

Critical Evaluation of the Sediment Effect Concentrations for Polychlorinated Biphenyls

D Scott Beckert[†] and Thomas C Ginn^{*‡}

[†]Integral Consulting, 7900 SE 28th Street, Mercer Island, Washington 98040, USA

[‡]Exponent, 1040 E Park Ridge Drive, Sedona, Arizona 86336, USA

(Received 31 July 2007; Accepted 11 November 2007)

ABSTRACT

In 2000, a set of sediment effect concentrations (SECs) was published for evaluating the toxicity of polychlorinated biphenyls (PCBs) in freshwater, estuarine, and marine sediments. According to the developers, these consensus-based SECs reconcile existing sediment quality guidelines (SQGs) that have been developed using various approaches, reflect causal rather than correlative effects, and can be used to determine the spatial extent of injury to sediment-dwelling organisms. In the present study, a critical evaluation of the SECs was conducted based on the original documents and databases used to develop the underlying SQGs for the SECs, as well as the original documents and data sets used to determine the predictive ability of the SECs. Results of the critical evaluation indicated that the SECs are simple mathematical constructs that share the same limitations as their underlying SQGs. The SECs are questionable “consensus” values, because many of their underlying SQGs are dissimilar, misclassified, or redundant with other SQGs. Because nearly all of the data sets included in the databases used to calculate the underlying SQGs, or to validate the SECs, were affected by elevated concentrations of multiple co-occurring chemicals, it was not possible to conclusively identify PCBs as the cause of any of the observed sediment toxicity. The SECs, and most of their underlying SQGs, are likely biased by the fact that their underlying databases are composed primarily of PCB concentrations less than 0.5 mg/kg dry weight. Comparisons between the SECs and bioaccumulation-based SQGs calculated using the equilibrium partitioning approach provide no information on whether the SECs are causally related to sediment toxicity. The primary available median lethal concentration (LC50) value for PCBs, determined using spiked-sediment toxicity tests, has limited applicability to most contaminated aquatic environments, because it was determined using an unusually low total organic carbon content. Finally, site-specific application of the SECs indicated that their predictive ability was very low, that concentration–response relationships were not found for a variety of test species and toxicity endpoints at PCB concentrations greater than the SECs, and that some of the highest survival and growth values in the toxicity tests were found at PCB concentrations considerably greater than the SECs. Based on the results of this study, we conclude that the SECs for PCBs should be used only in the screening-level evaluations that typically precede more direct assessments of sediment toxicity at individual study sites, and should not be used to predict the presence of sediment toxicity. Contrary to the conclusions of the SEC developers, the SECs do not reconcile existing SQGs, do not reflect causal effects, and should not be used to determine the spatial extent of injury to sediment-dwelling organisms.

Keywords: Polychlorinated biphenyls Sediment effect concentrations Sediment quality guidelines Sediment toxicity

INTRODUCTION

In 2000, MacDonald, DiPinto, et al. (2000) developed a set of sediment effect concentrations (SECs) for evaluating the toxicity of polychlorinated biphenyls (PCBs) in freshwater, estuarine, and marine sediments. According to the authors, these consensus-based SECs reconcile existing sediment quality guidelines (SQGs) that have been developed using various approaches, reflect causal rather than correlative effects, and can be used to determine the spatial extent of injury to sediment-dwelling organisms, as well as to evaluate the need for sediment remediation. However, Fuchsman et al. (2006) have recently criticized the SECs because they cannot be used to identify cause–effect relationships between PCB concentrations and sediment toxicity. Those authors proposed use of the equilibrium partitioning (EqP) approach as a better method of developing SQGs for PCBs.

As codevelopers of the apparent effects threshold (AET) approach for developing SQGs (Barrick et al. 1988; Becker et al. 1989), we have been involved with sediment quality issues

for more than 20 y and we agree with Wenning et al. (2005) that the use of SQGs in screening-level assessments can provide useful initial information on which sediments are likely to be nontoxic. However, we also believe that definitive evaluations of sediment toxicity should be based primarily on site-specific assessments of biological effects that address such factors as site-specific bioavailability and the presence or absence of concentration–response relationships. Laboratory studies that address potential causality, such as toxicity identification evaluations and spiked sediment studies, may also be useful.

