• *Benthic Macroinvertebrates*

Workshop held November 12–13, 1998, at the Florida DEP Central Laboratory in Tallahassee, FL.

Co-Chairs:

Steve Grabe	City of Tampa, Bay Study Group; Tampa, FL
Vicki Whiting	FDEP, Bureau of Laboratories; Tallahassee, FL

Core Group:

David Camp	FDEP, Florida Marine Research Institute; St. Petersburg, FL; retired
Steve Kent	FDEP, Central District; Orlando, FL
Michael Milligan	Center for Systematics and Taxonomy; Sarasota, FL
Harvey Rudolph *	Palm Beach DERM; West Palm Beach, FL
Doug Strom *	FDEP, Southeast District; Fort Pierce, FL
Ford Walton	FDEP, South District; Punta Gorda, FL

* unable to attend workshop

Attendees:

Dr. Dan Dauer #	Old Dominion University; Norfolk, VA
Dr. Virginia Engle #	U.S. EPA, Gulf Breeze ERL; Gulf Breeze, FL
Dr. Sara LeCroy	Univ. S. Miss, Gulf Coast Research Laboratory; Ocean Springs, MS
Dr. Chet Rakocinski #	Univ. S. Miss, Gulf Coast Research Laboratory; Ocean Springs, MS
Kelly Schratwieser	FDEP, Bureau of Laboratories; Tallahassee, FL
Steve Wolfe	FDEP, Bureau of Laboratories; Tallahassee, FL

gave workshop presentation

INTRODUCTION

The goal of the FDEP estuarine/marine bioassessment program is to develop methods to determine the relative condition of Florida's marine and estuarine ecosystems. The emphasis is on designating methodologies which are able to identify and accurately characterize disturbance along a gradients. This effort includes issues related to study design, gear type, etc.

Three scientists with considerable experience in designing, conducting, and assessing benthic macroinvertebrate investigations encompassing a regional geographic scale were invited to give presentations at an FDEP invertebrate workshop held in November, 1998. These were: Dr. Virginia Engle (U.S. EPA), who has been instrumental in conducting the EMAP benthic programs in Florida; Dr. Dan Dauer (Old Dominion University), who has been conducting investigations of the benthic habitats of Chesapeake Bay, as well as Tampa Bay; and Dr. Chet Rakocinski (Gulf Coast Research Laboratory),who has been involved in northern Gulf of Mexico benthic studies, including open coastal habitats, and has extensive experience in multivariate analyses of benthic communities.

EPA EMAP BENTHIC PROGRAMS

Dr. Virginia Engle stated that EPA's goal, with respect to EMAP, was to develop indicators of ecosystem status which could be usable by both scientists and the public. The premise was that the source of an impact (e.g. stormwater) may stress a given waterbody (e.g., decrease DO). This stressor is then manifested as an ecological effect (e.g., a decline in diversity). In addition, the indicators of ecological condition should be applicable over a large scale.

The indicators should not be considered a substitute for a compendium of ecological information. Rather, the indicators should be treated as "screening tool" useful in the preliminary identification of problem areas, and should quantify the extent of the problem, as well as identify biotic responses associated with the specific environmental exposures.

The protocol embraced the random, probability-based sampling strategy initiated and popularized by the EMAP program—although there were also specifically selected test sites. With respect to the benthos, triplicate samples were collected and samples were cored for sediment characterization (i.e., % silt+clay and RPD).

Criteria for "degraded" and "undegraded" (reference) sites necessary for the construction of a benthic index were:

Degraded sites had near-bottom dissolved-oxygen concentrations <2 mg/L **or** a sediment contaminant(s) concentration ≥ERL ("Effects Range–Low" of Long and Morgan 1990) **or** toxicity >20%. Reference sites were required to have: DO>3 mg/l **and** no contaminants >ERL **and** toxicity <15%. Within the Louisianian Province, all sites with a sand substrate (<20% silt+clay) were not degraded. In order to develop an index capable of discriminating degraded from undegraded sites, a suite of candidate measures were assessed. These included S (number of species), H' (Shannon-Weaver diversity), # of polychaete species, total abundance, and the mean abundance and percent composition of amphipods, bivalves, capitellid polychaetes, decapods, gastropods, mollusks, deposit feeders, omnivores, carnivores, and suspension feeders.

Each candidate measure was evaluated to determine whether there were associations with salinity, longitude, %silt+clay and sediment TOC. S and H'() were found to be correlated with salinity. To rectify this possibly confounding association, two solutions were explored: different indices could be constructed for the different salinity regimes; or, an adjustment could be made for salinity. In practice, the decision was made to use the 90th percentile values of H'() vs. salinity, calculate a polynomial equation describing the relationship and then adjust H' for inclusion in the index.

Discriminant function analysis ultimately revealed that adjusted H', tubificid abundance, % capitellids, % bivalves, and % amphipods were best able to discriminate between degraded and non-degraded sites. The criterion for acceptable classification error was <10%.

