Development of Sediment Quality Objectives for California Bays and Estuaries

Project Update-December 2005

Steven Bay

Southern California Coastal Water Research Project (SCCWRP)

steveb@sccwrp.org

TOPICS FOR TODAY

- Key program elements
- Results presented at SSC meeting
- SSC comments
- Recent activities
- Schedule

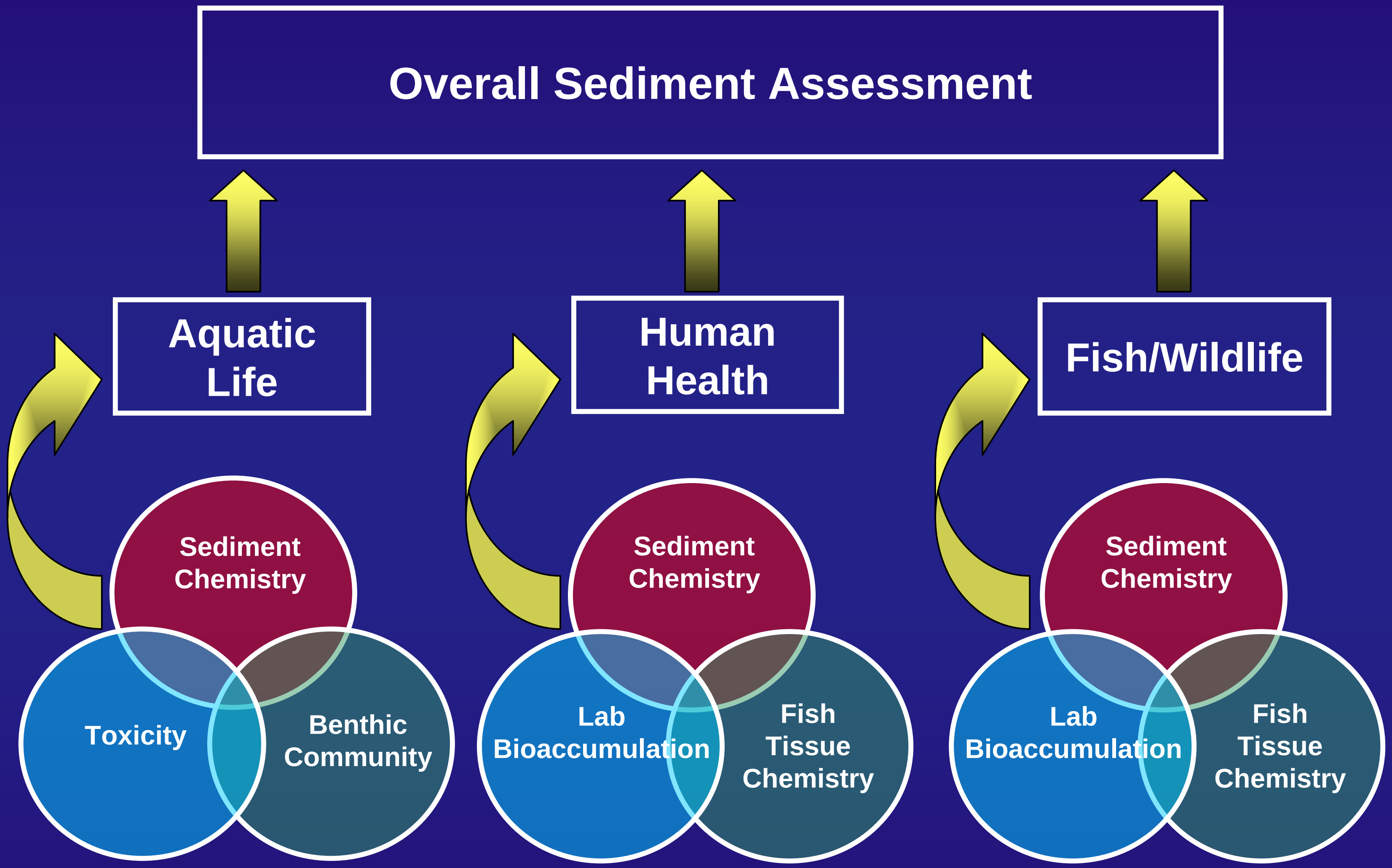
BACKGROUND

- **For many years, scientists have advocated a triad approach for evaluating sediment quality**
  - Individual lines of evidence; each have potential limitations
- **The triad has been widely used in site-specific assessments, but has not found its way into most statutory frameworks**
  - Most applications are based on best professional judgment
- **The State of California is developing sediment quality objectives based on multiple lines of evidence (MLOE)**
  - There are many challenges in translating scientific concept into regulatory framework

The Basic Framework

- **Three beneficial uses to be protected**
  - Aquatic life
  - Human health
  - Fish and wildlife
  - Each will be assessed separately
- **Within each beneficial use, a multiple line of evidence (MLOE) approach will be used**
  - MLOE involves demonstration of both exposure and effect
  - No single line of evidence is sufficient
- **More complex than water column criteria because chemical bioavailability in sediments is poorly understood**

SQO ASSESSMENT FRAMEWORK



1Dec05

CHALLENGES

- **Developing methods/assessment consistency across the state**
  - Multiple ecoregions
  - Numerous habitats
    - Initial focus on marine embayments
- **Standardizing data interpretation among individuals with varying expertise**
  - Engineers vs. Biologists
- **Developing assessment thresholds**
  - Fine line between science and policy

SCIENTIFIC ACTIVITIES

- **Select indicators for individual lines of evidence**
  - Evaluate multiple candidate indicators for each LOE
  - Base recommendations on performance, conceptual basis, and practicality
- **Establish thresholds for each indicator**
  - Quantitative
  - Understand linkage to presence and severity of effects
- **Develop a framework for integrating across lines of evidence**
