

## 8.4 Sediment quality guidelines

### 8.4.1 Sources of sediments and sediment contaminants

Aquatic sediments are principally derived from weathering processes, with major transportation from terrestrial sources under high runoff from storms and floods. In addition, discharges from urban, industrial and mining activities are potential sources of particulates. Anthropogenic contaminants, including metals, organics and nutrient elements are associated with particulate and dissolved inputs to natural waters. It is important to distinguish between point source and diffuse inputs. The former includes effluent streams, drains or licensed discharges, which can, if required, be the target of management actions. Diffuse sources include aerial deposition and land runoff, particularly from rural areas.

Particulate matter can act as binding sites for contaminants in soluble forms. Biological processes add particulate matter in the form of algal mats, dead cells, degradation and excretion products of animals, and living and dead plant biomass. Suspended particles gradually settle and accumulate as part of the bottom sediments. Rates of sedimentation vary from as low as 1 mm/y in coastal marine waters, to 10–20 mm/y in some riverine and estuarine systems. Highest values are found in settling basins removed from high currents and close to point sources, and rates in the range 3–7 mm/y are typical.

Contaminants are also associated with natural colloids, which can precipitate with aging or with changes in water chemistry. The change in salinity from fresh to saline waters will induce the precipitation of iron and manganese oxyhydroxides from both soluble ions and colloids, carrying with them other metals and organics.

### 8.4.2 Sediment properties

#### 8.4.2.1 Physics of aquatic sediments

Physical properties, such as grain size and density, are important in sedimentation and transport processes. Sediments are a heterogeneous mixture of particles ranging from millimetre to sub-micron in size. A classification of particles on the basis of grain size is shown in table 8.4.1. Typically, sediments are characterised as coarse material, clay/silt and sand fractions, on the basis of separations using 2 mm and 63  $\mu\text{m}$  sieves. Particles  $>2$  mm may consist of shells, rocks, wood, and other detrital materials, and are usually not a source of bioavailable contaminants (Mudroch et al. 1997). The clay/silt fraction has a high surface area and because of its surface chemistry is more likely to adsorb organic and heavy metal contaminants. Particles  $<63$   $\mu\text{m}$  are more common in the gut of sediment-ingesting biota (Tessier et al. 1984). It is not unusual to normalise contaminant analyses on the basis of the clay/silt fraction.

The sand and coarse silt fractions are generally dominated by quartz, sometimes by carbonates (shell, coral etc.), and occasionally by other silicates such as feldspar, or rock fragments. Primary silicates may also be present in the sand fraction, but are less evident in silt particles. Clay particles tend to be dominated by secondary silicates. Other secondary minerals such as oxides of aluminium and iron are prominent in the fine silt and clay fractions. This holds for most terrigenous sediments, and is not dissimilar in coastal marine sediments. Most anthropogenic contaminants (i.e. those associated with human activity) are associated with the clay and silt fractions.

**Table 8.4.1** Grain size classification of sediments

Grain size	Classification
<0.06 µm	Fine clay
0.06–0.2; 0.2–0.63 µm	Medium clay
0.63–2 µm	Coarse clay
2–6.3 µm	Fine silt
6.3–20 µm	Medium silt
20–63 µm	Coarse silt
>63 µm	Sand
>2 mm	Coarse material, rocks, detritus

(Mudroch et al. 1997)

Sediments are in intimate contact with the water that fills the voids between particles and within the pores of sediment particles. The volume of this interstitial or pore water will be governed by sediment porosity and is higher with the coarser sand fraction than with the finer clay/silt fraction. Sediment particle size is also critical to the ease and therefore the depths to which organisms can burrow. This is also dictated by the acceptability to the organism of the chemical environment of the sediment and its associated pore waters. Silty sand is a more acceptable medium for many benthic biota than is more compressible clay.

Physical processes in sediments influence the chemistry of sediments and their associated contaminants. Sediment resuspension can result from wind stirring, tidal currents and boating activities, as well as by biological activities (bioturbation). These processes can lead to particle sorting on the basis of density or size. They also expose particles to a different chemical environment, overlying water vs pore water and usually oxic vs anoxic. In the absence of any physical or biological sediment disturbances, contaminant movement occurs via diffusional processes in sediment pore waters, controlled by factors such as porosity.

### 8.4.2.2 Sediment chemistry

Sediment chemistry is controlled by redox conditions (dissolved oxygen, sulfides), pH, and the geochemistry of sediment particles. Contaminants are distributed over a range of geochemical phases, as well as being dissolved in the associated sediment pore waters, and the nature of these associations and sediment/pore water equilibria will determine their ultimate bioavailability.

In oxic sediments, the most important phases for metals are those containing hydrous iron and manganese oxides, although fractions will be present as weakly-adsorbed ion-exchangeable metals, or metals bound in carbonate, organic or sulfide phases. A significant metal fraction may be present in a detrital, mineralised form, but this is of little ecological importance as it is unavailable for bioaccumulation. Selective extraction schemes have been devised to attempt to chemically separate metals in these phases (Campbell et al. 1988, Kersten & Forstner 1989, Allen 1993). In anoxic marine and estuarine sediments, pyrite and other sulfide phases dominate.

Organic contaminants can be divided into hydrophobic (non-polar, water-insoluble) and hydrophilic (polar, soluble) species, and further subdivided as acidic, basic or neutral compounds. The former distinction, based on water solubility, can be related to the compound's octanol:water partition coefficient. Organic matter, either as discrete particles or

as coatings on inorganic substrates, is the primary adsorbing phase for hydrophobic organics, but depending on their charge, inorganic phases may be able to bind some compounds.

While metals may exist in both complexed and labile forms, they are not subject to the same degradation processes that are common to many organic molecules. In considering the environmental risk posed by organics and metals, the chemical form or speciation will be important, as will be the half-lives of chemical, physical and microbial degradation processes (Peterson & Batley 1993). In practice, hydrophilic organics are typically less persistent than hydrophobic compounds, because they are more amenable to hydrolysis and other solution degradation processes.

Contaminants bound to sediment phases are likely to be in thermodynamic equilibrium with the associated pore waters. This equilibrium will involve, almost exclusively, contaminants bound to adsorption sites on the sediment. Adsorption occurs during sedimentation and resuspension of particulates. Models to describe the binding of metals by sediments have been discussed by several authors (Oakley et al. 1981, Luoma & Davis 1983, Jenne et al. 1986, Campbell et al. 1988).

The redox state of sediments, i.e. whether they are in an oxidising or reducing environment, will be defined by the oxygen content of the pore waters. It is possible for sediments to be oxygen-deficient several millimetres below the surface. Oxygen deficiency will alter the chemistry of metals such as iron and manganese which in turn will affect the behaviour of other heavy metals that were previously bound to oxides of iron and manganese. Iron (III) hydrous oxides will be reduced to more soluble iron (II) species, while hydrous manganese oxides will be reduced to soluble manganese (II) species. Manganese, being more readily reduced than iron, appears in the pore water column at a higher zone in the sediment. The redox boundary is not necessarily stationary, and steady-state conditions may not apply because the boundary may move up and down through the sediments more quickly than the chemistry can respond. For example, the rate of oxidation of manganese is slower than that of iron, hence it is more readily transported through oxic environments and is relatively depleted in sedimentary rocks.

Most organic contaminants are not directly susceptible to redox changes, but indirectly, the presence of bacteria under specified redox conditions will affect the stability of such contaminants to microbial degradation processes.

The major nutrient elements of environmental concern in sediments are nitrogen and phosphorus. Both are present in organic and inorganic forms. Inorganic forms of nitrogen include nitrate, nitrite and ammonia. Organic nitrogen undergoes bacterial degradation and denitrification via ammonia, nitrite, and nitrate, ultimately to elemental nitrogen, as  $N_2$ . In oxygen-limited systems, these reactions can stop at ammonia. Phosphorus exists as phosphates, both monomeric and polymeric, and in sediments is usually bound with iron. Considerable phosphorus and nitrogen can also be bound by bacteria and it is important to consider living microscopic benthos as part of the sediment structure.

Sediments represent a potential source of contaminants to the overlying water and hence can influence water quality. The natural release of sediment contaminants is controlled by their dissolution into the sediment pore waters. Diffusion of these contaminants to the water column will occur if the pore water concentration exceeds that of the overlying water. The measurement of the fluxes of contaminants can be obtained using dialysis samplers (pore water peepers), benthic chambers or corer reactors.

### 8.4.2.3 Sediment biology

Interactions of biota with sediments and sediment contaminants occur at several trophic levels. Microbial processes are important in degrading organic matter and its associated nitrogen, via aerobic and anaerobic respiration, and nitrification and denitrification reactions. These influence sediment redox potential and pH, significantly altering metal bioavailability. Biomethylation is an important process affecting the bioavailability of mercury through the formation of methylmercury.

The exposure route to sediment contaminants for other benthic organisms has been assumed principally to involve pore waters, and hence water quality guideline values can be applied. Ingestion of sediment particles and dermal exposure can also be important exposure mechanisms. Benthic biota can include filter feeders (mussels, oysters) and deposit feeders (fish, crabs, etc.). Some are grazers (fish) and others are burrowers (crabs, polychaetes, shrimp, mussels). In addition there are organisms living in intimate contact with the sediment, such as benthic algae or rooted plants, that are incapable of ingesting particulate metals. An analysis of freshwater benthic organisms suggests that, with the exception of oligochaetes and some chironomids, they are not sediment ingesters (Adams 1987), whereas marine species include many burrowers that are able to ingest sediment. This latter observation is based solely on US data and its applicability to Australian freshwater organisms has not been tested.

Burrowing organisms have a significant impact on sediment chemistry and physics. Bioturbation, or burrowing activities, affects the sediment profile, by physically translocating contaminated sediments, mixing and redistributing the contamination. The change in profile can alter its use by biological communities through the formation of destabilised zones. The storage of food and faeces in burrows can result in new microhabitats that can promote microorganism growth. The types of burrows have been described by Aller (1982) and Rhoads (1976). Burrowing can involve particle sorting. Irrigation of the burrows can occur at quite extraordinary rates (1–750 mL/h) depending on the organism, and this has the ability to alter the redox environment at depth and mobilise contaminants. It is important to note also that the irrigation of burrows means that organisms are exposed to overlying water more than interstitial water (Boese et al. 1990).

