LONDON CONVENTION AND PROTOCOL: GUIDANCE FOR THE DEVELOPMENT OF ACTION LISTS AND ACTION LEVELS FOR DREDGED MATERIAL

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This document provides guidance to regulators and policy makers on the selection of Action Lists and the development of Action Levels for dredged material proposed for disposal at sea. An Action List is a set of chemicals of concern, biological responses of concern, or other characteristics that can be used for screening dredged material for their potential effects on human health and the marine environment. Action Levels establish thresholds that provide decision points that determine whether sediments can be disposed of at sea.

While the guidance is designed to assist with implementation of requirements under the *Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter, 1972* (London Convention) and its *1996 Protocol* (London Protocol), the guidance is general and could be applied to the assessment of dredged material under other instruments. The guidance does not, however, cover the assessment of other wastes or other matter allowed under the London Convention and Protocol.

There is no universal mechanism for the selection of Action Lists and the development of Action Levels. As such, the document is not prescriptive. Instead the document provides guidance on the process of selecting Action Lists and Action Levels and proposes options that are available to regulators and policy makers.

The process begins with an identification of the chemical, biological, or physical characteristics that will make up the Action List. This can be done by surveying relevant sources of contaminants to dredged material, inventorying valued resources in proximity to disposal sites and their risk factors, and tracking best practices in the science of sediment assessment and management. Also, for the purposes of the London Convention and Protocol, *priority shall be given to toxic, persistent and bioaccumulative substances from anthropogenic sources*.

Next, benchmarks must be set for each characteristic on the Action List. Benchmarks are often developed using a reference-based approach (comparing to background or ambient conditions) or an effects-based approach (based on knowledge or direct observation of the effects of exposure). Appendix 1 goes into further detail providing options that can be used in setting benchmarks.

Finally, Action Levels are set by integrating the relevant characteristics and benchmarks to form a decision rule. This can be as simple as a pass/fail based on a single benchmark or it can be more complex such as combining multiple lines of evidence in a weight-of-evidence approach. An Upper Action Level can be created above which there would be concern due to increased potential for effects on human health and the marine environment, and if desired, a Lower Action Level can be created below which there would be little concern. The document concludes by providing guidance on formulating Action Levels in language suitable for developing National regulations and dealing with sediments that fall between the Upper and Lower Action Levels. Appendices 1A to 1D provide descriptions of different approaches and country examples.

1 INTRODUCTION

1A. PURPOSE OF GUIDANCE

1.1 This document is intended to assist regulators and policy makers in developing National **Action Lists** and **Action Levels** for the assessment of dredged material proposed for disposal at sea.

1B. WHAT ARE ACTION LISTS AND ACTION LEVELS?

1.2 In the context of disposal at sea:

.1 An **Action List** is defined as *a "mechanism for screening candidate wastes and their constituents on the basis of their potential effects on human health and the marine environment".* The Action List can consist of chemicals of interest, biological responses of concern, or other characteristics that can provide insight into the potential for dredged material to cause adverse effects in the marine environment. An Action List can also be used as a trigger mechanism for further waste prevention considerations and could therefore have a role in controlling pollution at its source, in promoting cleaner technology, or in improving the efficiency of dredging to reduce the need for disposal.

.2 **Action Levels** are established as decision rules that identify dredged material that may be disposed because the risk for adverse effects is low and acceptable, those that may not be disposed without management controls because the risks for adverse effects would be considered too high, or to identify cases where additional information may be required to make a sound judgement about the potential for the dredged material to cause adverse effects. If developed for the purposes of meeting the requirements of disposal at sea treaties, the Action Level will specify an Upper Level and may also specify a Lower Level. The Upper Level should be set so as to avoid acute or chronic effects on human health or on sensitive marine organisms representative of the marine ecosystem. Below the Lower Level, there should be little concern for disposal at sea.

1C. WHY ARE ACTION LISTS AND LEVELS IMPORTANT?

1.3 The Action List is important to Contracting Parties, and prospective Contracting Parties, to the London Convention and Protocol as it is a key decision-enabling component of the Generic Guidelines. "*The **Guidelines for the Assessment of Wastes or Other Matter that May Be Considered for Dumping** guide national authorities in evaluating applications for dumping of wastes in a manner consistent with the provisions of the London Convention or Protocol*." A shortened form of these Guidelines appears in Annex 2 to the London Protocol and a more specific version of the Guidelines has been developed particularly for dredged material assessment. In each version of these Guidelines, the application of an Action List and its levels is used to enable authorities to categorize the dredged material as being: 1) of little concern for disposal at sea; 2) as requiring more detailed assessment; or 3) as not being suitable for disposal at sea without the use of management techniques or processes. A jurisdiction that has developed a National Action List and Action Levels will be in a better position to make sound permit decisions and to be in compliance with the requirements of these treaties. Action Levels can provide feedback for compliance efforts, for further assessment or for monitoring.

1.4 Action Lists and Levels may also be of use to non-parties that require a consistent and transparent scientific basis by which to categorize or assess dredged material based on the level of risk they may pose to the marine environment upon disposal.

1.5 This document provides guidance on the selection of an Action List and considerations for the development of Action Levels for dredged material. It is not a detailed technical manual, but rather provides an overview of the options for development and adoption of Action Lists and Levels, as well as examples from various jurisdictions. There is some discussion of the implementation of an Action List as part of the decision-making process for permitting disposal of dredged material at sea. Those jurisdictions with limited experience that wish to adopt an Action List and Levels will likely require additional guidance and support to select the most suitable approach and to adapt it as needed to their legal and environmental circumstances. Also, the guidance is given with a view to achieving a balance between the best level of assessment possible and the availability of resources and capacity in different countries. The jurisdiction is encouraged to begin with practices that are achievable in the short term, with a view to continuing improvement as capacity and expertise are acquired.

1.6 This is the first iteration of this document and it is recognized that as global experience with Action Lists and Levels increases, there will be a need to update and elaborate the information and examples provided. Comments on this guidance are welcomed and should be addressed to the IMO Secretariat for the London Convention and Protocol.

1D. OTHER WASTES OR MATTER

1.7 Fish waste, organic material of natural origin, inert, inorganic geological material, vessels and platforms, bulky items, etc., and carbon dioxide streams may also be considered for disposal at sea permits and the content of the Action List and the way that it is used may be different in each case. In future, separate documents will be produced for the development of Action Lists and Action Levels relevant to these wastes or other matter.

1E. ACTION LISTS AND LEVELS AS PART OF THE FRAMEWORK FOR ASSESSING DREDGED MATERIAL FOR DISPOSAL AT SEA

1.8 The generic waste assessment guidance (WAG) framework under the London Convention and Protocol is shown in Figure 1. It is an iterative process meaning that steps do not necessarily have to be taken in order. Action Lists make use of physical, contaminant, or biological testing data collected during the characterization step. Following this characterization, Action Levels, which are one of the key decision-making points in the framework, are used to determine whether dredged material are acceptable for disposal at sea. For general understanding of Action Lists in the context of the framework, refer to the WAG Tutorial.

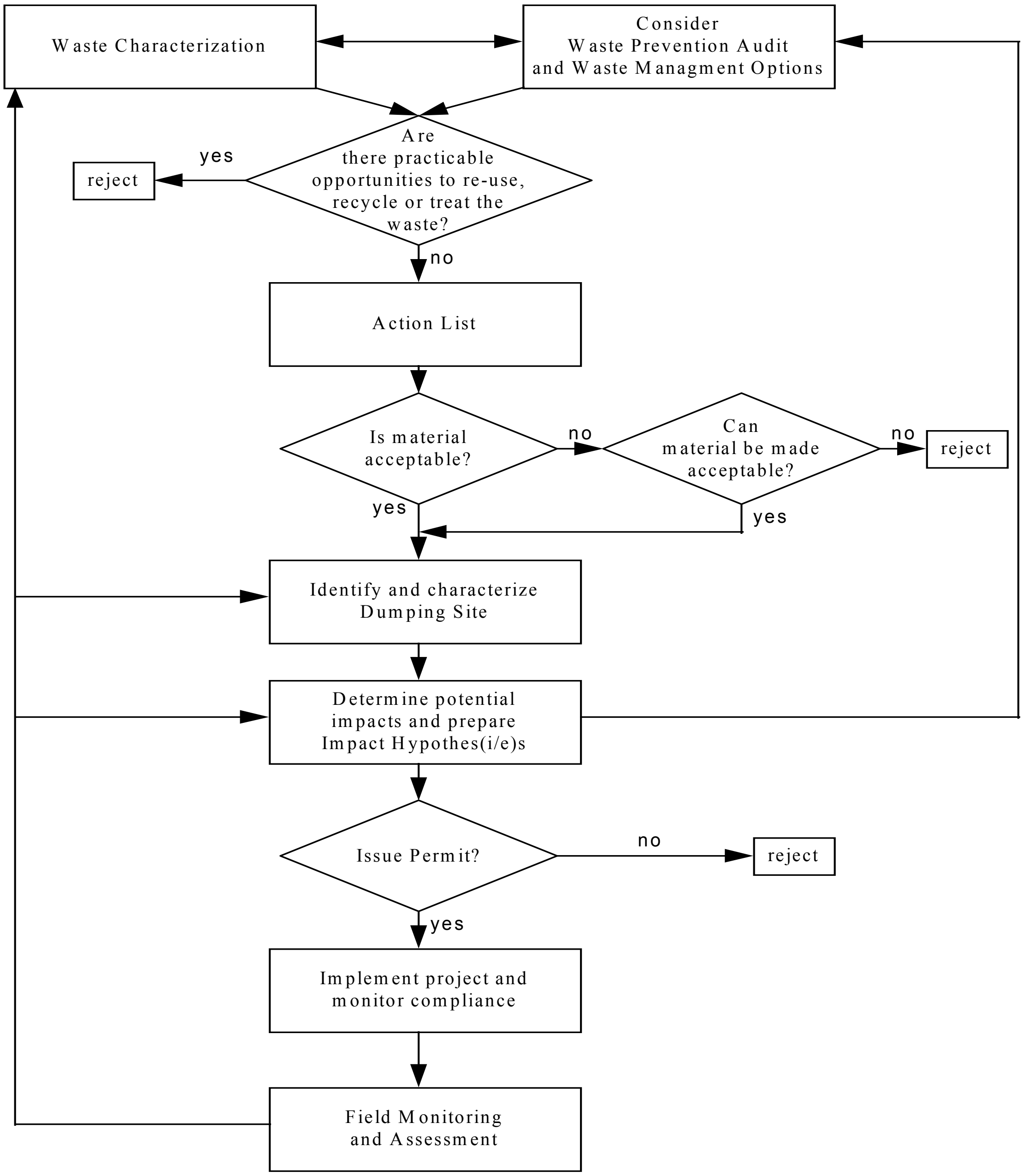

Figure 1. **Assessment Framework for the London Convention and Protocol**

2 SELECTION OF A NATIONAL ACTION LIST

2.1 A Dredged Material National Action List is a list or inventory of dredged material **characteristics** and their **metrics** that a jurisdiction decides are important to consider in order to make permit decisions. To arrive at this Action List, authorities will need to consider what potential concerns are created by the disposal of dredged material in their jurisdiction and what assets and resources need to be protected. This consideration should lead to a determination of what needs to be measured and assessed. In practice an Action List will be developed by assembling a list of characteristics that will be used to perform a regulatory evaluation of dredged material. **Benchmarks** for each characteristic are used for developing decision rules to define the Upper and Lower Action Levels.

