Chapter 5

Derivation of Numerical Sediment Quality Assessment Guidelines for Florida Coastal Waters Using the Weight of Evidence Approach

5.0 Introduction

The National Oceanic and Atmospheric Administration (NOAA) National Status and Trends Program (NSTP) weight of evidence approach (WEA; Long and Morgan 1990) is chosen as a basis for developing sediment quality assessment guidelines (SQAGs) for Florida coastal waters. This approach relies on the collection, evaluation, collation and analysis of data from a wide variety of sources in North America to establish relationships between concentrations of sediment-associated contaminants and their potential for adverse biological effects. A modified version of the WEA is recommended for deriving numerical SQAGs.

5.1 Modification of the Original Weight of Evidence Approach for Deriving Florida Sediment Quality Assessment Guidelines

A modified version of the WEA is recommended to derive preliminary SQAGs due to its practicality for developing guidelines, its limited requirement for additional resources, its overall scientific defensibility, and its applicability to all aspects of sediment quality assessment. This approach is closely related to the original WEA, however, the modifications increase the relevance of the resultant guidelines to Florida coastal sediments. Specifically, the modifications to this approach increase the level of internal consistency in the database (by establishing additional screening criteria), verify and expand the information contained in the original NSTP database, and use all of the information in the database to derive SQAGs (in contrast, only data which had concordance between sediment chemistry and biological effects were used to derive the original NSTP guidelines). In addition, user access to the information from individual studies has been improved by providing expanded data tables (see Volumes 3 and 4).

5.1.1 Procedures and Criteria for Screening Candidate Data Sets

The modified WEA integrates a diverse assortment of data to derive numerical SQAGs. As such, data from spiked-sediment bioassays, sediment toxicity bioassays, and assessments of benthic invertebrate community characteristics were merged, along with the sediment quality assessment values developed in other jurisdictions (e.g., Puget Sound AETs, SQCs derived using the EqPA, etc.) into a single database. These data were fully evaluated prior to inclusion to assure internal consistency in the database.

The screening procedures used to support the development of this database were designed to ensure that only high quality data was used to derive SQAGs for Florida. The screening criteria used to evaluate spiked-sediment bioassay data and other matching sediment chemistry and biological effects data (i.e., co-occurrence data) are described in Appendix 1. These screening criteria were established to evaluate the acceptability of the experimental designs, test protocols, analytical methods, and statistical procedures used in each study. To ensure internal consistency in the database, only those studies that met these screening criteria were considered appropriate for inclusion. The sediment quality assessment values that have been derived by other jurisdictions were either incorporated directly into the database (if the concentrations of contaminants were originally expressed on a dry weight basis) or converted to concentrations expressed on a dry weight basis at 1% total organic carbon (TOC; if the assessment values were originally expressed on a TOC basis). Conversion of contaminant levels to dry weight concentrations at 1% TOC rendered these data consistent with the other information included in the database (which had an average TOC of 1.2 +/- 1.8%).

5.1.2 Expansion of the National Status and Trends Program Database

One of the principal limitations of the original NSTP database on the biological effects of sediment-associated contaminants, with respect to deriving SQAGs for Florida, is its bias toward data derived from studies in the northeastern and western coastal areas of the United States. At the time the original database was assembled, few data were included on the biological effects of sediment-associated contaminants from sites located in the southeastern United States. Therefore, collection of acceptable data from Florida and other areas in the southeast was a priority in the present study.

To address the need for additional information on the biological effects of sediment-associated contaminants in general, and from sites in the southeastern United States in particular, a major initiative was undertaken to expand the original NSTP database. The first stage of the database expansion involved identifying and retrieving candidate data sets from sites in the southeastern United States. To this end, investigators in the field of sediment quality assessment located in the Gulf coast and southern Atlantic coast states were contacted and asked to identify studies they had conducted or participated in which contained matching sediment chemistry and biological effects data. Data sets were requested if the descriptions of these studies indicated that the data were likely to be acceptable. Contacts in the southeast included representatives from United States Environmental Protection Agency, United States Army Corps of Engineers, Florida Department of Environmental Protection, National Biological Survey, United States Fish and Wildlife Service, National Marine Fisheries Service, various academic institutions, and regionally-based consulting firms.

Significant effort was also expended to obtain additional data from other locations in the United States and Canada. In addition to the agencies identified above, contacts were made at Washington Department of Ecology, Oregon Department of Environmental Quality, California State Water Resources Control Board, Maryland Department of Environment, Port Authority of New York and New Jersey, Environment Canada, Public Works Canada, and the National Oceanic and Atmospheric Administration. Each of the 350 members of the ASTM Subcommittee E47.03 on Sediment Toxicology was contacted to obtain relevant data. Furthermore, a total of 12 bibliographic databases were searched electronically to obtain information that was published in the primary scientific literature.