It has often been suggested that if the organism is in equilibrium with the pore water, then the pore water concentration reflects the sum of all contaminants obtained via feeding (Maughan 1993), and that equilibrium is maintained by loss of higher accumulated concentrations to the pore water. This might be so for non-ionic organics whose partitioning to lipids is readily modelled, but is probably not the case for metals, where binding and immobilisation is less reversible. The concept that contaminant availability involved equilibrium dissolution from the solid phase was recently tested (Mayer et al. 1996). The digestive fluids of several marine invertebrate deposit feeders, including a polychaete (lugworm) and a holothuroid (sea cucumber), extracted more copper, lead and PAHs from coastal marine sediments than would be predicted by equilibrium partitioning. Nevertheless, less than 10% of the total contaminant load in each case was mobilised. There were a number of caveats on the results relating to the kinetics of contaminant release and the fact that animals are selective with respect to ingested sediment particles. The general low bioavailability observed was species-specific, but a relative but not absolute agreement with results from equilibrium sediment-water partitioning was conceded.

For plants, both water column and sediment uptake routes are possible. The latter is often masked by uptake from the water column. In sediments, pore waters are the major contaminant source and sediment oxygenation via the root system alters metal bioavailability in a similar manner to that induced by bioirrigation.

## 8.4.3 Review of approaches used to derive sediment quality guidelines

The many current approaches to the derivation of sediment quality guidelines can be broadly classified as being based on:

1. an effects or weight-of-evidence database from laboratory or field exposures to contaminated sediments;
2. an equilibrium partitioning approach and the application of existing water quality guidelines to sediment pore waters;
3. background levels, or some multiple of background levels, in the affected region.

The first two of these have been comprehensively reviewed elsewhere (e.g. Adams et al. 1992, Jones et al. 1996) and have formed the basis, either individually or collectively, of regulatory frameworks in many states of the US, and in Canada, the UK, Hong Kong and the Netherlands, as will be discussed.

The third approach, although lacking any good theoretical foundation is still in widespread use, principally for the derivation of sediment quality guidelines (e.g. the Oslo and Paris Commissions (1991, 1993) use this approach).

### 8.4.3.1 Effects-based guidelines

#### Spiked-sediment toxicity

An obvious approach to the development of sediment guidelines is to use concentration-response data for one or more benthic organisms, to establish cause and effect relationships in the same manner as used for water quality assays (USEPA 1989a). Data are obtained from laboratory testing of sediments spiked with known concentrations of contaminants, and can be used to generate quality criteria or to validate those generated by other methods. The tests generate unequivocal site-specific data which are highly defensible.

The technique requires mixing and equilibration of the sediment with a contaminant spike, added either to a sediment slurry or to the overlying water (Lamberson & Swartz 1991). The variable properties of sediments, in particular grain size, make the process more complex than the usual application of this technique for the generation of water quality criteria. Where the concentration of contaminant-absorbing phases present in the sediment is high, as for example in a fine-grained, clay-silt sediment, the equilibrium pore water concentration of the contaminant spike will be lower than in the absence of such phases. The guidelines will be sediment-specific unless some normalising procedure is used, as in the equilibrium partitioning approach described below (Section 8.4.3.2). As with all toxicity testing, acute toxicity will not be as reliable a basis for establishing acceptable criteria as chronic or life-cycle testing.

To date, the procedure has been applied to a limited number of contaminants usually in isolation, and to only a few benthic species. This is due in part to the high cost associated with the generation of a single sediment criterion. Furthermore, such data are generated from controlled laboratory tests only, usually under aerobic conditions, and field validation is required.

Data from spiked sediment testing have been incorporated in the derivation of guidelines using the biological effects database compiled by NOAA (Long et al. 1995). Acceptance criteria for these data are described in Section 8.4.5 and in the Canadian protocol (CCME 1995).

#### **Apparent effects threshold**

A more fundamental approach to the setting of sediment quality guidelines has been the use of an apparent effects threshold (AET). This is defined as the sediment concentration above which statistically-significant ( $p < 0.05$ ) biological effects are always observed for a given dataset. The method involves collection of matched chemical and biological effects data from tests carried out on sub-samples of the same field sample. Impacted and non-impacted sites are measured and the statistical significance of adverse biological effects is tested. Using only non-impacted sites, the AET is determined as the highest detected concentration of a given contaminant among sediment samples that does not exhibit a statistically-significant effect. The technique was first applied on datasets developed for sediments in Puget Sound, Washington (USA) (Beller et al. 1986, Barrick et al. 1988).

Criteria were derived for a number of metals and organics for sediments from Puget Sound (USA) (Barrick et al. 1989). For organic components, chemical data normalised to organic carbon were, surprisingly, no more predictive of observed biological effects than dry-weight-normalised data (Barrick et al. 1989).

The results obtained from the AET approach provide non-contradictory evidence of biological impacts, but have the disadvantages of being site specific, requiring a large database of chemical variables, and failing to separate combined from single contaminant effects. It was an early form of integrative sediment assessment (Chapman et al. 1992), where both chemical and biological effects were considered. Its use to derive numerical data for assessing sediment guidelines, has been largely overtaken by other more popular effects-based methods.

#### **Sediment quality triad**

The sediment quality triad concept was another of the early approaches to sediment quality assessment, involving data from three separate measurements: sediment chemistry, sediment bioassays and *in situ* biological effects (Chapman 1986). This concept is also an integrative assessment approach, generally conducted at the community or ecosystem level, to determine sediment quality in terms of ecosystem health (Chapman et al. 1992). Chemical and physical measurements assess the level of contamination and other parameters that might influence the abundance of infaunal species. Bioassay data provide information on the toxicity of contaminants, while *in situ* biological assessments look for histopathological abnormalities, community structure and other parameters that can be related to the sediment chemistry.

The consideration of data from each measurement enables separation of natural variability in biotic characteristics, that may result from differences in factors such as sediment particle size, from variability that might be due to sediment contaminants.

#### **Screening-level concentrations**

This approach uses field data and patterns of co-occurrence in sediments of specific benthic biota and particular concentrations of contaminants (Neff et al. 1986). The screening level concentrations are the estimated highest concentrations of selected contaminants that co-occur with approximately 95% of the infauna. It has been specifically applied to non-polar organic chemicals, where a cumulative frequency distribution of all stations at which a particular species of infaunal invertebrate is present, is plotted against the sediment concentration of the particular contaminant normalised to sediment organic carbon. Screening-level concentrations are obtained for a number of species, and from this, the concentration at which 95% of the species are found is determined as the sediment screening-level concentration.

This approach has been more widely pursued. The Ontario Ministry of the Environment (Persaud et al. 1990) developed sediment quality guidelines based on screening level concentrations from data for a range of local sediments and benthic biota. Two levels were reported, a low level which is the lowest that toxic effects become apparent, and a severe level, representing concentrations that could effectively eliminate most of the benthic organisms. For organics, the values are normalised to 1% organic carbon. Values are shown in table 8.4.2.

**Table 8.4.2** Ontario Ministry of Environment Screening Level Guidelines

<b>Contaminant</b>	<b>Low</b>	<b>Severe</b>
<b>Metals (mg/kg dry wt.)</b>		
Cadmium	0.6	10
Chromium	26	110
Copper	16	110
Lead	31	250
Manganese	460	1110
Mercury	0.2	2
Nickel	16	75
Zinc	120	820
<b>Metalloids (mg/kg dry wt.)</b>		
Arsenic	6	33
<b>Organics (<math>\mu\text{g}/\text{kg}</math> dry wt.)<sup>a</sup></b>		
Aldrin	2	80
BHC	3	120
a-BHC	6	100
b-BHC	5	210
Chlordane	7	60
Dieldrin	2	910
Total DDT	7	120
Endrin	3	1300
Heptachlor epoxide	5	50
Lindane	3	10
Mirex	7	1300
<b>Total PAHs</b>	<b>2000</b>	<b>110000</b>
<b>Total PCBs</b>	<b>70</b>	<b>5300</b>

a Normalised to 1% organic carbon

### Effects range approach

The use of large effects databases is now the most widely-accepted approach to sediment guideline development. The first approach of this type was reported by Long and Morgan (1990), based on an assessment of the potential for biological effects of sediment-sorbed contaminants in several hundred sites sampled as part of the National Oceanographic and Atmospheric Administration (NOAA) National Status and Trends Program in the US. The study examined data obtained from the equilibrium partitioning approach, the spiked sediment bioassay approach and other various approaches to sediment quality criteria. The chemical concentrations observed and predicted by the different methods to be associated with biological effects were sorted, and the lower 10 percentile and median concentrations were identified along with an apparent effects threshold. The lower 10 percentile data were identified as Effects Range-Low (ERL), and the median as Effects Range-Median (ERM). Data were used to rank sites in relation to the effects range values.

It should be emphasised that this approach was not originally intended to develop criteria, but rather to compare and rank sites, to identify priority contaminants, to estimate the relative potential for toxic effects and to itemise and describe the kinds of toxic effects previously observed in association with specific contaminant concentrations. Approaches to matching of biological and chemical data were discussed in a subsequent paper by Long et al. (1995).

MacDonald and co-workers (including Long) (1992, 1996) developed an expanded biological effects database for sediments (referred to by the acronym BEDS) for the Florida Department of Environmental Protection (FDEP), which is one of the most comprehensive documentations of sediment quality assessment yet reported. A limitation of guidelines derived from the original Long and Morgan approach was the size and the difficulty with user access of the database, and a lack of internal consistency in the data. MacDonald et al. (1992, 1996) converted the original database to a PARADOX<sub>TM</sub> format to allow greater ease of access and analysis, as well as compatibility with other common databases. Procedures were included for screening data from spiked-sediment bioassays and so-called co-occurrence data involving matching sediment chemistry and biological effects, prior to inclusion in the database.

The recommended guidelines were proposed on the basis that they could be implemented in the near-term, using existing data. Ranges of contaminant concentrations were used to evaluate sediment quality data, as these were considered more practical than single values covering diverse sediment conditions. The guideline values were considered to be preliminary and refinable as new data became available.

The setting of numerical guidelines involved sorting all of the information relating to a particular contaminant into two separate datasets, those that produced biological effects, and those that did not. These were sorted into ascending contaminant concentrations, and provided the datasets contained at least 20 entries, were used to derive a no-effects range, a possible effects range and a probable effects range, analogous to the ranges used by Long and Morgan (1990).