  - Clear decision points
  - Utility for prioritization
  - Simple, yet retain scientific content

July SSC MEETING

- Mid-course progress report
- Are our plans and results to date likely to result in clear recommendations regarding indicators and thresholds when the work is completed?
- Are our preliminary interpretations and recommendations appropriate?
- Are critical elements missing?

CHEMISTRY INDICATORS

- **There are numerous candidate approaches and indicators for interpreting sediment chemistry data**
  - Biggest dichotomy is empirical approach vs. equilibrium partitioning
  - Individual empirical chemical guidelines
    - ERM, PEL, AET
  - Mixture approaches
    - Mean quotient
    - Logistic regression models

- **Our approach is to develop a California-specific data base for evaluating multiple possible approaches**
  - Includes data from more than 150 studies
  - Evaluate performance of candidates to predict toxicity and benthic community impacts

CANDIDATE CHEMISTRY INDICATORS

- **Existing national Sediment Quality Guidelines**
  - Effects range median quotient (ERMq)
  - Consensus midrange effects concentration (Consensus MEC)
  - Sediment quality guidelines quotient (SQGQ1)
  - Logistic regression (Pmax)
  - Chronic equilibrium partitioning (EqP)
  - Acute equilibrium partitioning (EqP)
- **National SQGs recalibrated to California data**
  - ERMq
  - Pmax
- **New approaches**
  - Mean weighted kappa

Kappa Statistic

- **Developed in 1960**
  - Extensive peer-reviewed literature describes derivation and interpretation
- **Used in medicine, epidemiology, & psychology to evaluate observer agreement/reliability**
  - Similar problem to SQG development and assessment
  - Accommodates multiple categories of classification
  - Multiple thresholds can be adjusted by user
  - Categorical or ordinal data
  - Result reflects magnitude of disagreement (can be used to weight values)
- **Sediment quality assessment is a new application**

Kappa

Evaluates agreement between 2 methods of classification

- Magnitude of error affects kappa
- Optimization routine is used to identify thresholds that can be used for indicator development