Over the course of this study, more than 700 publications was retrieved and evaluated to determine their suitability for use in the derivation of SQAGs. More than 120 of these publications were used to verify and expand the original NSTP database. Roughly 35% of the publications used in the present study were from studies conducted in the southeastern and Gulf of Mexico regions of the United States (i.e., North and South Carolina, Georgia, Florida, Alabama, Mississippi, Louisiana, and Texas).

Each of the data sets obtained during the course of the study were reviewed and evaluated using the screening procedures outlined in Appendix 1. Acceptable data sets were subsequently analyzed and information pertaining to the potential biological effects of sediment-associated contaminants was integrated into the database. Following input into the database, every data entry (including each of the original NSTP database entries) was examined and verified against the original data source. This quality assurance procedure was designed to ensure that the database would meet Florida's requirements for consistently high quality data. This comprehensive, high quality sediment toxicity database provides a basis for deriving SQAGs for priority substances in Florida.

5.2 Derivation of Numerical Sediment Quality Assessment Guidelines

The expanded NSTP database consists of information from three types of studies, including equilibrium partitioning (EqP) modelling, laboratory spiked-sediment bioassays, and field investigations of sediment toxicity and benthic community composition. Equilibrium partitioning concentrations, if expressed in units of organic carbon, were converted to units of dry weight assuming a TOC concentration of 1%. Data from spiked-sediment bioassays were incorporated into the database directly. Field-collected data were treated with a variety of methods. Apparent effects thresholds (AET) and national screening level concentrations (SLC), both of which were based on evaluations of large, merged data sets, were entered directly into the NSTP database. Raw data from other individual surveys were evaluated using co-occurrence analyses (COA), using one of two procedures (Long 1992). If the authors of the reports identified samples that were statistically significantly different from the other groups of samples or from controls, then the mean chemical concentrations in the statistical groups were compared. If no such comparisons were reported, the frequency distributions of the biological data and the mean concentrations in subjectively determined groups of samples were compared (e.g., relatively highly toxic versus least toxic). Data entries were prepared for each endpoint measured in the study (e.g., survival, growth, reproduction, etc.), so that multiple entries for a single geographic area are common in the database.

The expanded NSTP database is a comprehensive source of information on the potential effects of sediment-associated contaminants. Each record in the database contains detailed information on the location of the study, species affected, endpoint measured, particle size distribution, factors that could affect bioavailability of the contaminants (such as TOC and acid volatile sulfides; AVS), and concentrations of the contaminants. Importantly, each entry in the database was assigned an 'effects/no effects' descriptor, based on the degree of concordance between the concentration of the chemical and the endpoint measured in the investigation. Those entries in which the chemical concentrations were considered to be associated with the biological effect measured were designated with an asterisk or "hit" (*; see Volumes 3 and 4). The descriptors, 'no gradient (NG), small gradient (SG), no concordance (NC), or no effect (NE)', were assigned when either the chemical concentrations were not strongly associated with the biological effect measured or no adverse effects were observed (see below). The data on each substance were then sorted, in ascending order of concentration, to create two separate data sets, which incorporated the entries associated with biological effects and the entries associated with no observed biological effects, respectively.

The 'effects data set' (EDS) was comprised primarily of information from COA in which specific adverse biological effects (as indicated from the results of sediment toxicity bioassays or benthic invertebrate community assessments) were observed at some of the sites sampled. However, results of the COA were only included in the EDS if concordance between the concentration of the chemical analyte and the observed biological response was apparent. In this respect, a contaminant was considered to be associated with the observed toxic response if the mean concentration at the sites at which significant biological effects were observed was a factor of two or more greater than the mean concentration at the sites at which no biological effects were observed (this criterion was adopted directly from Long and Morgan 1990). Data obtained from other types of studies (i.e., spiked-sediment bioassays) and sediment quality assessment values (i.e., from the SLCA, EqPA, SQTA, etc.) were also included in the EDS.