The effects data were sorted and the lower 15th percentile (ERL) and median or 50th percentile (ERM) calculated. From the no-effects data, the 50th percentile (No Effect Range Median, NERM) and the 85th percentile (No Effect Range High, NER-H) were determined. The threshold effects level (TEL) defines the upper limit of sediment contaminant concentrations of no-effects data (i.e. >75%, no-effects data) and was calculated as the geometric mean of the ERL and NERM. A safety factor of 2 was applied to the TEL values to define a no-observed-effects level (NOEL).

$$\text{TEL} = (\text{ERL} \times \text{NERM})^{1/2}$$

The probable effects concentrations (PEL) defining the lower limit of the range of contaminant concentrations that are usually associated with adverse biological effects (i.e. >75% effects data), were defined as the geometric mean of the ERM and NER-H values:

$$\text{PEL} = (\text{ERM} \times \text{NER-H})^{1/2}$$

There has been considerable debate over the derivation and use of effects range values (ERM and ERL) as numerical guideline values (Sampson et al. 1996a,b, MacDonald et al. 1996a,b). These values were never intended for use as stand-alone criteria; however, this caveat of the authors has frequently been disregarded by regulatory agencies.

Major criticisms include the fact that the derivation of ERL and ERM values were primarily derived from effects data only (Sampson et al. 1996a,b), a claim denied by MacDonald et al. (1996a,b). Sampson et al. (1996a,b) also suggested that the effects database includes results where mixtures of chemicals have resulted in the observed effect (Sampson et al. 1996a,b). One or more of these chemicals may have produced the effect, but it is ascribed to all chemicals in the mixture. Effects levels entered for some chemicals may therefore be well below actual effects thresholds. This is in fact overcome by a co-occurrence analysis so that only those chemicals exhibiting a concentration-dependent relationship with observed toxicity were included. The data use a variety of organisms and end-points, and therefore it was claimed that the derived guidelines would be more broadly applicable to a range of geographic areas (MacDonald et al. 1996b, Ingersoll et al. 1996a).

It was also suggested that there was an apparent biasing of the database towards effects rather than no-effects data. These factors result in large percentages of false positives, especially for metals (Sampson et al. 1996a). However, some 80% of the data are no-effects data. The data are designed to be predictive of both effects and no-effects. The ERL values are protective against false negatives and the ERM values protective against false positives (MacDonald et al. 1996a).

Ingersoll et al. (1996b) compared the effectiveness of ERL, ERM, TEL, PEL and no-effect concentrations (NEC). The latter are analogous to the apparent effect thresholds (AET) used by Barrick et al. (1988) and are defined as the concentration above which statistically significant toxic effects are always observed. They considered the ability of these sediment effects criteria to correctly classify toxicity or no toxicity and the respective abilities to classify non-toxic samples as toxic (Type I error, false positive) or toxic samples as non-toxic (Type II error or false negative). They concluded that ERMs and ERLs were generally as reliable as PELs and TELs in respectively classifying samples as toxic or non-toxic, but stressed the need to use field generated data, noting the problems with other contaminants in contributing to the observed effect.

A comparison of the NOAA and FDEP guidelines (table 8.4.3) shows that they are remarkably similar in most cases. The Ontario screening level values (table 8.4.2) for organics are mostly 2 to 10 times higher than the ERL or TEL, and are possibly under-protective, but for metals are comparable. The reliability of the NOAA values has been raised in Section 3/5/4 (Volume 1). A more detailed evaluation of the reliability of the FDEP guidelines has been provided by Jones et al. (1996) and there appears to be a similar degree of confidence in both approaches. The fact that the guidelines were primarily developed from estuarine and marine data was not seen as a limitation to their application and in the case of water quality, it has been suggested (Klapow & Lewis 1979) that the statistical difference between marine and freshwater guidelines is insufficient to preclude their combination.

There is clearly merit in the use of effects databases, provided their limitations are acknowledged and they are applied more as screening tools to delineate areas of concern. It is important that data are continually updated and revised, and guideline values that are inconsistent with other findings should be the subject of more detailed investigations.

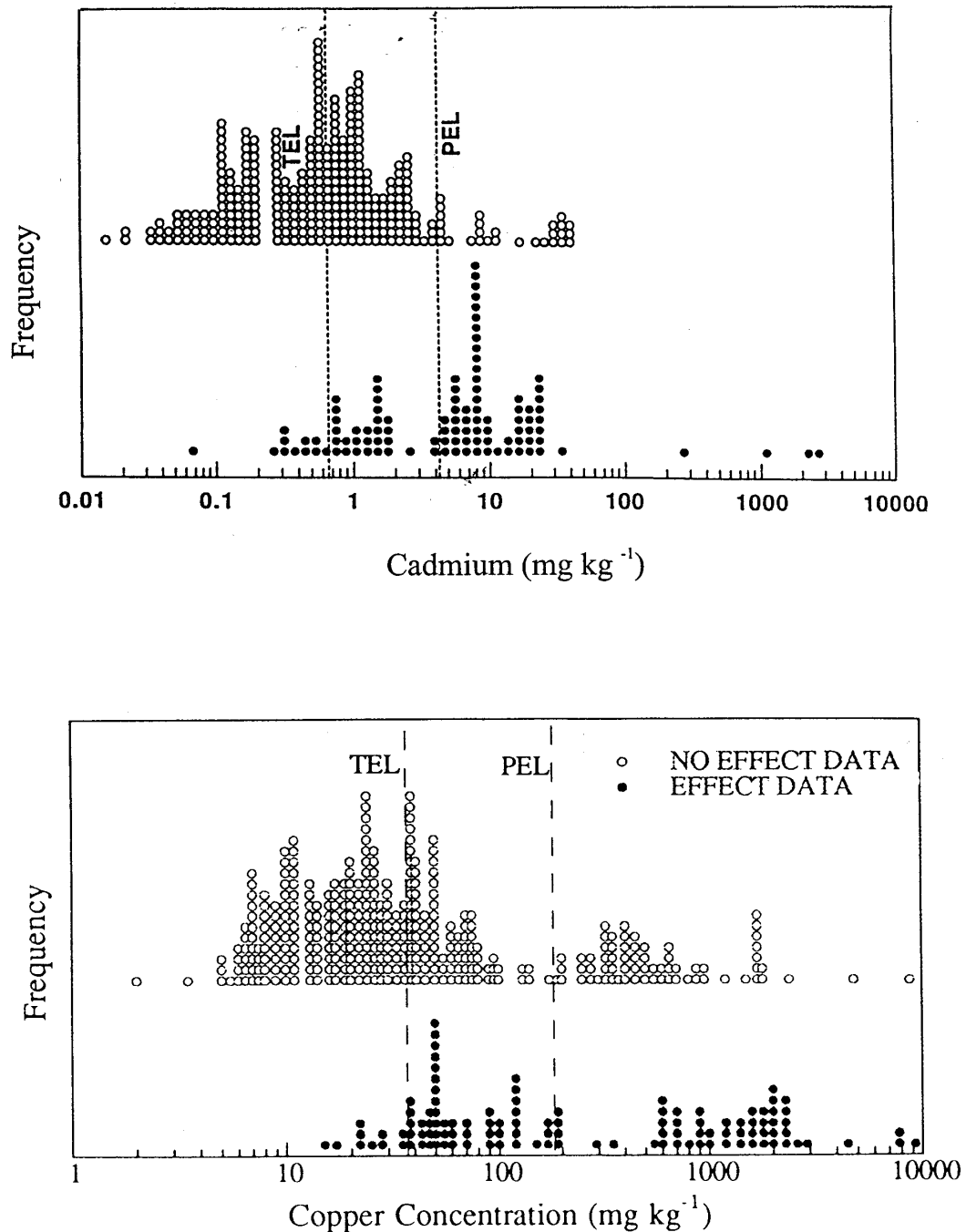
It is worth remarking that the number of significant figures used in the guideline values shown in table 8.4.3 is not justified given the appreciable errors not only in the analyses, but in the general level of confidence in the effects data. Appropriately rounded off numbers are used in the new guidelines (table 3.5.1, Volume 1).

**Table 8.4.3** Summary of Effects-Range Guidelines

Contaminant	NOAA Guidelines		FDEP Guidelines	
	ERL	ERM	TEL	PEL
<b>Metals (mg/kg dry wt.)</b>				
Cadmium	1.2	9.6	0.68	4.21
Chromium	81	370	52.3	160
Copper	34	270	18.7	108
Lead	46.7	218	30.2	112
Mercury	0.15	0.71	0.13	0.7
Nickel	20.9	51.6	15.9	42.8
Silver	1.0	3.7	0.73	1.77
Zinc	150	410	124	271
<b>Metalloids (mg/kg dry wt.)</b>				
Arsenic	8.2	70	7.24	41.6
<b>Organics (<math>\mu\text{g}/\text{kg}</math> dry wt.)<sup>a</sup></b>				
Acenaphthene	16	500	6.71	88.9
Anthracene	85	1100	46.9	245
Fluorene	19	540	21.2	144
Naphthalene	160	2100	34.6	391
Phenanthrene	240	1500	86.7	544
Low Molecular Weight PAHs	552	3160	312	1442
Benzo(a)anthracene	261	1600	74.8	693
Benzo(a)pyrene	430	1600	88.8	763
Dibenzo(a,h)anthracene	63.4	260	6.22	135
Chrysene	384	2800	108	846
Fluoranthene	600	5100	113	1494
Pyrene	665	2600	153	1398
High Molecular Weight PAHs	1700	9600	655	6676
Total PAHs	4022	44792	1684	16770
Total DDT	1.58	46.1	3.89	51.7
Dieldrin	0.02	8	0.72	4.3
Chlordane	0.5	6	2.26	4.79
Total PCBs	22.7	180	21.6	189

a Normalised to 1% organic carbon

Much of the concern and apparent inconsistencies in the data, especially for metals, relate to the use of total rather than bioavailable concentrations. Typical datasets for cadmium (IMO 1997) and copper used in the derivation of PEL and TEL values for Canadian Sediment Quality Guidelines are shown in figure 8.4.1.



**Figure 8.4.1** Distribution of (a) cadmium and (b) copper concentrations in marine and estuarine sediments that are associated with adverse biological effects (filled circles) and no adverse effects (open circles) (from IMO (1997) (cadmium) and Smith et al. (1996b) (zinc))

These illustrate the appreciable overlap between effects and no-effects data, no doubt biased by the use of total metals concentrations. In the case of copper (Allen 1996), many copper concentrations in the range 500–2000 mg/kg were present amongst the no-effects data, obviously representing non-bioavailable copper in either mineralised or sulfide-associated forms. The need to take this into account has been considered in the new guidelines by using acid-soluble rather than total metals concentrations.