Box 1. Definition of Major Terms

In this document the terms **characteristic, metric** and **benchmark** define the tools that are used to evaluate some aspect of the environment.

A **characteristic** is an attribute of the dredged material (e.g., copper, mercury, silt, petroleum compounds, pathogens) or a biological response to the dredged material (e.g., mortality, growth, bioaccumulation).

A **metric** is a measurement that can be made on the characteristic (e.g., concentration, percent survival).

A **benchmark** is a point on the range of the metric (e.g., 4 mg/kg copper, 20% amphipod mortality) that is used to identify where environmental concern may be low or high for that characteristic. These can be referred to as the lower benchmark and upper benchmark.

An Action List therefore comprises a number of characteristics to be considered for measurement in the dredged material.

An Action Level is a decision rule based on the findings of one or more characteristics in comparison to the respective benchmarks.

2.2 Selection of the characteristics and metrics in a National Action List should be based on knowledge concerning the nature of dredged material in the country where the list is to be used. For chemical characteristics, contaminants known to be in the material and those likely to have been deposited in the material from, *inter alia*, known point-source effluent discharges, tributaries, diffuse runoff, atmospheric deposition, accidents and spills, operational discharges and losses and direct dumping may need to be considered. Characteristics that give useful information on the potential for acute or chronic effects on sensitive marine organisms or on health should also be considered. For example, characteristics could include substances (particular chemicals of concern or interest); microbes, viruses or pathogens; biological responses or phenomena such as toxicity or bioaccumulation, and physical characteristics of the dredged material. See Appendix 2 for examples of lists developed by some of the London Convention and Protocol Parties.

Box 2. The Generic Guidelines specify the following with respect to developing Action Lists:

Action Lists of the London Convention and Protocol

Action Lists are an important part of Annex 2 to the London Protocol and may also be used in meeting the requirements of Annexes I and II to the London Convention. Details on such lists are further provided in the Revised Generic Guidelines (2008):

"National Action Lists will provide the mechanism for screening candidate wastes and their constituents on the basis of their potential effects on human health and the marine environment"… "In selecting substances for consideration in an Action List, priority shall be given to toxic, persistent and bio-accumulative substances from anthropogenic sources (e.g., cadmium, mercury, organohalogens, petroleum hydrocarbons and, whenever relevant, arsenic, lead, copper, zinc, beryllium, chromium, nickel and vanadium, organosilicon compounds, cyanides, fluorides and pesticides or their by-products other than organohalogens)."

"For an individual waste category, it may be possible to define national action levels on the basis of concentration limits, biological responses, environmental quality standards, flux considerations or other reference values."

2.3 Box 2 above outlines examples of chemical, biological and physical characteristics listed in the Revised Generic Guidelines which may be appropriate for inclusion in an Action List.

2.4 For those jurisdictions where little chemical or biological effects data are available, the authorities may wish to begin with an interim National Action List, either selected from another jurisdiction, or based on the guidance provided above within the London Convention and Protocol Guidelines. The commitment to refine an interim Action List could involve such follow-up actions as:

.1 *Conducting a survey of relevant sources of contaminants to the dredged material, including sources of industrial, agricultural, and urban run-off.* The purpose of such a survey is to ensure that the Action List is sufficiently comprehensive to support credible assessments of the potential for adverse effects. Sediment surveys can be used to confirm the presence and prevalence of characteristics.

.2 *Developing an inventory of valued resources in proximity to known or intended disposal sites* to provide the basis for Action Lists and Levels that will support sustainable management practices.

.3 *Tracking the development of relevant science*. The science that supports sediment assessment and management will evolve with time. Tracking advances in the relevant fields of study will enable authorities to benefit from updating their approaches.

2.5 The experience gained by authorities over time in applying Action Lists and levels, including the use of confirmatory monitoring, will support making updates and refinements to Action Lists and their application. Over time, the analysis of this information will help confirm or further refine the List.

2.6 Given that dredged material will often be influenced by site-specific sources of contamination, arriving at a National Action List that is representative of national concerns but not so large as to reduce the ability to conduct time- and cost-efficient assessments, is challenging for all jurisdictions. One approach is to set a smaller National Action List that includes only the most prevalent and critical characteristics that must be evaluated as minimum information in all cases and allow for the development and application of regionalized Action Lists that incorporate regional, local and site-specific knowledge of dredged material characteristics and valued resources. For example, when cadmium is widespread in the jurisdiction and of concern, it would appear on the National List. Whereas, chromium might only be prevalent in a limited number of areas associated with specific industries, and for that reason, would be included on relevant regional Action Lists.

3 ESTABLISHING UPPER AND LOWER BENCHMARKS

3.1 Following Section 2, a National Action List can be chosen. Each characteristic (e.g., cadmium, survival, etc.) will have a metric (what is being measured: mg/kg dry wt, % survival). The benchmarks are the levels for a particular characteristic below which there would be little concern (lower benchmark), or above which there would be concern due to increased risk or increased probability of effects (upper benchmark). Once benchmarks are established for the characteristics on the List they are used to establish the Upper, and if desired, may be used to establish Lower Action Levels (refer to section 1.2.2).

Relationship between Benchmarks and Action Levels

3.2 The application of Action Levels can range from relatively simple approaches to more complex formulations. In the simple approach (Table 1), the Action List consists of a series of contaminants (characteristics) that may be present in the material. By some means (see sections 3A, B and C below), lower and upper benchmarks are established for each characteristic on the List. Using the simple approach, exceedance of any *single* upper benchmark would be considered an exceedance of the Upper Action Level. In a complementary manner, following the simple approach, all characteristics of the sediment must be below the lower benchmarks to reach the conclusion that the material poses a low and acceptable level of risk to the marine environment and does not exceed the Lower Action Level. Sediments meeting neither of those situations would require additional investigation or evaluation before a decision could be reached.

Table 1. An Example of a Simple Action Level Approach

Single Characteristic Action Level Model				
Dredged material Characteristic	Dredged material passes Lower Action Level when:	Lower benchmark (LB) (mg/kg)	Upper benchmark (UB) (mg/kg)	Dredged material exceeds Upper Action Level when:
Contaminant A	All values below LB	120	340	Exceedance of any UB
Contaminant B		25	88	
Contaminant C		75	420	
Contaminant D		0.5	2.7	
Contaminant E		50	170	

3.3 More complex approaches use decision rules that rely on the exceedance of benchmarks by **multiple characteristics** to reach a determination that the Upper Action Level has been reached (Tables 2 and 3). In the first of these two examples (Table 2) the jurisdiction decided that the Action List of characteristics would consist solely of contaminants. In this case, the jurisdiction has determined that certain characteristics are of greater relevance to the decision process based on the nature of the information they provide. For example, some contaminants are of greater toxicological significance than others, or they may be more persistent and those factors may influence the use made of the benchmark in decision-making.

Table 2. Example 1 of a More Complex Action Level Approach

Weight of Evidence Action Level Model 1				
Dredged material Characteristic	Dredged mat. passes Lower Action Level	Lower benchmark (LB) (mg/kg)	Upper benchmark (UB) (mg/kg)	Dredged material exceeds Upper Action Level
Persistent Organic	No exceedance of any UB, all Organic values below LB, no more than 1 Metal between LB and UB			Exceedence of any Organic UB or 2 Metal UB
Organic A		2.5	15	
Organic B		0.5	7.5	
Organic C		0.5	2.7	
Organic D		0.1	2.2	
Heavy Metal				
Metal A		50	125	
Metal B		140	330	
Metal C		85	210	
Metal D		14	40	

3.4 In the second example (Table 3), the selected Action List of characteristics includes both contaminants and biological responses.

Table 3. Example 2 of a More Complex Action Level Approach

Weight of Evidence Action Level Model 2				
Dredged material Characteristic	Lower Action Level	Lower benchmark (LB)	Upper benchmark (UB)	Upper Action Level
Persistent Organic	No exceedance of any UB, no more than 1 Organic between LB and UB	(mg/kg)	(mg/kg)	Exceedence of any Organic UB or two Bioassay UB
Organic A		2.5	15	
Organic B		0.5	7.5	
Organic C		0.5	2.7	
Organic D		0.1	2.2	
Benthic Bioassay			% Mortality	
Species A			25	
Species B			30	
Species C			20	
Species D			30	

APPROACHES

3.5 This section discusses strategies to establish upper and lower benchmarks for the characteristics chosen to be part of the National Action List. Some jurisdictions have already used these approaches and set numerical levels or decision-making criteria for the characteristics that are relevant to their situations. When formulating or revising practice with the intent of establishing Action Levels, Contracting Parties are well served to review the practices of others and draw from existing approaches.

3.6 Benchmarks should be developed and applied with an understanding of what valued resources they are intended to protect and the technical argument linking the specific benchmark and the protection objective.

3.7 When reviewing the approaches it will be important to gain an understanding of the:

.1 *Theory and method of derivation* – Will this approach generate levels that are consistent with the objective?

.2 *Assumptions* – Will any of the assumptions built into the approach make it less relevant for use in developing the benchmarks for this jurisdiction?

.3 *Data needs and the uncertainties* – Does this approach require local data in order to be relevant? How much data would be needed to set a level for a given characteristic? What is the level of uncertainty associated with this approach for this characteristic? Is this approach equally useful for all characteristics and their metrics?

3.8 The Generic Guidelines are relatively clear in the description of what Upper and Lower Action Levels are intended to do:

"The Upper Level should be set so as to avoid acute or chronic effects on human health or on sensitive marine organisms representative of the marine ecosystem."

and for the lower level the description is:

"... wastes which contain specified substances, or which cause biological responses, below the relevant Lower Levels should be considered to be of little environmental concern in relation to dumping."