To validate the index, 13 sites were visited during two sample periods; sites were generally classified similarly in both periods.

Cumulative distribution function plots were used to estimate the percent of the province considered to be degraded and non-degraded.

Index scores were tested for association with specific stressors, 64% of the variability could be explained by contaminants alone and 17% of the variability was associated with contaminants and toxicity.

Dr. Engle also summarized the results of studies in specific estuaries in the northern Gulf of Mexico. Without going into the specifics, these studies employed step-wise logistic regression as a tool for extracting the best exploratory variables which may be related to the degradation of the benthos. In each of the systems studies, such variables were, in fact, identified.

The next step is for EMAP to work on developing an overall index of estuarine “health” by combining the benthic index with other measures (e.g., fish, submerged aquatic vegetation-SAV).

With respect to the FDEP’s needs, one benthic index may not prove realistic for state-wide application. There may need to be separate indices for estuaries and coastal areas, at a minimum.

MACROBENTHIC COMMUNITIES AND ESTUARINE MONITORING: THE CHESAPEAKE BAY EXPERIENCE

Dr. Dan Dauer summarized the approach used in Chesapeake Bay to set restoration goals for benthic habitat and described how an index was developed for this purpose.

The first challenge was to compile all available benthic data from the bay. They were fortunate in that a large number of data sets were available—although not all contained matching sediment chemistry, DO, sediment composition data, etc. Data sets were judged suitable if they were from 0.5-mm mesh sieve collections, identification was to the lowest practicable taxon, and data were available via electronic media. The major challenge was to make the data sets comparable—especially with regards to nomenclature. Biomass data were also included and the observation was made that biomass analysis was relatively inexpensive if done at the species level as identifications and counts are being done. Only data collected by Young samplers during summer surveys were included.

Once the suitable data sets were identified, numerical classification (cluster analysis) was used to identify strata by salinity and sediment composition type. The strata identified by this procedure included 6 salinity, 5 sediments (%silt+clay), and 3 depths. Of the 90 possible strata, only 8 habitat classes were ultimately considered for further analysis (some habitats, while "possible" may not have been represented in the data or were "rare").

Once stratification was performed, degraded and reference sites were identified. Degraded sites were defined as DO <2 mg/L 80% of the time (≥5 site visits) **or** any measured chemical contaminant exceeding the ERM ("Effects Range-Median" of Long & Morgan 1990) [n.b. contrast with ERL criteria in Engle's presentation above]. Once "reference" (non-degraded) sites were identified, criteria were established. Of the >4,000 samples in the original compilation of databases, reference conditions for Chesapeake Bay ultimately were defined by only 99 samples. One problem is the confidence in tracking changes in "marginal" locations; i.e., is "health" increasing or decreasing? Marginal sites tend to have more "noise" or variability in the measured variables.

With the identification of reference sites, 24 measures of benthic habitat status were evaluated. These included diversity parameters (S,H''), abundance and biomass (overall and for various taxonomic groups), life history strategies (pollution-sensitive vs. pollution-tolerant), measures of activity beneath sediment surface, and feeding guild measures. Ten attributes were eventually selected. Without listing them individually (see Dr. Dauer's handout), it should be noted that feeding guild measures generally did not provide useful information and measures of subsurface benthic activity are logistically difficult to quantify. Note also that mean S is correlated with the surface area of the sampling gear so either one gear type should be considered to be "standard" or mean S must be expressed per unit area.

When the various measures were evaluated by habitat type, different habitats had a different number of metrics suitable for inclusion in an index. Generally, H', total abundance, biomass, and relative abundance of "pollution indicative taxa" were suitable metrics for each habitat type. The minimum number of metrics suitable to characterize all habitats needs to be ascertained—if possible.

For some metrics, there is a distribution of measures indicative of degraded habitat and for measures indicative of the reference condition. For most metrics, the response to increased stress is monotonic—either high or low values represent degradation and vice versa for reference sites. Finally, other metrics may show a pattern where extreme values (high and low) are indicative of a degraded condition. A scoring system is applied to each of these types of distributions and their associated metrics. These types of distributions are depicted in the handouts along with the scoring system.

Scoring a site is accomplished by adding or, better, averaging scores across all metrics suitable for that habitat stratum. Scores <3 represent "degraded" habitat. Between 93% and 100% of all reference sites were properly classified with this approach and between 86% and 100% of degraded sites were correctly classified.

It was emphasized that the "threshold" values may require periodic recalibration.

INDICATOR SPECIES/ASSEMBLAGES

Dr. Chet Rakocinski provided an overview of the benefits, as well as the problems, associated with the use of benthic communities as indicators of environmental stress. In particular, it is often difficult distinguishing natural variability from variability caused by environmental stressors.