More recent considerations of effects range data have introduced the concept of ERM quotients, i.e. the measured concentration divided by the ERM concentration (Long et al. 1998, Long & MacDonald 1998). Evaluations of sediments have included the use of a mean ERM quotient for a number of contaminants (e.g. five metals), with acceptability being if this value is below 1. As discussed earlier, the focus of these studies has been the ability of the measures to predict toxicity, and this is not the objective of the guidelines. The use of mean quotients is of concern if the exceedance of a single contaminant can be masked by quotients less than 1 for a number of other contaminants.

#### **8.4.3.2 Equilibrium partitioning approach**

The equilibrium partitioning approach is based on the assumption that the critical factor controlling sediment toxicity is the concentration of contaminant in the sediment pore water. Water quality guidelines can be applied to pore water contaminants, and the sediment quality guideline value can be defined by the concentration of contaminant in the sediment that is in equilibrium with the water quality guideline concentration in the pore water. The ratio of the contaminant concentration in the sediment ( $C_S$ ) and its concentration in the surrounding water ( $C_W$ ) is defined as the partition coefficient,  $K_D$ .

The approach is attractive to many regulators because toxicity can be predicted using  $LC_{50}$  values obtained from water-only toxicity tests. The major research effort has been devoted to attempts to predict the factors controlling the partitioning of contaminants within the sediment solid phases, into pore waters. The approach is most readily applicable to hydrophobic organic chemicals and has been incorporated for these chemicals in the derivation of the NOAA and FDEP guidelines. Jones et al. (1996) have summarised equilibrium partitioning guidelines for a range of these organics.

#### **Organics**

For non-ionic organic chemicals, it has been well-established that the partitioning is dominated by sediment organic carbon (Di Toro et al. 1991). For sediments having >0.2% organic carbon (dry weight), sediment quality guidelines normalised to mg/kg organic carbon, have been shown to be valid for a range of sediment types. This cut-off in organic carbon content is judged to be necessary because at lower organic carbon contents, second-order effects such as particle size and adsorption to non-organic mineral fractions become more important. Recent studies (Luthy et al. 1997, Le Boeuf & Weber 1999) have challenged the appropriateness of normalisation to organic carbon. However, these studies have been restricted to higher carbon concentrations than normally considered in sediments.

The sediment/pore water partition coefficient,  $K_D$ , is related to the organic carbon partition coefficient,  $K_{OC}$ , and  $f_{OC}$ , the fraction by weight of organic carbon,  $f_{OC}$ :

$$K_D = f_{OC} K_{OC}$$

Note that  $K_{OC}$  is empirically related to the readily-determined octanol/water partition coefficient  $K_{OW}$ .

The equilibrium partitioning model predicts that sediments will be toxic when the pore water concentration exceeds the water-only toxic concentration. Thus if WQG ( $\mu\text{g/L}$ ) is the no-effect concentration in water, then the sediment quality guideline value, SQG ( $\mu\text{g/kg}$ ), is given by the partition coefficient,  $K_D$  (L/kg), between pore water and sediment according to the equation:

$$\text{SQG} = K_D \text{WQG},$$

and in terms of measurable parameters:

$$\text{SQG} = f_{\text{OC}} K_{\text{OC}} \text{WQG}$$

Using published data for the toxicity to marine and freshwater amphipods, Di Toro et al. (1991) demonstrated that 100% mortality occurred when the ratio of the pore water concentration to water-only  $\text{LC}_{50}$ , exceeded a value of 1. This agreement implied that benthic organisms were as sensitive as water-column organisms. They rationalised that, if the pore water and sediment are in equilibrium, then the effective exposure concentration is the same regardless of the exposure route.

Using a predicted sediment toxic unit (STU) as given by:

$$\text{STU} = (C_S/f_{\text{OC}})/(K_{\text{OC}} \cdot \text{LC}_{50 \text{ WATER-ONLY}})$$

where  $C_S/f_{\text{OC}}$  is the organic-carbon-normalised concentration of contaminant in the sediment, a similar plot to that for pore water toxicity can be obtained for a range of organics, with 100% mortality occurring where  $\text{STU} = 1$ .

The USEPA has accepted the equilibrium partitioning approach for hydrophobic organics and have commissioned a number of studies addressing the derivation of criteria for specific organic chemicals (USEPA 1989b; 1993a-d). The criteria have been derived on an organic carbon basis by determining  $K_{\text{OC}}$  and the WQG according to:

$$\text{SQG}_{\text{OC}} = K_{\text{OC}} \text{WQG}$$

The current water quality criteria are used for WQG.

It is important to recognise that the normalisation to organic carbon is only valid when  $f_{\text{OC}} > 0.2\%$ , as discussed earlier. The studies usually incorporate uncertainty analyses, so that for dieldrin, for example (USEPA 1993d), the uncertainties associated with measurements in water-only and sediment over a range of organisms lead to upper and lower 95% confidence limits of 4.2 and 19 for a freshwater criterion of 9.0 mg/kg OC, although these did not include uncertainties in the water quality criteria. The claim is that a freshwater sediment concentration of dieldrin of  $< 9.0$  mg/kg organic carbon would offer acceptable protection to aquatic organisms. The upper confidence limit is the value above which impacts would be highly likely, and the lower the concentration below which impacts would be unlikely.

For polar organics, the  $K_{\text{OC}}$  model can overestimate the bioavailable concentration, because adsorption can be enhanced by factors other than hydrophobicity. The general behaviour of such organics will also differ because they are susceptible to a range of degradative and removal processes.

### Metals

The application of the equilibrium partitioning approach to metals is less advanced, and is confounded by the dependence of metal bioavailability on more than one phase in the sediments, and the fact that bioavailability can be ameliorated in the pore waters by complexation with dissolved organic matter.

Considerable research has been directed to defining those phases that control metal bioavailability. This has traditionally focussed on oxic sediments, where it is now acknowledged that the principal metal binding phases are hydrous iron and manganese oxide phases, and in some instances organic carbon as well. The current limitation to the development of predictive capabilities is the lack of reliable sorption constants.

Tessier and co-workers have been responsible for the most significant developments in this area (Tessier 1992, Tessier et al. 1984, 1989, 1993). Field-derived values for the partition coefficients have been obtained for a range of freshwater sediments, on the basis of iron oxyhydroxides being the major sorptive phase for cadmium, copper, nickel, lead and zinc. Pore water concentrations of each metal were measured, along with the concentrations of iron oxyhydroxide substrate and metals associated with this phase, following dissolution using a suitable reducing agent. The constant,  $K_{Fe}$ , will be pH dependent, and plots of  $\log K_{Fe}$  vs pH were generally linear over the range of pore water pH values, with a slope of one.

Deviations from linearity occur where more than one phase is controlling adsorption. The additional binding sites are likely to be organic ligands. Deriving constants for these, requires a selective extraction with hydrogen peroxide to determine the bound metal fraction, and total organic carbon measurements as an estimate of the concentration of organic binding sites. The extension of this approach to the generation of sediment quality criteria requires field measurements on a range of sediment types to derive appropriate constants. This necessary research is still in progress.

The toxicity of pore waters will be diminished if the released metals are bound by soluble complexing agents. Very little work has been carried out on metal speciation in oxic pore waters, although it would be expected that speciation effects would be similar to those seen in overlying waters. It has been reported (Mahony et al. 1991) that the toxicity of copper in pore waters to amphipods diminished as a function of increasing dissolved organic carbon. This needs to be considered in any predictions based on equilibrium partitioning.

In anoxic sediments, available sulfide will regulate the solubility of metals such as cadmium, copper, mercury, nickel, lead, silver and zinc, which form relatively insoluble sulfides. The significance of sulfide partitioning in controlling metal bioavailability in marine sediments spiked with cadmium was demonstrated by Di Toro et al. (1990, 1992). Normalising the sediment metal concentrations to the concentrations of acid-volatile sulfide (AVS) was shown to provide a reasonable predictor of when pore waters should exhibit toxicity from metals (Ankley et al. 1991a, Carlson et al. 1991). The ratio of AVS, operationally-defined as the sulfide liberated from wet sediment by treatment with 1M hydrochloric acid, to the concentration of metals (excluding iron) simultaneously extracted (SEM), is the key parameter. If this ratio exceeds one, then the excess of sulfide will imply no metal toxicity. If the ratio is less than one, the sediments may be toxic. This approach has been criticised as being overly simplistic. Hare et al. (1994) pointed out that it is the difference between SEM concentration and AVS that is important, because there may well be examples where the ratio exceeds unity, but there is no toxicity because the metal concentration is too low. The release of sulfides from forms that are unable to exchange with heavy metals in solution may well occur, resulting in an underestimation of toxicity.

More recently, Simpson et al. (1997, 1998) have pointed out the limitations to the application of AVS theory to metals such as copper, nickel and cobalt. The sulfides of these metals are insoluble in dilute HCl and so will not appear as AVS, although oxidative release of these metals can occur. The tendency may be to over-estimate SEM:AVS. They also found examples of metal oxide particles in sediments armoured with sulfide coatings, in which the

bioavailability of the metals was minimised by the sulfide coating, yet both oxides and sulfides appeared in the SEM fraction.

Ankley et al. (1994a, 1996) have proposed an approach to deriving sediment quality criteria for selected metals (copper, cadmium, nickel, lead and zinc) based on equilibrium partitioning-based estimates of metal concentrations associated with the lack of adverse biological effects. They use four procedures: (a) comparing the sum of their molar concentrations to the molar concentration of AVS; (b) comparing the interstitial water concentrations to the water quality criteria final chronic values (WQGs); (c) using organic carbon-based partition coefficients in addition to AVS and SEM to compute interstitial water metal concentrations; and (d) using minimum partition coefficients (e.g. generated from chromatographic sand (Hassan et al. 1996)) to compute sediment concentrations that would not result in interstitial water exceeding metal WQGs.

There have been a number of studies to date (Leonard et al. 1996, Hassan et al. 1996, Berry et al. 1996, Sibley et al. 1996) that indicate the ability of the above approaches as predictors of no-effects as distinct from effects of metal contaminants. This factor may in itself be a concern, but is generally consistent with the major objective for sediment quality guidelines in identifying sediments that are worthy of preservation in a non-toxic state, as discussed earlier.