Wenning et al. (2005) described how SQGs can be misused. Given the uses for which MacDonald, DiPinto, et al. (2000) conclude that the SECs for PCBs are applicable, the potential exists for the SECs to be given more validity than the SQGs on which they are based and, in turn, to be used in a more authoritative and conclusive manner than is justified. For example, MacDonald Environmental Sciences Ltd. (MESL 1999) stated that measured concentrations of PCBs in sediments can be compared to the SECs to determine whether the concentrations are sufficient to injure sediment-dwelling organisms. This statement implies that the SECs

* To whom correspondence may be addressed: ginnt@exponent.com

Published on the Web 11/11/2007.

have the scientific merit to be used as stand-alone criteria for identifying the presence of adverse biological effects related to PCBs.

In this paper, we present the results of a critical evaluation of the SECs for PCBs. Specifically, we show that the SECs represent questionable consensus values, are limited by the same biased databases as most other SQGs, are strictly correlative in nature (rather than causative), and can provide highly erroneous predictions when applied to site-specific conditions. All concentrations presented in this paper are expressed on a dry-weight basis.

METHODS AND MATERIALS

MacDonald, DiPinto, et al. (2000) developed the SECs for PCBs by compiling 30 existing SQGs—17 that are applied primarily to freshwater sediments, and 13 that are applied primarily to marine and estuarine (i.e., saltwater) sediments (Table 1). MacDonald, Ingersoll, et al. (2000) conducted a similar kind of analysis when developing consensus-based SQGs for 28 chemicals and chemical groups (including PCBs) for freshwater sediments. However, in that study, only 2 consensus-based SQGs were developed for each chemical: A threshold effect concentration (TEC) and a probable effect concentration (PEC). The TEC and PEC for PCBs identified by MacDonald, Ingersoll, et al. (2000) were 0.06 and 0.676 mg/kg, respectively.

In determining the SECs for PCBs, MacDonald, DiPinto, et al. (2000) subdivided the 30 existing SQGs into 3 groups based primarily on their narrative intent, as defined by their original developers and interpreted by MacDonald, DiPinto, et al. (2000; Table 1). The 3 groups of SQGs were then used to develop 3 kinds of SECs: A TEC, a midrange effect concentration (MEC), and an extreme effect concentration (EEC). Initially, the 3 SECs were calculated separately for freshwater and saltwater SQGs as the geometric means of the 3 SQG groups. However, because the 3 pairs of freshwater and saltwater SECs were found to be statistically similar, the underlying freshwater and saltwater SQGs were combined to form a single SEC for each of the 3 SQG groups. The final TEC, MEC, and EEC for PCBs were calculated as 0.04, 0.40, and 1.7 mg/kg, respectively.

MacDonald, DiPinto, et al. (2000) evaluated the reliability of the 3 combined SECs for PCBs by 1) determining their predictive ability to correctly identify sediments as toxic or nontoxic in 2 independent freshwater and saltwater sediment toxicity data sets (i.e., validation data sets), 2) comparing them to concentration–response data from spiked-sediment toxicity tests, and 3) comparing them with SQGs developed using the EqP approach, which is a theoretical approach that identifies toxicity thresholds for nonionic organic chemicals when they occur alone in sediments. The evaluation of predictive ability was conducted using matching sediment chemistry and biological effects data from 10 freshwater and 15 saltwater data sets from throughout the United States (Table 2; MESL 1999). The primary concentration–response information used by MacDonald, DiPinto, et al. (2000) was provided by a spiked-sediment toxicity study conducted by Swartz et al. (1988), and the SQGs developed using the EqP approach were those of Bolton et al. (1985) and NYSDEC (1999).

To conduct the critical evaluation of the SECs for PCBs in the present study, as many of the following items were assembled as were available: The original documents and

databases used to develop the 30 underlying SQGs for PCBs, and the original documents and data sets for the freshwater and saltwater data sets used by MacDonald, DiPinto, et al. (2000) to validate the SECs. All of the original SQG documents were obtained, and the only SQG databases that could not be obtained from the original authors were those for the freshwater and saltwater screening-level concentrations (SLCs) and those for the freshwater lowest effect level (LEL), severe effect level (SEL), minimal effect threshold (MET), and toxic effect threshold (TET; Neff et al. 1986, 1987; Environment Canada and Ministère de l'Environnement du Québec 1992). The only original document and data set used for validation that could not be obtained was the freshwater data set for the Trinity River (Dickson et al. 1989). However, most of the data for that study were available in Ingersoll and MacDonald (1999).