A separate data set was also established, the 'no effects data set' (NEDS), to include the balance of the assembled data. Data on the concentrations of specific substances that were not associated with adverse effects provides important information for defining the relationships between contaminant challenges and biotic responses. For this reason, it was considered appropriate to include the no effects data set in the guidelines derivation process. Several types of information were included in this data set. In general, these entries consisted of data from bioassays in which exposure of aquatic organisms to test sediments did not result in significant biological effects (i.e., no effect). In addition, the descriptors, 'no gradient, small gradient, or no concordance,' were assigned when no differences in the concentration of a particular chemical were reported between stations, the mean chemical concentrations between groups of samples differed by less than a factor of two, or there was no concordance between the severity of the effect and the chemical concentration, respectively. Data from field surveys of benthic invertebrate community indices were designated in a similar manner. Indeterminate AET values were reported in the data tables (MacDonald *et al.* 1994) but were not included to derive the SQAGs.

Both the effects and the no effects data sets were used to derive numerical sediment quality assessment guidelines for Florida coastal waters, using the steps presented in Figure 2. The arithmetic procedures used in the guidelines derivation process were designed to define three distinct ranges of contaminant concentrations; a minimal effects range, a possible effects range, and a probable effects range. A conceptual representation of the three ranges of contaminant concentrations defined by the guidelines is provided in Figure 3. This figure illustrates the concept that the probability of observing adverse biological effects increases with increasing contaminant concentration.

The range of sediment contaminant concentrations that are not likely to be associated with adverse biological effects on aquatic organisms (i.e., the minimal effects range) was defined using a two step process. First, a threshold effects level (TEL) was calculated. The TEL represents the upper limit of the range of sediment contaminant concentrations dominated by no effects data entries (i.e., the minimal effects range). Within this range, concentrations of sediment-associated contaminants are not considered to represent significant hazards to aquatic organisms. The TEL was calculated as follows:

$$\mathbf{TEL} = \sqrt{\mathbf{EDS\text{-}L} \quad \mathbf{NEDS\text{-}M}}$$

where:

TEL = Threshold effect level;
EDS-L = 15th percentile concentration in the effects data set; and,
NEDS-M = 50th percentile concentration in the no effects data set.

Figure 2. An overview of the modified NSTPA for deriving numerical sediment quality assessment guidelines in Florida.