There are a number of limitations to the equilibrium partitioning approach for metals that have been identified so far, not the least of which is the labile nature of acid volatile sulfide. Oxidation can occur readily during sample handling, or as a result of aeration of anoxic sediment zones by burrowing organisms, altering the metal chemistry and toxicity from that predicted by the SEM/AVS ratios (Aller 1982). Often the burrows are lined with mucus. A recent study by Peterson et al. (1996) found that the burrowing oligochaete, *Lumbriculus variegatus*, significantly reduced AVS concentrations in surficial sediments, as well as increasing the concentration of bioavailable cadmium in the sediment pore waters at all depths disturbed by the organisms. In evaluating AVS, it is obviously important to specifically consider surficial sediments, where AVS will be reduced, as well as analysing deeper sediments, rather than a pooled sample over a range of depths. Seasonal changes can also influence the AVS/SEM relationship.

More importantly, the implicit assumption in the equilibrium partitioning process is that pore waters represent the major uptake route for sediment contaminants. This will not always be the case, although it may well be that the most sensitive organisms are those that respond to pore water concentrations only, despite the greater 'available' contaminant pool in the sediments.

### **Pore water guidelines**

In some instances pore waters may represent the dominant phase in which a contaminant is found, usually as a consequence of its formation in this phase as a result of chemical and microbiological processes, and/or because of its high aqueous solubility. Ammonia is a case in point, as are nutrients such as nitrate and nitrite. As discussed in Section 3.5.4.2 (Volume 1), in such cases it is appropriate to apply the water quality guideline values, or equivalent values derived using pore water toxicity testing with benthic organisms.

Ammonia is, potentially, a highly toxic naturally-occurring constituent of sediment pore waters, and is generally not considered a contaminant of concern in the regulation and management of sediments (e.g. dredged material). The toxicity of ammonia is influenced by the temperature and pH of the water and at elevated levels has the potential to confound interpretation of sediment toxicity tests using sensitive species, and influence the distribution of infaunal species measured in ecological impact studies.

Measurements of total and un-ionised ammonia concentrations in sediment pore waters from the Mississippi River frequently exceeded water quality criteria for ammonia, with strong seasonal and spatial concentrations which were positively correlated with silt and volatile solids content (Frazier et al. 1996). High spatial and temporal variability of pore-water ammonia has also been found in stream sediments (Sarda & Burton 1995). A review of pore water ammonia concentrations in 322 estuarine/marine sediments showed a log-normal distribution of data with mean concentration of 9.03 and 40.74 mg NH<sub>3</sub>-N/L respectively for natural and dredged sediments, and comparison of the potential exposure concentrations with measured amphipod sensitivities indicated a significant potential for mortality due to ammonia in 10-day tests (Moore et al. 1997).

Comparison of the toxicity of ammonia in spiked-sediment versus water-only exposures has shown good correspondence between the LC<sub>50</sub> values for the infaunal *L. variegatus* and the chironomid *C. tentans*, indicating that ammonia bioavailability and toxicity may be accurately predicted from pore-water concentrations for some species (Whiteman et al. 1996). However, the epibenthic *Hyalella azteca* exhibited a behavioural response and apparently avoided the spiked sediments being frequently observed in the overlying waters. Ammonia toxicity tests on New Zealand invertebrate species suggest that they may be among the more sensitive species (Hickey & Vickers 1994). Generally, the information available suggests that measurement of pore water ammonia levels and comparison with guideline values for water-only exposure will provide adequate prediction of potential sediment effects.

#### 8.4.3.3 Guidelines based on bioaccumulation

Where aquatic organisms are harvested for human consumption, there is the potential for bioaccumulation of contaminants to concentrations that exceed health standard values. This may occur at concentrations in the sediment that are below those that exert chronic or acute toxic effects on the organism. High bioconcentration factors are most likely encountered for hydrophobic organic contaminants which partition to the high lipid-containing sites in the organisms. Many metal concentrations are often regulated or excreted by an organism.

It is possible to derive sediment quality guidelines on the basis of a consideration of bioconcentration factors (BCFs) and standards for human consumption of aquatic organisms (Van der Kooij et al. 1991). For example, the criterion for dissolved concentration of an organic contaminant ( $C_w$ ) can be calculated from the known bioconcentration factor and the health standard for fish ( $C_{org}$ ):

$$C_w = \frac{C_{org}}{BCF}$$

The related criterion for sediment ( $C_{sed}$ ) is related to the criterion for suspended matter ( $C_{susp}$ ) by:

$$C_{sed} = \frac{C_{susp}}{r}$$

where  $r$  is an empirical concentration ratio of suspended matter:sediment ( $r = 2$  for organics and 1.5 for metals).

$$C_{sed} = \frac{(C_{org} \times K_D)}{r \times BCF}$$

Using the relationship:  $K_D = f_{OC} K_{OC} = 0.6 f_{OC} K_{OW}$  (Karickhoff et al. 1979), where  $K_{OW}$  is the octanol:water partition coefficient, it is possible to derive  $C_{sed}$  from:

$$C_{sed} = \frac{C_{org}}{BCF} \times 0.06 f_{oc} K_{ow}$$

In applying such calculations for selected hydrophobic organic contaminants, it was possible to show that in some instances the health standard-based guideline values were lower than the values calculated on the basis of equilibrium partitioning calculations (Van der Kooij et al. 1991). There are, however, many uncertainties in the derived numbers. Bioconcentration factors, for example, need to be appropriate to sediment-ingesting biota. For metals, the health-based guidelines were typically higher than that based on toxicity. Apart from the Netherlands, no other country has actively considered health-based guidelines.

For further information on the derivation of water quality guidelines for protection of human consumers of aquatic foods, see Section 9.4.3 (Volume 3).

#### 8.4.3.4 Guidelines for the ocean disposal of dredged sediments

Interim ANZECC guidelines for the assessment of dredged sediments for acceptability for ocean disposal, were released in 1998 (ANZECC 1998). The document outlines methods for sediment sampling and analysis, sediment quality assessment and biological testing. The interim guideline values are identical to those proposed in this document, with the lower and upper guideline values being referred to as Screening and Maximum levels respectively.

The sediment quality assessment procedure used in the sea disposal guidelines is as follows:

- i) Where data were available to establish the regional concentrations in the sediments of the receiving area, the mean value of such concentrations was used as the background level for naturally occurring substances. A lower (reference) screening level is developed by multiplying the background level by two to account for sampling and analytical variability and the range of natural values in the area. Where background data were not available, the screening values in table 8.4.3 were used. Dumping is permitted where the mean of all contaminants to be dredged is below the Screening Level. It is likely that where clay/silt sediment is being dumped on sand that even uncontaminated sediment would fail the twice background criteria, and then the Screening Guidelines are applied, as outlined in table 8.4.3.
- ii) Where the mean value of one or more contaminants is between the Screening and Maximum Levels, further assessment is required, including the determination of acute sediment toxicity on suitable test organisms. Where one or more contaminants is above the Maximum Level, the sediment is unsuitable for disposal at sea, although this may yet be permissible if the results of further sediment bioassays, including an evaluation of sub-lethal toxicity and bioaccumulation, show that the material is non-toxic.

The Screening Level values are considered to be tentative and are to be revised as international criteria are updated, and/or Australian criteria are developed. The value for radionuclides is the maximum specified by Australian ocean dumping legislation. For organochlorine pesticides where reliable detection levels are close to screening levels, it is recommended that a case-by-case assessment be made.

Sediment toxicity testing using protocols such as those developed by the USEPA (1991a, 1994) or the American Society for Testing and Materials (ASTM 1997a,b) were considered

to be the most appropriate for predicting the bioavailability, toxicity and bioaccumulation potential of contaminants in sediments.

For the interim guidelines (ANZECC 1998), elutriate testing (USEPA 1991a, 1994) was used to determine the water quality impacts of disposal. Using a 1:4 dilution and a four-hour mixing, the results are compared against the marine water quality guidelines criteria, taking into account appropriate dilution factors.

#### **8.4.3.5 Other international approaches to sediment quality**

##### **Canada**

The Canadian approach to the derivation of sediment quality guidelines has been outlined in a publication by the Canadian Council of Ministers of the Environment (CCME) in 1995. In earlier studies the Ontario Ministry of the Environment has been prominent in the development of dredged sediment quality criteria, published in 1988, and subsequent provincial sediment quality guidelines, which have been summarised by Bennett and Cabbage (1991). The latter underwent several revisions, but basically defined three levels of long-term chronic effects on benthic organisms, a no-effect level, a lowest-effect level and a screening level. The screening level represents a concentration that would have a pronounced effect on sediment-dwelling organisms and would be detrimental to most benthic species, while the lowest effect level could be tolerated by most benthic species.

Environment Canada in 1992 commissioned a study of marine environmental quality guidelines, which reviewed sediment quality guidelines, and provided recommendations for the setting of guidelines (MacDonald et al. 1992b). In the short term, the report endorsed the use of information that is currently available and in use. Dose/response data were acknowledged as being the most defensible, and more ecologically relevant than indirect measures of biological effects, such as those derived from equilibrium partitioning. The authors were comfortable with the three-tiered approach recommended to the Ontario Ministry of the Environment (Hart 1988). The procedure was as follows.

- i) Select the lowest of the effects-based guidelines if any of these have been or can be calculated.
- ii) If no effects-based data are available, for organic contaminants, use the lowest value obtained from the equilibrium partitioning and water quality guidelines approaches, and for metals the equilibrium partitioning approach is recommended subject to acceptance of the AVS normalising procedure (Allen et al. 1993).
- iii) For site-specific objectives, the background limit should be used if the interim guideline from (i) or (ii) is below the upper background limit of the contaminant.

This approach was largely pursued in the 1995 publication (CCME 1995) which set out in detail a protocol for the derivation of guidelines. For the longer term, the recommendation was basically to derive guidelines on the basis of studies to meet specific contaminant challenges.

The new Canadian guidelines (CCME 1999) are shown in table 8.4.4. Although these data are based on the FDEP approach of TEL and PEL values, there is a clear difference, both in the case of metals and especially in the case of organics, from those values used elsewhere. The difference between marine and freshwater values could in most cases be considered insignificant (Smith et al. 1996a). However, it is surprising that the data are tabulated to an apparently unjustifiable number of significant figures.

## The Netherlands

The Netherlands has played a prominent role in the past in the development of soil quality guidelines, and in recent years has been actively examining sediment quality. An approach based on equilibrium partitioning has been outlined by Van der Kooij et al. (1991). Two guideline values are derived, one based on measured aquatic toxicity data, and the other on bioaccumulation data as discussed earlier. The direct effects scheme bases its estimation of  $K_D$  on mean coefficients from a water quality database for Dutch surface waters. Since these are based on suspended sediment, and not bottom sediment, a factor is used to convert the data. For metals, an empirical concentration ratio of suspended matter/sediment is taken as 1.5, while for organics, the ratio is taken as 2.0, based on a measured difference in organic matter content.