3.9 Given that the objective for Upper and Lower Action Levels is different, it is not essential that the same approach be applied to each upper and lower benchmark or to all the characteristics on a National Action List. It should be recognized that the use of different approaches might mean that benchmarks for the same characteristic may not be comparable across lists or jurisdictions. Approaches that have been used to derive benchmarks are described briefly below and in more detail in Appendices 1A to 1D.

3A. REFERENCE-BASED APPROACHES

3.10 Benchmarks for physical, chemical or biological characteristics can be set based on knowledge of background or ambient conditions in comparable areas that have not been impacted by dumping. This is a reference-based approach. For example, the Lower Action Level may be set at the background concentration for the chemical of interest. One approach for establishing lower benchmarks for an Action List would be to establish the lower benchmark as the 50th percentile of the background concentration distribution for each contaminant on the Action List. Alternatively, lower benchmarks could be established using the results of sediment toxicity tests by using reference conditions to compare the responses of test animals exposed to the dredged material and to reference sediment.

3.11 When using reference-based approaches to develop benchmarks for individual characteristics on the Action List, it is important to distinguish between man-made substances and naturally occurring substances, e.g., PCBs vs. ammonia. While PCBs are present in the environment as a direct result of industrial activity, ammonia is the natural product of protein diagenesis in sediments. Ammonia levels in sediment can be affected by human activity, e.g., through the introduction of nutrients and fertilizers; however, in most circumstances its presence in sediments does not evoke a level of concern comparable to PCBs. As a means of focusing regulatory attention on anthropogenic activities, chemical benchmarks using the reference approach have made use of information on contaminant levels in heavily industrialised harbours, lightly industrialized harbours, and recreational harbours to establish the distribution of data used to establish the Action Level for a contaminant.

3.12 Regardless of the characteristic chosen, the resulting benchmarks need to be indicative of the potential for effects in the field, which is the basis of the Generic Guidelines. An example of reference-based benchmarks is described in Appendix 1A of this document. The Generic Guidelines specify that Upper Action Levels should be set so as to avoid acute or chronic effects on human health, or on sensitive marine organisms. Therefore, any benchmarks used to establish Upper Action Levels should minimize, to the extent practical, likelihood that dredged material could exceed such values but produce no effects at a disposal site (false negatives). Reference-based levels are commonly used for setting lower benchmarks and Lower Action Levels, as it is reasonable to expect that levels that are similar to background levels would be unlikely to cause unacceptable effects.

3B. EFFECTS-BASED APPROACHES

3.13 Benchmarks for physical, chemical or biological characteristics can also be based on knowledge of effects that can be produced following exposure to dredged material. Such limits can be based on information concerning the likelihood or magnitude for an effect.

3.14 The physical characteristics of the dredged material can be used to reach conclusions about whether the dredged material is unlikely to cause adverse effects on the environment, i.e. to establish lower benchmarks. For example, sediments found in areas of high current or wave energy and composed predominantly of coarse-grained sediments (e.g., rock, cobble and sand) have a low potential to carry significant amounts of chemical contaminants because of the relatively small surface area available for sorption of contaminants. Based on past experience, regulatory authorities may set quantitative or qualitative criteria to define when sediment will be judged to be predominantly composed of such coarse-grained material. Lower benchmarks for other physical characteristics that can be used in combination with the geotechnical data to establish Action Levels include the depth of dredging (e.g., will the material be dredged from sediment horizons that have had no contact with industrial chemicals) and geographic proximity to known or suspected sources of contaminants (PIANC 2006). Physical factors are also important additional pieces of information that can be used to adjust benchmarks set using other approaches. For example, if toxicity tests are used as one of the characteristics in the List, it will be important to know how physical characteristics can affect or confound the results of the toxicity test (PIANC 2006).

3.15 Chemical benchmarks are developed using an effects-based approach by making use of calculated or measured relationships between the concentration of the chemical(s) and some form of biological response. Such levels can be established using a variety of empirical and theoretical approaches and many examples are available. The chemical concentration that establishes the limiting benchmark can be based on concentrations in whole sediment, a sediment fraction, porewater, or the tissues of organisms exposed to sediments in a biological test (PIANC 2006). There may also be a desire to set levels that guard against unacceptable effects on human health (e.g., safe fish consumption levels). Appendix 1B to this document describes some of the major approaches to developing effects-based chemical benchmarks.

3.16 Biological benchmarks can be set using information to establish the likelihood that effects would be observed in the field, or to distinguish an acceptable from an unacceptable magnitude of effect. Biological benchmarks are generally expressed as some type of biological response (e.g., rates of survival, growth or reproduction in the test organism used in a toxicity test, changes in benthic community structure, etc.). Biological benchmarks have also been set by establishing a threshold for the magnitude of response that must be observed in a toxicity test before the Action Level is determined to have been exceeded, e.g. > 20% more mortality observed in a dredged material in comparison to a reference sediment (this specific example illustrates the use of an approach that combines both reference and effects-based approaches). Examples of biological tests which can be relevant to the development of biological effects based Action Levels are described in Appendix 1C to this document.

3.17 Where biological effects are used to set Action Levels, benchmarks concerning the likelihood for effects can be derived from the results of a battery of toxicity tests performed on dredged material, whereby the larger the number of tests in the battery that show evidence of toxicity the greater the confidence that effects are likely to occur. It is important to note that typical sediment toxicity testing with benthic organisms is not an appropriate means of assessing risk from chemicals whose primary effects are mediated through bioaccumulation, trophic transfer and subsequent effects in higher-level predators (e.g., dioxin-like chemicals). Where such chemicals are on the Action List, assessment should be based on methods directly addressing bioaccumulation pathways (Wenning *et al.*, 2005).

3C. SETTING BENCHMARKS

3.18 One of the considerations in evaluating the approaches described above will be meeting the data requirements and the cost, time, and capacity considerations associated with developing the benchmarks. Frequently, there will be insufficient data, time or funding to ensure that benchmarks are set on purely scientific grounds and that all uncertainties in the methods and the data can be addressed. In order to proceed with a functional decision-making system in a reasonable time it is often necessary to take interim measures. Many jurisdictions may have limited information and simply decide to apply safety factors to benchmarks derived for other purposes, or set one benchmark as a multiple of another benchmark in an arbitrary fashion to help overcome a lack of data, or allow consistent decisions to be made.

3.19 When data are insufficient within a jurisdiction to calculate or derive benchmarks for specific characteristics on the National Action List, upper and lower benchmarks can also be adopted directly from other jurisdictions as an interim measure.

3.20 However, such action should be combined with a broader strategy to evaluate the reliability of these levels within the subject jurisdiction and/or to derive levels that are more nationally or regionally applicable. A number of factors should be considered when evaluating the applicability of another jurisdiction's benchmarks. Most will have been developed using one or more of the approaches described above, but slight variations are common, so a full review should include those factors in section 3.7 as well as:

.1 Was the role of mineralogy or geochemistry considered in the development of naturally occurring substances such as metals? *(e.g., the spatial variation in metal concentrations caused by natural factors, i.e. unrelated to industrial activities, can be considerable)*;

.2 What ecological considerations played a role in the derivation of the level*? (e.g., how sensitive were the test organisms used and what is their relevance to the location and environmental conditions under consideration?)*;

.3 What types and sources of anthropogenic pollution were important in the area for which the level was designed? *(e.g., the relevance of contaminants can be expected to vary across regions)*; and

.4 What physical oceanographic conditions dominate the area for which the guideline was designed? (*e.g., the extent of exposure to any hazards will be related to the size of the area over which the material is dispersed and the concentrations of relevant substances*).

3.21 When considering these factors, it is important to determine the relevance of each to differences between local conditions and conditions in the jurisdiction where the benchmarks were developed. Careful thought and analysis should be undertaken to support decisions about whether and how to make use of levels developed by other jurisdictions. Technical expertise in sediment geochemistry, toxicology, statistics, as well as other disciplines will be needed to guide such decision-making.

3D. STRENGTHS AND WEAKNESSES OF VARIOUS DERIVATION APPROACHES

3.22 The use of any physical, chemical and biological benchmark to build Action Levels, whether reference or effects-based, will involve varying degrees of uncertainty. Stated plainly, no level is perfect. Credible use of each benchmark and resulting Action Level will require giving consideration to the uncertainties associated with its derivation and use to reach conclusions about the presence or absence of risks:

.1 The physical/chemical analyses needed to apply physical/chemical benchmarks can be relatively straightforward to conduct and the requirements for conducting these analyses can be readily met in many countries. These analyses are also amenable to inter-laboratory calibration and standardized quality assurance/quality control (QA/QC). However, using the results of such analyses as a basis for reaching conclusions about the potential for adverse effects involves uncertainties related to the fact that these metrics are not themselves a measure of effect but in some way related to the potential for an effect. For example, measurement of the presence and concentration of a specific contaminant can be related to a specific effect through empirical or mechanistic means. The role of unmeasured contaminants also presents a source of uncertainty in the application of chemical levels. It is not possible to analytically quantify the concentration of every chemical constituent in a sediment sample. This is a source of uncertainty when chemical characteristics are used to set Action Levels, whether the derivation of the benchmarks is based on empirical or mechanistic methods.

2. Biological benchmarks are intended to be indicators of potential impacts and provide for integrating across exposures and effects (i.e., the combined effects of mixtures of chemicals and/or effects of chemicals not measured/determined). Uncertainties related to the use of such levels include the relationship between exposure conditions in the laboratory, for the reference condition, and at the disposal site; inter-specific variation in tolerance or sensitivity to contaminants; and the relationship between effects on individual organisms, populations, and communities (PIANC 2006). The expertise and facilities required to set Action Levels based on biological characteristics may not be as commonly available as is the case for more routine physical and chemical analyses.

4 APPROACHES TO SETTING ACTION LEVELS

4.1 It is important to recognize the difference between Action Levels, which are intended to represent regulatory decision points, and the upper and lower benchmarks set for the individual characteristics. Some of the benchmarks, depending upon the derivation approaches selected, will have been developed to accomplish screening or to monitor environmental trends detached from a specific regulatory intent. The Upper Action Level is intended to provide a definitive decision point where the dredged material may not be dumped except in cases where control measures can be taken to manage the risks at acceptable levels. The Lower Action Level is that level below which a dredged material would be expected to have little potential to produce an adverse effect in the marine environment and for this reason can be disposed without the need for special management controls.

4.2 Action Levels should meet a number of general criteria including:

.1 they should be meaningful for the dredged material characteristics and valued resources at issue;

.2 should focus on characteristics caused by anthropogenic impacts;

.3 should be sufficiently protective to minimize the probability of false negatives at the Lower Level, i.e. reaching a conclusion that the dredged material poses no risk when in fact it does; and

.4 they should be sufficiently accurate to minimize the probability of false positives at the Upper Action Level, i.e. reaching the conclusion that a dredged material poses a risk when, in fact, it does not.