Toxicity Result	SQG Result or Chemical Concentration			
	High	Moderate	Marginal	Reference
High				
Moderate				
Low				
Reference				

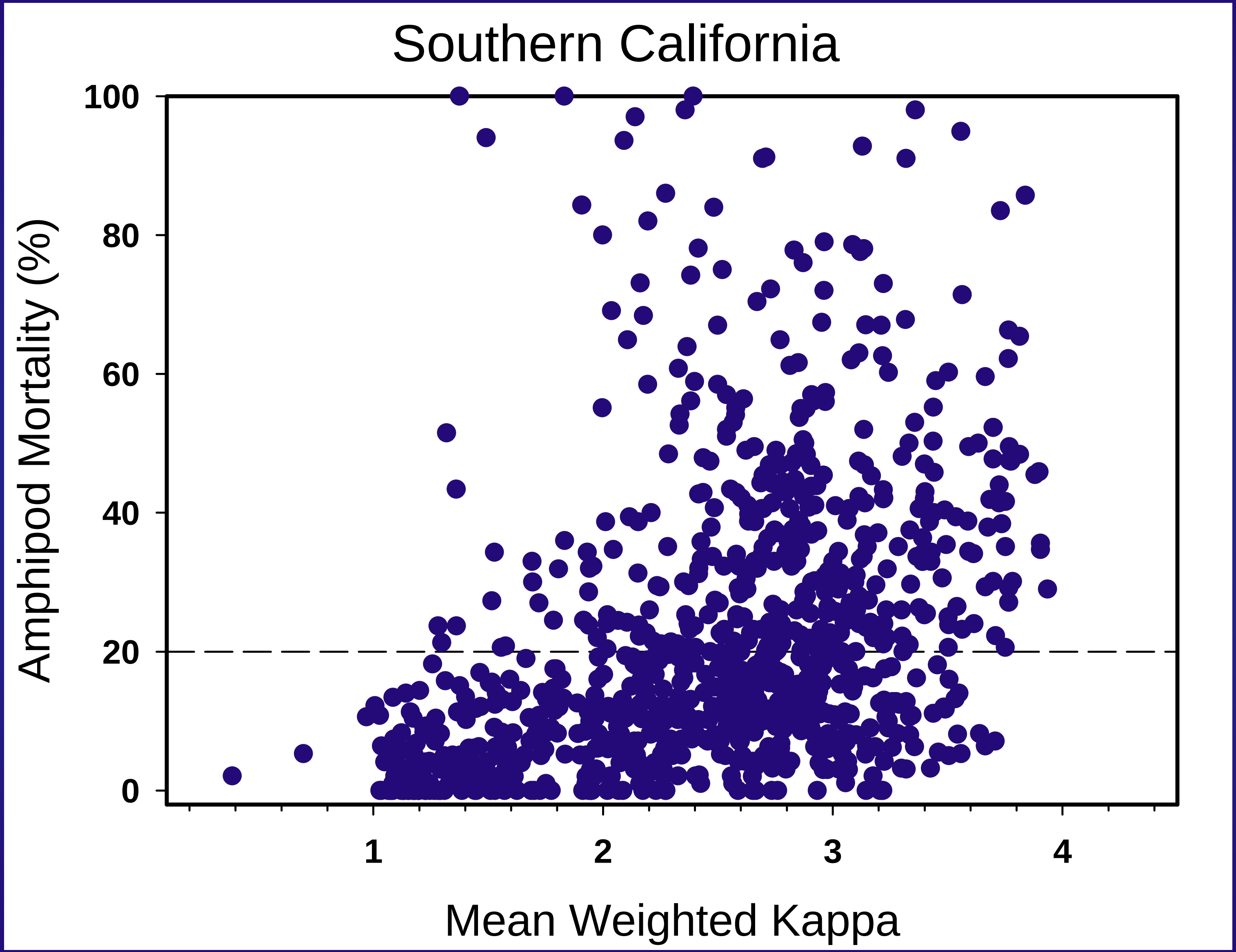
T3 T2 T1

Kappa SQGs

- **Derived Kappa and thresholds for target chemicals using amphipod mortality data**
  - Kappa (k) : indicates strength of relationship
  - Thresholds: used to predict sample classification (cat)
- **Calculated Kappa score for each chemical in sample**
  - *k* x *cat*
- **Mean weighted Kappa score**
  - Average of *k* x *cat*
  - Each constituent contributes to final classification in a manner proportional to reliability of relationship
  - Mixture joint effects model