RESULTS AND DISCUSSION

Based on the evaluations of the consensus-based SECs for PCBs, their 30 underlying SQGs, the 2 SEC validation data sets, and the documents associated with the previous items, we identified 7 major factors that limit the usefulness of the SECs:

- Limitations of underlying SQGs,
- Inconsistent grouping of SQGs,
- Presence of co-occurring chemicals,
- Use of unrepresentative databases,
- Inferences related to predictive ability,
- Inferences related to causation, and
- Site-specific contradictions.

These factors are discussed in the following sections.

Limitation of underlying SQGs

Because the SECs are simple mathematical constructs based on their underlying SQGs, they are affected by the same limitations as the various SQGs on which they are based. In this section, the major limitations of the SQGs used to develop the SECs for PCBs are identified, along with the manner in which the original authors recommended that the SQGs be used in sediment quality assessments, if such recommendations were made. The objective of this section is not to provide an exhaustive review of the approaches used to develop the SQGs, but to estimate the degree of confidence that can be placed in each SQG, as well as to identify the kinds of uses for which each SQG is considered appropriate.

Freshwater and saltwater SLCs—Neff et al. (1986, 1987) developed the freshwater and saltwater SLCs for PCBs, as well as analogous values for 14 other chemicals or chemical groups. These studies were conducted largely as exploratory evaluations of whether the SLC approach had merit. Neff et al. (1986) noted that, because of the preliminary nature of the effort, the databases used in the study were not subjected to an intensive quality assurance review. Neff et al. (1987) subsequently recalculated the saltwater SLCs after conducting a quality assurance review of the underlying data, but did not revise the previous freshwater SLCs, including the one for PCBs. Neff et al. (1986) noted that the SLC approach assumes no cause-and-effect relationships between chemical concentrations and adverse biological effects. They also stated that the observed biological effects were likely the result of multiple chemicals, rather than the chemical of interest, resulting in SLCs that tend to be lower than they would be if

Table 1. Summary of the freshwater and saltwater SQGs used to develop the SECs for PCBs^a

SQG	Total PCBs (mg/kg dry wt)	SEC category ^b	Reference
Freshwater SQGs			
SLC	0.003	TEC	Neff et al. (1986)
ERL	0.050	TEC	Long and Morgan (1991)
ERM	0.40	MEC	
LEL	0.070	TEC	Persaud et al. (1993)
SEL	5.3	EEC	
MET	0.20	TEC	Environment Canada and Ministère de l'Environnement du Québec (1992)
TET	1.0	EEC	
TEL	0.034	TEC	Smith et al. (1996)
PEL	0.277	MEC	
TEL	0.032	TEC	Ingersoll et al. (1996)
ERL	0.050	TEC	
PEL	0.24	MEC	
ERM	0.73	MEC	
NEC	0.19	MEC	
LAET-WA	0.021	TEC	Cubbage et al. (1997)
PAET-WA	0.45	MEC	
HAET-WA	0.82	EEC	
Saltwater SQGs			
SLC	0.043	TEC	Neff et al. (1987)
Microtox AET-PS	0.13	TEC	Barrick et al. (1988)
Benthic AET-PS	1.0	MEC	
Oyster AET-PS	1.1	MEC	
Amphipod AET-PS	3.1	EEC	
Bivalve AET-CA	0.088	TEC	Becker et al. (1989)
Benthic AET-CA	0.36	MEC	
Amphipod AET-CA	0.96	EEC	
ERL	0.023	TEC	Long et al. (1995)
ERM	0.18	MEC	
TEL	0.022	TEC	MacDonald et al. (1996)
PEL	0.189	MEC	
SEC-SC	0.835	MEC	MacDonald (1997)

^a AET = apparent effects threshold; CA = California; EEC = extreme effect concentration; ERL = effects range-low; ERM = effects range-median; HAET = highest apparent effects threshold; LAET = lowest apparent effects threshold; LEL = lowest effect level; MEC = midrange effect concentration; MET = minimal effect threshold; NEC = no-effect concentration; PAET = probable apparent effects threshold; PCB = polychlorinated biphenyl; PEL = probable effect level; PS = Puget Sound = WA; SC = Southern California; SEC = sediment effect concentration; SEL = severe effect level; SLC = screening-level concentration; SQG = sediment quality guideline; TEC = threshold effect concentration; TEL = threshold effect level; TET = toxic effect threshold; WA = Washington State.

^b From MacDonald, DiPinto, et al. (2000).