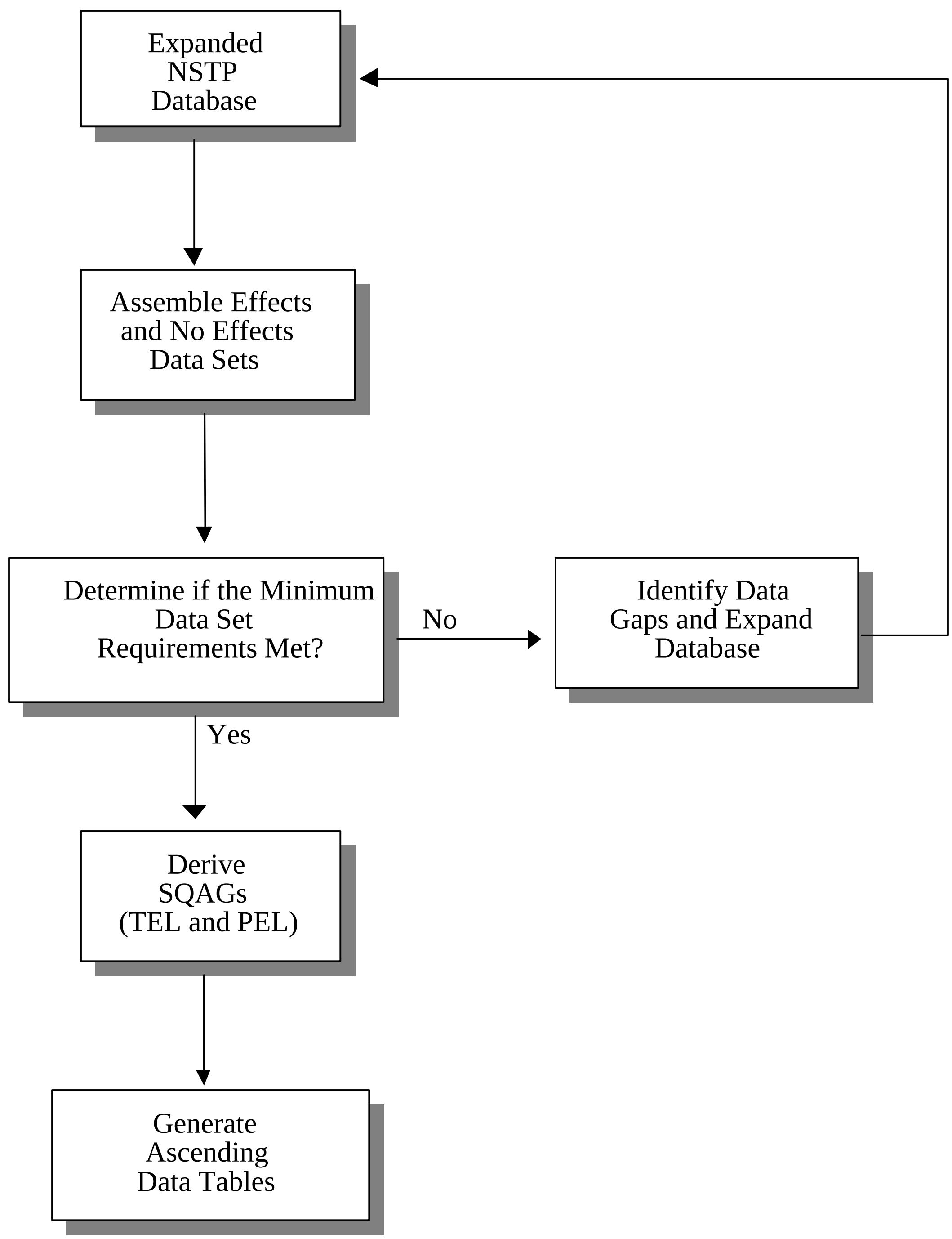
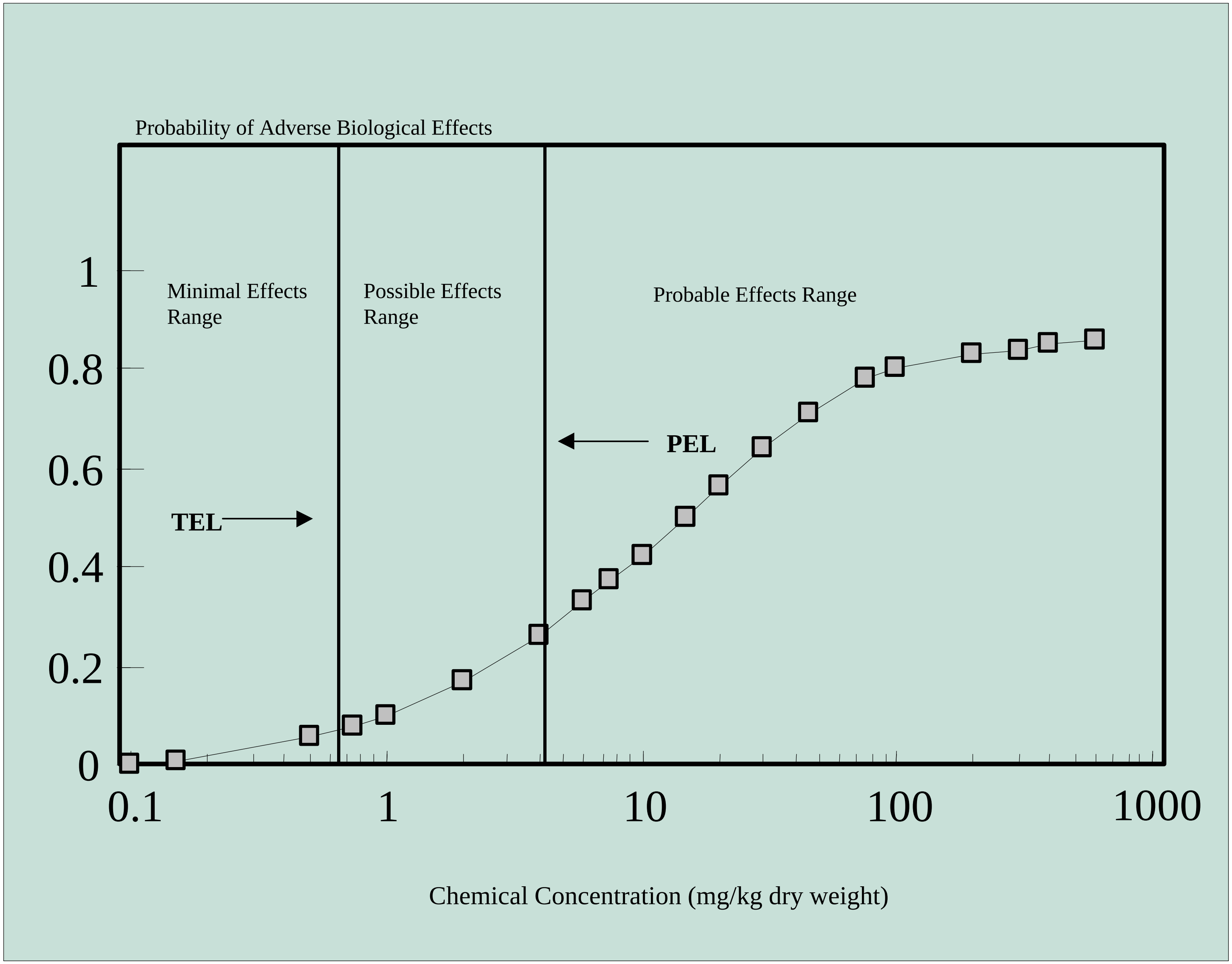