CORRELATION WITH CA SEDIMENT TOXICITY (preliminary)

SQG	NORTH	SOUTH
Mean Weighted Kappa	0.54	0.46
CA ERMq	0.37	0.28
ERMq	0.44	0.30
Consensus MEC	0.35	0.37
SQGQ1	0.37	0.32
CA Pmax	0.35	0.32
National Pmax	0.27	0.22
Chronic EqP	-0.08	-0.06
Acute EqP	-0.09	-0.08


Southern California
Amphipod Mortality (%)
100
80
60
40
20
0
1
2
3
4
Mean Weighted Kappa

SSC COMMENTS ON CHEMISTRY LOE

- **Expand list of target chemicals for new indicator development**
  - PAH classes or types
  - DDT metabolites and other pesticides
- **Additional development/validation of kappa statistic**
  - Interpretation for sediment quality assessment
  - Sample size effects
- **Evaluate and develop chemical indicators for impacts on benthic communities**
  - Use results of benthic indicator development activities
- **Development of indicators for northern and southern portions of state is supported**
  - Provided accuracy and reliability is improved

NEXT STEPS FOR CHEMISTRY LOE

- **Refine chemistry:toxicity analyses**
  - Investigate kappa procedure
  - Revise selection of chemicals
  - Complete calibration and alternative approach development
- **Evaluate candidate indicators against benthic response**
  - Do the candidate chemistry indicators rank the same?
  - Does the correlation improve?
- **Select the best chemical indicator**
  - Possibly more than one
- **Determine thresholds for levels of effect**
  - Reference condition
  - Marginal deviation from reference
  - Moderate potential for effect
  - Severe potential for effect

TOXICITY INDICATORS

- **There are many types of toxicity tests**
  - Various species
  - Variable exposure methods and endpoints
- **Three types of tests were evaluated**
  - Survival of amphipods (four species)
  - Short-term fertilization and embryo development
    - Sea urchins and mussels
  - Long-term sublethal response
    - Amphipods, copepods, worms, clams, oysters
- **Conducted comparison tests and reviewed existing data**
  - Consistent with program objectives
  - Established methods and technically feasible
  - Likely to provide useful information

SHORT-TERM SURVIVAL

- **Recommended**
  - *Eohaustorius estuarius*
  - *Leptocheirus plumulosus*

- **Not recommended**
  - *Rhepoxynius abronius*
    - Limited availability
    - Grain size sensitivity
  - *Ampelisca abdita*
    - Low sensitivity
    - Low test success rate

SUBLETHAL TESTS

- **Recommended**
  - Polychaete growth test (N. arenaceodentata)
    - 28-day exposure to whole sediment
  - Embryo development test using mussels
    - 2 or 3-day exposure at sediment-water interface

- **Other methods not recommended mostly based on feasibility**
  - Potential confounding factors
  - Organism supply issues
  - No standard method
  - Low test success rate
  - Lack of capacity in California

SSC COMMENTS ON TOXICITY LOE

- **Use of recommended acute and sublethal tests supported**
  - Encourage development of additional sublethal test methods in future
- **Toxicity classification thresholds should be test-specific and incorporate statistical comparisons**
  - Use minimum significant difference and results of statistical comparison to control
- **Integration strategy for multiple tests should not discourage use of more than minimum number of tests**
  - Classification based on relative proportion of “hits” not just presence
  - Equal weighting for acute and sublethal tests

