UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION I
ONE CONGRESS STREET SUITE 1100
BOSTON, MASSACHUSETTS 02114-2023

Date: June 29, 2009
From: Edward Hathaway, RPM
ME/VT/CT Superfund Section
To: Site File
Re: Callahan Mine Ecological Risk Assessment Benthic Community Assessment

Introduction:

A benthic community assessment was one of the lines of evidence used to evaluate the ecological threats that may result from the contamination at the Callahan Mine Superfund Site. The overall conclusion of the benthic community assessment was a finding of no significant difference between the Site and local reference locations. This put the benthic community assessment in conflict with other lines of evidence, specifically the sediment toxicity testing and food chain modeling. EPA has assigned a lower weight to the benthic community evidence due to concern with the overall study design, implementation challenges, and the lack of a good biological conceptual model upon which to base the interpretation. In addition, the Remedial Investigation Report documented that the Site has several distinct “hot spot” areas that should have been addressed with a stratified sampling plan.

To gain additional insight regarding the degree to which the benthic community data should be relied upon in the risk management decision, EPA sought additional feedback from internal experts and external experts. Marguerite Pelletier of EPA’s Office of Research and Development provided an initial, more generic, review of the benthic community assessment in December 2008. To supplement this review, EPA tasked the Environmental Services Assessment Team to obtain additional expert feedback. Two individuals with substantial expertise in marine ecology were retained to provide the additional expert feedback. The review comments from each of these individuals is attached to this memo.

Conclusions:

The attached comments from Marguerite Pelletier, EPA Office of Research and Development, Brain Beal, Univ of Maine Machias, and Tom Trott of Suffolk University support EPA’s position that the benthic community assessment is a low weight line of evidence given the uncertainties associated with the collection and interpretation of the data.

Attachments:

1. TechLaw June 29, 2009: Review of the Benthic Community Study Performed in Support of The Baseline Ecological Risk Assessment at the Callahan Mine Superfund Site, Brooksville, ME (interim deliverable by Brian Beal). Includes attached

SDMS DocID 452692

2. TechLaw June 29, 2009: Review of the Benthic Community Study Performed in Support of The Baseline Ecological Risk Assessment at the Callahan Mine Superfund Site, Brooksville, ME (interim deliverable by Thomas Trott).
3. EPA Memorandum dated Dec 19, 2008 to Bart Hoskins from Marguerite (Peg) Pelletier.

ATTACHMENT 1

TechLaw
Quality & Integrity

175 Cabot Street, Suite 415
Lowell. MA 01854-3650
978-275-9730
978-275-9489 FAX
www.techlawinc.com

June 29, 2009

Office of Environmental Measurement and Evaluation
US EPA - Region I
11 Technology Drive
North Chelmsford, Massachusetts 01863-2431

To: Mr. Bart Hoskins, EPA TOPO
Via: Mr. Louis Macri, ESAT Program Manager

TDF No. 1457A
Task Order No. 26
Task No. 01

Subject: *Review of the Benthic Community Study performed in Support of the Baseline Ecological Risk Assessment at the Callahan Mine Superfund Site, Brooksville, ME (Interim Deliverable by Brian Beal).*

Dear Mr. Hoskins:

The environmental protection agency requested that the environmental services assistance team perform the following tasks in support of the baseline ecological risk assessment for the Callahan Mine Superfund Site, located in Brooksville, ME:

- Review site history, toxicity test results, and benthic community survey reports.
- Participate in conference calls with the Task Order Project Officer (TOPO) and the remedial project manager to discuss first impressions and plan for the deliverable.
- Participate in a site visit on June 16 and 17, 2009.

This deliverable represents the comments prepared by Dr. Brian Beal (University of Maine, Machias; TechLaw consultant). A second deliverable prepared under the same TDF by Dr. Tom Trott (Suffolk University; TechLaw Consultant) is provided under separate cover.

The task was requested by Mr. Hoskins, the TOPO, under technical direction form No. 1457A. The interim completion date is June 29, 2009, whereas the final completion date is July 15, 2009

Do not hesitate to contact myself at (207)255-1314 or Stan Pauwels at (617) 918-8669 with any questions or comments.

Sincerely,

Brian Beal
Senior Staff Consultant
TechLaw, Inc.

Review of the Benthic Community Study
performed in Support of the Baseline Ecological Risk Assessment
at the Callahan Mine Superfund Site, Brooksville, ME
(Interim Deliverable by Brian Beal).

TDF No. 1457A
Task Order No. 26
Task No. 01

Submitted to the:

Task Order Project Officer
Office of Environmental Measurement and Evaluation
USEPA - New England Regional Laboratory
11 Technology Drive
North Chelmsford, MA 01863-2431

Submitted by:

ESAT - Region I
TechLaw, Inc.
The Wannalancit Mills, 175 Cabot Street, Suite 415
Lowell, MA 01854

June 29, 2009

EPA Contract EP-W-06-017

1.0 GENERAL INTRODUCTION

1.1 Task Description

The Environmental Protection Agency (EPA) issued Technical Direction Form (TDF) No. 1457 on May 27, 2009. The TDF requested that the Environmental Services Assistance Team (ESAT) provide the following support:

- Review site history, toxicity test results, and benthic community survey reports.
- Participate in conference calls with the Task Order Project Officer (TOPO) and the remedial project manager to discuss first impressions and plan for the deliverable.
- Participate in a site visit on June 16 and 17, 2009.

EPA modified the TDF on June 25, 2009 (TDF No. 1457A). The task was modified as follows:

- Attend several more conference calls with the TOPO and RPM to discuss elements for inclusion in the final deliverable.
- Attend a Callahan Mine Superfund Site public meeting in Brooksville, ME on July 9, 2009 to support EPA in answering technical questions that may arise on ecological conditions on the Site.
- Include a qualitative assessment of the benthic invertebrates collected during the June 16 and 17, 2009 Site visit.

1.2 Site history

The Site operated from 1968 to 1972 as an open-pit copper (Cu) mine located in a small estuary called Goose Pond. The pond is connected to Goose Cove in nearby Penobscot Bay via Goose Falls, which is a reversing tidal fall. The mining company built two small dams, one at Goose Falls to cut off incoming tides and one upstream in Goose Pond to block an unnamed stream. These two dams allowed the pond to be drained. The unnamed stream was diverted to a nearby cove (Weir Cove), also on the Penobscot Bay.

The mining company blasted a 600- to 1000-foot wide and 320-foot deep pit into the bedrock below the pond to reach the ore. Five million tons of rock and 800,000 tons of ore rich in Cu and zinc (Zn) were removed from this pit. Ground water was pumped out of the pit and released untreated in Goose Cove. The waste rock was dumped in three large waste rock piles at the Site. The ore was staged at the on-site ore pad and finely crushed. This material was placed in floatation cells which used chemicals to float and remove the metal-rich particles. These particles were concentrated and shipped to an out-of-state smelter. The left-over tailings were discarded in an 11-acre tailings pile built next to the unnamed stream. Some of these materials have washed out into the aquatic habitats at the Site.

Mining activity ended in 1972. The mining company removed most of the infrastructure and seeded the waste rock piles and tailings pile with grasses. The two dams were breached to allow water to refill Goose Pond, including the open pit.

The State of Maine has monitored sediment, surface water, and aquatic biota at Goose Pond and Goose Cove over the last three decades. EPA listed the Site on the national priorities list on September 2002 based on these data. Ground water enriched with heavy metals enters the pond from several seeps along the shoreline down-gradient from the rock piles and tailings pile. The terrestrial and aquatic habitats were sampled between 2004 and 2007 to estimate the potential for ecological risk from exposure to Site-derived contamination. These and other data were used to write a Baseline Ecological Risk Assessment (BERA) for the Site and to develop preliminary remediation goals.

Several lines of evidence were used to evaluate the potential for ecological risk in the aquatic habitats at the Site. These included laboratory toxicity tests with tailings and sediment samples, wildlife food chain modeling, and contaminant measurements in invertebrates and fish. Appraisal of the benthic invertebrate community at the Site and nearby reference locations provided one more line of evidence.

The benthic community study, as reported, appeared to be at odds with several of the other lines of evidence. The toxicity tests showed enough sediment toxicity to warrant remedial action. The food chain modeling suggested the potential for impacts to piscivorous birds and shore birds feeding in the wetlands. The biota residue data documented that some Site-related contaminants accumulate in clams, fish and vegetation. In contrast, the field study concluded that the benthic macroinvertebrate communities at the Site did not differ significantly from those found at off-Site reference locations. No remedial action would be warranted to protect the benthic community if this line of evidence was considered of high confidence.

The review focused on an analysis of the field study conducted by MACTEC for the Maine Department of Transportation. The comments address three main questions: (a) were the study design and field methods rigorous enough to support the conclusions of "no effect", (b) what were the weaknesses or strong points of the study design, and (c) how well did the study succeed in representing the benthic communities at the Site? Additional thoughts are provided about a potential future monitoring program for Goose Pond.

This technical memorandum is organized as follows: Section 2.0 provides general comments and observations on the benthic community study; Section 3.0 outlines the structure of a future benthic invertebrate monitoring program; Section 4.0 provides a summary and conclusions, and Section 5.0 lists references.