**Table 8.4.4** Interim Canadian sediment quality guidelines

Contaminant	Interim Guideline Marine	PEL Marine	Interim Guideline Freshwater	PEL Freshwater
<b>Metals (mg/kg dry wt.)</b>				
Cadmium	0.7	4.2	0.6	3.5
Chromium	52.3	160	37.3	90.0
Copper	18.7	108	35.7	197
Lead	30.2	112	35.0	91.3
Mercury	0.13	0.70	0.17	0.486
Zinc	124	271	123	315
<b>Metalloids (mg/kg dry wt)</b>				
Arsenic	7.24	41.6	5.9	17.0
<b>Organics (<math>\mu\text{g}/\text{kg}</math> dry wt)</b>				
Phenanthrene	86.7	544	41.9	515
Benzo(a)anthracene	74.8	693	31.7	385
Benzo(a)pyrene	88.8	763	31.9	782
Chrysene	108	846	57.1	862
Fluoranthene	113	1494	111	2355
Pyrene	153	1398	53.0	875
DDD	1.22	7.8	3.54	8.51
DDE	2.07	374	1.42	6.75
DDT	1.19	4.77	1.19	4.77
Dieldrin	0.71	4.30	2.85	6.67
Chlordane	2.26	4.79	4.50	8.87
Total PCBs	21.5	189	34.1	277

The lower of the values obtained from each of the two schemes is adopted as the guideline value. To date, guidelines for some 120 contaminants have been derived for both sediments and water (Stortelder et al. 1989), and have been adopted by the Netherlands Ministry of Transport and Public Works. Not surprisingly, perhaps, the guidelines derived from these rather empirical approaches differ significantly from other recommended criteria. The guideline for cadmium calculated by Van der Kooij et al. (1991) as 8.7 mg/kg, is almost an order of magnitude above the value derived by MacDonald et al. (1992a) and higher than most other criteria.

The environmental quality objectives and the policy framework in which they are set, is updated regularly by the Ministry of Housing, Spatial Planning and the Environment (Van Der Weiden et al. 1994). A summary of the latest recommended values is shown in table 8.4.5 (MHSPE 1999). Guidelines are specified in terms of a target value that is set at a negligible concentration, usually 1/100th of the maximum permissible concentration, or the NOEL. If the negligible level is lower than the natural background concentration, then the target value will be set to that level. As the target values can often only be achieved in the longer term, intermediate objectives have in some cases been defined in terms of limit values (Van Der Weiden et al. 1994). These may never be exceeded. They are in a grey area and are determined by considering environmental, economic and social interests, and technical options, and follow the ALARA (as low as reasonably achievable) approach. Limit values were only set for the upper sediment layer, in direct contact with water, and were derived using the equilibrium partitioning approach.

**Table 8.4.5** Sediment quality objectives in the Netherlands<sup>a</sup>

Contaminant	Target Value	Maximum Permissible Concentration
	mg/kg dry weight	mg/kg dry weight
Arsenic	29	55
Cadmium	0.8	12
Copper	36	73
Lead	85	530
Mercury	0.3	10
Zinc	140	620
	µg/kg dry weight	µg/kg dry weight
Chlordane	0.03	3
Dieldrin	0.5	450
Lindane	0.00005	2
Total PCBs	0.02	–
DDT	0.09	2
DDD	0.02	1
DDE	0.01	1
atrazine	0.2	26
chlorpyrifos	0.01	1
Endosulfan	0.01	1
diazinon	0.01	1
naphthalene	0.001	0.1
fluoranthene	0.03	0.4
Benzo(a)pyrene	0.003	3
Total PAHs	1	-
TBT	0.007	0.7

<sup>a</sup> Values apply to a sediment having 10% organic matter and 25% clay

The sediment quality objectives proposed for organics were in many instances based on measurements of the current quality of sediment which is considered to be relatively

unpolluted. The list provided is extremely comprehensive. For some of the substances a risk evaluation was carried out, and the findings were included as target values.

The maximum permissible concentrations (MPC) are ones above which there is serious pollution, but the recommendations stress the importance of other sediment parameters. For example, in the case of metals, high concentrations under anaerobic conditions will pose low immediate risks, but if the conditions should change, the risk could obviously increase. The MPCs for organics were based on surface water values, using substance specific equilibrium partition coefficients.

The Dutch approach closely relates both soil and sediment, and formulae are given for relating soil (sediment) composition and percent organic matter and percent clay. The value of this is questionable in the case of metals, but for organics, the importance of organic carbon as a normalising factor has already been identified. The values for organics in table 8.4.5 are based on a standard sediment having 10% organic matter. It is assumed that a linear relationship exists between the organic matter content and the concentration at which adverse effects occur.

The target values listed for metals in table 8.4.5 are not drastically different from those used elsewhere. The concentrations for organics are, in general, far more conservative than other guidelines.

### **United Kingdom**

In common with many other countries, the push for sediment quality guidelines in the United Kingdom has come from a concern for ocean disposal, not only of dredged sediments, but also of sewage sludge. This has been addressed by the Marine Pollution Monitoring Management Group of the Directorate of Fisheries Research, Ministry of Agriculture Fisheries and Food. Their research has been the subject of five reports since 1989 (MAFF 1989, 1991a,b, 1992, 1993). They have pursued the equilibrium partitioning approach for setting provisional environmental quality criteria for organic components (Webster & Wridgeway 1994). Some values have been proposed, based on US data, but some revision has been undertaken to take into account the water quality guidelines in operation in European waters. Proposed action levels for metals have also been reported (MAFF 1992).

### **Hong Kong**

The approach recommended for Hong Kong sediment quality values (Chapman et al. 1999) was to adopt a set of international effects-based numbers, which are compared with a limited range of existing Hong Kong data for consistency. These numbers, called interim sediment quality values (ISQG) would be refined on the basis of local site-specific information. Values were required below which biological effects were unlikely and above which adverse effects were very likely, to deliver information on three classes of sediment: uncontaminated, contaminated and highly contaminated.

The international data chosen for metals and metalloids were the NOAA co-occurrence values (Long & Morgan 1991, Long et al. 1995), because these showed a reasonable relationship in most cases for concentrations and effects, and include data from tropical areas. The ISQG-low values for metals used previous Hong Kong criteria for which no effects had been observed to date, but for mercury and arsenic the NOAA values were used. The ISQG-high values for arsenic and metals other than mercury used the NOAA effects range median, with the exception of a value for nickel which was considered unreliable. The ISQG-high value for mercury used previous Hong Kong criteria to fit the international range of values and because the ERM was not reliable.

For organics, values for PAHs were based on NOAA values, which were lower and more conservative than the criteria for only three PAHs proposed by the USEPA. Total DDT and total PCBs were based on NOAA values, but no high values were set. For other organics such as phenols and TBT, no reliable values were available and consequently none were set.

In general, where NOAA values were not appropriate, due either to an absence of values or a poor agreement with incidence of effects, then either no value was listed or the USEPA value was chosen.

Being based primarily on temperate data, ISQG values are not ideal, but are not inappropriate when compared to the range of chemical concentrations in Hong Kong sediments. Site specific data are needed to confirm or revise them. As with all effects-based data they do not consider synergism between contaminants, and are based on toxicity to biological receptors only, although these are usually more stringent than human health values. They do not consider biomagnification or secondary toxicity.

The accepted interim values provide an effective basis for screening sediments, and where sediments are found to be contaminated, biological effects data are needed as part of the final decision making. The accumulation of new effects data will ultimately lead to refinement of interim values and the development of accepted sediment quality criteria.

#### **8.4.3.6 Reliability of the ANZECC/ARMCANZ effects-based guideline values**

The effects-based guidelines recommended for Australia and New Zealand are primarily based on a single, large biological-effects dataset of North American sediment data, with appropriate rounding off in keeping with their precision. These are supplemented by site specific screening data and some predictions based on equilibrium partitioning

Ideally, the derived guidelines should accurately predict the toxicity or absence of toxicity in a sediment, but there are many qualifiers on these data, as we have seen via the decision-tree approach, which need to be taken into consideration in their application. One of the important considerations is the question of background concentrations.

We have chosen to select a database that considers effects data only, although conscious of the fact that this may result in some bias to the guidelines. This is not considered as serious as the bias that might be introduced by the inclusion of no-effects data, particularly those from samples of coarser grain size, where contaminant concentrations would be expected to be insignificant. Using metals data that are more closely related to bioavailability in the derivation of future guidelines should result in a more meaningful effects dataset. It may be that a similar approach to no-effects data and closer scrutiny of the data to avoid potential bias will allow for the combination of effects and no-effects data in later guidelines development.

A further deficiency is that the guideline end-point is toxicity. At some stage, consideration must be given to addressing other undesirable end-points, such as excessive algal growth induced by high nutrient concentrations.

As stressed earlier, it is important to recognise the uncertainties in the derived guideline values. A review of the values used in the NOAA listing which underpin the recommended guidelines have been critically reviewed by Jones et al. (1996). Their evaluations of the effects range-low (ERL) and effects range-median (ERM) metals data, equivalent to ISQG-Low and ISQG-High of the current guidelines, are as follows:

**Antimony:** Data are available from only two geographic regions (Puget Sound/Commencement Bay and San Francisco Bay); therefore, the degree of confidence in the NOAA values is moderate (Long & Morgan 1990).

**Arsenic:** Confidence in the ERL is relatively high. Confidence in the ERM is somewhat lower, therefore ERL appears to be a reliable indicator of the threshold for effects, and the ERM appears to be the better indicator of the level above which real effects are likely.

**Cadmium:** A relatively large amount of data exists for cadmium including spiked-sediment toxicity tests and equilibrium partitioning-based assessments (MacDonald et al. 1994). Klapow and Lewis (1979) calculated a statistically significant difference in the medians of acute aqueous toxicity data from saltwater and freshwater organisms. This supports the findings of resistance to cadmium being higher among marine than freshwater species (Long & Morgan 1990). However, the degree of confidence in the lower and upper NOAA values is high, and these values are considered to be reliable predictors of effects.

**Chromium:** There are some inconsistencies in the data available for chromium, possibly due to lack of speciation information. All data were reported as total chromium, whereas the hexavalent form is more toxic than the trivalent form. There are also no supporting data from single-chemical, spiked-sediment toxicity tests or from the EqP approach (MacDonald et al. 1994). Although overall confidence in the NOAA values is relatively high, Long et al. (1995) cautioned that the incidence of effects may be unduly exaggerated by data from multiple tests performed in only two studies. Based on the available evaluations, the ERL appears to be a reliable predictor of the threshold for effects.