NEXT STEPS FOR TOXICITY LOE

- **Develop thresholds for these tests**
  - Need to develop comparability of scoring across tests

- **Develop method for integrating multiple tests into a LOE score**
  - Scientific Steering Committee recommended use of both acute and a sublethal tests

PROPOSED TOXICITY CLASSIFICATION

- **Nontoxic**
  - Response not different from that observed in uncontaminated control sediments
    - Response within range of controls
- **Slight toxicity**
  - Relatively low magnitude response that may not be greater than test variability
    - Response less than 90^{th} percentile minimum significant difference
- **Moderate toxicity**
  - High confidence that a statistically significant response is present
    - Response greater than 90^{th} percentile minimum significant difference
- **High toxicity**
  - The strongest effects observed for the test, indicative of substantially greater exposure to toxicants
    - Probably a combination of very high statistical confidence, response to increased dose, and subset of highest observed responses

BENTHIC ASSESSMENT CHALLENGES

- **Interpreting species abundances is difficult**
  - Samples may have tens of species and hundreds of organisms
  - Indices provide a means of summarizing complex information
- **Benthic species and abundances vary naturally with habitat**
  - Reference condition needs to vary by habitat
- **Sampling methods vary among programs**
  - Gear type sampling area and sieve size affect species and individuals captured

BENTHIC INDICATOR DEVELOPMENT ACTIVITIES

- **Determine the number of biogeographic provinces in California**
  - Index calibration/validation to be conducted separately for each
  - Six habitats; defined by salinity, grain size, latitude
    - Euhaline bays
    - Shallow estuaries and wetlands
    - Very coarse sediments
    - Polyhaline San Francisco Bay
    - Mesohaline San Francisco Bay
    - Tidal freshwater

BENTHIC INDICATOR DEVELOPMENT ACTIVITIES

- **Select from among several possible index approaches**

Index	Approach
Index of Biotic Integrity (IBI)	Community measures
Relative Benthic Index (RBI)	Community measures
Benthic Quality Index (BQI)	Community measures
Benthic Response Index (BRI)	Species types
River Invertebrate Prediction and Classification System (RIVPACS)	Presence/absence of expected species

BENTHIC INDICATOR DEVELOPMENT APPROACH

- **Give each index developer a development data set**
  - Refine or develop index for each habitat

- **Withhold data for independent index evaluation**
  - Classification of “known” good and bad sites
    - “Gold Standard” based on independent assessment by experts
  - Repeatability across replicates
    - Same day, months, years
  - Independence from natural habitat gradients
    - Grain size, TOC

INITIAL CLASSIFICATION ACCURACY

Index	Overall (n=35)
RIVPACS	83
BRI	77
IBI	70
BQI	63
RBI	51

Low accuracy may be due to incorrect identification of “good” and “bad” sites based on chemistry and toxicity

SSC COMMENTS ON BENTHOS LOE

- **Index evaluation should be based on agreement with best professional judgment by benthic ecologists**
  - Correspondence with chemistry and toxicity is problematic
- **Use more than one benthic index for station evaluation**
  - Agreement among independent indices will improve reliability of classification
- **Classification thresholds should be based on degree of biological impairment**
  - Not correlations with chemistry or toxicity
- **Development of benthic indices for other habitats within the state is not feasible at this time**
  - Collection of additional data and continued index development activities are recommended

Next Steps for Benthic LOE

- **Redefining a validation data set based on expert opinion**
  - Have recently given 36 new sites to the experts for their assessment
- **Continue with repeatability and gradient evaluation**
- **Select recommended indices and thresholds**
  - May be more than one
- **Develop data integration strategy for multiple indices**

REVISED CLASSIFICATION ACCURACY (preliminary)
Based on consensus rankings by nine benthic ecologists

INDEX	Southern Euhaline	Northern Polyhaline
BRI	0.82	0.95
BQI	0.81	0.92
IBI	0.58	0.85
RBI	0.73	0.90
RIVPACS	0.60	0.88