2.0 GENERAL COMMENTS AND OBSERVATIONS

2.1 Introduction

The following documents were read as part of the review effort.

a) "Appendix J & K" Maine DOT Callahan Mine Superfund Site – Baseline Ecological Risk Assessment (March 2008). MACTEC Engineering and Consulting, Inc. Project Number 3612-06-2047-27. Appendix J included Benthic Community Survey Reports in the form of two pages giving sampling sites, dates, species collected, and numbers of each species collected. This portion of the document included a letter from Michael P. Johnson, Project Manager at Stantec (Portland, Maine) to Peter Barker at MACTEC Engineering and Consulting, Inc. (Portland, Maine). The letter indicated that Stantec had completed at MACTEC's request a taxonomic identification of 29 "macroinvertebrate samples." Stantec collected the samples from 27 August to 5 September 2007 at two locations in Goose Pond, and at 27 reference locations outside of Goose Pond. The letter indicated that each benthic sample had been sieved in the field using a 500-micron sieve. Invertebrates were hand-picked from each sieved sample, preserved in Carolina Perfect Solution (a preservative containing formaldehyde and phenol), and returned to Stantec's wet lab for identification and counting.

Appendix K included a statistical evaluation of the macroinvertebrate community structure, sediment toxicity test results, and potential chemical stressors. This Appendix referred to benthic samples taken at Goose Pond and at reference locations both in 2005 and 2007. None of the descriptions of Goose Pond samples from 2005 were available in this document; however, most of the data used to compare various community indices (Shannon-Weiner, Evenness, Species Richness, etc.) from Goose Pond to the reference locations came from the 2005 sampling. Only two samples (one each from station GP-11 and GP-14) from the 2007 sampling event were used in the analyses. No reason was given for this replacement of data. The statistical tests that compared "Site vs. Reference" were subdivided into three supposedly independent analyses based on sediment type (i.e., marine, estuarine [permanently flooded], and brackish [irregularly flooded]). These parametric and non-parametric tests indicated statistically similar means (or medians) between the Goose Pond and the reference locations ($P > 0.05$). In addition, cluster analysis was performed for samples from the permanently flooded sites which showed that a group of Goose Pond sampling stations (i.e., GP-12, GP-14, GP-15, GP-9, GP-10, and GP-22) clustered together with no reference sites clustering in that subset. Subsequent analyses (t-tests) compared each of these six sites vs. the ten reference sites and showed similar means. The report concluded that benthic community structure at those six Goose Pond sites did not appear to be impaired relative to the reference sites.

b) "Callahan Benthic Aquatic Infauna table 08" was an Excel spreadsheet with information on the faunal composition of benthic samples taken at Goose Pond and Goose Cove. No dates were given for these samples, even though they appear to be the 2005 data referred to above. Numerous graphs were presented relating levels of heavy metals in sediments and pore water vs. number of taxa and organisms per sample.

c) "Benthic Data for Goose Pond May 2009" was an Excel spreadsheet with the same faunal data presented in the "Callahan Benthic Aquatec Infauna table 08". This spreadsheet also grouped samples from various locations around Goose Pond designated as "similar" by the investigators. For example, GP-13, GP-14, GP-15 were designated "Dyer Cove," GP-22 and GP12 were designated "Mine Waste," etc.

d) "memo 062308" was a Word document comprised of two parts. The first was a TechLaw technical memorandum dated 25 June, 2008 regarding the draft review of sediment toxicity test data presented in the final draft of the BERA for the Site. The second was the actual draft review.

e) "memo_062308 Attachment 3" was a memorandum (no date given) by Jim Heltshe from the Office of Research and Development (ORD) to the TOPO describing problems inherent with the experimental design and subsequent analysis of sediment toxicity tests involving two marine invertebrates.

f) "ORD benthic survey comments" was a Word document dated 19 December 2008 from Marguerite Pelletier (ORD) to the TOPO. The document commented on the draft BERA for the Site. It raised several questions, including the use of a Ponar vs. Van Veen grab, why samples were sorted in the field, how site areas and sampling stations were chosen, the comprehensiveness of the data analyses, and why epifauna (amphipods, and mobile gastropods) were not included in some of the analyses.

2.2 <u>Site Visit</u>

ESAT accompanied the TOPO on June 16 and 17, 2009 for a site visit to Goose Pond, Goose Cove, and reference locations in Orcutt Harbor and Upper/Lower Horseshoe Cove. Sediment samples were taken using one or two coring devices (surface area of small coring device = 0.0033 m^2; surface area of large coring device = 0.0182 m^2) to provide an independent understanding of the benthic community structure at these locations. The devices were inserted approximately 15 cm into the sediment or until they hit a substrate (clay hard pan, cobble, ledge, etc.) which prevented further coring. Two replicate cores were taken at each sampling location.

Samples were taken on 16 June in an intertidal muddy area near the upper intertidal at Goose Cove, a subtidal location at Goose Cove referred to as GC-05, a subtidal location in Goose Pond near Dyer Cove referred to as GP-14, and at a very shallow subtidal location in Goose Pond about half-way between Dyer Cove and the earthen dam next to a stand of *Phragmites* spp. The smaller coring device was used at the GC-05 location, whereas the larger coring device was used at the three other locations. A total of eight samples were taken.

Six attempts were made from a small boat in about 3 m of water (at low tide) to replicate the sampling by MACTEC at GC-05 using a Petite Ponar grab sampler. None of the six attempts produced sediment or fauna.

ESAT and the TOPO visited reference stations R-11 and R-12 (on the western and eastern side of Outer Horseshoe Cove, respectively) and R-14 (above the falls at Horseshoe Cove) on 17 June 2009. The small coring device was used at R-11 and R-12, whereas both coring devices were used at R-14. The goal was to see if increasing the surface area of the samples would result in greater species diversity and/or total number of organisms because of the unsuccessful attempts on 16 June to take a "good" sample using the Petite Ponar grab, and because all the samples taken by MACTEC used this particular device. ESAT and the TOPO also visited reference station R-21 on the western shore of outer Orcutt Harbor to take two small cores.

Three other stations in Goose Pond were also visited on 17 June 2009. The first was near Goose Pond Falls, about 30 meters above the road. The second was on the opposite shore (east) from Dyer Cove, and the third was at GP-11. Small cores were taken at the first two sites, and both coring devices were used at GP-11. (Total number of small cores taken = 6; total number of large cores taken = 2). A total of 16 small cores and 10 large cores were collected during the two-day Site visit.

All benthic samples taken during the 16 and 17 June, 2009 Site visit were each placed separately into labeled plastic bags, placed on ice in a cooler, and stored in a walk-in cooler (5^oC) at the University of Maine at Machias until processing on 18 June 2009. The samples were processed separately by washing the sediments through a 500-micron sieve. Residue containing benthic infauna, epifauna, detritus, and other sedimentary material was preserved in a plastic jar using 10% buffered formalin. Rose Bengal, a stain that helps detect living organic material, was added to each preserved sample to increase efficiency of sorting organisms from each sample.

The samples were processed in the laboratory on 18 and 19 June, 2009. All the preserved (formerly live) infauna and epifauna organisms were removed, identified to the lowest taxonomic level when possible, and counted. A Shannon-Weiner diversity index was calculated for each sample, as follows:

$$H' = \Sigma\, p_i \ln p_i,$$

where p_i is the proportion of organisms belonging to the i^{th} species, and ln is the natural logarithm (i.e., base e).

In addition, evenness was calculated for each sample using:

$$\text{Evenness} = H'/\ln S,$$

where S = species richness, or the number of species in the sample.

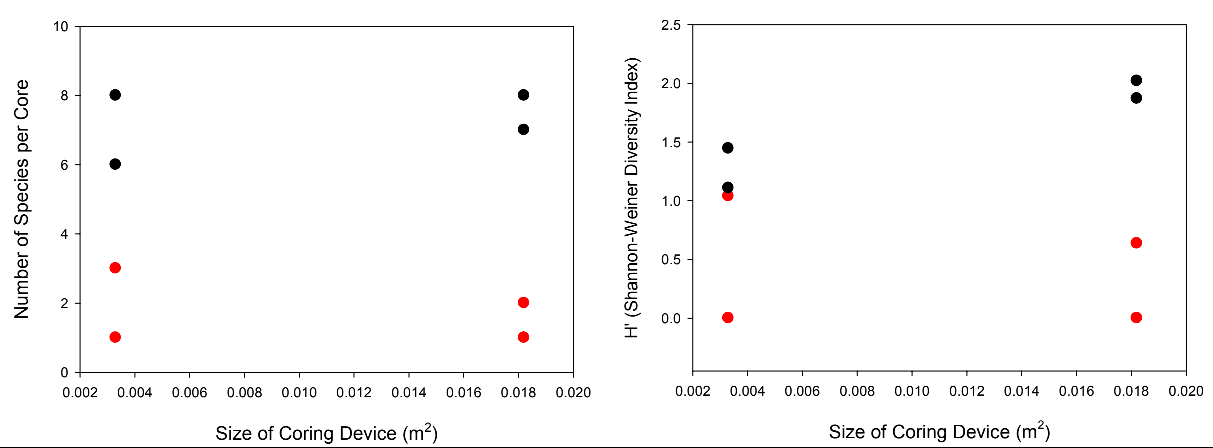
2.3 Data analysis and interpretation

The sampling was performed to corroborate information about samples taken in 2005 and 2007, not to provide a complete characterization of the benthic fauna at the Site or its reference locations. Therefore, this rather limited sampling effort cannot be considered representative of either Goose Pond or any of the reference locations. Certainly more samples would have had to be collected both at the Site and reference locations, and sampling should be done during different times of the year, to accurately determine seasonal trends in diversity and total number of organisms per sample or per species.