4.3 Once benchmarks are established for the characteristics on the Action List, their use to construct Action Levels must be defined.

4.4 Figure 2 below shows some of the different types of information that can be used in a complementary and additive manner to set benchmarks for the characteristics on the Action List. As more complementary information from different benchmarks is integrated into the decision rule for the Action Level, the confidence in the decision should improve as the weight of evidence accumulates to support a specific conclusion.

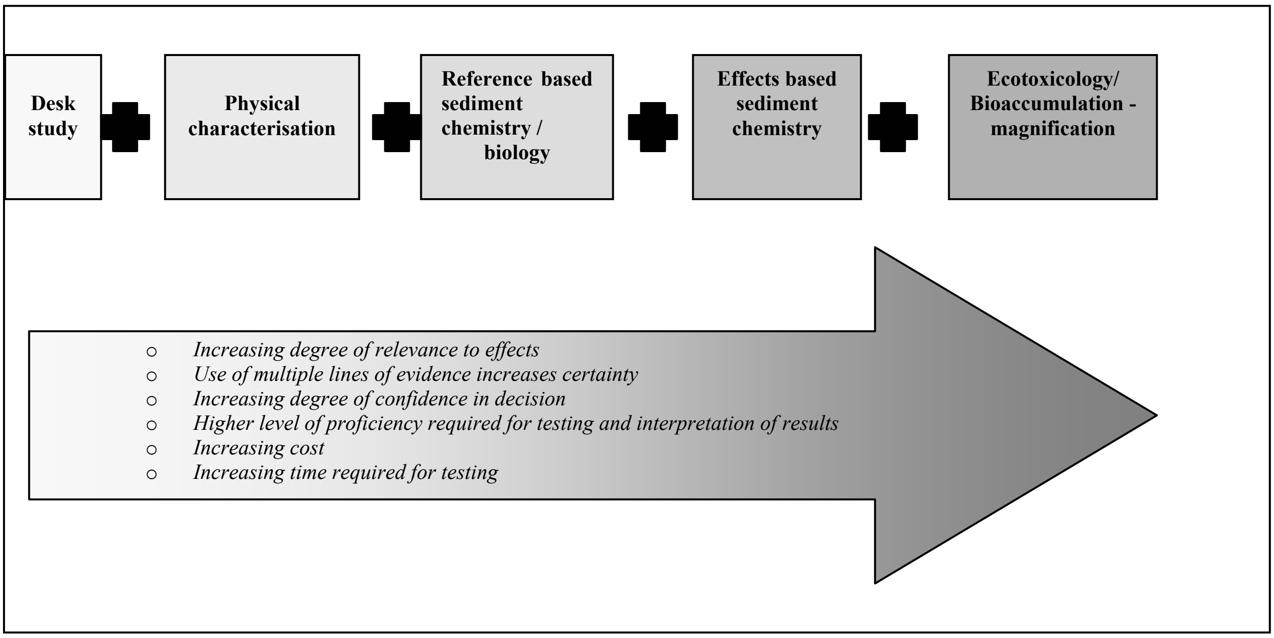

Figure 2. Action Level complexity vs. certainty

4.5 Action Levels can be integrated within a decision framework in a number of possible ways. It should be noted that there is no perfect mechanism for selecting Action Levels. Jurisdictions are encouraged to set Levels appropriate to their capacity to implement and administer them. A gradual increase in the level of sophistication and coverage can be achieved as experience and capacity increase.

4.6 The simple approach (see Table 1) is a simple pass/fail method. The more complex approaches (see Tables 2 and 3) include elements of weight-of-evidence approaches.

4A. PASS/FAIL APPROACH TO INTERPRETING ACTION LEVELS

4.7 Using Upper and Lower Action Levels as part of a simple pass/fail approach generally involves setting strict limits based on upper and lower benchmarks for each characteristic in the Action List. For example, if the dredged material is below all Lower benchmarks the material would be considered to pose a negligible risk to the marine environment and human health and is below the Lower Action Level (see Table 1 for an example). Exceeding any one upper benchmark would result in the material being classified as unsuitable for sea disposal without management. If the material is lower than all upper benchmarks, but exceeds any one lower benchmark, then additional assessment would be required to determine whether the material presents a negligible or significant risk – that is, it would fall between the Upper and Lower Action Levels.

4.8 Simple pass/fail Action Levels offer the advantage of enabling clear, transparent and repeatable decisions that can be implemented with relatively little training and experience by a permitting authority. The regulated community also has clarity about what is considered acceptable and can plan accordingly. The simplicity and standardization also mean that the Action Levels may not always fully describe the potential for impacts or adverse effects, or may be overprotective or under-protective in some cases because the real environment is complex, variable and uncertain. The simple pass/fail approach can lead to cases where increased costs are incurred by the dredger when dredged material posing a minimal risk to the environment is restricted from sea disposal or to cases where the approach fails to provide protection because all the available and relevant information on a material was not considered. Using a simple pass/fail approach would give no consideration to the magnitude of an exceedance; the case where a dredged material exceeds an Action Level by 0.01% would have an equivalent outcome to the case where the Action Level is exceeded by 1000%.

4B. WEIGHT-OF-EVIDENCE APPROACH

4.9 Another means of implementing Action Levels is through the use of a weight-of-evidence approach. This method can be more complex to implement, and requires a substantial amount of professional judgment in reaching decisions, but has the advantage of potentially integrating all the measured characteristics into the final decision.

4.10 In a weight-of-evidence approach, interpretation rules would be based on results from a number of "lines of evidence" (i.e. physical, biological and chemical data) including, if appropriate, consideration of relevant characteristics of a proposed disposal site. In such an approach, no one single benchmark would ordinarily determine that an Upper Action Level is exceeded (unless that single measurement was, in itself, of sufficient "weight" to indicate substantial concern for adverse effects (see Tables 2 and 3)). Rather, an Action Level would be considered to have been exceeded when a combination of several pre-determined "interpretation criteria" are met. The information being combined to reach this determination would consider the likelihood that valued resources would be exposed to characteristics of the material with the potential to cause harm and the nature and likelihood of the effects that could be produced. Such assessments would consider relevant characteristics of the material, including physical information, chemical properties of the sediment, as well as its biological attributes (e.g., a combination of contaminant concentrations, contaminant loads, toxicity tests, biomarkers, measures of bioaccumulation, etc.). Such information would be considered in light of the suitability and capacity of the disposal site, the patterns of sediment movement from the disposal site and the location of areas of conservation status or importance from the perspective of fisheries resources or fisheries activity.

4.11 In short, a number of the Action List chemical, biological and physical characteristics and their benchmarks can be considered in parallel, such that each characteristic provides part of the information required to determine whether a particular material is above or below an Action Level. For example, physical characteristics (e.g., grain size) could be used to predict the likelihood of a sediment retaining contaminants, chemical analysis provides information about rates of exposure that organisms could experience, and biological tests provide measures of the bioavailability and toxicity of the sediment. It must be stressed, however, that while a weight-of-evidence approach can utilize many additional lines of evidence and provide a means of reducing uncertainty concerning conclusions about environmental risks of disposal at sea, it does not eliminate uncertainty.

4.12 Providing that the necessary levels of professional expertise and experience are available, a weight-of-evidence approach can allow more informed and case-sensitive decision-making than the application of more simplified pass/fail Action Levels, and can also better serve the consideration of alternative management options. Nevertheless, the robustness of any weight-of-evidence-based decision to allow dumping, despite one or more upper benchmarks being exceeded, will depend on both a solid understanding of the methods used to derive the individual benchmarks and the ability to justify and defend the professional judgments made in coming to the decision that an Action Level is exceeded or is not exceeded.

4.13 A number of approaches have been applied to conducting sediment assessments using a weight-of-evidence approach. Describing the details of these approaches lies outside the scope of this document. However, those interested in learning more about such approaches are referred to other resources where these specifics are described and discussed (Burton *et al*., 2002; Adams *et al*., 2005; Bridges *et al*., 2005, PIANC 2006).

5 POSSIBLE FORMATS FOR UPPER AND LOWER ACTION LEVELS

5.1 The Upper or Lower Action Levels can be formulated in a number of ways. Below are several possible formats for formulating Action Levels. All these examples are for simple Pass/Fail Action Levels.

5A. LOWER ACTION LEVELS

5.2 Materials *below the relevant lower levels should be considered to be of little environmental concern in relation to dumping.* The purpose of establishing Lower Action Levels is to efficiently screen out materials that pose a negligible risk to the marine environment and human health. Lower Action Levels can be established using physical, chemical or biological data by the approaches discussed above.

5.3 Lower Action Levels can be simple and based on physical characteristics, e.g., "the material comprises greater than X% rock and cobble and was dredged from areas distant/remote from known sources of contamination".

5.4 Action Levels and their potential formulations are presented in the following sections. These examples are merely intended to represent possible types of Lower Action Level formats. They do not constitute either a definitive or a complete list of possibilities.

Lower Action Levels that are formulated as a fixed number

5.5 Lower Action Levels can be based on a set of numbers that are fixed or pre-defined. These fixed (pre-defined) limits should be developed so that they can take into account site-specific levels and/or natural background conditions.

5.6 For chemical characteristics, Lower Action Levels can be formulated as:

.1 "*The Lower Action Level is not exceeded if the mean concentrations in sediment of all the following are below the lower benchmarks: Cd X.X mg/kg, Hg X.X mg/kg, X.X ug/kg PAH, XX ug/kgPCB*;"

5.7 For biological response characteristics (e.g., toxicity), the Lower Action Levels could be formulated as:

.1 "*The Lower Action Level is not exceeded and disposal is not likely a concern if x% of a sensitive marine species used in an assay survive.*"

Lower Action Levels that depend on comparison with a reference site or value

5.8 Rather than a pre-defined set of numbers, Lower Action Levels may also be formulated as a comparison to a reference value.

5.9 These are formulated such that the Lower Action Level is not exceeded as long as each measured characteristic (chemical or biological metric) is no different than or below that of the designated reference site (or average ambient or background concentrations):

.1 "*The Lower Action Level is not exceeded and disposal is not likely a concern if chemical concentrations in the dredged material are not significantly different than concentrations in an appropriate reference sediment.*"

.2 "*The Lower Action Level is not exceeded and disposal is not likely a concern if the % survival of a sensitive marine species is:*

(a) less than 20% different than the reference, and
(b) not significantly different from the reference."