Table 2. Summary of the freshwater and saltwater data sets used to validate the SECs for PCBs^a

Location	Nr. of samples	TOC (% dry wt)	Reference
Freshwater data sets			
Grand Calumet River and Indiana Harbor, IN	13	14	Hoke et al. (1993) USEPA (1996)
Lower Fox River and Green Bay, WI	13	4.5	Call et al. (1991)
Potomac River, DC	15	4.3	Schlekat et al. (1994) Velinsky et al. (1994) Wade et al. (1994)
Saginaw River, MI	9	1.9	USEPA (1996)
Trinity River, TX	72	ND	Dickson et al. (1989)
Upper Mississippi River, MN/MO	51	1.9	Kemble et al. (1998) USEPA (1996, 1997)
Waukegan Harbor, IL	22	3.9	USEPA (1996, 1999)
Saltwater data sets			
Hudson-Raritan Estuary and Newark Bay, NY/NJ	235	2.1	Rice et al. (1995) Long et al. (1996a)
Long Island Sound, NY	63	1.7	Wolfe et al. (1994)
Biscayne Bay, FL	105	2.5	Long et al. (1998)
Boston Harbor, MA	30	1.6	Long et al. (1996b)
Virginian Province (EMAP)	388	1.7	Strobel et al. (1995) Schimmel et al. (1994)
Narragansett Bay, RI	19	ND	Munns et al. (1991)
Puget Sound, WA	15	1.5	Pastorok and Becker (1990)
San Diego Bay, CA	119	ND	Fairey et al. (1996)
San Francisco Bay, CA	9	1.7	Chapman et al. (1987)
San Pedro Bay, CA	44	ND	Sapudar et al. (1994)
South Carolina and Georgia	63	ND	Long et al. (1998)
Tampa Bay, FL	61	2.5	Long et al. (1994)

^a EMAP = Environmental Monitoring and Assessment Program (US Environmental Protection Agency); ND = no data available; SEC = sediment effect concentration; TOC = total organic carbon.

only the chemical of interest were present. Although the SLC approach was also used to develop the freshwater LEL, SEL, MET, and TET (Environment Canada and Ministère de l'Environnement du Québec 1992; Persaud et al. 1993), it has rarely been used since that time. In addition, Von Stackelberg and Menzie (2002) described how statistical artifacts associated with use of presence and absence information on benthic macroinvertebrate taxa can provide misleading conclusions regarding relationships between chemical concentrations and biological effects.

Freshwater LEL, SEL, MET, and TET—The freshwater LEL, SEL, MET, and TET were developed for 35 chemicals by Persaud et al. (1993) and Environment Canada and Ministère de l'Environnement du Québec (1992), using the SLC approach originated by Neff et al. (1986), and a database composed of benthic macroinvertebrate community data from the Great Lakes. Environment Canada and Ministère de l'Environnement du Québec (1992) concluded that the SQGs are tools for

assessing sediment contamination that provide indications of the necessity for conducting more detailed analyses. They also specified that when the MET is exceeded, sediment toxicity tests should be conducted to determine whether the affected sediments are toxic. Persaud et al. (1993) concluded that if the SEL is exceeded, sediment toxicity tests are required to assess whether the sediment is acutely toxic.

Great Lakes SQGs—Ingersoll et al. (1996) developed 5 kinds of SQGs for PCBs (referred to herein as the Great Lakes SQGs), and for 26 other chemicals, for use in evaluating sediment samples from the Great Lakes, including a threshold effect level (TEL) and probable effect level (PEL) analogous to those developed by Smith et al. (1996) and MacDonald et al. (1996), an effects range-low (ERL) and effects range-median (ERM) analogous to those developed by Long and Morgan (1991) and Long et al. (1995), and a no-effect concentration (NEC) analogous to the AETs developed by Barrick et al. (1988). Although the sample sizes used to

calculate these SQGs for 21 of the 26 other chemicals were greater than 50, only 29 sediment samples were used to calculate the SQGs for PCBs. In addition, of the 29 PCB concentrations used to calculate the SQGs, only 9 were greater than the TEC of 0.04 mg/kg. Furthermore, because only 5 of the 29 sediment samples were designated as toxic after an initial screening step, the 5 Great Lakes SQGs were developed using only 5 data points where biological effects were associated with detected PCB concentrations. This value is well below the minimum level of 20 specified by Smith et al. (1996) and MacDonald et al. (1996) for development of the freshwater and saltwater TELs and PELs. Because Ingersoll et al. (1996) arbitrarily specified that a minimum of 5 toxic samples had to be present for the SQGs to be calculated for a particular chemical, the SQGs for PCBs were calculated using the minimum amount of effects data specified by the authors.