Attachment 1 and 2 summarize the counts of the benthic samples taken on 16 and 17 June, 2009. The 2009 sampling effort identified 21 species of marine invertebrates (infauna and epifauna) in Goose Pond. This total compares to 22 species identified in samples collected from Goose Pond in 2005 and 2007. However, the species composition differed. Thirty marine invertebrate species (infauna and epifauna) have been identified in Goose Pond based on combining all samples collected in 2005, 2007, and 2009. Goose Pond did not represent a very diverse area compared to other locations along the eastern Maine Coast. For example, 15 to 25 species per core (0.0182 m^2) have been reported from intertidal samples collected from Cobscook Bay (B. Beal, unpubl., Maine Oil Spill Advisory Council-funded research, 2008/2009). Also, the samples contained relatively low numbers of organisms per species. None of the samples collected outside of Goose Pond in 2005, 2007, and 2009 were particularly diverse though. The 2009 sampling did not target highly mobile marine invertebrates such as crabs. Evidence of crab damage (chipped and crushed juvenile and adult individuals of *Mya arenaria*) was observed at all but GP-11. This damage could have resulted from green crabs (*Carcinus maenas*) observed in Goose Cove, or rock crabs (*Cancer irroratus*), or lobsters (*Homarus americanus*). Killifish (*Fundulus* spp.), which are also predators of juvenile *Mya*, were observed near GP-11.

Small and large core samples were taken at GP-11 and R-14 to determine if sampling area could help explain variation in species number, H', or evenness per sample. A two-way, fixed-factor Analysis of Variance (ANOVA) was used with sampling location and corer size as the main factors. No interaction term was significant ($P > 0.20$) for any of the three analyses, and neither was corer size ($P > 0.50$; **Figure 1**).

Figure 1. Relationship between size of coring device and number of species per core and H'. Red circles represent samples taken at GP-11. Black circles represent samples taken at R-14.



This analysis was initially performed because it was unclear what size grab samples had been used in 2005 and 2007. A review of the available information showed that a 15 cm x 15 cm Petite Ponar grab sample was used in the previous sampling efforts. The available information shows that investigators in 2005 and 2007 took ten samples from a given location, and combined the contents of all ten samples to form a single, composite sample per location.

Attachments 1 and 2 show that samples taken from Goose Pond on 16 and 17 June, 2009 had high spatial variability in total number of marine organisms and diversity (H'). For example, samples taken at GP-11 (above the earthen dam in Goose Pond) in a dense stand of *Ruppia maritima*, and where salinity on the afternoon of 17 June 2009 was zero, contained a few small gastropods, *Hydrobia minuta*, a few annelids, and many Coleoptera larvae (total number of marine invertebrates in the four samples taken at GP-11 = 4; **Attachment 2**). In contrast, benthic samples taken at GP-2 near the northern end of Goose Pond in a gravelly substrate about 20 meters from the falls contained the most species (13) of any sample taken on those two dates, as well as the highest number of individuals per sample (**Attachment 2**).

A single-factor ANOVA was performed on the number of species per core and the diversity index (H') to determine if the observed spatial variation in the June 2009 samples was statistically significant. This analysis gives an understanding of location-to-location variability within Goose Pond. It cannot be performed if a single sample is taken per location because no within-site variation can be estimated. When more than one sample is taken per location, and processed/analyzed separately, the single-factor ANOVA tells if within-site variability exceeds between-site variability, or vice-versa. It provides an insight into community dynamics that would otherwise not be measured.

The overall F-test for number of species per core was significant ($P = 0.0075$), and the comparison of all samples taken at GP-11 (this included samples taken with both sizes of coring devices) to those taken elsewhere in Goose Pond (which also included samples taken with both sizes of coring devices) was also significant ($P = 0.0021$). In fact, 65% of the spatial variability associated with number of species per core could be explained by this particular contrast. The overall F-test for mean diversity was not significant ($P = 0.0761$), but a similar contrast comparing the mean from the four GP-11 samples to the mean of the eight samples taken elsewhere in the Pond was highly significant ($P = 0.0082$). The percent of variability explained by this particular contrast was 98%. Neither of these results are surprising, however, since GP-11 occurs at or near the upper fringe of marine habitat in Goose Pond. In summary, this test showed that a characterization of the benthic fauna at Goose Pond should include a number of samples taken within and between many locations to better understand the nature of spatial variability with respect to the dependent variables of interest.

The Site visit on 16 and 17 June, 2009 also included excursions to several of the reference locations used by MACTEC. Those locations included the mouths of Horseshoe Cove and Orcutt Harbor, Goose Cove, and the upper portions of Horseshoe Cove in area near Bell Marsh. It was not until the Final Remedial Investigation Report was read that the logic of why certain reference sites were chosen was understood. The mouth of Orcutt Harbor and Horseshoe Cove were used in the past as reference areas for the subtidal location at Goose Cove (GC-05) rather than Goose Pond proper. With hindsight, the time spent investigating R-11, R-12, and R-14 may have been better spent in and around the Bagaduce River, which served as reference areas for Goose Pond. It is difficult to report if the Bagaduce River locations were comparable to Goose Pond because they were not observed during the Site visit due to time limitations.

The briefness of the Site visit and the limited sampling effort also made it a challenge to fully understand the conditions at the Site and surrounding area, and to provide an fully-supported opinion on what benthic community type and structure ought to be present. Goose Pond should be considered unique in terms of habitat characteristics. First, the tidal amplitude observed on the neap tides during the 16 and 17 June visit was less than 0.5 m, and probably closer to 0.25 m. This is in sharp contrast with any other area in and around Penobscot Bay. Second, Goose Pond shows salinity gradients which vary both spatially and (probably) temporally. Third, it was unusual to find soft-shell clams (*Mya arenaria*) and sand worms (*Neanthes virens*) next to and just below the earthen dam given that salinities in this area were at or near zero during the Site visit. Finally, it was a surprise to find substantial populations of large *Mya* (50-75 mm in shell length, SL) from the earthen dam north to a small marsh located near GP-1.

A review of the tables in Appendix J (Benthic Community Survey) of the BERA showed that the fauna at the Bagaduce salt pannes (which was used as a reference for GP-11) was very similar to the fauna at GP-11 (i.e., mostly insect larvae, a few amphipods, and one *Neanthes diversicolor*). It would appear from the few photos of this salt panne in Appendix N that this area is similar in habitat with GP-11. That the two areas have similar marine invertebrate fauna suggests that the benthic community is not severely impacted at GP-11. This assertion, however, is based on minimal data and only relates to a single location in Goose Pond. It was not possible to discern from photos in Appendix N or descriptions of habitats in the Final Remedial Report Investigation (pp. 3-60 to 3-63) if the other reference locations in the Bagaduce River have habitat comparable to Goose Pond.

The available information showed that the sampling techniques and designs used in the 2005 and 2007 benthic community surveys were unsuitable to fully characterize Goose Pond relative to other shallow subtidal or marginally intertidal sites in the Cape Rosier-Brooksville-Castine area. Several limitations in the sampling were discussed in the 19 December 2008 memorandum from Marguerite Pelletier (ORD) to the TOPO. Those are not repeated here, but additional issues are discussed below:

- Samples were taken from Goose Pond and the reference sites during one sampling season (late summer/early fall) in both 2005 and 2007. Although this allows comparisons between sites, the information is only specific to a certain season. Due to the nature of organisms having different reproductive patterns/cycles, samples restricted to one season of the year limit the conclusions one can draw about the marine benthic community.

- As many as ten samples were taken at each of the Site and reference locations in 2005 and 2007. Those samples appear to have been homogenized into single samples. The investigators ignored within-site variability by homogenizing samples. **Attachments 1 and 2** provide several examples of how variable two samples from single locations in this area can be. As a result, the 2005 and 2007 sampling effort may or may not be truly representative of the location from which the samples were taken. These studies also cannot assert that any comparison (statistical, or otherwise) of the benthic fauna in Goose Pond versus the reference sites is unbiased. Note that the grab sampling in Goose Cove (GC-05; 16 June 2009) and Horseshoe Cove (R-12; 17 June 2009) using the Petite Ponar failed to produce a single “good” sediment sample after eight or so attempts. The substrate at both locations was soft enough to sample the same sites using a round coring tube (surface area = 0.0033 m^2).

- It is unclear how locations within a given site were chosen (this issue is different from how reference sites were chosen). That is, once a decision was made to take samples from GP-14 or GP-11, for example, it is not known what criteria were used to sample the benthos from a particular area at those locations. These remarks are similar to the second objection listed above, and could only have been dealt with objectively by using a stratified random sampling approach within a given location or some blocking. In other words, deciding that a given sample is truly representative of a given site or location requires more than one sample; otherwise, assumptions must be made that the area from which the sample was collected was homogeneous over some spatial scale not identified in the 2005 and 2007 studies.