Figure 3. Conceptual example of sediment quality assessment guidelines for cadmium



The geometric mean, rather than the arithmetic mean, of EDS-L and NEDS-M is calculated because these data are not, necessarily, normally distributed (Sokal and Rohlf 1981; see Appendix 2 for an example of this calculation).

A probable effects level (PEL) was calculated to define the lower limit of the range of contaminant concentrations that are usually or always associated with adverse biological effects (i.e., the lower limit of the probable effects range). The procedure utilized to calculate the PEL is designed to define a range of concentrations that is dominated by entries from the EDS. Within the probable effects range, concentrations of sediment-associated contaminants are considered to represent significant and immediate hazards to aquatic organisms. The PEL was calculated as follows:

$$\mathbf{PEL} = \sqrt{\mathbf{EDS\text{-}M} \quad \mathbf{NEDS\text{-}H}}$$

where:

PEL	=	Probable effect level;
EDS-M	=	50th percentile concentration in the effects data set; and,
NEDS-H	=	85th percentile concentration in the no effects data set.

The range of concentrations that could, potentially, be associated with biological effects (i.e., the possible effects range) is delineated by the TEL (lower limit) and the PEL (upper limit). Within this range of concentrations, adverse biological effects are possible; however, it is difficult to predict the occurrence, nature, and/or severity of these effects. Site-specific conditions at sites with contaminant concentrations within this range are likely to control the expression of toxic effects. When contaminant concentrations fall within this range, further investigation is recommended to determine if sediment-associated contaminants represent significant hazards to aquatic organisms. Such investigations would focus first on the determination of the probable origin of the contaminant (i.e., through the use of the metals interpretive tool; Schropp *et al.* 1989), and then on the toxicity of *in situ* sediments (i.e., using bioassessment techniques), as required. See Volume 2 for further guidance on applying the SQAGs. ***It should be noted that guidelines, developed using the recommended procedures, do not address the potential for bioaccumulation of persistent toxic chemicals and potential adverse effects on higher trophic levels of the food web.***

5.3 Rationale for Establishing the Procedure for Deriving the Guidelines

Several arithmetic and statistical procedures were considered for deriving numerical guidelines from the expanded NSTP database. One of the statistical procedures that was considered involved the determination of guideline values based on the percent incidence of effects associated with increasing contaminant concentrations, which could be determined directly from the database. For example, Long *et al.* (In press) suggested that the ER-M for a specific substance would be considered reliable if there is a relatively high incidence of effects above that concentration (i.e., > 75% hits). Therefore, the percent incidence of effects was considered for deriving the SQAGs, as well. While this option was initially attractive, an evaluation of the relative distribution of the effects and no effects data entries indicated that its implementation would be challenging. Specifically, calculations using the data available on several substances (cadmium and chromium) revealed that, for example, a 75% incidence of adverse effects occurs within concentration ranges that could be delineated by several different guideline values. For this reason, it was decided that an arithmetic procedure that could be applied more consistently to each chemical should be established. While numerous arithmetic procedures are possible, the methods utilized by Long and Morgan (1990) appeared to be the most promising.

Long and Morgan (1990) used the 10th (ER-L) and 50th (ER-M) percentile values in the effects data set to establish informal guidelines for evaluating sediment chemistry data collected under the NSTP. This method was similar to the procedure used by Klapow and Lewis (1979) to establish marine water quality standards in California. Klapow and Lewis (1979) reasoned that the use of percentiles of aquatic toxicity data minimized the influence of single (potentially outlier) data points on the resultant assessment values. For this reason, a percentile-based procedure was selected to support the derivation of SQAGs for Florida coastal waters.