5B. UPPER ACTION LEVELS

5.10 "*The Upper Level should be set so as to avoid acute or chronic effects on human health or on sensitive marine organisms representative of the marine ecosystem.*" Upper Action Levels are intended to indicate the point above which materials will pose an unacceptable risk to the marine environment and human health. Materials that exceed Upper Action Levels cannot be disposed of at sea without the application of management techniques and processes.

5.11 It should be noted that the following section is not intended to recommend formats, but merely to serve as examples of the types of format that Upper Action Levels might take.

Upper Action Levels that are formulated as a fixed number

5.12 Fixed Number Upper Action Levels can be formulated as:

.1 "*The Upper Action Level is exceeded and disposal is not permitted if the sediment concentration exceeds any effects-based upper chemical benchmark on the National Action List e.g., Cd X.X mg/kg, Hg X.X mg/kg, X.X ug/kg PAH, XX ug/kg PCB*"; or

.2 "The *Upper Action Level is exceeded and disposal is not permitted if the percent survival in a 10-day amphipod toxicity test is less than 70%.*"

Upper Action Levels that depend on comparison with a reference site or condition

5.13 When using chemical characteristics, Upper Action Levels can be constructed such that the Upper Action Level is exceeded when a measured characteristic is above that of a known reference condition. The most commonly used reference condition refers to a site or sediment that has not been significantly impacted by past dredging activities or other sources of contaminants. However, when being applied to an Upper Action Level, a reference condition could be used that represents a limit beyond which conditions would be considered degraded and causally linked to adverse effects. Used in this manner, a chemical reference condition for an Upper Action Level could be derived as a specific percentile from a distribution of chemical survey data from coastal and near-shore sediments, in urban as well as other environments.

5.14 When using biological response characteristics, reference-based Upper Action Levels can be constructed as follows:

.1 "*The Upper Action Level is exceeded and disposal is not permitted if the percent survival in a 10-day amphipod toxicity test is statistically lower in the dredged material, compared to the reference sediment, and more than 20% different.*"

5C. BETWEEN THE UPPER AND LOWER ACTION LEVELS

5.15 In the case when a dredged material falls between the Upper and Lower Action Levels, additional information would be required before a decision permitting disposal could be made. This information would be produced through further assessment. Alternatively, a decision could be made to seek a disposal option other than sea disposal, for example, in circumstances where the costs associated with additional assessment are expected to be larger than the differential between sea disposal and the next, least costly option.

5.16 The nature of follow-on assessments that could be conducted in cases falling between Upper and Lower Action Levels will depend on the nature of the existing results. At this stage of the process, the purpose of additional assessment would be to address specific sources of uncertainty that prevent classifying the sediment as either suitable or unsuitable for sea disposal.

5.17 In some cases additional sampling may be required to accomplish further assessment. Additional sampling could be undertaken to increase spatial coverage (i.e. a larger number of samples per unit area), to increase the depth of coring to examine the vertical distribution of characteristics, to expand the list to chemicals being analysed, etc. This may show that some discrete areas within the dredging zone may be suitable for disposal at sea while others are not. Alternatively, an investigation of the source(s) of contaminants could also be undertaken. Additional bioassays with different endpoints could be used to better determine the effects associated with identified contaminants. Again, the specific nature of follow-on assessments will depend on the specific features of the site and the results of the initial assessment.

6 CONCLUSIONS

6.1 An Action List is a set of chemicals of concern, biological responses of concern, or other characteristics that can be used for screening dredged material for its potential effects on human health and the marine environment. Action Levels establish thresholds that provide decision points that determine whether sediments can be disposed of at sea.

6.2 There are a number of approaches for selecting Action List characteristics and to derive levels for dredged material assessment. Jurisdictions will need to be clear on the level of protection they require and on their ability and capacity to administer a permit system using Action Levels to facilitate transparent and consistent decision-making.

6.3 There is no universal mechanism for the selection of Action Lists and the development of Action Levels.

6.4 The process begins with an identification of the chemical, biological, or physical characteristics that will make up the Action List. Next, benchmarks must be set for each characteristic on the Action List. Benchmarks are often developed using a reference-based approach (comparing to background or ambient conditions) or an effects-based approach (based on knowledge or direct observation of the effects of exposure). Finally, Action Levels are set by integrating the relevant characteristics and benchmarks to form a decision rule. This can be as simple as a pass/fail based on a single benchmark, or it can be more complex such as combining multiple lines of evidence in a weight-of-evidence approach.

6.5 An Upper Action Level can be created above which there would be concern due to increased potential for effects on human health and the marine environment, and if desired, a Lower Action Level can be created below which there would be little concern.

6.6 Appendix 1 goes into further detail providing options that can be used in setting benchmarks and provides some country examples.

6.7 Readers are encouraged to provide additional examples and it is expected that this document will evolve over time, as greater experience with the development and application of Action Levels is obtained.

7 REFERENCES

Adams, W., Green, A., Ahlf, W., Brown, S., Burton G.A., Chadwick, B., Crane, M., Gouguet, R., Ho, K.T., Hogstrand, C., Reynoldson, T., Ringwodd, A., Savitz, J., Sibley, P. (2005). "Using sediment assessment Tools and a weight-of-evidence approach." In *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments*. R. Wenning, C. Ingersoll, G. Batley, D. Moore, editors. SETAC Press, Pensacola, FL.

Ankley, G. T., Berry, W. J., DiToro, D. M., Hansen, D. J., Hoke, R. A., Mount, D. R., Reiley, M. C., Swartz, R. C., Zarba, C. S. (1996). "Use of equilibrium partitioning to establish sediment quality criteria for non-ionic chemicals: A reply to Iannuzzi *et al*." *Environ. Toxicol. Chem.* 15(7): 1019-1024.

Apitz, S., Chadwick, D., Germano, J., Gieskes, J., Kirtay, V., Maa, G., Montgomery, M., Paulsen, R., Smith, C., Ziebis, W. (2003). "In situ evaluation of contaminant behaviour in nearshore sediments: Approach for the pathway ranking for in situ sediment management (PRISM) programme." The Society for Underwater Technology (SUT), Aberdeen, UK.

Bridges, T., Berry, W., Ells, S., Ireland, S., Maher, E., Menzie, C., Porebski, L., Stronkhorst, J., Dorn. P. (2005). "A risk-based assessment framework for contaminated sediments." In *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments*. R. Wenning, C. Ingersoll, G. Batley, D. Moore, editors. SETAC Press, Pensacola, FL.

Buchman, M. F. (1999). *NOAA Screening Quick Reference Tables*. Seattle WA, Coastal Protection and Restoration Division, National Oceanic and Atmospheric Administration. Report nr NOAA HAZMAT Report 99-1.

Burton, G. A., Chapman, P. M., and Smith, E. P. (2002). "Weight-of-evidence approaches for assessing ecosystem impairment." *Human and Ecological Risk Assessment* 8(7), 1657-1673.

Cairns, J., Jr. (1986). "The myth of the most sensitive species." *Bioscience* 36, 670-672.

Chapman, P. M. (1989). "Current Approaches to Developing Sediment Quality Criteria." *Environ. Toxicol. Chem.* 8(7): 589-599.

Clarke, J.U. and McFarland, V.A. (2000). "Uncertainty analysis for an equilibrium partitioning-based estimator of PAH bioaccumulation potential in sediments," *Environ. Toxicol. Chem.* 19(2): 360-367

Di Torro, D. M., Zarba, C. S., Hansen, W. J., Berry, W. J., Swartz, R.C., Cowan, C. E., Pavlou, S. P., Allen, H. E., Thomas, H. A., and Paquin, P. R. (1991). "Technical basis for establishing sediment quality criteria for non-ionic organic compounds using equilibrium partitioning." *Environ. Toxicol. Chem.* 10(12), 1541-1583.

Fairey, R., Long, E. R., Roberts, C. A., Anderson, B. S., Phillips, B. M., Hunt, J. W., Puckett, H. R., Wilson, C. J. (2001). "An evaluation of methods for calculating mean sediment quality guideline quotients as indicators of contamination and acute toxicity to amphipods by chemical mixtures." *Environ. Toxicol. Chem.* 20(10): 2276-2286.

Field, L. J., MacDonald, D. D., Norton, S. B., Ingersoll, C. G., Severn, C. G., Smorong, D., and Lindskoog, R. (2002). "Predicting amphipod toxicity from sediment chemistry using logistic regression models or evaluating sediment chemistry and toxicity data using logistic regression modeling," *Environ. Toxicol. Chem.* 21(9): 1993-2005.

GIPME (2000): *Guidance on Assessment of Sediment Quality*. IMO Publication No.439/00

IMO (2006): *Guidelines on the Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter, 1972*. IMO Publication No. 1531E

Long, E. R., Chapman, P. M. (1985). "A Sediment Quality Triad – Measures of Sediment Contamination, Toxicity and Infaunal Community Composition in Puget-Sound." *Marine Pollution Bulletin* 16(10): 405-415.

Long, E. R., Macdonald, D. D., Smith, S. L., Calder, F. D. (1995). "Incidence of Adverse Biological Effects within Ranges of Chemical Concentrations in Marine and Estuarine Sediments." *Environmental Management* 19(1): 81-97.

MacDonald, D. D., Carr, R. S., Calder, F. D., Long, E. R., Ingersoll, C. G. (1996). "Development and evaluation of sediment quality guidelines for Florida coastal waters." *Ecotoxicology* 5(4): 253-278.

MacDonald, D. D., Dipinto, L. M., Field, J., Ingersoll, C. G., Long, E. R., Swartz, R. C. (2000). "Development and evaluation of consensus-based sediment effect concentrations for polychlorinated biphenyls." *Environ. Toxicol. Chem.* 19(5): 1403-1413.

PIANC (2006): *Biological Assessment Guidance for Dredged Material*, report of Working Group 8 of the PIANC Environmental Commission. ISBN 2-87223-155-2

Swartz, R. C. (1999). "Consensus sediment quality guidelines for polycyclic aromatic hydrocarbon mixtures." *Environ. Toxicol. Chem.* 18(4): 780-787.

Swartz, R. C., Schults, D. W., Dewitt, T. H., Ditsworth, G. R., Lamberson, J. O. (1990). "Toxicity of Fluoranthene in Sediment to Marine Amphipods - a Test of the Equilibrium Partitioning Approach to Sediment Quality Criteria." *Environ. Toxicol. Chem.* 9(8): 1071-1080.

U.S. Environmental Protection Agency/U.S. Army Corps of Engineers (USEPA/USACE), (1991). *Evaluation of Dredged Material Proposed for Ocean Disposal – Testing Manual.* EPA 503/8-91-001, Washington, DC.

