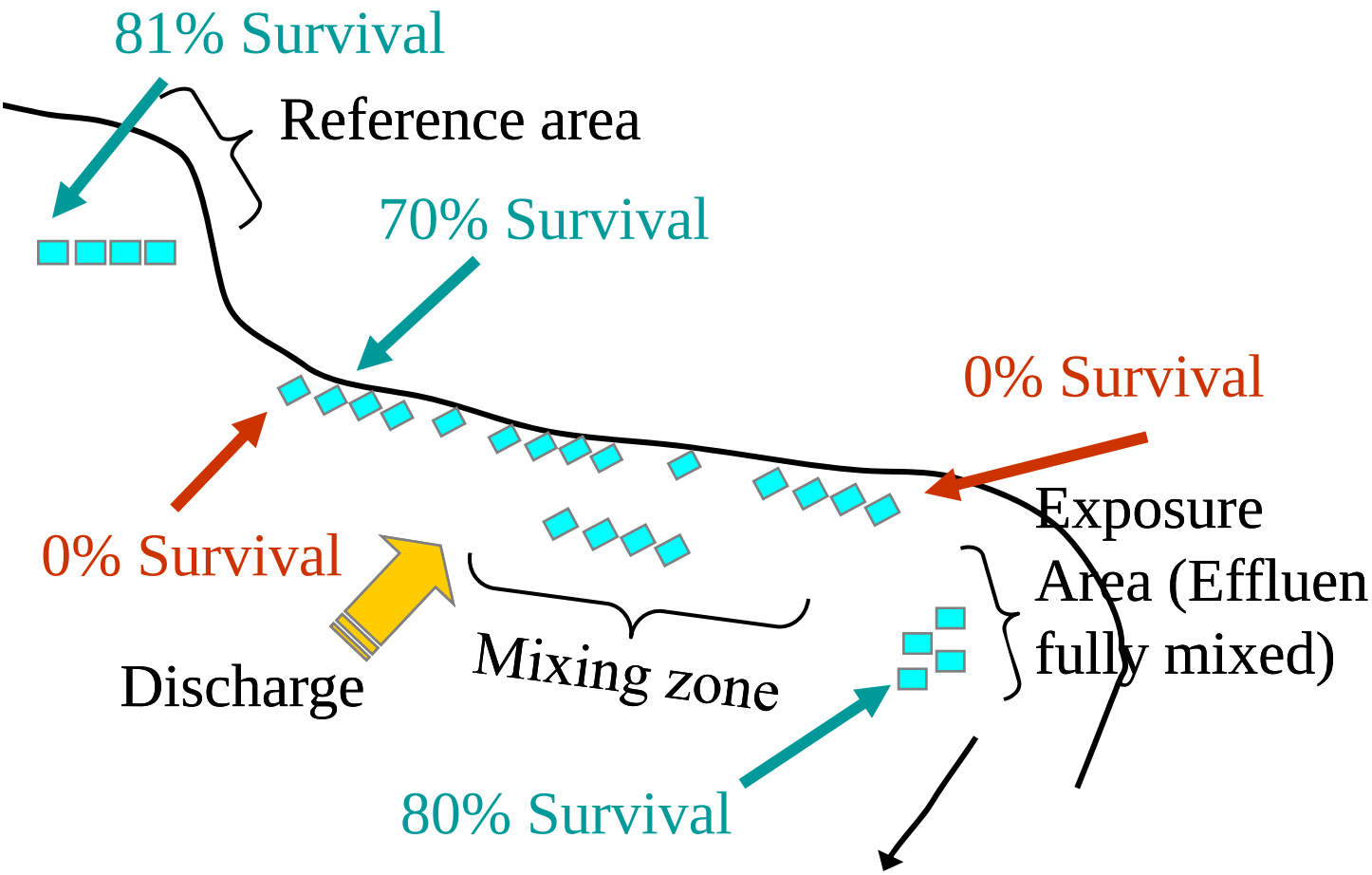
- No measurable impact at

Exposure site compared with

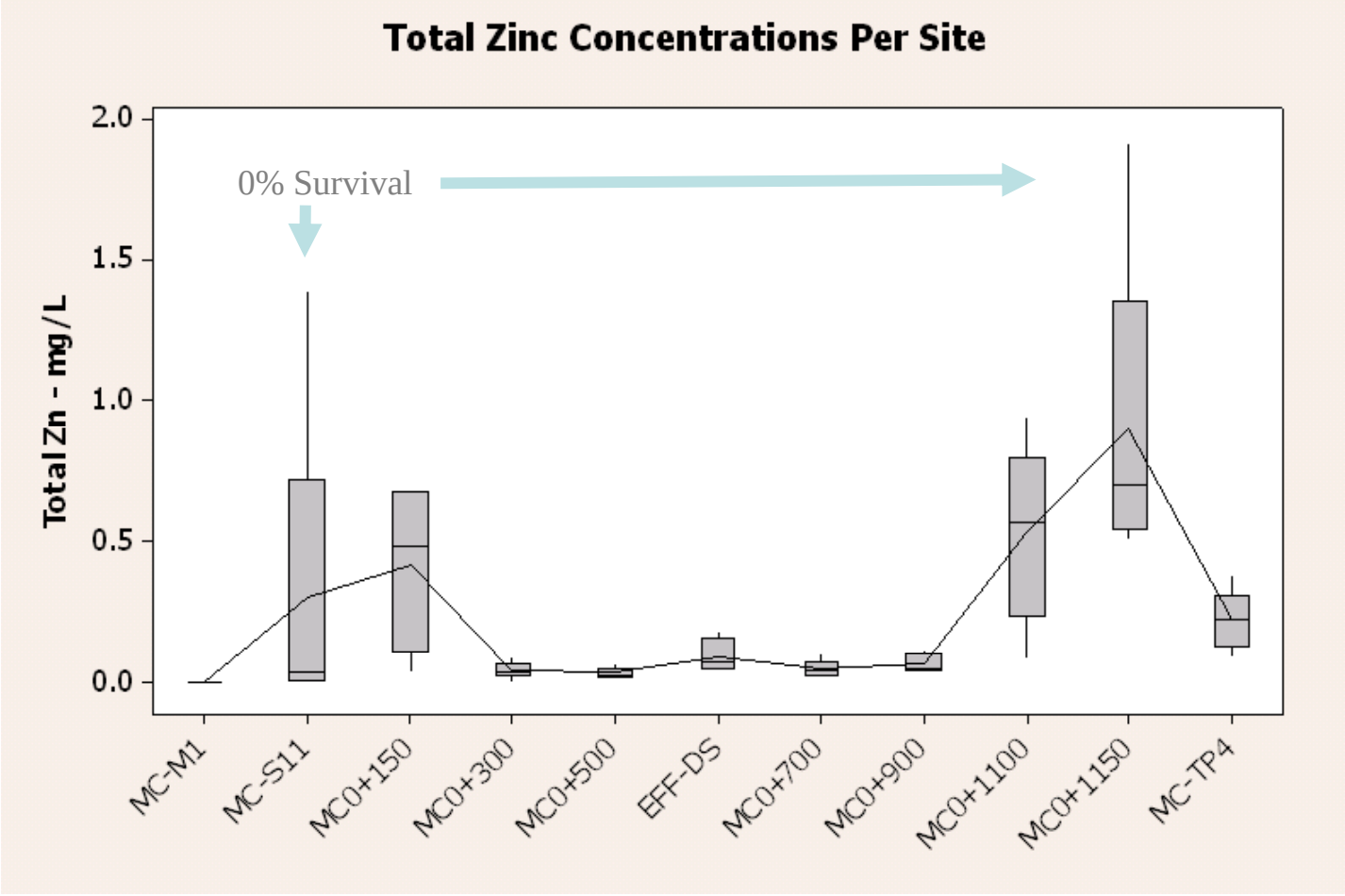
Reference site

- Identified toxicity associated with groundwater seeps along tailings impoundment area
  - Areas of potential impact highly localized.