- The investigators used 2005 data in their statistical comparison of Goose Pond community indices (number of species present, H’, evenness, etc.), except that at least two data points were taken from the 2007 sampling at G-11 and G-14 within Goose Pond (see Appendix K of the Draft Final BERA). The investigators provided no logical reason for this “substitution” of data, and as a result, the statistical comparisons that show no adverse effects on community structure are suspect.

- The available information suggests that samples were not stained prior to picking. Staining is a common technique to make picking fauna from benthic cores more efficient. Not staining could have led to missing various fauna. For example, some of the samples taken in June 2009 from GC-Upper intertidal contained >250 *Tubificoides benedeni*. These are small oligochaetes that are difficult to see in sieved and stained samples, let alone from sieved and unstained samples.

- The statistical comparisons for the benthic community indices in Appendix K of the Draft Final BERA do not have enough statistical power (i.e., 1- beta is less than 0.8; below this minimal power level, a statistical test does not have enough power to detect differences if they are in fact present). The information provided by MACTEC shows that its statistical tests do not meet this minimal power requirement. As such, one cannot know if the failure to reject the null hypothesis of no difference between means characterizing the community is a “Type II error” (i.e., failing to reject a null hypothesis when the null hypothesis is incorrect) or not. The only way to correct for the lack of statistical power is to take more samples in a given location. No statistical tests are available to improve the power of the hypotheses reported in the MACTEC report. Also, a fundamental assumption of statistical comparisons is that the Site and reference locations differ only in their level of “impactedness”, but not in their habitat characteristics. This assumption is severely challenged based on the information discussed earlier in this technical memorandum.

- It appears that no commercial or recreational harvesting of marine worms, soft-shell clams, or sea urchins takes place in Goose Pond. However, it is not known if harvesting has occurred in the past or is still occurring at one or more of the reference locations. Any comparison between the Site and reference locations may be biased without this information. For example, suppose that dragging for sea scallops and/or sea urchins occurred at one or more of the reference locations in the Bagaduce River during the winter before sampling in 2005 or 2007. Dragging or harvesting can disturb sediments and influence benthic community structure. This activity may reduce species diversity and total number of organisms so that the reference locations would appear more similar (generally depauperate) to the benthic community in Goose Pond.

The benthic fauna observed at Goose Pond during the 16 and 17 June 2009 Site visit was “typical” of what would be expected in this type of habitat. However, Goose Pond represented a unique habitat. It was unclear if truly representative “control” locations in that area were available to allow for “simple, unbiased” comparisons of various dependent variables between Goose Pond and these reference locations.

Other benthic habitats in Cobscook Bay (eastern Maine) have a more diverse fauna. However, sampling alone cannot discern the reason(s) why differences in sample means might exist between two or more locations. Sampling can yield quantitative information that should be considered observational for the purpose of constructing a conceptual model of community dynamics at a given location. It may be important to assign a reason or cause for a given observation, but only guesses are possible without knowing the mechanisms (e.g., predation, competition, weather, chemical toxicity) at play in a given system.

For example, *Fundulus* and other small fish such as sticklebacks (*Gasterosterus* spp.) were noted near the earthen dam at Goose Pond. It is possible that other fish species, such as silversides (*Menidia menidia*) are found in Goose Pond as well**.** Decapod crustaceans (*Crangon septemspinosa*) love in Goose Pond because these shrimp were sampled in one of the benthic cores during the Site visit. Green crabs (*Carcinus maenas*) have also been found in Goose Pond during past investigations. It is not unreasonable to assume that other, larger decapods such as rock crabs (*Cancer irroratus*) or lobsters (*Homarus americanus*) exist there, too.

The presence and density of mobile predators can affect benthic species composition, irrespective of other factors. Trapping these predators, followed by gut analysis, is needed because such organisms are not effectively sampled using bottom cores. This kind of information helps build a stronger biological model of the system. Ultimately, a study should be considered that combines a statistically valid sediment sampling design with experimental manipulations (i.e., removal of predators from a certain bottom; addition of chemically inert sediment; addition of prey items, etc.). Only then would it be possible to discern if the dynamics at Goose Pond differ significantly from those occurring in "similar" reference habitats.

A surprising observation during the 16 and 17 June, 2009 Site visit was the lack of larger populations of grazers such as Hydrobid and Littorinid snails in Goose Pond. This may be due to intense predation by decapods or fish, due to the lack of significant food for these organisms, due to poor natural recruitment in recent years, or because the sediments in and around Goose Pond are toxic in places. Any or all of these hypotheses are reasonable, but without some objective experimental manipulation, it is impossible to discern which, if any, can explain the mechanisms underlying these observations.

3.0 OUTLINE FOR A POTENTIAL FUTURE MONITORING PROGRAM

It is anticipated that a monitoring program may need to be implemented before and/or after any sediment remediation at Goose Pond. Provided below is a brief outline of a sampling design and study that could be implemented to help understand and characterize the benthos in Goose Pond and to uncover some of the mechanisms that affect distributional and abundance patterns over time.

The pond should first be sub-divided into strata using toxicological information, sediment grain size, or habitat (e.g., eelgrass vs. unvegetated soft sediments vs. unvegetated gravelly sediments). Next, within each strata, a sampling design at various locations would consist of randomly selecting five to ten "blocks" of about 2m x 1m per location. Two benthic cores would be collected from each block. This approach helps to understand spatial variability at a sampling location since the blocks are randomly allocated. It also allows one to ask if significant block-to-block variation (a measure of within-location spatial variability) exists at a given location rather than make assumptions. This information can be used in subsequent sampling or in an experimental setting if considerable spatial variability exists at a given location. The key point of this approach is that it can accurately and efficiently characterize the benthic infauna within and between strata at Goose Pond.

Much of the benthic community recruits to the sediment via the water column because most benthic invertebrates are broadcast spawners with a diphasic life-history, i.e., adults live in or on the sediment, larvae are planktonic in the water column, and juveniles settle out of the water column unto the sediment. As a result, the sampling design described above should be carried out at least in the spring and fall. This is the minimum sampling intensity that could be undertaken to describe temporal variability. A single sampling date cannot discern temporal patterns.

For example, suppose we wish to sample ten locations in Goose Pond and we want to take samples during spring and fall to get an estimate of temporal variability. Also, suppose that we randomly allocate five blocks (with two core replicates per block) at each of the ten locations. Total number of samples would be: 10 locations x 5 blocks/location x 2 cores/block x 2 times/year = 200 (or 100 samples in the spring and 100 samples in the fall). **Table 1** gives the sources of variation (null hypotheses) that can be tested with this design.

Table 1. Example of a sampling design to address spatial and temporal variation in the benthic community at Goose Pond.

Source of variation	df	Mean Square Estimate	F-ratio
Sampling Location	9	$\sigma^2_e + cn\sigma^2_B + bcn\alpha_A$	MS Location/MS Blocks(Location)
Blocks (Location)	40	$\sigma^2_e + cn\sigma^2_B$	MS Blocks(Location)/MSE
Time	1	$\sigma^2_e + n\,\sigma^2_{BC(A)} + abn\alpha_C$	MS Time/MS Time x Block(Station)
Time x Location	9	$\sigma^2_e + n\,\sigma^2_{BC(A)} + bn\alpha_{AC}$	MS Time x Location/MS Time x Block(Station)
Time x Block (Location)	40	$\sigma^2_e + n\,\sigma^2_{BC(A)}$	MS Time x Block(Location)/MSE
Error	100	σ^2_e	
Total	199		

Note: Each source of variation should be considered a statistical null hypothesis. Sampling location (a = 10) and Time (b = 2) would be considered "fixed factors," whereas Blocks (c = 5) would be considered a "random factor." (n = 2).

The following null hypotheses can be tested using this design:

- No difference exists in the dependent variable (e.g., sediment toxicity, species richness, sediment grain size, total number of organisms) between sampling locations;
- No variability exists within locations;
- No difference exists in the dependent variable between sampling dates;
- The pattern observed from location-to-location with respect to the dependent variable is the same from one sampling date to the other;
- The block-to-block variation observed in the dependent variable at each sampling location is the same from one sampling date to the other.

The proposed sampling program should be accompanied by several manipulative field studies to determine what factors may control distribution and abundance in Goose Pond. Many experimental approaches are available for this purpose. Two approaches are described below.

- Several experimental designs have been used to test hypotheses about the dynamics of juvenile *Mya arenaria* (e.g., Beal et al. 2001; Beal and Kraus 2002; Beal 2006) in the intertidal and shallow subtidal of eastern Maine. This information was then used to determine the relative importance of both biotic and abiotic factors affecting clam growth and survival. Cultured animals were used to ensure similar initial sizes and known ages. Juvenile clams can be deployed in experimental units (0.0182 m^2) and placed *in situ* easily. Units can be covered with plastic mesh netting to exclude/deter epibenthic predators such as crabs, fish, and birds. Arrays of units can easily be placed in most habitats at almost any time during the year, which permits spatial and temporal estimates of growth and survival and gives an insight into the dynamics not only affecting clams, but other residents of the marine benthos.
- Another approach is to fill similar-sized experimental units with defaunated sediments from the particular study location. Half of the units can be covered with mesh netting to reduce/deter

predators while the other half can be left uncovered. Units can be removed at selected times during the year to follow recruitment dynamics. Both approaches help understand data from systematic sampling programs. By themselves, sampling programs can only generate hypotheses about why certain species occur in selected areas or why so few or so many individuals of a species are present in a location at a specific time. Manipulative studies performed at the same time as sampling programs can provide insights into how communities function and what factors affect the dynamics of populations of interest.