Environmental quality guidelines are, typically, derived from the LC_{50}s, EC_{50}s, or lowest observed effect levels that are determined from acute and chronic bioassays (Stephan *et al.* 1985; CCME 1991). In addition to this type of information, the expanded NSTP database also includes data from studies in which exposure of organisms to sediments did not result in significant biological effects. The CCREM (1987) recognized the value of using additional information for deriving water quality guidelines (WQGs) for irrigation and livestock watering. In this case, WQGs were derived by applying a safety factor to the geometric mean of the lowest and no observed effects levels that are obtained from acceptable toxicological studies. Likewise, lowest and no observed effect concentrations from tests conducted on contaminant-spiked or amended soils have been employed to derive soil quality criteria for the protection of agricultural, residential, and other land uses (CCME 1993). Furthermore, Kenaga (1982) recommended the use of acute to chronic ratios (ACRs) for estimating long-term no effect concentrations of environmental contaminants in water from median lethal or effective endpoints. In this context, chronic no observed effect levels are used in conjunction with LC_{50}s or EC_{50}s to calculate the ACRs. Therefore, it was considered appropriate to modify the original procedure from Long and Morgan (1990) to facilitate incorporation of both the effects and the no effects data into the guidelines derivation process.

A variety of arithmetic procedures, which rely on calculations of specific percentiles of the effects and no effects data distributions, could be used to derive SQAGs. To develop a screening tool, the Florida Department of Environmental Protection identified the need to develop two guideline values, which would define ranges of concentrations that would infrequently, sometimes, and frequently be associated with adverse effects. First, a low incidence of adverse effects was required in the minimal effects range. This objective could be achieved by determining chemical concentrations falling within the low end of the effects range (i.e., effects data set) and the middle of the no effects range (i.e., no effects data set). Subsequently, the mean of these values would establish a threshold effects level for each chemical. The geometric mean of these two values should be calculated because the two data distributions are not likely to be normally distributed. Using similar logic, the upper guideline value could be derived using a procedure that would establish a probable effects range that has a high incidence of adverse effects. This objective could be met by determining chemical concentrations falling within the middle of the effects range and the upper end of the no effects range. Likewise, the geometric mean of these values could be calculated to establish a probable effects level.

A sequential testing approach was used to determine which percentile values should be used to derive the numerical SQAGs. For example, the 5th, 10th, 15th, and 20th percentiles of the effects data set, in conjunction with the 50th percentile of the no effects data set, were all used to derive candidate TEL values for several indicator substances (cadmium, copper, fluoranthene, and phenanthrene). Several procedures for deriving the PEL were also proposed and tested in a similar manner. Subsequently, each of the candidate procedures was evaluated to determine which technique would be most appropriate for use in Florida.

The incidence of effects within each of the ranges of contaminant concentrations provides a powerful tool for evaluating the reliability of SQAGs (Long *et al.* In press). Therefore, the percent incidences of adverse effects within the minimal and probable effects ranges were calculated for each of the TELs and PELs, respectively, derived using the candidate procedures. The TELs were considered to be reliable if the incidence of adverse effects was less than 10% within the minimal effects range, while an incidence of adverse effects of 65% or more within the probable effects range was considered indicative of reliable PELs. Only those procedures that satisfied the criteria for reliability were further considered.

Next, the remaining procedures were evaluated in terms of their overall applicability. In this respect, candidate procedures were considered to be broadly applicable if they supported the derivation of guidelines that could be used to effectively screen sediment chemistry data at a large number of sites. That is, the guidelines would be most useful if they defined only a limited range of contaminant concentrations between the TEL and PEL. By tightening the possible effects range (i.e., reducing the range of uncertainty), the guidelines could be used to clasify a larger number of samples in terms of their potential toxicity. Of all the procedures tested, the methods described in section 5.2 appeared to provide the most reliable and appropriate means of deriving SQAGs for Florida coastal waters.

There is a great deal of variability in the quantity of information available for each chemical, ranging from less than twenty data entries for 2,3,7,8-T_4CDD to several hundred data entries for cadmium. Due to the uncertainty associated with the evaluation of matching sediment chemistry and biological effects data, minimum data requirements were established to ensure that any guidelines developed were supported by the weight of evidence. To adhere to this principle, SQAGs were derived only for those contaminants which had at least 20 entries in both the effects and no effects data sets.

The number of data entries required to support the derivation of preliminary SQAGs was established based on the results of sequential calculations of guidelines for a total of four substances (cadmium, chromium, fluoranthene, and PCBs) using data sets of various sizes. Using the guideline derivation procedure described above, guidelines were sequentially calculated using randomly selected data sets of various sizes (ranging from 2 to 30 data entries). This procedure was repeated 10 times for each chemical to support the calculation of the mean guideline value and its standard deviation for each data set size. The results of this investigation indicated that the estimate of the guideline value stabilized when the data set contained 15 to 20 entries. The variability in this estimate was not significantly reduced over the range of 20 to 30 entries. Therefore, it was concluded that at least 20 entries from each data set were required to support the derivation of SQAGs.