Next Steps for Benthic LOE

- **Redefining a validation data set based on expert opinion**
  - Have recently given 36 new sites to the experts for their assessment
- **Continue with repeatability and gradient evaluation**
- **Select recommended indices and thresholds**
  - May be more than one
- **Develop data integration strategy for multiple indices**

MLOE ASSESSMENT FRAMEWORK

- Approach to validating the MLOE framework
- A new issue associated with missing data
- Expansion of the MLOE framework to indirect effects

VALIDATION

- **Repeatability**
  - Are replicate samples at a site classified the same?
- **Consistency in distribution among individual LOE**
  - Is there general agreement among each LOE
- **SQO application in water bodies of “known” condition**
  - Does the MLOE framework perform as expected
  - Adjustment of integration strategy or thresholds may be needed

KNOWN CONDITION

- **Use Bay Protection and Cleanup Program assessment to identify compromised sites**
  - That program had targeted data collection to identify hot spots

- **Use chemistry and toxicity to identify cleanest sites**
  - Use data from adjacent sites where only part of the triad is available
  - Screen these waterbodies further based on land use

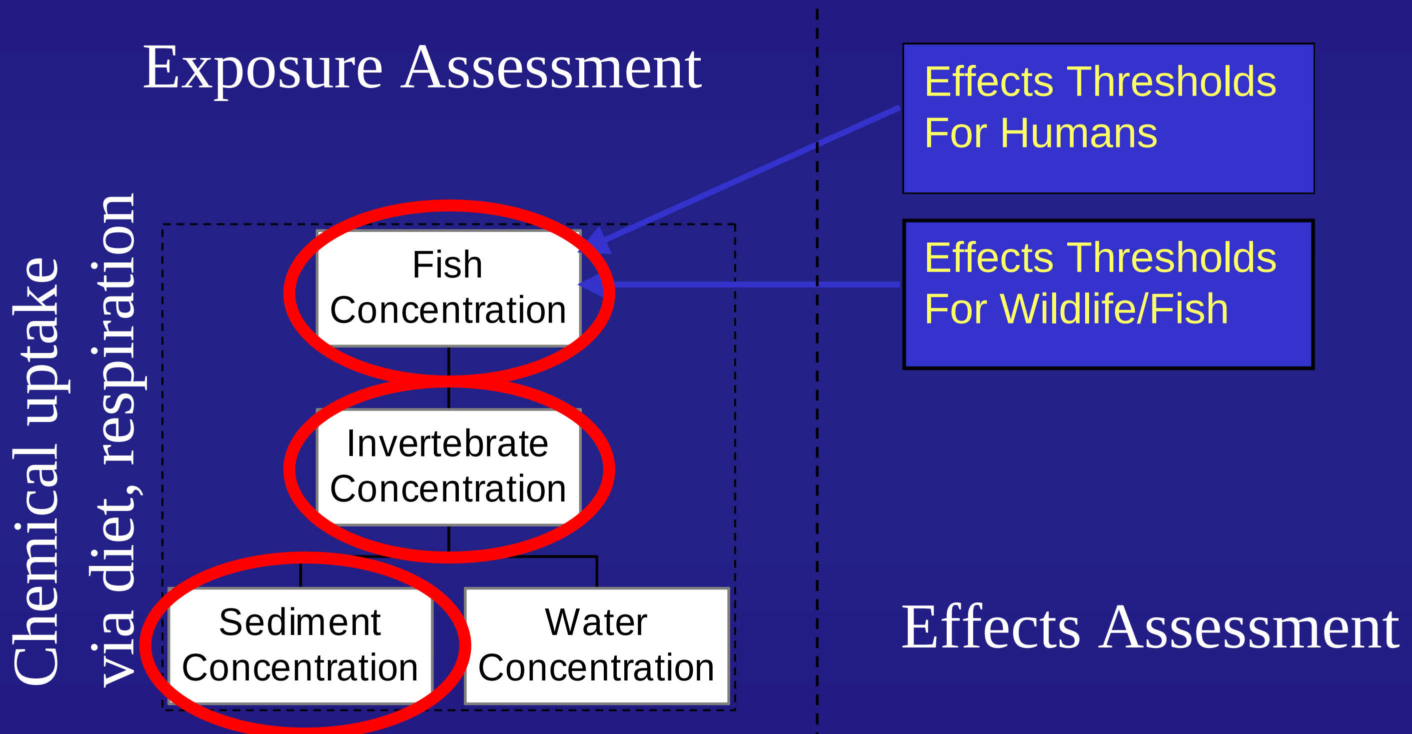
MISSING DATA

- **How should the MLOE framework be applied in habitats where we have not yet developed interpretational tools?**
  - Differs from situation where data were not previously collected, but could be collected

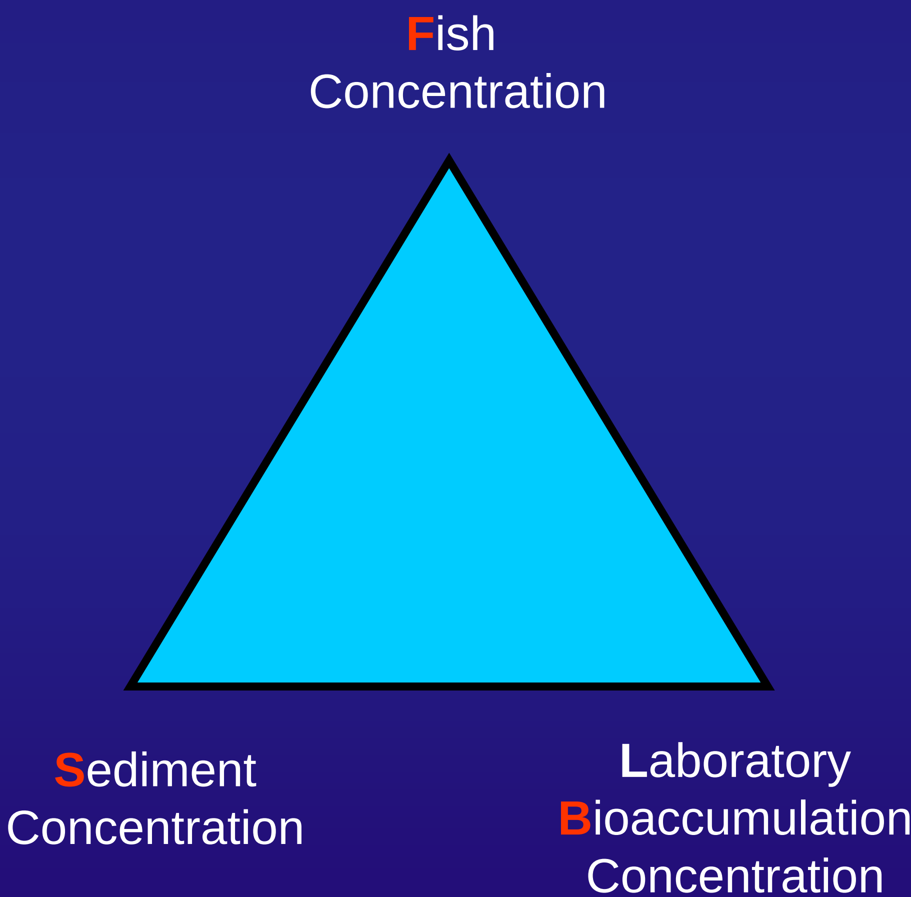
- **When tools are available for two LOE, apply the modified assessment framework that was developed for the unavailable data situation**

- **When tools are for only available for one LOE, use that LOE in a screening mode**
  - Apply BPJ to other LOE when screening indicates the necessity
  - BPJ may require additional data collection to establish site-specific reference condition