Finally, suppose the following situation occurs: (a) a decision is made to dredge mine waste "hot spots" in Goose Pond, followed by restoration of the disturbed areas, and (b) a viable marine benthic community currently exists in Goose Pond. A concern may exist that dredging the hot spots may cause a negative impact on the existing benthic community. This situation presents an opportunity to use the "Before-After-Control-Impact," or BACI (Underwood 1991, 1994) sampling protocol to examine impact effects.

The first objective of BACI is to find a control area, or reference area, that closely resembles the impacted areas in as many ways as possible (i.e., similar habitats, similar tidal amplitudes, similar faunal characteristics, etc.). Then, a sampling study as described above is started at each location BEFORE dredging. The sampling continues AFTER dredging is completed at both locations for at least one year using the same sampling approach. One then compares the relative change in a dependent variable at the dredged site over time to the change in the same dependent variable over the same time interval at the control location (multiple controls are better than a single control). If the relationship or pattern between the dependent variable and sampling time is similar in both locations, then no "impact" is detected. Conversely, if the pattern differs between control and dredged site, then an "impact" is detected (the "impact" may be positive or negative depending on the slope of the relationship at the dredged location relative to the control site[s]).

4.0 SUMMARY AND CONCLUSION

EPA requested that ESAT (a) review site history, toxicity test results, and benthic community survey reports for the Callahan Mine Superfund Site; (b) participate in conference calls with the task order project officer and the remedial project manager to discuss first impressions and plan for the deliverable; and, (c) participate in a Site visit which took place on June 16 and 17, 2009.

The review of the benthic community field study, plus supplemental data collected during the Site visit, identified numerous limitations in sample design and data analysis which invalidated the conclusions of the study. The major limitations which precluded an accurate assessment of the benthic community were as follows:

- The 2005 and 2007 benthic community sampling design was flawed.
- The sample processing may have introduced some degree of error that reduced the number of species per sample or number of organisms per sample.
- The reference locations may have differed from Goose Pond in ways other than habitat congruity.
- The statistical comparisons of benthic community structure are suspect.

5.0 REFERENCES

Beal, B.F. 2006. Relative importance of predation and intraspecific competition in regulating growth and survival of juveniles of the soft-shell clam, *Mya arenaria* L., at several spatial scales. J. Exp. Mar. Biol. Ecol. 336:1-17.

Beal, B.F. & Kraus, M.G. 2002. Interactive effects of initial size, stocking density, and type of predator deterrent netting on survival and growth of cultured juveniles of the soft-shell clam, *Mya arenaria* L. in eastern Maine. Aquaculture 208:81-111.

Beal, B.F., Parker, M.R. & Vencile, K.W. 2001. Seasonal effects of intraspecific density and predator exclusion along a shore-level gradient on survival and growth of juveniles of the soft-shell clam, *Mya arenaria* L., in Maine, USA. J. Exp. Mar. Biol. Ecol. 264:133-169.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Aus. J. Mar. Freshwat. Res. 42:569-587.

Underwood, A.J. 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol. Appl. 4:3-15.

Attachment 1. Benthic infauna sampled on 16 June 2009 at several locations within the Callahan Mine site (Goose Pond = GP) and in Goose Cove (= GC) in Cape Rosier, Maine. (GCUL = Goose Cove, Upper Intertidal, Large Corer [0.0182 m^2]; a,b refer to replicate samples; GC-05 = subtidal cores, Small Corer [0.0033m^2]; GP-14 = Goose Pond, near Dyer Cove, large corer; GP-1 = Goose Pond, half-way between earthen dam and Dyer Cove, adjacent to large stand of *Phragmites* spp.)

Locations:	GCULa	GCULa	GC-05a	GC-05b	GP-14a	GP-14b	GP-1a	GP-1b
Annelids								
Oligochaetes								
Tubificoides benedeni	> 250	3		2				
Tubificidae spp.					13	8		
Polychaetes								
Aglaophamus neotenus			4					
Capitella capitata							3	7
Cossura longocirrata								
Eteone longa								
Fabricia sabella								
Glycera dibranchiate								
Heteromastus filiformis								
Neanthes diversicolor					10	6		
Neanthes virens		1			8	12	5	7
Nepthys ciliata								
Nicomache lumbricalis								
Pholoe minuta								
Polydora quadrilobata								
Pygiospio elegans								
Scoloplos fragilis	2				11	30	8	14
Streblospio benedecti					1			
Tharyx acutus								
Echinoderms								
Strongylocentrotus droebachiensis			1					
Mollusks								
Bivalvia								
Gemma gemma								
Macoma balthica								
Mya arenaria	2	2	1			2	4	2
Tellina agilis								
Gastropoda								
Hydrobia minuta		1		3		2	30	16
Nemerteans								
Lineus ruber	1							
Crustaceans								
Amphipoda								
Ampelisca abdita								
Corophium volutator								
Gammarus mucronatus		2						
Decapoda								
Crangon septemspinosa								1
Insects								
Coleoptera larvae								
Total number of organisms	> 255	9	6	4	43	60	50	47
Total number of species	4	5	3	2	5	6	5	6
Shannon-Weiner Diversity H'	0.1172	1.5231	0.8676	0.6732	1.4500	1.3940	1.2010	1.5111
Evenness	0.0845	0.9463	0.7897	0.9710	0.9009	0.7781	0.7461	0.8433

Attachment 2. Benthic infauna sampled on 17 June 2009 at selected reference sites (R-11, R-12, R-14, R-21) near Cape Rosier, Maine. a,b refer to replicate samples. Small cores (Surface area = 0.0033m^2) were taken at each reference site. (R-14aS = Reference site 14, small core a; R-14aL = Reference site 14, large core [Surface area = 0.0182m^2].)

Locations:	R-11a	R-11b	R-12a	R-12b	R-14aS	R-14bS	R-14aL	R-14bL	R-21a	R-21b
Annelids										
Oligochaetes										
Tubificoides benedeni									1	
Tubificidae spp.					20	13	2	4	4	1
Polychaetes										
Aglaophamus neotenus			2	6					4	3
Capitella capitata			1				2		1	1
Cossura longocirrata				4						2
Eteone longa										
Fabricia sabella										
Glycera dibranchiata									1	
Heteromastus filiformis						1		5		
Neanthes diversicolor						1				
Neanthes virens	1				3		2	2		
Nepthys ciliata							1			
Nicomache lumbricalis		1								
Pholoe minuta										
Polydora quadrilobata										
Pygiospio elegans						1				
Scoloplos fragilis					1	1	1	3		
Streblospio benedecti					1	4				
Tharyx acutus										
Echinoderms										
Strongylocentrotus droebachiensis										
Mollusks										
Bivalvia										
Gemma gemma					1	1		5		
Macoma balthica							1	2		
Mya arenaria										
Tellina agilis			1							
Gastropoda										
Hydrobia minuta					4	1		2		
Nemerteans										
Lineus ruber										
Crustaceans										
Amphipoda										
Ampelisca abdita							1			
Corophium volutator							1			
Gammarus mucronatus										
Decapoda										
Crangon septemspinosa										
Insects										
Coleoptera larvae										
Total number of organisms	1	1	4	10	30	23	11	23	11	7
Total number of species	1	1	3	2	6	8	8	7	5	4
Shannon-Weiner Diversity H’	0	0	1.0400	0.6734	1.1090	1.4450	2.0200	1.8710	1.3900	1.2770
Evenness	0	0	0.9464	0.9710	0.6191	0.6947	0.9713	0.9613	0.8635	0.9212

Attachment 2 (cont.). Benthic infauna sampled on 17 June 2009 at several locations within the Callahan Mine site (Goose Pond = GP) in Cape Rosier, Maine. a,b refer to replicate samples. Small cores (Surface area = 0.0033m^2) were taken at each site. (GP-11aS = Goose Pond site 11, small core a; GP-11aL = Goose Pond site 11, large core [Surface area = 0.0182m^2].) GP-2 (near Falls); GP-3 (across from Dyer Cove)