Webster, J. and Ridgway, I. (1994). "The Application of the Equilibrium Partitioning Approach for Establishing Sediment Quality Criteria at 2 UK Sea Disposal and Outfall Sites." *Marine Pollution Bulletin* 28(11): 653-661.

Wenning, R., Ingersoll, C., Bateley, G., and Moore, D. (2005). *Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments.* SETAC Press, Pensacola, FL. 783 p.

8 GLOSSARY

Action List is defined as a "*mechanism for screening candidate wastes and their constituents on the basis of their potential effects on human health and the marine environment*". The Action List can consist of chemicals of interest, biological responses of concern, or other characteristics that can provide insight into the potential for dredged material to cause adverse effects in the marine environment. An Action List can also be used as a trigger mechanism for further waste prevention considerations and could therefore have a role in controlling pollution at its source, in promoting cleaner technology, or in improving the efficiency of dredging to reduce the need for disposal.

Action Levels establish decision rules to identify dredged materials that may be disposed because the risk for adverse effects is low and acceptable, those that may not be disposed without management controls because the risks for adverse effects would be considered too high, or to identify cases where additional information may be required to make a sound judgement about the potential for the dredged material to cause adverse effects. If developed for the purposes of meeting the requirements of disposal at sea treaties, the Action Level will specify an Upper Level and may also specify a Lower Level. The Upper Level should be set so as to avoid acute or chronic effects on human health, or on sensitive marine organisms representative of the marine ecosystem. Below the Lower Level, there should be little concern for disposal at sea.

Ambient Conditions: The conditions observable in the vicinity of a site, e.g., a disposal site.

Background: the conditions observable in the vicinity of the site that are due to natural conditions, i.e. not due to anthropogenic activities.

Benchmark: is a point on the range of the metric (e.g., 4 mg/kg copper, 20% amphipod mortality) that is used to identify where environmental concern may be low or high for that characteristic. These can be referred to as the lower benchmark and upper benchmark.

Characteristic: is an attribute of the dredged material (e.g., copper, mercury, silt, petroleum compounds, pathogens) or a biological response to the dredged material (e.g., mortality, growth, bioaccumulation).

Metric: is a measurement that can be made on the characteristic (e.g., concentration, percent survival).

Reference site or sediment: is used as a basis for comparison to a disposal site or dredged material. The reference represents ambient conditions in the vicinity of the disposal site, absent any influence from past disposal activities. It is desirable that the reference should be substantially free of contaminants, but pristine conditions are generally not achievable.

APPENDIX 1: APPROACHES TO SETTING BENCHMARKS AND ACTION LEVELS

1A: REFERENCE-BASED APPROACH

Benchmarks for physical, chemical or biological characteristics can be set based on knowledge of background or ambient conditions in comparable areas that have not been impacted by dumping. This is a reference-based approach.

If reference-based benchmarks are to be used as Action Levels (simple approach), it is important that they be based upon some measure of potential risk. This way, resources can be applied to the highest risk sites.

Benchmarks derived from knowledge of background concentrations of substances may be termed "background-based" levels. These may be particularly relevant to the derivation of lower benchmarks for naturally occurring substances. If metals in the dredged material, for example, would not be expected to elevate metal concentrations in the receiving environmental media (e.g., sediment and water) above the range in natural background concentrations then there should be "*little environmental concern in relation to dumping*" of that dredged material.

Some dredged material-associated substances that can produce adverse effects in the marine ecosystems also occur naturally within the environment. This can be the case for both organic (e.g., hydrocarbons) and inorganic substances of regulatory interest, but it is particularly relevant for metals. Metals exist naturally as components of minerals, as ions, and in complexes with other materials, including organic compounds. Lower benchmarks for metals can be based, for example, on the natural abundances of the subject metals in soil, crustal rock or sediment in the relevant region. The GIPME (2000) Guidelines describe the use of geochemical markers (e.g., Al, Li, Fe) to account for spatial variations in regional formations and geology.

The results of an analysis of sediment samples will depend in part on the methods used to collect, store, prepare and analyse those samples. It is important to bear this in mind when developing and applying Action Levels.

It should also be noted that if these approaches are used to set Lower Action Levels, they should trigger additional assessment when exceeded. In practice, there may be many cases in which almost all dredged material within an urbanised or industrialized region will exceed lower "background-based" benchmarks for naturally occurring substances, particularly if pre-industrial or natural background concentrations are used.

In addition, many substances, whose inputs arise solely from anthropogenic sources, have no natural background levels, including, for example, PCBs, TBT and many pesticides. The presence of compounds such as PCBs in trace quantities may be unrelated to local, or even regional activities, as these compounds can be transported through global circulation in the atmosphere and oceans. For such compounds, reference-based Benchmarks can be derived from ambient concentrations in sediment. Reference sites used as data sources for determining ambient concentrations should be remote from the influence of local waste streams and past dredged material disposal activities. GIPME (2000) and PIANC (2006) provide some guidance for selecting and using data from reference locations.

Lower benchmarks set using reference-based approaches, either alone or in combination with other approaches, are applied in many countries to make up Lower Action Levels, including Australia, Canada, Denmark, Finland, France, Germany, Hong Kong, China, Ireland, the

Netherlands, Spain, Sweden, the United Kingdom and the United States. The example below describes the use of reference-based benchmarks for this purpose in Ireland.

Country: Ireland

Overview of framework – Ireland applies a 3-phase weight-of-evidence approach to the assessment of the suitability of dredged material for disposal at sea; Phase I – screening based on a critical assessment of the available literature; Phase II – further assessment using sediment chemistry criteria; Phase III – further testing, including appropriate toxicity tests, should either lower or upper benchmarks be exceeded.

Derivation of the Action List – historic monitoring results for harbours in Ireland were used in order to prioritize those contaminants of greatest concern and to exclude from routine consideration those substances which had not been detected over the previous 10 years. The Action List can be reviewed and revised as necessary in light of new information.

Contribution of reference-based benchmarks – reference-based benchmarks are used for the derivation of lower action levels for a range of metals and organic contaminants (HCB, HCH and PCBs), in particular as part of Phase I and Phase II. Upper benchmarks are set using a chemical effects-based approach and bioassays if necessary.

Derivation of lower benchmarks – for metals and PAHs, reference-based benchmarks are set at the 95 percentile of background data, where data are available. In the case of arsenic, the ERL was used in the absence of relevant background data for Irish sediments. For several organic compounds, the 95 percentiles of ambient background levels were taken.

Definition of the Lower Action Level – if none of the lower benchmarks are exceeded then the Lower Action Level is not exceeded and the sediment quality is considered acceptable for dumping at sea (subject to dump-site and operational approval). If, however, one or more benchmarks are exceeded then, unless exceedance is marginal, further assessment of sediment quality would normally be required before any decision on acceptability for dumping or other management options can be made.

Definition of the Upper Action Level – if any of the upper benchmarks are exceeded then the Upper Action Level is considered to have been exceeded and further testing, including the use of bioassays, is used to inform the management decision.

Benchmarks

Characteristic	Units (dry weight[a])	Lower benchmark	Upper benchmark[b]
Arsenic	mg kg^{-1}	9[c]	70*
Cadmium	mg kg^{-1}	0.7	4.2
Chromium	mg kg^{-1}	120	370
Copper	mg kg^{-1}	40	110[d]
Lead	mg kg^{-1}	60	218
Mercury	mg kg^{-1}	0.2	0.7
Nickel	mg kg^{-1}	21	60
Zinc	mg kg^{-1}	160	410
Σ TBT & DBT	mg kg^{-1}	0.1	0.5
γ – HCH (Lindane)	μg kg^{-1}	0.3	1
HCB	μg kg^{-1}	0.3	1
PCB (individual congeners)	μg kg^{-1}	1	180
PCB (Σ ICES 7)	μg kg^{-1}	7	1260
PAH (Σ 16)	μg kg^{-1}	4000	
Total extractable hydrocarbons	g kg^{-1}	1.0	

[a] total sediment <2mm
[b] ERM (rounded up)
[c] ERL (rounded up) – No background Irish data available
[d] PEL as ERM considered high

Further information available from – www.marine.ie

Box 1: The use of reference-based approaches to set lower benchmarks in Ireland

1B: EFFECTS-BASED CHEMICAL ACTION LEVELS

Benchmarks for physical, chemical or biological characteristics can also be based on knowledge of effects that can be produced following exposure to dredged material. Such benchmarks, which may be termed "effects-based" benchmarks, can be based on information concerning the likelihood or magnitude for an effect.

Several approaches suitable for establishing effects-based benchmarks have been proposed in the scientific literature, including those outlined below, though in practice some are more widely used than others. Overall, the different effects-based approaches may be subdivided into empirical, mechanistic and consensus approaches.

Empirical approaches to deriving Levels

1. Co-occurrence approach. A number of benchmarks have been developed empirically by comparing large databases of sediment chemistry and effects. They were not originally intended as clean-up or remediation targets, as discharge targets, as pass/fail criteria for dredged material disposal or for any other regulatory purpose (Buchman, 1999).

These levels, set by evaluating the impact of real contaminant mixtures in real sediments, indirectly account for issues of bioavailability and sediment geochemistry and they account for the synergistic and other effects of contaminant mixtures but report it for a single contaminant. They do so only in an average way, however, and cannot account for site-specific geochemical conditions, atypical bioavailability or the effects of unusual mixtures. Furthermore, these empirical approaches do not imply causality, but simply describe the co-occurrence of contaminants and observations of toxicity.

All the benchmarks described below are based upon compilations of many literature reports in which sediment chemistry and toxicity were reported, rather than one large, coordinated study. As such, values are generated from studies with potentially different sampling and analytical methods.

The Effects Range-Low (ERLs) and Effects Range-Median (ERMs) of Long *et al.*, (1995) and the marine Threshold Effects Levels (TELs) and Probable Effects Levels (PELs) of MacDonald *et al.*, (1996) are based upon similar data compilations, but are generated using different calculations. For example:

– **ERL values** (calculated as the lower 10th percentile concentration of the available sediment toxicity data, using only data for those samples identified as toxic by original investigators) are at the low end of a range of levels at which effects were observed in the studies compiled, and therefore represent **the values at which toxicity may begin to be observed in sensitive species** – in contrast, **ERM values** (median concentrations of the compilation of toxic samples) represent **chemical concentration ranges usually associated with toxicity in marine and estuarine sediments**.

In a similar manner:

– **TEL values** (calculated as the geometric mean of the 15th percentile concentrations of the toxic effects data set and the median of the no-effect data set), in a similar way to ERLs, represent **the concentrations below which adverse effects are expected to occur only rarely**.