5.4 Strengths and Limitations of the Recommended Approach for Developing Sediment Quality Assessment Guidelines

The following discussion provides a general critique of the recommended approach for deriving SQAGs in the context of Florida's requirements.

5.4.1 Strengths of the Approach

The WEA is characterized by a number of attributes that make it an attractive choice for deriving SQAGs for Florida coastal waters. First, the approach is supported by a comprehensive database on the biological effects of sediment-associated contaminants. Interpretation of the information contained in the expanded NSTP database provides relevant tools for evaluating the potential for biological effects at various contaminant concentrations. Such interpretations are supported by detailed summaries of a large volume of data linking contaminant concentrations to biological effects (MacDonald *et al.* 1994). In this way, the WEA provides an effective basis for evaluating the degree of confidence that should be placed on the resultant guidelines. Confidence in these data is enhanced by the rigorous screening procedures that were used to evaluate candidate data sets.

Unlike other approaches to the development of SQAGs, the WEA does not attempt to establish absolute sediment quality assessment values. Instead, the approach delineates ranges of contaminant concentrations that are probably, possibly, and not likely to be associated with adverse biological effects. This approach explicitly recognizes the uncertainty associated with the prediction of biological effects from chemical concentration data, thereby enhancing the defensibility of the guidelines for identifying priority conditions with respect to contaminated sediments.

One of the more important attributes of the WEA is its overall practicality. The expanded NSTP database supports the derivation of numerical SQAGs for a variety of chemical substances. Many of these chemicals are known to occur in Florida's coastal sediments at elevated levels. Sediment quality assessment guidelines are required to support assessments of the potential for biological effects in these sediments. In addition, by considering matching sediment chemistry and biological effects data from studies conducted in the field, the influence of mixtures of chemicals in sediments is incorporated in the resultant SQAGs. Furthermore, the information in the expanded NSTP database is highly relevant to the guidelines derivation process because it applies to a wide range of biological organisms and endpoints, and incorporates a large number of direct measurements on organisms that are normally associated with bed sediments. The database also includes a significant quantity of data from studies conducted in the southeastern United States (including Florida). These attributes are likely to give the SQAGs derived using the WEA broad applicability. As such, there is a high probability that these guidelines will be appropriate for implementation in Florida.

In addition to the other advantages of the approach, the procedure recommended for calculating SQAGs considers both the EDS and NEDS for each chemical constituent. And, in contrast to the AETA, this procedure does not rely heavily on individual data points. Therefore, outliers do not carry much weight in the overall guidelines derivation process. Integration of the effects and no effects data sets into a single data set allows preparation of ascending data tables for each contaminant. These data tables provide detailed information on specific biological responses that have been observed at various concentrations of the contaminant and, in themselves, can contribute significantly to sediment assessments. As such, these data tables are useful tools for evaluating the potential biological significance of contaminant concentrations that fall within the three ranges described above (i.e., no effects, possible effects, and probable effects ranges). Many reviewers of the original NSTP document (Long and Morgan 1990) indicated that the data tables (with contaminant concentrations arranged in ascending order) were extremely useful tools for evaluating sediment quality data from specific sites (E. Long, NOAA, Seattle, Washington. Personal communication).

In addition to supporting the derivation of numerical SQAGs, the expanded NSTP database provides a basis for evaluating the guidelines. In this respect, it is possible to use the database to calculate the distribution of effects and no effects entries within each range of contaminant concentrations (i.e., the minimal effects, possible effects, and probable effects ranges). These distributions describe the percent incidence of effects (i.e., number of biological effects entries divided by the total number of entries times 100) within each range of concentrations and, as such, provide an estimate of the probability of observing adverse biological effects when the concentration of a contaminant falls within a specific range of concentrations (Long *et al.* In press). For example, if the incidence of effects within the probable effects range for cadmium was 68%, then there is a high likelihood that adverse biological effects would be observed at sites with concentrations of cadmium equalling or exceeding the PEL. This feature of the approach provides environmental managers a measure of confidence for ranking the relative priority of contaminated sediments. This tool can also be used within a risk assessment framework for assessing contaminated sediments.