Locations:	GP-2a	GP-2b	GP-3a	GP-3b	GP-11aS	GP-11bS	GP-11aL	GP-11bL
Annelids								
Oligochaetes								
Tubificoides benedeni	38	76	1			2		
Tubificidae spp.	1	26						
Unidentified oliogochaete species		28						
Polychaetes								
Aglaophamus neotenus			1					
Capitella capitata		1						
Cossura longocirrata						1		
Eteone longa	1	1						
Fabricia sabella		1						
Glycera dibranchiata								
Heteromastus filiformis	1	1	1					
Neanthes diversicolor	1	1						
Neanthes virens							2	
Nepthys ciliata								
Nicomache lumbricalis								
Pholoe minuta	1							
Polydora quadrilobata	5	17						
Pygiospio elegans		8						
Scoloplos fragilis			2	2				
Streblospio benedecti		7		1				
Tharyx acutus				1				
Echinoderms								
Strongylocentrotus droebachiensis								
Mollusks								
Bivalvia								
Gemma gemma								
Macoma balthica								
Mya arenaria								
Tellina agilis								
Gastropoda								
Hydrobia minuta				4	3	1	4	5
Nemerteans								
Lineus ruber		1	1					
Crustaceans								
Amphipoda								
Ampelisca abdita								
Corophium volutator								
Gammarus mucronatus								
Decapoda								
Crangon septemspinosa								
Insects								
Coleoptera larvae* Terrestrial – not included in summary statistics					10	5	22	27
Total number of organisms	48	168	6	8	3	4	6	5
Total number of species	7	12	5	4	1	3	2	1
Shannon-Weiner Diversity H’	0.8238	1.6140	1.5610	1.2130	0	1.0400	0.6365	0
Evenness	0.4233	0.6495	0.9697	0.8750	0	0.9464	0.9183	0

ATTACHMENT 2

TechLaw
Quality & Integrity

175 Cabot Street, Suite 415
Lowell. MA 01854-3650
978-275-9730
978-275-9489 FAX
www.techlawinc.com

June 29, 2009

Office of Environmental Measurement and Evaluation
US EPA - Region I
11 Technology Drive
North Chelmsford, Massachusetts 01863-2431

To: Mr. Bart Hoskins, EPA TOPO
Via: Mr. Louis Macri, ESAT Program Manager

TDF No. 1457A
Task Order No. 26
Task No. 01

Subject: *Review of the Benthic Community Study Performed in Support of the Baseline Ecological Risk Assessment at the Callahan Mine Superfund Site, Brooksville, ME (Interim Deliverable by Thomas Trott).*

Dear Mr. Hoskins:

The environmental protection agency requested that the environmental services assistance team perform the following tasks in support of the baseline ecological risk assessment for the Callahan Mine Superfund Site, located in Brooksville, ME:

- Review site history, toxicity test results, and benthic community survey reports.
- Participate in conference calls with the Task Order Project Officer (TOPO) and the remedial project manager to discuss first impressions and plan for the deliverable.
- Participate in a site visit on June 16 and 17, 2009.

This deliverable represents the comments prepared by Dr. Thomas Trott (Suffolk University; TechLaw consultant). A second deliverable prepared under the same TDF by Dr. Brian Beal (University of Maine, Machias; TechLaw Consultant) is provided under separate cover.

The task was requested by Mr. Hoskins, the TOPO, under technical direction form No. 1457A. The interim completion date is June 29, 2009, whereas the final completion date is July 15, 2009

Do not hesitate to contact myself at (617)305-1968 or Stan Pauwels at (617) 918-8669 with any questions or comments.

Sincerely,

Thomas Trott
Senior Staff Consultant
TechLaw, Inc.

Review of the Benthic Community Study
performed in Support of the Baseline Ecological Risk Assessment
at the Callahan Mine Superfund Site, Brooksville, ME
(Interim Deliverable by Thomas Trott).

TDF No. 1457A
Task Order No. 26
Task No. 01

Submitted to the:

Task Order Project Officer
Office of Environmental Measurement and Evaluation
USEPA - New England Regional Laboratory
11 Technology Drive
North Chelmsford, MA 01863-2431

Submitted by:

ESAT - Region I
TechLaw, Inc.
The Wannalancit Mills, 175 Cabot Street, Suite 415
Lowell, MA 01854

June 29, 2009

EPA Contract EP-W-06-017

1.0 GENERAL INTRODUCTION

1.1 Task Description

The Environmental Protection Agency (EPA) issued Technical Direction Form (TDF) No. 1457 on May 27, 2009. The TDF requested that the Environmental Services Assistance Team (ESAT) provide the following support:

- Review site history, toxicity test results, and benthic community survey reports.

- Participate in conference calls with the Task Order Project Officer (TOPO) and the Remedial Project Manager (RPM) to discuss first impressions and plan for the deliverable.

- Participate in a site visit on June 16 and 17, 2009.

EPA modified the TDF on June 25, 2009 (TDF No. 1457A). The task was modified as follows:

- Attend several more conference calls with the TOPO and RPM to discuss elements for inclusion in the final deliverable.

- Attend a Callahan Mine Superfund Site public meeting in Brooksville, ME on July 9, 2009 to support EPA in answering technical questions that may arise on ecological conditions on the Site.

- Include a qualitative assessment of the benthic invertebrates collected during the June 16 and 17, 2009 Site visit.

1.2 Site history

The Site operated from 1968 to 1972 as an open-pit copper (Cu) mine located in a small estuary called Goose Pond. The pond is connected to Goose Cove in nearby Penobscot Bay via Goose Falls, which is a reversing tidal fall. The mining company built two small dams, one at Goose Falls to cut off incoming tides and one upstream in Goose Pond to block an unnamed stream. These two dams allowed the pond to be drained. The unnamed stream was diverted to a nearby cove (Weir Cove), also on the Penobscot Bay.

The mining company blasted a 600- to 1000-foot wide and 320-foot deep pit into the bedrock below the pond to reach the ore. Five million tons of rock and 800,000 tons of ore rich in Cu and zinc (Zn) were removed from this pit. Ground water was pumped out of the pit and released untreated in Goose Cove. The waste rock was dumped in three large waste rock piles at the Site. The ore was staged at the on-site ore pad and finely crushed. This material was placed in floatation cells which used chemicals to float and remove the metal-rich particles. These particles were concentrated and shipped to an out-of-state smelter. The left-over tailings were discarded in an 11-acre tailings pile built next to the unnamed stream. Some of these materials have washed out into the aquatic habitats at the Site.

Mining activity ended in 1972. The mining company removed most of the infrastructure and seeded the waste rock piles and tailings pile with grasses. The two dams were breached to allow water to refill Goose Pond, including the open pit.

The State of Maine has monitored sediment, surface water, and aquatic biota at Goose Pond and Goose Cove over the last three decades. EPA listed the Site on the national priorities list on September 2002 based on these data. Ground water enriched with heavy metals enters the pond from several seeps along the shoreline down-gradient from the rock piles and tailings pile. The terrestrial and aquatic habitats were sampled between 2004 and 2007 to estimate the potential for ecological risk from exposure to Site-derived contamination. These and other data were used to write a Baseline Ecological Risk Assessment (BERA) for the Site and to develop preliminary remediation goals.

Several lines of evidence were used to evaluate the potential for ecological risk in the aquatic habitats at the Site. These included laboratory toxicity tests with tailings and sediment samples, wildlife food chain modeling, and contaminant measurements in invertebrates and fish. Appraisal of the benthic invertebrate community at the Site and nearby reference locations provided one more line of evidence.

The benthic community study, as reported, appeared to be at odds with several of the other lines of evidence. The toxicity tests showed enough sediment toxicity to warrant remedial action. The food chain modeling suggested the potential for impacts to piscivorous birds and shore birds feeding in the wetlands. The biota residue data documented that some Site-related contaminants accumulate in clams, fish and vegetation. In contrast, the field study concluded that the benthic macroinvertebrate communities at the Site did not differ significantly from those found at off-Site reference locations. No remedial action would be warranted to protect the benthic community if this line of evidence was considered of high confidence.

The review focused on an analysis of the field study conducted by MACTEC for the Maine Department of Transportation. The comments address three main questions: (a) were the study design and field methods rigorous enough to support the conclusions of "no effect", (b) what were the weaknesses or strong points of the study design, and (c) how well did the study succeed in representing the benthic communities at the Site? .

This technical memorandum is organized as follows: Section 2.0 provides general comments on the benthic community study; Section 3.0 outlines the structure of a future benthic invertebrate monitoring program; Section 4.0 provides a summary and conclusions, and Section 5.0 lists references.

2.0 GENERAL COMMENTS

General Comment 1:

The benthic community survey for the BERA depended heavily on direct comparison of locations based on matching habitats. Benthic samples collected using a Ponar grab sampler were sieved on a 0.5 mm screen to separate organisms for identification and counting. Sampling locations were selected around Goose Pond, outside the pond entrance in Goose Cove, and at several reference locations around Cape Rosier. Samples were processed in the field and the collected invertebrates were archived. The benthic communities from the affected areas at the Site were compared to those from reference locations using descriptive community parameters, such as species richness, diversity, and evenness.

The study concluded that the species diversity of the benthic community sampled at the Site was not different from that found at the reference locations. This conclusion was invalid because of a fundamental limitation in the sample design, i.e., the lack of closely matching habitats between the Site and the reference locations.

While the sample design is simple in principle, the major difficulty with this method is finding reference locations that match the habitats at the Site. The benthic community study did not achieve this important goal, except for finding reference habitats that were closely matched with Goose Cove.