– while **PEL values** (geometric means of the 50% of impacted, toxic samples and the 85% of the non-impacted samples) are **levels above which adverse effects are frequently expected** (Buchman, 1999).

ERLs and TELs have been used in some countries to set lower benchmarks, while recognising that error rates associated with their use are on the order of 10-20 per cent. Adjustments can be made for cases in which the background values are higher than the ERLs or TELs.

ERMs have been used to inform the setting of upper benchmarks (e.g., Denmark, Ireland and the United Kingdom), while recognizing that their use in isolation as upper benchmarks may not always be appropriate. The example below describes their use by the United Kingdom.

Country – United Kingdom

Overview of framework – Action Levels are used as part of a weight-of-evidence approach to assessing dredged material and its suitability for disposal at sea. This considers balancing multiple lines of evidence concerning ecological assessment as an aid to decision-making. New Benchmarks have recently been proposed.

Derivation of the Action List – Reference was made to the OSPAR compilation document on Action Levels to prioritise those contaminants of greatest concern in the United Kingdom.

Contribution of empirically derived effects-based chemical benchmarks – Effects-based data have been used in conjunction with benchmarks from other jurisdictions to inform the setting of the proposed new benchmarks.

Derivation of upper benchmarks – Existing benchmarks were derived from historical information, including existing data on contaminants from ports and harbours, combined with expert scientific judgement. Ecotoxicological data, based largely on data sets from the United States, in conjunction with benchmarks from other jurisdictions were used to guide the setting of the proposed new upper benchmarks for metals and PCBs and TBT. Proposed new lower benchmarks for PCBs were also set on the basis of ecotoxicological data, however, for metals these have been derived using nominal background concentrations. The United Kingdom is currently in the process of establishing upper benchmarks for PAHs.

Definition of the Lower Action Level – In general, dredged material containing concentrations below lower benchmarks are of no concern and are unlikely to influence the licensing decision. Dredged material with contaminant levels between lower and upper benchmarks require further consideration and testing before a decision can be made.

Definition of the Upper Action Level – Dredged material with one or more contaminant levels above upper benchmarks would generally be considered unsuitable for sea disposal.

Benchmarks

Characteristic	Existing lower benchmark* mg.kg-1 (ppm)	Existing upper benchmark* mg.kg-1 (ppm)	Proposed lower benchmark mg.kg-1 (ppm) (dry weight)	Proposed upper benchmark mg.kg-1 (ppm) (dry weight)
As	20	50-100	20	70
Cd	0.4	5	0.4	4
Cr	40	400	50	370
Cu	40	400	30	300
Hg	0.3	3	0.25	1.5
Ni	20	200	30	150
Pb	50	500	50	400
Zn	130	800	130	600
Tributyltin	0.1	1.0	0.1	0.5
PCBs	0.02	0.2	0.02	0.18
PAHs				
Acenaphthene			0.1	
Acenaphthylene			0.1	
Anthracene			0.1	
Fluorene			0.1	
Naphthalene			0.1	
Phenanthrene			0.1	
Benzo[a]anthracene			0.1	
Benzo[b]fluoranthene			0.1	
Benzo[k]fluoranthene			0.1	
Benzo[g]perylene			0.1	
Benzo[a]pyrene			0.1	
Benzo[g,h,i]perylene			0.1	
Dibenzo[a,h]anthracene			0.01	
Chrysene			0.1	
Fluoranthene			0.1	
Pyrene			0.1	
Indeno(1,2,3cd)pyrene			0.1	
Total hydrocarbons	100		100	

Further information available from – www.cefas.co.uk

*** The United Kingdom refers to upper and lower benchmarks as Action Level 1 and Action Level 2**

Box 2: The use of empirically derived effects-based chemical benchmarks in the UK

Apparent Effect Thresholds (AETs) relate chemical concentrations in sediments to synoptic biological indicators of injury (i.e. sediment bioassays or diminished benthic infaunal abundance). Individual AETs represent the concentrations observed in the highest *non-toxic* samples. As such, they represent the concentrations above which adverse biological impacts would *always* be expected by *that* biological indicator due to exposure to that contaminant alone (in the data set used). It should be noted that adverse impacts could also occur at levels below the AET. AET values were developed for use in Puget Sound (Washington DC, United States) and are not easily compared directly to other benchmarks based on single-chemical models and broader data sources.

2. **Triad Approach**. This involves the concurrent collection of sediment chemistry, benthic community, and sediment toxicity samples from field stations representing the range of regional sediment contamination (Long & Chapman 1985). The data from these samples are then evaluated using statistical approaches that establish levels at which biological effects (such as changes in the benthic community or some toxicity response) might be expected to occur for the various contaminants found in the sediments. Thus, benchmarks are *inferred* from the cumulative assessment of data.

3. **Spiking approach**. Spiking involves the deliberate introduction of a range of concentrations of a selected contaminant into uncontaminated sediment samples, to which test organisms can then be exposed in order to estimate concentrations that cause toxic effects. Such an approach, though more commonly used to establish water quality criteria, has been used by Canada, for example, as one component in the development of sediment quality guidelines used for assessment. A number of caveats apply, including that contaminant bioavailability may differ from a natural situation and that the mixing process may change equilibrium characteristics of the sediment. For example, available literature suggests that significant ageing (on the order of 6 months to a year for some substances) may be necessary to give realistic results and that the time to reach equilibrium can be different for each contaminant of interest. Data on spiked sediments are not widely available and those interested in using the approach would likely need to invest in research and field validation.

4. **Quotient Approach**. Recent work has focused on developing quotient methods for applying empirical benchmarks that are derived by summing the "toxic" contributions of a number of contaminants of concern (Wenning *et al*. 2002). Regression analysis that permits considering effects from several contaminants at once along a continuum of concentrations has also been developed (Field *et al*. 2002).

Mechanistically Derived Levels

Mechanistic approaches can be used in the development of benchmarks through the use of theoretical relationships based on knowledge of mechanics of action. **Equilibrium Partitioning (EqP)**, in which sediment benchmarks are established using water quality criteria as a starting point (Ankley *et al.* 1996, Chapman 1989, Swartz *et al.* 1990, Webster and Ridgway 1994), is an example. This approach involves calculating the sediment concentration that would be necessary to create a toxic water quality level when the sediment is assumed to be in equilibrium with the pore water contained within it.

This assumes that the distribution of contaminants among different compartments in the sediment matrix (i.e. solids and interstitial water) is predictable from physicochemical properties and that it is the interstitial water value that defines the contaminant risk.

In the EqP approach, benchmarks are calculated using water quality criteria, usually the final chronic values or equivalent criteria, in conjunction with sediment/water partition coefficients for the specific contaminants. The final chronic value is derived from the species mean chronic values that have been calculated using published toxicity data. The EqP approach provides a theoretical basis for identifying chronic effects thresholds for contaminants when they occur alone in sediments. However, this approach does not address the potential synergistic effects caused by contaminant mixtures and neglects a number of potential pathways in which organisms can be exposed to sediment contaminants. These values are probably of greatest value for the examination of the probability of toxicity of a single non-ionic organic, rather than for contaminant mixtures (Wenning *et al*., 2005).

The example below describes the application of EqPs in Belgium.

Country: Belgium

Overview of framework – The assessment of dredged material is based upon quantitative action levels that must be manageable, scientifically founded and realistic in practice. Chemical analyses are followed by bioassays, if necessary, to aid in the decision-making process.

Derivation of the Action List – Selected to fulfil mandate under OSPAR

Contribution of mechanistically derived benchmarks – Equilibrium Partitioning was used, among other methods, to set Action Levels.

Derivation of benchmarks – Equilibrium partition coefficients were derived from a combination of sediment analysis and data available from published literature and used to determine 'safe concentrations' as benchmarks. These values were subsequently adjusted for local conditions and, if deemed necessary, also for bioaccumulation and biomagnification potential.

Definition of the Lower Action Level – if all analyses yield values below lower benchmarks, then the material may be dumped at sea. If any lower benchmark is exceeded, additional sampling and analysis are required. If these results confirm the original, then bioassays are required to aid the final decision.

Definition of the Upper Action Level – if the analyses yield values exceeding any three of the upper benchmarks, the Upper Action Level is exceeded and dumping at sea is not permitted.

Benchmarks

Characteristic (dry weight for whole sediment)	Lower benchmark	Upper benchmark
Hg	0,3 ppm	1.5 ppm
Cd	2,5 ppm	7 ppm
Pb	70 ppm	350 ppm
Zn	160 ppm	500 ppm
Ni	70 ppm	280ppm
As	20 ppm	100 ppm
Cr	60 ppm	220ppm
Cu	20 ppm	100 ppm
TBT	3 ppb	7 ppb
Mineral oil	14 mg/g_{OC}	36 mg/g_{OC}
PAKs	70 µg/g_{OC}	180 µg/g_{OC}
PCBs	2 µg/g_{OC}	2 µg/g_{OC}

Further information available from – www.mumm.ac.be

Note: Belgium refers to upper and lower benchmarks as Action Level 1 and Action Level 2.

Box 3: Use of mechanistically derived effect-based chemical benchmarks in Belgium

In undisturbed anoxic sediments, the chemistry of many trace metals is dominated by reactions with sulfide. Di Toro *et al.* (1991) developed an EqP-type approach to trace metals in anoxic sediments, known as the acid volatile sulfide (AVS) model. According to this model, the iron in sedimentary iron monosulfide, FeS(s) (defined as AVS), can be exchanged with a divalent trace metal to form a solid sulfide less soluble than FeS(s), thus releasing equivalent amounts of iron into pore waters. As long as the trace metal concentration in sediments is less than the concentration of AVS, free-metal ion activity in the porewater is maintained at very low levels and the sediment is not toxic. When the metal concentration added becomes greater than that of AVS, free-metal ion activity increases sharply in pore waters and the sediment can become toxic. The measurements needed to apply this model are the AVS concentration and the sum of the molar concentrations of divalent trace metals forming less soluble sulfides than iron, referred to as simultaneously extracted metals (SEM).

While this approach may be useful for predicting metal availability in static, reduced sediments, it has a number of problems that limit its usefulness. For example, while the bulk of undisturbed sediments may be anoxic, disturbance of sediments by shipping activity, storms, the activities of benthic organisms (bioturbation) and, in particular, by dredging and disposal, markedly affect redox conditions. It has also been observed that metals often flux out of sediments at higher rates than would be predicted by porewater gradients, possibly due to bioirrigation and other sources (*e.g.*, Apitz and Chadwick, 2003).