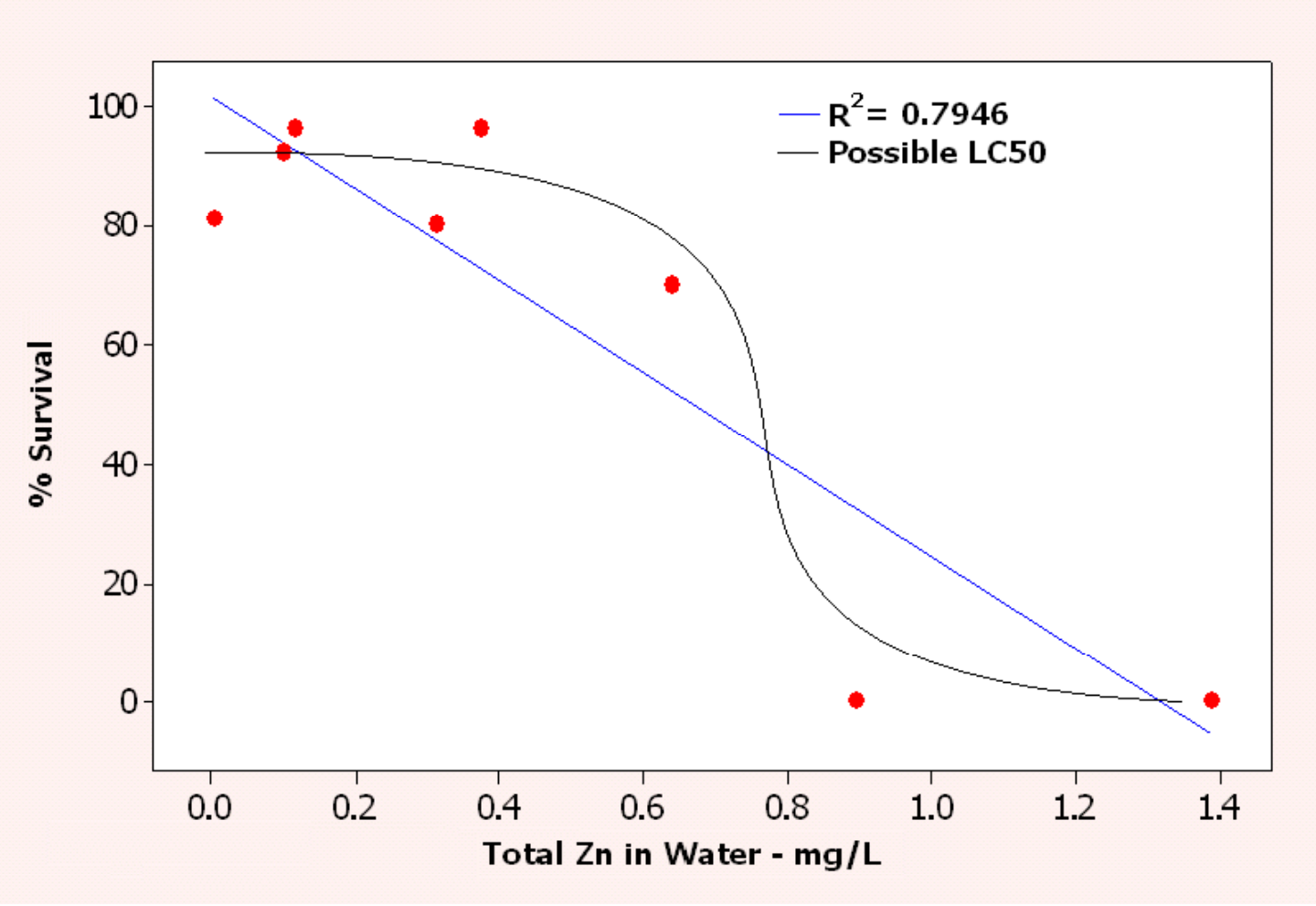
Survival and Mortality



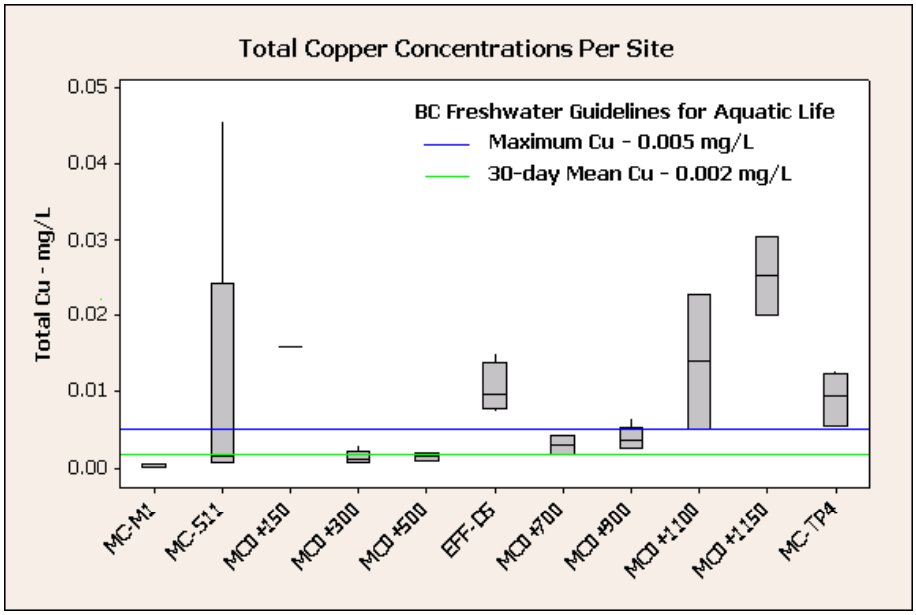
Zinc Concentrations

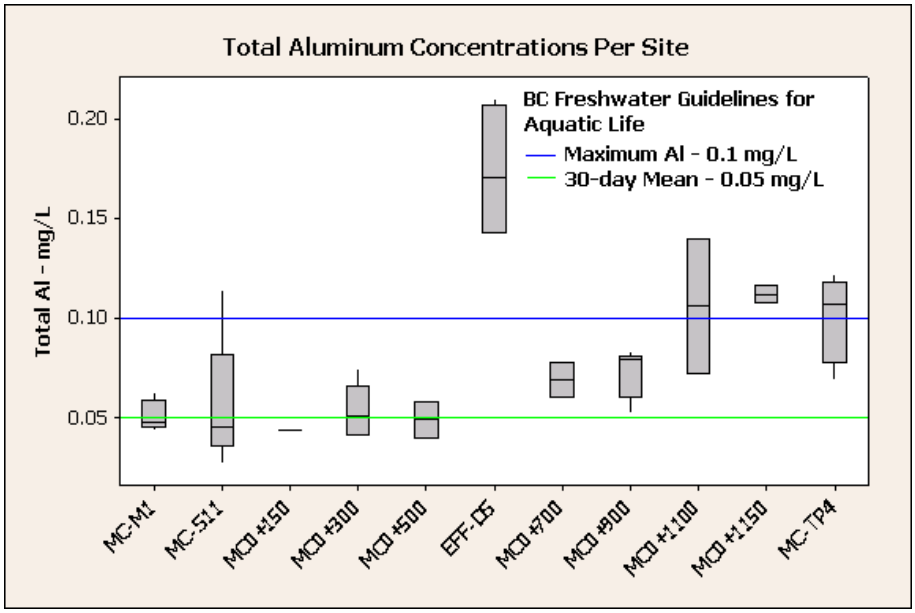


Survival vs. Zn in Water



Other Metal Concentrations of Interest





Statistical Sensitivity

- Minimum Significant Differences
  - Survival – 14%
  - Length – <5%
  - Wet Weight – <5%
- Statistically Significant Effects Detected
  - Wet Weight – 5% difference
    Reference vs. Effluent Discharge Site
  - Length – 5% difference
    Reference vs. Effluent Discharge Site