Goose Pond, the body of water of primary concern, is a brackish coastal pond which receives tidal saltwater intrusions from Penobscot Bay and varying amounts of freshwater from Marsh Creek. None of the reference sites included a coastal pond, even though Mill Pond near Flat Landing would have been an appropriate match. The reference locations in Bell Marsh and the Bagaduce River are not close approximations. The following illustrates this point. The rationale for comparing Goose Pond to Bell Marsh might have been based on the presence of marsh in both locations. However, the marsh in Goose Pond consists of a fringing salt marsh lining the shore opposite of the mine and a freshwater marsh of limited marine influence primarily above the earthen dam. In contrast, Bell Marsh is a comparatively expansive salt marsh typically influenced by semidiurnal tides. These three habitats differ greatly. Different species assemblages can be expected *a priori*, regardless of the presence of mine waste, based on factors such as area, salinity, tidal influence, sediment characteristics and plant assemblages.

General Comment 2:

The study concluded that the benthic communities in Goose Cove and Goose Pond were no different from those found at the surrounding reference locations. This conclusion was supported by measures of community complexity, such as species richness, Shannon Wiener Index of diversity, Peilou's evenness, percent dominance, and abundance. Under-sampling was a key limitation in the sampling design used to assess the benthic community. As a result, the number of samples was not large enough to characterize the benthic community using these indices. This issue could have been avoided by preliminary sampling to construct a species-area curve and estimate the minimum number of samples needed to find most species in the sample locations. Measures like species richness (total number of species present), and diversity cannot be accurate without this approach.

A second limitation in the sample design is that sample stations were not sufficiently distributed spatially within a particular location to yield a robust measure of variation in community structure. Sampling protocols have been designed for this purpose using random sampling or blocked designs, neither of which were used. Community indices that reflect spatial variation, i.e., Peilou's index of evenness of community structure, are erroneous without adequate sampling. The effects of under-sampling and insufficient spatial distribution of sampling station interact to produce, at best, poor estimations of abundance and dominance based on percentage of total organisms from the most frequently identified taxon. These two measures need enough samples distributed across an area for a robust estimation to accurately reflect the sampled community.

General Comment 3:

The benthic community study based its conclusions on samples taken only in the fall during one year. This approach did not account for temporal variability. Temporal variation in benthic community structure is significant from events such as recruitment, survival, and reproduction. The sensitivity of invertebrates to environmental influences, including contaminants, is greatest during the larval phases of their life histories. Sampling during one season only, particularly in the fall which is out of the window of recruitment for most benthic invertebrates, can not fully assess what could live in a particular location. Note that while temporal variability applies to all locations, the sampled habitats were different enough that temporal effects would vary across sample locations.

General Comment 4:

The study assessed the benthic communities based on sediment samples taken with a Ponar grab sampler. Using this sampling device without testing the effectiveness of alternative sampling apparatus compromised the benthic community assessment. The effectiveness with which the Ponar grab sampler penetrates bottom sediment directly affects the sampling efficiency. The ability of the grab to penetrate the substrate and obtain a sample diminishes with harder substrate (e.g., sand and gravel versus mud).

The reliability of Ponar grab samples for the quantitative study of benthic invertebrates is questionable in freshwater habitats (Nalepa et al., 1988; Panis et al., 1995). The reliability may be even less in marine habitats because the diversity of biota found in coarse sediments is greater than in freshwater. The unreliability of the Ponar grab was underscored during the June 16 and 15, 2009 site visit. Five attempts to sample the benthos at three locations using the Ponar grab failed to collect a bottom sample. However, subsequent sampling at the same three locations using a coring device yielded 28 strictly benthic taxa. Note that the benthic survey report prepared for the BERA only found nine strictly benthic taxa.

General Comment 5:

The meiofauna is one group of organisms that was not sampled but are known to be sensitive to copper-rich mine tailings (Lee and Correa, 2005, 2006). This species assemblage, which contains benthic organisms from 0.1 to 1 mm in size, was not sampled due to sieve size. Many of these organisms support species higher in the food chain and their absence can have obvious consequences. Many 0macroinvertebrates, such as the larval stages of polychaetes, molluscs and crustaceans, have meiofaunal life history stages. Clear differences between reference and contaminated sites would have been highly probable had these organisms been included in the benthic community assessment.

General Comment 6:

The conclusions derived from the benthic community data were based on procedures that violated assumptions of the statistical tests used. It is incorrect to use a t-test on derived variables. In other words, t-tests can be used on means of counts or measurements such as number of individuals or number of species in a sample, but not on derived indices like Shannon-Weiner Index and Peilou's Evenness Index. Performing statistical tests on derived variables rather than using the raw data from which they are calculated ignores variation, which is important for calculating t values.

It was appropriate to use the nonparametric Kruskal-Wallis test. In fact, this test should have been used for all statistical testing of derived variables, despite its insensitivity in detecting differences. This would have avoided violating the assumptions of the parametric tests used, which consequently makes the results of such tests less robust. For example, counts of organisms were made from subsamples from a single sample. The result of each subsample was used in the calculations. This approach is known as pseudoreplication, and violates the assumption of parametric tests that samples need to be independent of one another.

General Comment 7:

Qualitative sediment sampling was performed during the June 16 and 17, 2009 Site visit. This effort produced a sketch of the local community structure. Some basic observations are discussed below, even though the low sample size limits the extent that community structure of Goose Pond can be understood with these data.

Most apparent is the absence and/or limited presence of particular taxa that would be expected in a coastal pond. High current velocity is usually associated with high species diversity. This relationship was supported by the diverse assemblage of macroalgae observed directly under the road bridge in the high flow area of the reversing falls between Goose Cove and Goose Pond. However, sampling and walk-abouts showed low species richness in Goose Cove (which is under strong marine influence). Similar locations along the Maine coast are home to sea anemones, hydroids, sponges, mussels, starfish, crabs, and sea squirts. None of these were found in Goose Cove. Several species of gastropods would also be expected, even though only one gastropod, the common periwinkle, was observed.

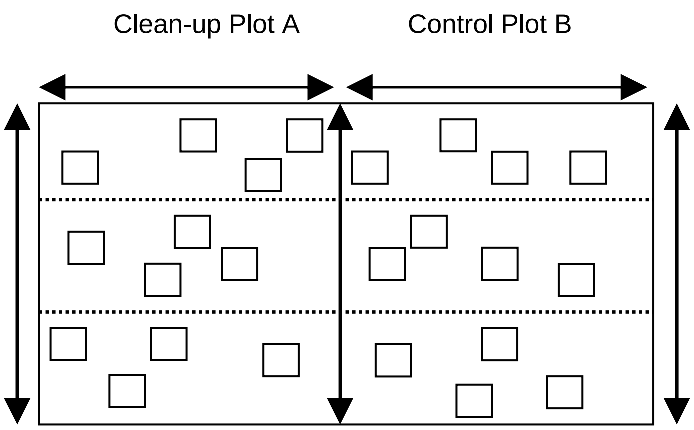
The Goose Pond sediment samples were rich in annelids (21 species) but poor in molluscs (4 species) and arthropods (4 species, excluding insects). Coastal ponds can be low in species richness, though sponges, which were absent, and more gastropods (only one species, *Hydrobia minuta*, was found) would be expected. Mobile species have a high probability of being under-sampled with the coring device and Ponar grab. Such species include green crabs and isopods, neither of which was found during the Site visit. Amphipods were found, but at low species richness (3 species) and abundance (1-2 individuals of each).

Finally, Atlantic coastal ponds are often home to horseshoe crabs that move to shore to spawn and lay eggs. While these were not found in Goose Pond, the benthic survey in support of the BERA observed horseshoe crabs in the Bagaduce River. Their absence may have been due to the season, although this species was not reported to be present in Goose Pond. This species has been given special status in Maine and is monitored through various public and state sponsored programs because the state represents the northernmost limit of its distribution along the Atlantic coast.

3.0 OUTLINE FOR A FUTURE MONITORING PROGRAM

The goal of a future monitoring program is to gauge the effectiveness of remedial actions on the benthic infaunal community in Goose Pond. The logic behind the proposed method is to compare areas before and after clean-up to detect and monitor changes. Control areas need to be set aside to account for environmental influences beyond those created by clean-up operations. These areas are selected to reflect locations similar to the impact site. One way to accomplish this goal is by choosing control areas near impact areas. The spatial locations of control and impact sample stations are randomly assigned within each of these areas. The stations are sampled before and after remedial action at pre-determined, nonrandom intervals, often monthly. Spatial and temporal variation is accounted for through random assignment of sample locations within control and impacted areas plus sampling over time.

SCHEMATIC OF SAMPLING DESIGN



Squares represent sample quadrats randomly assigned in clean-up and control plots. The plots are sampled at pre-determined times, often monthly.

Various additions can be made to this protocol to increase the effectiveness of monitoring. One addition might be to conduct exclusion studies using cages in both experimental (impacted/remediated) and control plots to separate the effects of biotic interactions, such as predation and competition, from abiotic effects, such as contaminants. Parameters such as growth, survival, and fecundity could be monitored for population growth within the caged/un-caged stations. Another addition could use traps or plates to study recruitment of invertebrate larvae at control/impacted plots to compare with Goose Cove. Many other possibilities can be included as add-ons to the basic monitoring design.

A different approach is needed to effectively monitor mobile vertebrates (i.e., fish) and invertebrates (i.e., snails, crabs); some of the latter are part of the benthos. Capture/recapture studies of tagged animals could give population estimates for some of these species. Since Goose Pond is not a closed system for highly mobile animals, tagging studies would also provide an estimate of retention within the Goose Pond complex.