**Copper:** Considerable data exist for copper in sediments, and a relatively high degree of overall confidence exists for the NOAA values.

**Lead:** Considerable data exist for lead in sediments, and a relatively high degree of overall confidence exists for the NOAA values. This is consistent with expected reliability of the ERL and ERM values. Although relatively large amounts of data exist for lead, there were no spiked-sediment toxicity test data to confirm the toxic concentrations (MacDonald et al. 1994).

**Mercury:** Considerable data exist for mercury, though only total mercury concentrations were reported in the dataset (MacDonald et al. 1994). Confidence in the ERL is relatively high. Therefore, the lower screening value should be at least a conservative predictor of the threshold for effects. The ERM may significantly overpredict the likelihood of real effects, given that confidence in these values is low.

**Nickel:** Toxicity of nickel is greatly influenced by water hardness and salinity (Long & Morgan 1990). Data were from marine and estuarine field studies only, and no spiked-sediment toxicity tests or EqP approaches were used (MacDonald et al. 1994). Confidence in the ERL is relatively high. Therefore, the lower screening values should be at least conservative predictors of the threshold for effects. The ERM may significantly overpredict the likelihood of real effects, given that confidence in the value is low.

**Silver:** A moderate amount of data is available for silver in sediments, though there are no data from spiked-sediment toxicity tests or from EqP approaches (MacDonald et al. 1994). The NOAA values hold relatively high overall confidence. These data suggest that the ERL is likely to be a reliable predictor of the threshold for effects. The ERM is likely to be a good predictor of real effects.

**Zinc:** Considerable data exist for zinc in sediments, including spiked-sediment toxicity tests and equilibrium partitioning-based assessments (MacDonald et al. 1994). There is an expected reliability of the NOAA values.

**Polycyclic Aromatic Hydrocarbons (PAHs):** The reliability of the ERLs for anthracene and fluorene is low, while reliability of the ERM for dibenzo(a,h)anthracene is relatively low. The reliability of the ERLs and ERMs for the remaining PAHs is relatively high.

**Total Polychlorinated Biphenyls (PCBs):** The reliability of the ERM is considered to be relatively low. Concordance of the concentrations and effects was not high. This may be the result of insufficiently detailed chemical information. That is, the mixture of PCB congeners may have varied considerably among sites and studies, but this information was unavailable or not included in the NOAA analyses. Caution should be used when screening PCBs with these benchmarks.

**Pesticides:** NOAA values are available for only two pesticides, p,p'-DDE and total DDT, for which poor concordance of effects and concentrations was observed. This may be due to the inclusion of relatively low EqP values which were not based on toxicity to benthic organisms (Long et al. 1995). Therefore, the NOAA values may tend to over-predict the likelihood of effects.

On the basis of the above comments, several alterations to the NOAA values were made before inclusion as a recommended guideline for Australia and New Zealand. Given the unreliability of the NOAA ERM value for mercury, the upper sediment quality guideline value of 1.0 mg/kg adopted for sediments in Hong Kong by Chapman et al. (1999), has been used as a more realistic number. Similarly the unreliable values for nickel have been noted, but rather than delete these as was done in the Hong Kong case, they are included until more reliable local data can be obtained. Also noted is the low ERL value for arsenic, which is below the background values in many Australian sediments. The revised number of 20 mg As/kg used in the ocean disposal guidelines (ANZECC 1998) has therefore been used. Note that the lower guideline value chosen for lead is the rounded-off NOAA value. The Hong Kong lower guideline value for copper of 65 mg/kg, based on previous values used by Hong Kong and not no-effects data, was also accepted for these guidelines.

No value is available for selenium, although there are sites in Australia and New Zealand where interim guidelines would be of value. Recent studies by Van Derveer and Canton (1997) have suggested a threshold level of 4 mg/kg for freshwater sediments based on the lower 10th percentile of effects data, although these are limited to tests on fish.

For radionuclides, the values used in the ocean disposal guidelines (ANZECC 1998) have been used. These are the maximum specified by Australian legislation. In terms of the impact on aquatic biota, a more detailed evaluation is required.

Guidelines for tributyltin have been estimated on the basis of equilibrium partitioning, based on data summarised for the USEPA (Weston 1996). Given the uncertainties in the appropriateness of both organic-carbon-based partition coefficients and water quality guideline values used in the calculation of these guidelines, further validation through toxicity testing is desirable.

In the listing of organics, values for chlordane, dieldrin, endrin and DDD, from the Long et al. (1995) paper, have been included in our guidelines, as was done by Jones et al. (1996). No values were available for lindane, so the Florida State values (MacDonald et al. 1994) are included as interim values. Note that for all organics data, the original tabulation covered a

range of organic carbon concentrations from 1% to 3%. However, normalising to 1% was seen as acceptable to better define acceptable values for higher carbon percentages.

The guidelines for many organic contaminants, in particular, are derived on the basis of pore water toxicity testing, with the sediment guideline concentration being calculated from measured sediment:water partition coefficients ( $K_D$ ). The latter value may vary over several orders of magnitude, being highest for clay/silt sediments. Guidelines should be appropriate to the sediment type being considered, otherwise meaningless and unachievable guideline values may be derived.

It is tempting to use water quality guidelines to derive sediment quality guidelines on the basis that:

$$SQG = K_D \cdot WQG$$

This is only acceptable if it can be demonstrated that benthic organisms are as sensitive as water-column organisms. Given the uncertainties in the derivation of sediment guidelines, the approach is possibly justified for the provision of interim guideline values.

The original US databases were developed for estuarine and marine sediments but have been applied to freshwater sediments. Recent development in Canada of guidelines for freshwater sediments showed good agreement ( $r = 0.98$ ) with the marine-based guidelines (Environment Canada 1995, Smith et al. 1996a). For the NWQMS, it is desirable that separate datasets be developed for saline and freshwaters. Often the chemicals of concern are different (e.g. pesticides such as endosulfan in rural environments vs tributyltin in marine and estuarine waters), and replication in both environments might not be considered necessary.

The use of an effects-only database, while appropriate for setting interim guidelines, can have limitations if all data are biased to contaminated sediments where the effects are dominant, in the same way that the use, as well, of no-effects data can add abnormally low values because of coarser grain-sized sediments. These factors should be considered along with the other physico-chemical parameters in determining the suitability of data for inclusion in the database. The use of acid-soluble rather than total metals may also help overcome a distortion to the ultimate guideline value.

Once an acceptable dataset is obtained for any single chemical, its calculated guideline value will replace the existing interim value.

#### **8.4.3.7 Relevance to Australia and New Zealand**

At the current stage of their development, the relevance of the adopted guidelines to Australia and New Zealand has yet to be determined, but evaluation of Australian aquatic toxicity data to date has not shown the response of organisms to particular contaminants to lie outside the range of values normally found overseas. Given the paucity of sediment toxicity data at this time, the values are considered only to be interim values until verified by new local data, as described above. New data obtained in the interim, as part of the refinement of sediment quality guidelines for other temperate or sub-tropical countries in the region such as Hong Kong, will also be assessed for ANZECC/ARMCANZ, along with refinements of North American guidelines. The latter could be added as part of a revised interim dataset.

## 8.4.4 Australian and New Zealand data

There are currently few available sediment quality data for either Australian or New Zealand sediments or organisms. While there have been a large number of studies where the chemical concentrations of contaminants have been measured in sediments, very few have been related to biological effects, either in the nature of descriptions of the natural benthic populations or laboratory-based bioassays.

The chemical data are useful in indicating the natural ranges of concentrations found, and if examined carefully could provide information on background contaminant concentrations. Very few studies have examined acid-soluble metals, so the usefulness of metals data, in terms of the recommendations of this document, is questionable. These data remain largely in investigation reports and few find their way into published papers. Attempts to collate available data have been limited to activities by selected bodies such as the NSW EPA, or NIWA in New Zealand.

There are some data for marine sediment contamination by organics in New Zealand (C Hickey, pers. comm.) covering a range of 'clean' sites, harbour wharves and marinas, and which shows that the proposed guideline values for DDT will be exceeded at a large number of sites. Guidelines for PAHs, PCBs, chlordane, lindane and dieldrin will be exceeded at relatively few sites (<20%). The generally high levels of DDT suggest that determination of the ecological significance of this contaminant should be a priority in sediment toxicity research.

Available metals contaminant data for estuarine and marine sediments has not been summarised. Some data are available for sediments from the Waikato River in New Zealand (Hickey et al. 1995) showing exceedance of sediment quality guidelines for mercury and arsenic at most sites and for copper, chromium, iron and manganese at some sites. The implications for the upper guideline exceedance by arsenic are being investigated using a suite of sediment toxicity tests. Establishing the appropriateness of the new ANZECC/ARMCANZ guideline values for these contaminants is particularly important as a large number of New Zealand's freshwater lakes have natural and industrial geothermal inputs.

An ANZECC/ARMCANZ database on existing sediment toxicity data for particular contaminants is being assembled, based largely on a limited number of studies by Dr John Chapman's group at the Centre for Ecotoxicology, University of Technology, Sydney, and Dr Chris Hickey's group at NIWA in Hamilton, New Zealand.

Sediment toxicity testing in Australasia is in its infancy. In freshwaters, the suitability of five sediment-dwelling species (amphipod, *Chaetocorophium lucasi*; freshwater clam, *Sphaerium novaezelandiae*; oligochaetes, *Lumbriculus variegatus*; freshwater tanaid, *Tanais standfordi*; and the burrowing mayfly, *Ichthybotus hudsoni*) for 10-day sediment toxicity testing has recently been assessed by comparison of their sensitivity to reference toxicants (phenol and pentachlorophenol) and contaminated sediments (Hickey & Martin 1995). The results showed good survival and high sensitivity to reference toxicants for the amphipods and clams. Further development of sub-lethal growth tests for amphipods and clams, and reproduction tests for oligochaetes is underway.

In Australia, a 10-day whole sediment test has been developed (Hyne & Everett 1998) with a *Corophium* species, similar to *C. volutator* (Pallas), which has been selected as a sediment test organism in Europe (OECD 1995). This species is euryhaline and is suitable for testing both fresh and estuarine sediments. The species was sensitive to copper and ammonia in both water exposures and whole sediment and it has been used to assess sediment quality in sites in the Hawkesbury-Nepean River, NSW, where reduction in field abundances of the species

was correlated with known pollution. A chronic test is currently being developed (Hyne pers. comm. 1997).