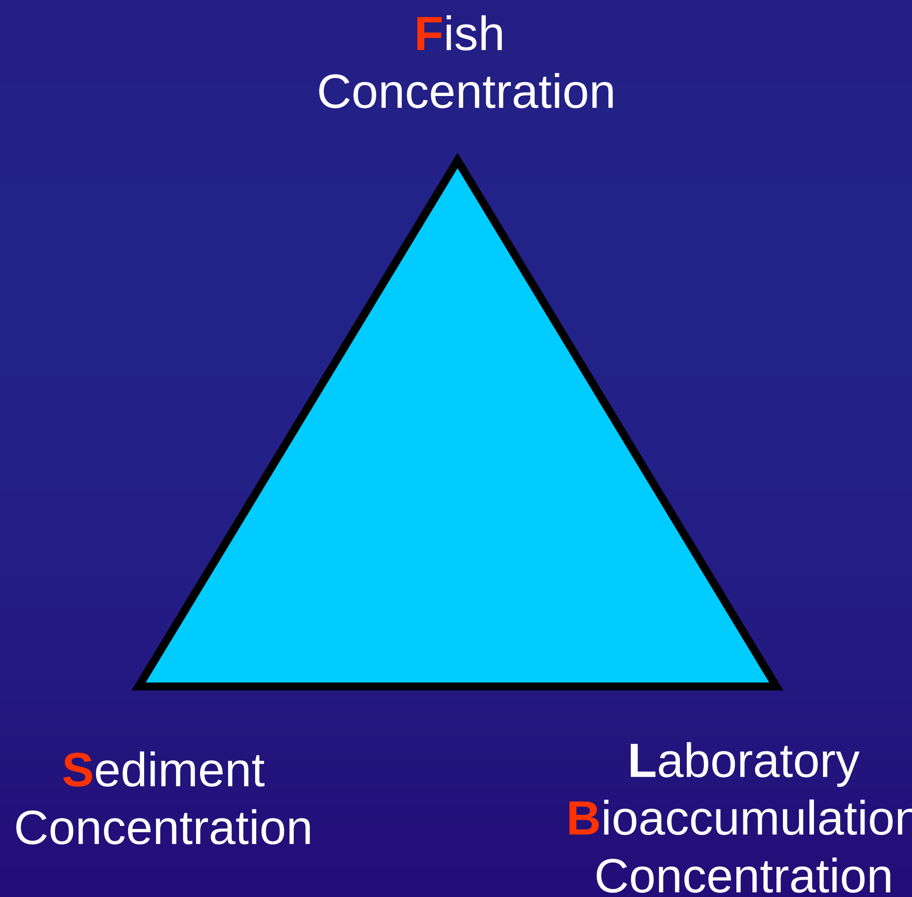
INDIRECT EFFECTS MLOE APPROACH



Human Lines of Evidence



Fish and Wildlife Lines of Evidence



SSC COMMENTS ON OVERALL MLOE FRAMEWORK

- **MLOE approach is suitable for all habitats within policy boundaries**
  - Greater clarity is needed regarding boundaries
- **Assessement strategy should give equal weight to each LOE**
  - All three LOE should be measured for complete assessment
- **Use of a single LOE is discouraged in habitats not having fully developed tools**
  - Use of 2 LOE is appropriate
  - Should develop needed tools for other habitats
  - Interim assessment guidance based on partial tools is discouraged

SSC COMMENTS ON INDIRECT EFFECTS MLOE FRAMEWORK

- **Greater detail needed regarding assessment framework**
  - Conceptual basis
  - Application method
  - Scale and sequence of evaluation
- **Assessement should be based on individual compounds, not mixture effects**
- **All three proposed LOE should be evaluated**
  - Use both empirical data and mechanistic models to develop sediment chemistry thresholds
- **Fish tissue LOE thresholds should be established by SWRCB and used by science team in developing other LOE**

SCHEDULE

- **Indicator selection**
  - December 2005
- **MLOE validation**
  - Jan-Feb 2006
- **Next SSC meeting**
  - Feb. 28 – March 2, 2006
- **Indicator development reports**
  - March-June 2006