Consensus SQGs

Consensus levels, currently under review by west coast regions in the United States and Canada, are the mean value of the benchmarks derived from a variety of methods (Swartz 1999, MacDonald, 1999). Since benchmarks are often given for low, intermediate and high effects levels, consensus levels have been proposed for Threshold, Median and Extreme Effects Concentrations (TEC, MEC, EEC), or for Threshold and Probable Effects Concentrations (TEC, PEC). The consensus benchmarks for mixtures of PAHs (Swartz 1999) and PCBs (MacDonald *et al.* 2000) were used in the successful formulation of SQGQ1 by Fairey *et al.* (2001). The motivation behind the consensus method is not simply to make a list of available benchmarks for a particular chemical and then to calculate the average. The premise behind the consensus approach is that if different methods for deriving benchmarks result in quantitatively similar concentrations, then the validity of the common result is greatly enhanced. Only then is the calculation of a consensus guideline justified. Even if consensus of different benchmarks is not evident for a particular chemical, the method is expected to serve the function of identifying potential errors with one or more benchmarks. In the calculation of consensus-based benchmarks for total PAHs, Swartz (1999) reported that many of the benchmarks derived with a similar narrative intent, but derived using different empirical and theoretical approaches, resulted in very similar concentrations and that this similarity probably was not coincidental.

1C: BIOLOGICAL EFFECTS-BASED ACTION LEVELS

Plants and animals will come in contact with dredged material-associated substances through one of three primary exposure pathways: 1) through contact with bedded sediment particles; 2) through contact with water; and 3) through contact with contaminants through bioaccumulation and trophic transfer within a food chain.

Risk, i.e. the likelihood for adverse effects, is a function of the rates that organisms are exposed to contaminants and the relationship between rates of exposure and adverse biological effects. Biological tests may provide either measures of effect or exposure, or in some cases both. Laboratory and field-based methods are available for generating information about the potential for **effects** and **exposure**:

.1 Tests that provide measures of **effect** provide insight into risk by providing information about the toxicity of a substance and adverse responses in organisms exposed to the material. Laboratory-based toxicity tests are commonly used in dredged material evaluations. Methods for measuring effects in the field are also appropriate for some case-specific applications. Such field methods include measurements of benthic community structure or observation and measurement of effects on individual organisms (e.g., cancer in resident fish).

2. Other biological tests provide information about **exposure** conditions including measures of the bioavailability of the contaminants present in dredged material, or the concentration or dose received by the receptor. Bioaccumulation tests that

measure the movement of contaminants into the tissues of the test organism are the most commonly applied biological tests for collecting information about exposure. For an effects-based Action Level, such a measure of exposure might be compared to limits to ensure the protection of humans exposed through consuming fish or shellfish, or similar limits to ensure protection of wildlife.

This Appendix describes three approaches for establishing Biological Effects-Based Action Levels. These approaches are: Solid-phase toxicity tests, Water-column Toxicity Tests, and Bioaccumulation Tests.

Solid-Phase Toxicity Tests (Bioassays in whole sediments)

1 Effects-based Action Levels for dredged material can be based upon direct measures of toxicity using solid phase toxicity tests. These tests involve exposing test organisms to bedded sediments for a defined period and measuring the responses of those organisms (e.g., rates of survival, growth, reproduction) at the conclusion of the test. To ensure that test results will be protective with respect to the exposure conditions expected at a management site, the species used in such tests should be selected based on their close behavioural association with the sediment and their sensitivity to contaminants. Organisms that live in and/or ingest sediments (e.g., infaunal invertebrates) are expected to have high exposure to sediment-associated contaminants due to their intimate contact with sediment particles and pore water. Tests using infaunal amphipods, polychaetes, bivalve molluscs, urchins, and other taxa have been developed and commonly applied to assess dredged material (PIANC 2006).

2 Recognized differences among candidate test species, in terms of their behaviour within sediments, have resulted in a broad consensus on the need for testing using multiple species. Some taxa actively burrow through sediments while others live within semi-permanent burrows or even tubes they construct with mucus and sediment particles. Some species actively ingest sediment particles while others rely more on removing particles from suspensions in the overlying water. Species with these different behavioural characteristics will experience different exposures to contaminants adsorbed to sediment particles or dissolved within pore waters. Selecting a battery of tests that represents this diversity of behaviour will provide for more confidence that the assessment will be protective of exposure conditions at the management site. Efforts should also be made to ensure that the species used in such tests are sensitive to contaminants, i.e. they respond to the presence of contaminants. Taxa differ in their sensitivity to contaminants with respect to one another and among contaminants. Even though limited understanding of this variation in sensitivity currently prevents tailoring assessments for specific mixtures of contaminants or benthic communities at disposal sites, it must be acknowledged that using multiple tests with different species is a precautionary approach for assessing sediments (Cairns 1986).

Water-column Toxicity Tests

3 Substances may be released from a dredged material into the water column during or after disposal through diffusion, leaching, or other mechanisms. Effects-based Action Levels to assess the potential for effects being caused through these means of exposure can be developed using water-column toxicity tests. These tests generally make use of planktonic species including algae, copepods, and other arthropods (e.g., cladocerans), as well as larval molluscs, echinoderms, and fish. Such tests, described in detail in USEPA/USACE (1991), are commonly conducted using a dilution series of sediment-water mixtures (e.g. elutriates) in evaluations of dredged material. Tests are conducted by exposing test organisms to water extracts of dredged

material that are intended to represent the range of substance concentrations organisms would be exposed to in the field.

Bioaccumulation Tests

4 Addressing questions concerning the potential for contaminants in dredged materials to move into the food chain and produce effects in organisms, above and beyond the borders of a disposal site, begins with assessing bioaccumulation potential. Bioaccumulation, in this case, refers to the movement of contaminants from the dredged material into the tissues of exposed organisms. It is important to recognize that bioaccumulation tests provide a measurement of exposure rather than effect. Bioaccumulation of a compound will not always result in an adverse effect on the organism accumulating the compound. In the case of essential elements (e.g., zinc and copper), a certain amount of accumulation is required to support normal physiological function. In general, adverse effects from any contaminant will only be manifest after the concentration exceeds a specific tolerance level or toxicological threshold. For this reason careful attention must be given to interpreting bioaccumulation data.

5 Laboratory bioaccumulation tests are generally conducted by exposing the test organisms to the test material (e.g., dredged material) under controlled conditions and recovering the animals at the end of the exposure to measure the concentration of contaminants of concern in the tissues of the test organisms. Test organisms used in bioaccumulation tests are generally selected on the basis of their relative tolerance to contaminants (i.e. they survive the exposure) and their body size, such that there is sufficient tissue recovered at the end of the exposure for chemical analysis.

6 Because of the expense and time involved in conducting bioaccumulation tests, alternative approaches have been developed for assessing bioaccumulation potential. One of these approaches is called Thermodynamic Bioaccumulation Potential (TBP). This approach makes use of the principle of equilibrium partitioning of non-polar organic chemicals as a means of estimating the amount of chemical that will partition to the lipid phase within the organism from the organic carbon phase of the sediment at equilibrium (Clarke and McFarland, 2000).

7 Interpreting the consequences of bioaccumulation test data commonly involves the use of mathematical models and risk calculations in order to apply the data to a particular target of protection within the food web (e.g., a fish-eating bird or particular human population). These and other aspects of using biological tests to evaluate dredged material are discussed at length in PIANC (2006).

Country: Canada

Overview of framework – The Canadian framework adopts both chemical and biological Action Levels to determine sediment is suitable for ocean disposal. The framework employs, in tier 1, a chemical Lower Action Level (LAL) that can adjusted for background concentration and exposure and, in tier 2, a biological Upper Action Level (UAL) that conside lethal and sub-lethal toxicity and bioaccumulation. This approach focuses resources on materials that pose greater risk, or a associated with greater uncertainty.

Derivation of the Action List – the Action List comprises different characteristics for consideration at lower and upper action levels. Characteristics considered for Lower action levels are chemical contaminants, selected on the basis of the London Convention Black List and supplemented on a case-by-case basis by site-specific characteristics of the area, while those considered for upper action levels are biological endpoints.

Contribution of biological-effects based benchmarks – biological effects testing is used to set the Upper Action Level based on toxicity, persistence and bioaccumulation.

Derivation of upper benchmarks – upper benchmarks are set using a battery of standard bioassays, including metrics of survival, growth, bioaccumulation and metabolic response. In contrast, lower benchmarks were set using a combination of reference-based and effects-based chemical approaches.

Definition of Lower Action Levels – if all analyses yield values below lower benchmarks, then the material may be dumped at sea. If, however, any regulated lower benchmark is exceeded, the Lower Action Levels is exceeded and further assessment is required.

Definition of Upper Action Levels – the Upper Action Level is exceeded if the lethal bioassay is failed or if any two biological tests are failed. Open water disposal is then not permitted without the use of management techniques or processes.

Benchmarks

Characteristic	Units (dry weight[a])	Lower benchmark	Upper benchmark
Arsenic	mg kg^{-1}	7.2	
Cadmium	mg kg^{-1}	0.6*	
Chromium	mg kg^{-1}	52.3	
Copper	mg kg^{-1}	18.7	
Lead	mg kg^{-1}	30.2	
Mercury	mg kg^{-1}	0.75*	
Nickel	mg kg^{-1}	Na	
Zinc	mg kg^{-1}	124	
PCB Total	µg kg^{-1}	100*	
Total PAH (Σ 16)	µg kg^{-1}	2500*	
Amphipod survival			A decrease in survival of at least 20% is observed between the test sediment and a clean sediment used as a reference
Photoluminescent bacterial metabolic effect			Five minute IC50 is less than 1000 mg/kg.
Echinoid reproduction and development			A decrease in fertilization of at least 25% is observed between the test sediment and control water.
Bivalve bioaccumulation			Significant difference from reference/ control

[a] Levels are regulated and represent minimum action list at lower level

Further information available at – www.ec.gc.ca/seadisposal **Note: Canada refers to upper and lower benchmarks as upper and lower Action Levels**

Box 4: The use of biological effects benchmarks in Canada

1D: OTHER APPROACHES

Among the examples from jurisdictions, it was noted that some benchmarks and Action Levels were set based on an approach other than those described. Frequently, there will be insufficient data, time or funding to enable the setting of benchmarks on purely scientific grounds. In order to proceed with a functional decision-making system in a reasonable time it is often necessary to take interim measures. Many jurisdictions may have limited information and simply decide to apply safety factors to benchmarks derived for other purposes, or set one benchmark as a multiple of another benchmark using a policy basis, or in an arbitrary fashion to help overcome a lack of data, or allow consistent decisions to be taken.