Lastly, the WEA has been extensively reviewed by experts from across North America. Over 1000 copies of the original publication (Long and Morgan 1990) have been distributed to date. In addition, it has recently been peer reviewed and published in a primary journal (Long 1992; Long et al. In press). Further, it has been described and evaluated in the *Sediment Classification Methods Compendium* (EPA 1992). Since its development in 1990, this approach has received positive evaluations from a wide variety of user groups and has been adopted directly and/or modified for implementation by both California (Lorenzato and Wilson 1991) and Canada (Smith and MacDonald 1993) as part of their guidelines derivation process. These favourable assessments emphasize the importance and utility of this procedure for deriving numerical SQAGs.

5.4.2 Limitations of the Approach

In spite of the obvious benefits associated with the WEA for deriving SQAGs, a number of limitations are also evident which could restrict application of these guidelines in Florida. The most serious of these shortcomings is associated with the limitations on the data that describe the bioavailability of sediment-associated contaminants. Ancillary data on grain size, levels of TOC, and concentrations of AVS were not provided in a large number of the reports reviewed. As such, it is not currently possible to express the guidelines in terms of the factors that could influence the bioavailability of these contaminants. While reliance on ranges of concentrations instead of absolute values and consideration of the no effect data set serves to minimize this limitation, a potential for under- or over-protection of aquatic resources exists if guidelines are implemented that do not consider the bioavailability of sediment-associated contaminants.

The Florida Department of Environmental Protection (FDEP) recognizes the importance of accounting for the bioavailability of sediment-associated contaminants and is participating with NOAA in bio-effects studies to help address this issue. In addition, FDEP has developed a companion tool for use with the biological effects-based guidelines. The metals interpretive tool is based on normalizing metal levels to concentrations of aluminum in sediment, and provides a means of assessing the probable origin of sediment-associated metals (Schropp and Windom 1988; Schropp *et al.* 1989; Schropp *et al.* 1990). This tool emphasizes the importance of 'normalizers' in the interpretation of sediment chemistry data and provides a practical tool that compliments the effects-based guidelines. A detailed discussion on how these tools may be used together to assess sediment quality is provided in Volume 2.

It is anticipated that the bioavailability of sediment-associated contaminants will be one of the principal issues addressed in refining the preliminary guidelines. Currently, there is little comprehensive information with which to reliably predict the bioavailability of sediment-associated contaminants. Nonetheless, EPA has developed sediment quality criteria for five non-polar organic substances that are expressed on an OC-normalized basis. Data from other studies conducted in Florida and elsewhere may provide more relevant information for determining the factors that influence bioavailability in Florida coastal sediments.

Another limitation of the WEA is that it does not fully support the quantitative evaluation of *cause and effect* relationships between contaminant concentrations and biological responses. Although information from spiked-sediment bioassays and EqP models is included in the expanded NSTP database, the recommended approach is considered to predict *associations* between contaminant concentrations and biological responses only. A wide variety of factors other than concentrations of the contaminant under consideration could have influenced the actual response observed in any given investigation. For example, Jones-Lee and Lee (1993) identified ammonia as a factor that could potentially affect the toxicity of aquatic sediments. More recent results from more intensive monitoring in areas such as Tampa Bay and Los Angeles Harbor indicate that unionized ammonia levels in porewater were rarely in the range that would be associated with biological effects (Long Unpublished data). Nonetheless, variables such as ammonia, hydrogen sulfide, and dissolved oxygen, have the potential to affect benthic organisms and should be evaluated in site-specific assessments of sediment quality. In the WEA, the assembly of extensive information from numerous estuarine and marine sites in North America into a single database serves to minimize the impacts of unmeasured contaminants on the guidelines; however, there is still an undefined level of uncertainty associated with the application of the SQAGs at specific sites.

Application of the recommended approach may also be restricted by other limitations on the available information. Currently, limited data are available on the chronic toxicity of sediment-associated contaminants to marine and estuarine organisms. However, information on the responses of *in situ* benthic invertebrate communities to contaminant challenges partially mitigates this limitation on the database. In addition, only limited data are available on some potentially important sediment contaminants in Florida (including a variety of pesticides, dioxins and furans, etc.).

The results of this preliminary evaluation indicate that SQAGs developed using the recommended procedure are likely to be appropriate tools for conducting assessments of sediment quality in Florida. However, care should be exercised in applying these guidelines. In particular, these guidelines may not be directly applicable to sediments with atypical levels of the factors that may influence the bioavailability of contaminants (e.g., very high or very low levels of TOC). Detailed discussions on the application of SQAGs in regional and site-specific sediment quality assessments are provided in Volume 2. In addition, a more thorough evaluation of the SQAGs comparability, reliability, and predictability is presented in Chapter 7.