In studies conducted in the northern Gulf of Mexico, a large suite of environmental variables was reduced by application of Principal Components Analysis to five new variables (i.e., components). These included metals (PC1), organics (PC2), the physicochemical variables, DO, temperature, salinity (PC3), sample depth, pesticides (PC4), RPD, and TOC (PC5).

The trophic structure of the benthos was analyzed by multiple regression to determine the association between various feeding types and the five PC's. Associations were found, for example, between percent of subsurface deposit feeders and the first three principal components. Trophic diversity decreased as PC's 1–3 increased.

There was also an interest in determining whether taxa-specific responses could be gleaned from the data as a means of identifying "indicator" species and "indicator" assemblages. The multivariate technique used for this purpose included Canonical Correspondence Analysis, although TWINSPAN (two-way indicator species analysis) could also be used for this purpose. Examples of some of the results from such analyses were that the polychaetes *Mediomastus californiensis* and *Cossura delta* were correlated with high concentrations of metals. *Corbicula* and *Limnodrilus* were associated with high contaminant concentrations in general. *Ampelisca holmesi* was associated with low concentrations of both metals and organics whereas another amphipod, *Corophium cf. Baconi,* was associated with medium-to-high concentrations of metals and organics. Trophic structure can be evaluted in the same manner, low trophic diversity was associated with higher contaminant loadings. Additionally, trophic diversity represented almost the mirror image of the EMAP Benthic Index for the Louisianian Province.

Dr. Rakocinski also presented a modification of the Species Abundance-Biomass curves vs. organic inputs as presented by Pearson & Rosenberg (1978). Curves were developed for metal contamination and organic contamination gradients. Similarly, curves were also developed for salinity, depth ,and % silt+clay.

DISCUSSION

Steve Wolfe reminded the group that one of the major factors driving this program was the need to be able to assess the effects of dischargers on the biological condition of the receiving water. Thus, to the extent possible, methods selected should be useful in discriminating effects due to both point-source and non-point-source discharges.

There was discussion of the appropriate sampling methods for various habitat types (see

Appendix)

There was also discussion about appropriate sieve mesh sizes (0.5 mm vs. 1.0 mm). Steve Grabe will provide comparative data from a subset of samples from ongoing sampling in Tampa Bay. The issue of how many "replicates" are required for various habitats and gear types will require field research). There was a recommendation to compile Florida datasets and identify those which would be useful for analysis of replicate variability as well as generating species-saturation curves to attempt to answer this question.

The question whether biomass was a desirable measure was addressed. Dr. Dauer noted that in Tampa Bay he had observed that abundance and biomass were correlated but that, because of the presence of large burrowing species, that relationship was not observed in Chesapeake Bay. This, in turn, led to questions about penetration depth of various grab samplers since some large bodied forms are found at depths >35 cm—below the penetration depth of typical grab samplers.

Steve Wolfe raised the question of whether it would be difficult to identify reference sites for the various habitat types present in Florida. The presumption appears to be that if habitats are stratified by salinity and sediment types (as % silt+clay), then the faunas should be similar. Dr. Dauer noted that it may, in fact, be difficult to find true "reference" conditions in estuaries *vis a vis* a point source discharger (gradients in salinity, flow, may confound).

There seemed to be consensus that the level of identification should be the "lowest practicable taxon" and that with adequate QA/QC and, perhaps, additional training (taxonomic workshops) this should be sufficient.

Both Steve Wolfe and Dr. Dauer made reference to the need to recalibrate thresholds for the various metrics over time and also to establish new reference sites as a system improves.

	Selected Benthic Macronvertebrate Sampling Gear:				
	Habitat	Qualitative	Quantitative	needs testing	proven
I. Grassbeds					
epifaunal	thrownet	X		X	
	trawl	X			X
	dipnet	X			X
infaunal	core		X		X
	ponar		X		X
II. Oyster beds					
epifaunal	nets	X		X	
	quad suction	X	X	X	
infaunal	core	X			X
III. Sandbottom					
epifaunal	trawl	X			X
	Yabby pump			X	
infaunal	core		X		X
	ponar		X		X
	Young grab		X		X
	Yabby pump			X	
IV. Coral Reef					
epifaunal	photos	X	X		X
	baited traps	X		X	
	suction dome	X		X	
infaunal	core		X		X
V. Mud/Muck/Silt					
A. Emergant veg.					
epifaunal	thrownet	X		X	
	trawl	X			X
	dipnet	X			X
infaunal	core		X		X
	ponar		X		X
B. Non-veg.					
infaunal	ekman		X		X
	Young		X		X
VI. Mangrove					
epifaunal	dipnets	X			X
	photo grids	X	X	X	
infaunal	cores		X	X	
VII. Hardbottom substrate					
epifaunal	photos-trans		X	X	
	picking bugs	X	X	X	
infaunal	diver core		X		X
	box dredge	X	X		X