4.0 SUMMARY AND CONCLUSIONS

EPA requested that ESAT (a) review site history, toxicity test results, and benthic community survey reports for the Callahan Mine Superfund Site; (b) participate in conference calls with the TOPO and the RPM to discuss first impressions and plan for the deliverable; and (c) participate in a site visit on June 16 and 17, 2009.

The review of the benthic community field study identified numerous limitations in sample design and data analysis which invalidated the conclusions of the study. The major limitations which precluded an accurate assessment of the benthic community were as follows:

- Mismatching the habitats between the Site locations and reference locations.
- Not adequately accounting for spatial and temporal variation.
- Failure to select the most effective sampler through preliminary testing.
- Focusing on organisms larger than 0.5 mm, which excluded numerous meiofauna that are known to be sensitive to mine wastes and to serve as food to higher trophic levels.
- Invalid statistical procedures, resulting in questionable conclusions based on improperly collected data.

Finally, a limited benthic community sampling effort during the June 16 and 17, 2009, site visit suggested that many expected macroinvertebrate taxa were missing from Goose Pond.

5.0 REFERENCES

Lee, M.R. & Correa, J.A. 2005. Effects of copper mine tailings disposal on littoral meiofaunal assemblages in the Atacama region of northern Chile. Marine Environmental Research 59: 1-18.

Lee, M.R. & Correa, J.A. 2006. An assessment of the impact of copper mine tailings disposal on meiofaunal assemblages using microcosm bioassays. Marine Environmental Research 64: 1-20.

Nalepa, T.F., M.A. Quigley & R .W. Ziegler. 1988. Sampling efficiency of the Ponar grab in two different benthic environments. J. Great Lakes Res. 14: 89-93.

Panis l, B. Goddeeris & R.F. Verheyen. 1995. On the reliability of Ponar grab samples for the quantitative study of benthic invertebrates in ponds. Hydrobiologia 312: 1 4 7-152.

ATTACHMENT 3

Date:	19 December 2008
To:	Bart Hoskins
From:	Marguerite (Peg) Pelletier
Cc:	Timothy Gleason, Wayne Munns
Re:	Callahan Mine Superfund Site

At your request I reviewed the draft Baseline Ecological Risk Assessment: Callahan Mine Superfund Site. I primarily focused on the marine/estuarine benthic invertebrate assessment and have several concerns. They focus on three issues: sampling methodology, site selection and completeness of the analyses.

The first question I have regards the invertebrate sampling. Why was a Ponar grab used instead of a larger grab (such as a Van Veen) which would have penetrated deeper and sampled more of the benthic community? Second, why were the invertebrates sorted in the field? In most marine studies, the sediment is sampled and perhaps sieved in the field, but the preserved samples are separated from any remaining sediment and identified and enumerated in the laboratory (see US EPA 2001, http://www.epa.gov /emap/nca/html/docs/c2kfm.pdf). By sorting organisms in the field it is likely that many small organisms would not be detected. It is also more likely that organisms would be damaged or destroyed.

It was also unclear how the site areas were chosen (i.e., Goose Cove, Goose Pond Permanently Flooded, and Goose Pond Irregularly Flooded). Were the different areas chosen based on clustering or ordination of biological communities? Or physical data (salinity, depth, etc)? Or was it based primarily on the morphology of the site? Although some physical data were available (Table 2-2), they were minimal, and did not seem to be extensive enough to support some of the site descriptions. For example, the Goose Pond estuary is characterized as having, "extreme swings in salinity and/or flow depending upon lunar phase, precipitation amounts, and the period of the daily tidal cycle." However, water quality was only measured a few times in the fall, with a minimal number of samples taken – with the data presented as one summary measure. Later in the report, the Goose Pond Irregularly Flooded was characterized as flooding completely during full moons and draining much more completely during neap cycles, implying the "extreme" fluctuations potentially occurring on a lunar cycle. Unfortunately, the reference sites seem to have been chosen to have "extreme" salinity fluctuations on the order of a tidal cycle rather than a lunar cycle typical of the Superfund area. In addition, dissimilar habitats seem to have been compared. For example, based on Table 2-1, Figure 2-5 and Appendix K, the salt marsh was included as part of the irregularly flooded area as well as being analyzed separately. The marsh is a much different environment than the subtidal and their benthic communities are not necessarily comparable.

It was unclear how the sampling areas (stations) were chosen. Were they randomly selected? And if not, what criteria were used to select them? Many of the samples appear to have been collected close to shore, which may not be representative of the estuary as a whole. Some reference areas were moved. In one case in Middle Horseshoe Cove, a reference site was moved because, "the overall species diversity in the Goose Pond Permanently Flooded area is more closely represented by R05 located in an intertidal mud flat." Since this study was designed to look for impacts, using diversity as one of the endpoints, matching diversity at a "reference" site to diversity found at an impacted site is problematic. It is especially problematic since the "reference" site is intertidal – yet allegedly represents a permanently flooded subtidal area.

Reference areas should be selected so that the environment is similar to the test area, but without the stressor of concern (in this case the Superfund site). Matching diversity is not appropriate, especially since an impacted site might be expected to have low diversity. If “reference” sites are chosen to match this low diversity, then no differences between the reference area and the site would be apparent or even expected. I don’t believe moving the initial reference site was appropriate. It is unclear whether the other reference sites were also chosen to match diversity found at the test site, which would potentially reduce the ability to detect impacts at the Superfund site.

In addition to my concerns about the selection of the individual sampling stations (detailed above), many of the areas selected as “reference” areas don’t appear to correspond well with the Superfund areas. Benthic communities can be structured by a variety of physical factors including hydrodynamics, salinity and grain size. Different communities are expected based on both salinity and grain size. However, these characteristics did not seem to be considered when selecting reference sites. For example, Goose Cove is sheltered, with silt and mud substrates. Lower Horseshoe Cove, its putative reference, is exposed, with sandy mud and gravel substrate. Similarly, Orcott Harbor is exposed with some sandy/gravel substrate. Goose Pond (permanently flooded) was compared with Middle Horseshoe Pond – an area which included the site that was moved to “match diversity.” There was not enough information to assess the other stations selected in this area, although they also appear to have been collected close to shore, so they may also be intertidal. The Goose Pond (irregularly flooded) area was described as having “wide swings in flow and salinity.” However, based on later description, this variation is seen primarily on a lunar cycle. At low neap tide, if there is a precipitation event, salinities can drop. This is a natural process to which estuarine benthic communities are adapted. Upper Horseshoe Cove, one of the proposed references is a tidal channel associated with a marsh. It dries out on a tidal cycle, with corresponding extreme salinity fluctuations. While estuarine invertebrates have adapted to varying salinity, extreme fluctuations (0 ppt to 25/28 ppt) on a tidal cycle (twice daily) is extremely stressful. In addition, these tidal channels accumulate rotting seaweed and vegetative material resulting in high oxygen demand and sulfides. This is not a reference site for the overall Goose Pond (irregularly flooded – although brackish might be a better descriptor). The second reference site, the Bagaduce Salt Pannes, is also not a reasonable reference area. Salt pannes are areas within a salt marsh where the vegetation has died off and rotted. These areas are characterized by high levels of organic matter and high amounts of toxic hydrogen sulfides. This is not an unimpacted reference area. Although this site was described in the text and used in analyses (Appendix K), it was not listed in Table 2-1, which listed the site and reference habitats. The final reference area was the Bagaduce River. The description of this site (widgeon grass with a large cattail marsh to the south) suggests that it may be fresher than Goose Pond, which had widgeon grass only growing near the freshwater earthen dam.

I also had some questions about the comprehensiveness of the data analyses. Although benthic invertebrate data were available for both 2005 and 2007, only data from 2007 were used in the analyses presented in Appendix K. This meant that only 1 station in each of the Goose Pond habitats, and 1 station in Goose Cove were used to represent the site. It is unclear whether these sites were representative of the site as a whole, and they were compared to sites which may not have been good choices as references. There were far more samples taken within the Superfund site which were not examined, except using diversity measures (Tables 3-36 to 3-38). From the text it appears that only infaunal data was presented. What did the epifaunal community indicate? How about using both infaunal and epifaunal organisms?

In conclusion, there are significant uncertainties associated with the benthic invertebrate survey as presented in this report. First, it is unclear whether the invertebrates were adequately sampled. By sorting in the field, it is very likely that only large and robust organisms were sampled. Smaller organisms were likely not sampled and other organisms were likely damaged. Second, it is unclear how stations were chosen to ensure that they were representative of the area. Many stations are located close to shore rather than being randomly distributed, and some habitat types were inappropriately combined (subtidal benthos and marsh benthos). Similarly, many of the reference sites do not appear to correspond well with the Superfund sites. I am particularly concerned that the reference sites were matched by benthic diversity rather than by physical factors. Finally, I am concerned that comparisons were made between one station/habitat type and multiple (perhaps inappropriate) reference stations. I believe that the estuarine benthic invertebrate survey results and the conclusions drawn from them should be viewed with caution.

Reference:

US EPA. 2001. National Coastal Assessment: Field Operations Manual. EPA 620/R-01/003. US Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL, 72 pp.