Nymphs of the epibenthic mayfly, *Jappa kutera*, have been used for *in situ* testing (Leonard et al. 1997) and testing of sediments spiked with endosulfan in the rivers of the cotton growing area of NSW. *J. kutera* was more sensitive to endosulfan in the pore water and overlying water than in the whole sediment. These tests demonstrated the desorption properties and the differential behaviour of the different isomers of endosulfan in contact with sediment (Leonard et al. 2000, R Hyne pers. comm. 1997).

In the marine environment, New Zealand ecotoxicity studies have mainly concentrated on sediment-dwelling amphipod and juvenile shellfish species. Hickey and Roper (1992) compared the acute sensitivity to cadmium of two sediment-dwelling amphipod species (tube-dweller *Paracorophium excavatum*, and burrower *Proharpinia hurleyi*) and reported values which were within the range expected for amphipod species. Sublethal behavioural responses, including avoidance and burial rate, have been compared with morbidity and mortality in bioassays using the marine bivalve, *Macomona liliana*, exposed to copper- and chlordane-dosed sediments (Roper & Hickey 1994, Roper et al. 1995). Avoidance was the most sensitive behavioural response, and occurred at 6- to 20-fold, respectively, lower levels than morbidity. Such behavioural responses may be particularly useful in understanding the ecological effects of contaminants by comparing laboratory results with field density measurements and diversity biomonitoring. The absence or reduced diversity of some species in receiving waters may be caused by an organism's behavioural preference for less contaminated areas. Chronic growth-rate tests have been developed with estuarine amphipods (*Paracorophium lucasi*) and have shown significant reductions at two contaminated (heavy metals and organics) estuarine sites (Mischewski 1994). Chronic studies for a range of sediment types (muds to sands) have been characterised for growth response of amphipods and juvenile bivalves to physical factors, in order to assist in the interpretation of contaminant effects studies (Nipper & Roper 1995).

Laboratory toxicity studies have complemented field contaminant effect studies. Field-dosing studies using chlordane applied to intertidal sandflats at a single low concentration have shown persistence over extended tidal cycles (Smith et al. 1992) and a reduced abundance of juveniles of two shellfish species after 44 tidal cycles (Pridmore et al. 1992). Chlordane uptake and depuration rates were shown to differ between the filtration-feeding cockle and the deposit-feeding wedge shell (Wilcock et al. 1993) and with depth in the sediments (Wilcock et al. 1994). Laboratory-measured behavioural (avoidance) responses to chlordane occurred at a concentration ~20-fold lower than that causing morbidity, at concentrations approaching field effect-levels (Roper & Hickey 1994).

Sediment tests in saline conditions have been developed by Everett (1997) using a burrowing amphipod, *Victoriopisa australiensis*, for whole sediment tests and a benthic amphipod, *Hyale crassicornis*, for pore water tests. These were assessed using copper and heptachlor, which, in the case of the burrowing species, were spiked into the sediment. *V. australiensis* was remarkably insensitive to copper, with a 10-d LC<sub>50</sub> of 1520 µg/L, on a dry weight basis. *H. crassicornis* was among the more sensitive species to water-borne copper (96-h LC<sub>50</sub> of 135–150 µg/L) but less sensitive to heptachlor (33 µg/L) than other marine species. Although a low degree of mortality was detected at some of the nine sites tested in the Hawkesbury estuary, this was not significantly different from the reference sediments.

The incorporation of Australian and New Zealand data into ANZECC/ARMCANZ guidelines requires a thorough review and evaluation of all existing published and unpublished sediment toxicity testing data.

## 8.4.5 Updating of Guideline Values

For future revision of guideline trigger values, the procedure adopted in Canada should be followed (CCME 1995), relying both on field studies (co-occurrence data matching sediment chemistry and biological-effect data) as well as spiked sediment toxicity testing. The addition of new data to the existing US biological effects (BEDS) database will have only a minor effect on the guideline values until the number of data entries becomes substantial. In the establishment of a NWQMS database, toxicological data must be sorted on a chemical by chemical basis, in order of ascending chemical concentrations. For inclusion, data must meet the following criteria (CCME 1995):

- i) The procedures used for collection, sampling, handling and storage of saline and freshwater sediments should be consistent with standardised protocols (e.g. ASTM 1997b).
- ii) Data must contain matching sediment chemistry and biological effect data collected from the same locations at the same time.
- iii) The concentrations of one or more analytes must vary by at least a factor of 10 at different sampling sites represented in a single co-occurrence dataset (at a particular location).
- iv) Toxicity tests should employ generally accepted laboratory practices of exposure and environmental controls. Tests should follow standardised protocols. Novel protocols need evaluation.
- v) Concentrations of the chemical in the sediment must be measured and not calculated.
- vi) Static, static-renewal, or flow-through tests may be employed, but tests should demonstrate that adequate environmental conditions for the test species are maintained throughout.
- vii) Preferred end-points include effects on embryonic development, early life-stage survival, growth, reproduction and adult survival, although other ecologically-relevant end-points may be considered.
- viii) Responses and survival of controls must be measured, and should be appropriate for the life stage of the species used.
- ix) Appropriate analytical procedures must be used to generate data on contaminant concentrations in the sediments.
- x) Measurements of abiotic variables should be reported. In the overlying water these should include pH, dissolved oxygen, total suspended solids, suspended and dissolved organic carbon, water hardness (and/or alkalinity) and salinity. In the sediment, variables include total organic carbon, particle-size distribution, acid volatile sulfide, pH, redox conditions and sediment type.
- xi) Appropriate statistical measures should be used and reported.

The statistical derivation of guideline values from the effects database is straightforward, provided the minimum toxicological data requirements are met. These ensure that there is adequate weight-of-evidence linking chemical concentrations to biological effects. Effects datasets for each chemical must include at least twenty (20) entries in a guideline derivation table for the chemical under consideration.

Spiked sediment toxicity testing can be used to supplement testing on contaminated sediments described above. A minimum set of requirements must again be satisfied (CCME 1995), namely:

- i) at least four studies are required on two or more sediment-resident invertebrate species, of which at least one must be a benthic amphipod species, for marine sediments, or one benthic crustacean species plus one benthic arthropod (other than a crustacean) for freshwater sediments.
- ii) at least two of the studies must be partial or full life-cycle tests that consider ecologically-relevant end-points such as growth, reproduction or developmental effects.

From these spiking studies, guideline values are based on the lowest-observed-effect level (LOEL) from a chronic toxicity assessment using a non-lethal end-point, multiplied by an appropriate safety factor. If an acute toxicity assessment for another species is the most sensitive, a safety factor is applied to the LC<sub>50</sub> or EC<sub>50</sub>. Further information on acceptable factors is provided in the protocol for the Canadian sediment guidelines (CCME 1995).

### 8.4.6 Sediment toxicity testing

The evaluation of sediment toxicity through laboratory or field bioassays is critical to the derivation of acceptable sediment quality guidelines. Biological assessment identifies whether any chemical contamination may be eliciting a biological response and assists with decisions for remediation or other action. In addition, it is an appropriate means of demonstrating conformity with the interim guideline values, and this may be undertaken at any stage during the risk-based decision-tree evaluation (Section 3.5, figure 3.5.1).

The science of sediment toxicology is very young, with most peer-reviewed papers on the subject being published since 1988. There are relatively few standardised test methods for evaluating sediments. Sediment toxicity assessment methods have been summarised in several recent reviews (Burton 1992, Keddy et al. 1994). These have included examining water leachates of sediments (elutriates), interstitial water or whole (bulk) sediment phases, using test species spanning the aquatic food chain from bacteria to fish (Burton 1992). The results of such testing have been used not only to develop the guidelines for sediment-effect thresholds, but also to manage dredge disposal and as a component in ecological assessments (ANZECC 1996).

Toxicity testing of contaminated sediments has focussed primarily on acute toxicity (lethality) effects of organisms, with highly contaminated material showing correlations between sediment contaminant concentrations and survival in some cases but not in others (Burton & Scott 1992). More recent work has been developing sublethal end-points for sediment tests. 'Whole sediment' testing with infaunal species has the greatest relevance for predicting ecologically-relevant end-points. However, natural variability in sediment particle size, natural contaminants (e.g. ammonia, hydrogen sulfide) and intraspecies competition may result in a number of factors which may confound interpretation of sediment assay results. To some extent, some of these can be predicted or modelled (e.g. AVS). The inclusion of reference sediments in tests is used to allow for interference by some of these factors. The variation in these parameters from site-to-site adds more weight to the need to conduct toxicity tests in conjunction with chemical measurements (Ankley et al. 1994b). Marine and freshwater sediment testing with the amphipods, *Rhepoxynius abronius* (Swartz et al. 1985) and *Hyaella azteca* (Nebeker & Miller 1988) respectively, have probably been the most widely used 10-day exposure tests. Guidelines also exist for whole-sediment assays

(acute and short-term chronic exposure) in freshwaters using midge larvae (chironomid species), other insects (mayfly, *Hexagenia* sp) and worms (oligochaetes, *Lumbriculus* sp); and marine waters using polychaete worms (*Neanthes* sp and *Capitella capitata*). Generally, inclusion of a number of species having a range of contaminant sensitivities and tolerance of sediment characteristics is desirable to provide a weight-of-evidence for measured toxic responses.

Environmental impact investigations involving sediments, commonly include a combination of toxicity tests, chemical contaminant measurements and infaunal macroinvertebrate analyses, together known as the triad approach (e.g. Chapman 1986) to provide a weight-of-evidence of adverse impacts. Such an approach involves integrating multiple measurement end-points to provide an assessment of whether significant risk of harm is posed to the environment. This is particularly applicable for sediments where a multitude of factors may result in community changes and cause/effect relationships may not be directly applicable to changes in chemical contaminant levels.

Elutriate and pore-water sediment toxicity tests are used mainly for dredge spoil assessment and for toxicity identification evaluation (TIE) procedures. These tests can utilise a wider range of species normally used for water assessments. The TIE procedures use a range of chemical treatments (e.g. pH adjustments, filtration, air sparge) to remove or modify the bioavailability of contaminants (Burkhard & Ankley 1989, USEPA 1991b). By repeating the toxicity tests on the modified sample, an indication of the type of toxicant causing the toxicity can be determined. Such investigations are particularly useful in identifying treatment or mitigation options for toxicity reduction.