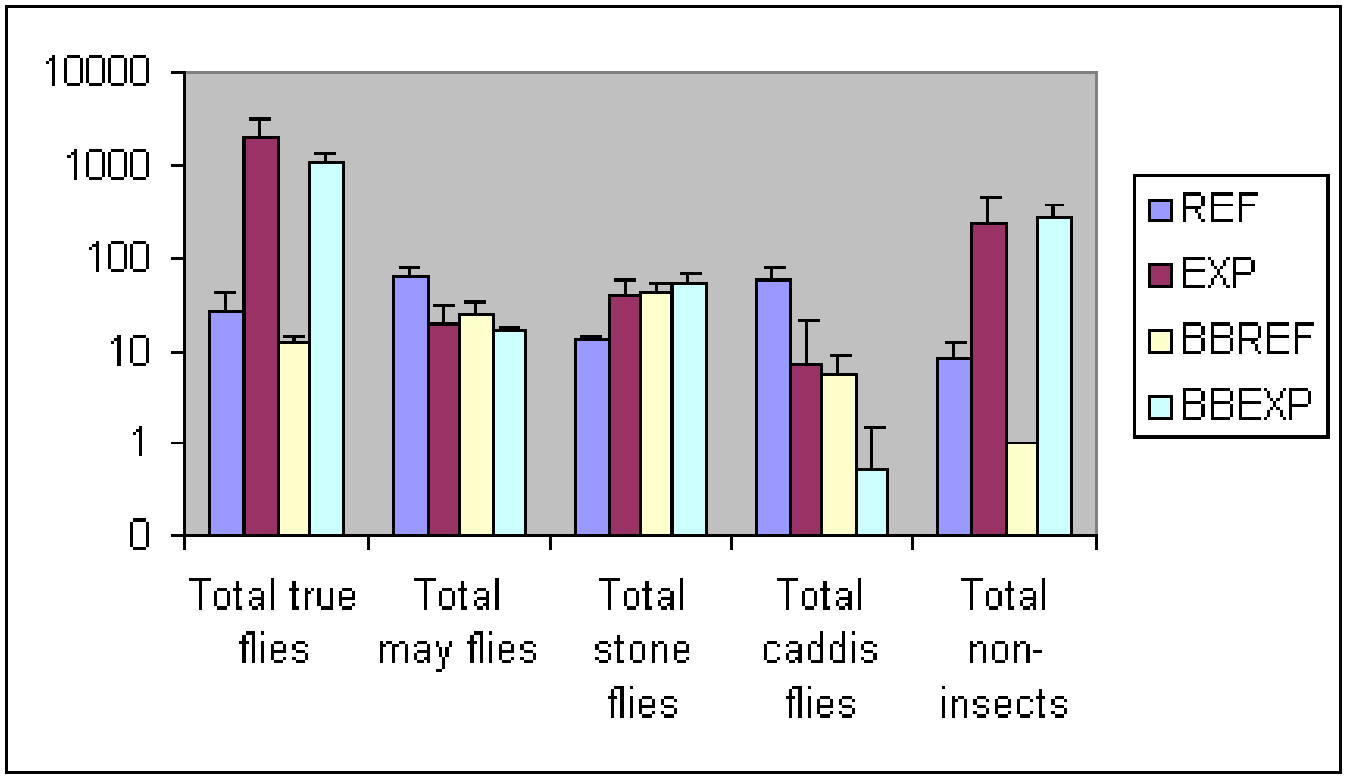
Advantages

- Evaluate adverse effects under site-specific conditions
- Known exposure
- Separate water quality from habitat effects
- Can be paired with bio-monitoring studies
- Inexpensive to conduct

Additional Studies

- Pair with other biological community sampling:
  - Stream macro-invertebrates sampled in Reference and Exposure areas; compared with colonization of substrate added to stream bed.

Benthic Invertebrate Results



Summary

- Powerful tool for evaluating potential effects and causes
- Controlled exposure minimizes confounding factors
- High statistical sensitivity for a field test
- *In situ* approach can be applied to other species and endpoints

Applications at Other Sites

- Readily applied in a variety of situations
- High statistical sensitivity (~ 10%)
- Trout in southern California?
  - Southern steelhead endangered ecological unit
  - This approach has high relevance for evaluating water quality in areas of critical habitat

Acknowledgements

Armando Tang and Andreas Diewald with Nautilus Environmental assisted in the field work.

Sharlene Henderson with NVI Mining Ltd, Myra Falls Operations assisted in Statistical Analysis.

Mike Hagen of Environment Canada provided useful comments on study design.

NVI Mining Ltd., Myra Falls Operations for supporting this study.