The small number of samples used to develop the Great Lakes SQGs for PCBs adds considerable uncertainty to the validity of these SQGs. This uncertainty is highlighted by the fact that the NEC of 0.19 mg/kg is less than the PEL of 0.24 mg/kg, and almost 4 times as low as the ERM of 0.73 mg/kg. Because the NEC is defined as the highest no-effect value in a database, it would typically be greater than the PEL and ERM. The fact that the NEC is less than the PEL and ERM for the Great Lakes database is likely an artifact of the small number of samples used to develop those SQGs.

Ingersoll et al. (1996) concluded that the Great Lakes SQGs should be used to predict the potential for sediment toxicity and to provide guidance for identifying sites that require further investigation. The authors also concluded that the SQGs should not be used independently to establish cleanup levels for sediments. Finally, Ingersoll et al. (1996) cautioned that the Great Lakes SQGs represent concentrations that are associated with adverse effects, but do not necessarily cause the effects. They noted that because field-collected sediments typically contain complex mixtures of chemicals, additional information is needed to identify the specific chemicals that were actually responsible for any observed toxicity.

Freshwater TEL and PEL—Smith et al. (1996) developed the freshwater TEL and PEL for PCBs, as well as analogous values for 22 other chemicals or chemical groups. The authors also conducted an evaluation of the predictive ability of the SQGs for the 23 chemicals. They found that the prevalence of toxicity above the PEL for PCBs was only 50%, which is well below the minimum value of 75% used by others (e.g., Long et al. 1995, 1998; MacDonald, Ingersoll, et al. 2000) to identify reliable PELs and ERMs. This indicates that the PEL has limited predictive value for identifying a station as toxic or nontoxic.

Smith et al. (1996) cautioned that the database used to develop the freshwater SQGs was biased toward reference and slightly contaminated areas, and that data from contaminated sites were generally underrepresented. They therefore recommended that the underlying database of the freshwater TEL and PEL be refined and expanded in the future to represent a broader range of contaminated conditions. Smith et al. (1996) concluded that their SQGs are screening tools that provide one type of information for sediment assessments, and emphasized the importance of using these values in conjunction with other kinds of information, such as biological assessments.

Freshwater ERL and ERM—Long and Morgan (1991) developed the freshwater ERL and ERM for PCBs, as well as analogous values for 30 other chemicals or chemical groups. Although the database used to calculate the ERL and ERM included data from both freshwater and saltwater sediments, these values have been used primarily to assess freshwater sediments, following the development of the saltwater ERL and ERM by Long et al. (1995). Long and Morgan (1991) did not conduct quantitative evaluations of the reliability of the SQGs. The authors stated that the various ERLs and ERMs were informal guidelines that were not intended for use in regulatory decisions or any other similar applications, and that they should not be construed as National Oceanic and Atmospheric Administration standards or criteria. The authors concluded further that more data are needed to reduce the uncertainty associated with most of the SQGs, including that for PCBs.

Saltwater TEL and PEL—MacDonald et al. (1996) developed the saltwater TEL and PEL for PCBs, as well as analogous values for 33 other chemicals or chemical groups. The authors also conducted reliability evaluations of the SQGs for the 34 chemicals. They found that the prevalence of toxicity above the PEL for PCBs was only 55%, which is similar to the value of 50% found by Smith et al. (1996) for the freshwater PEL for PCBs, and well below the minimum value of 75% used by others (e.g., Long et al. 1995, 1998; MacDonald, Ingersoll, et al. 2000) to identify reliable PELs and ERMs. This indicates that the use of the PEL to identify a station as toxic or nontoxic has limited predictive value. MacDonald et al. (1996) concluded that the overall reliability of the PEL for PCBs was low. In fact, the PELs for nearly all (i.e., 31) of the other 33 chemicals evaluated by MacDonald et al. (1996) had higher reliability scores.

MacDonald et al. (1996) cautioned that the weight-of-evidence approach used to develop the saltwater TEL and PEL does not fully support quantitative evaluation of cause-and-effect relationships between chemical concentrations and adverse biological effects, because it is largely based on associations between chemical concentrations and biological effects. They noted further that various other factors, including co-occurring chemicals, could have been responsible for the observed effects. Finally, the authors concluded that the saltwater TEL and PEL should not be used as stand-alone sediment quality criteria.

Saltwater ERL and ERM—Long et al. (1995) developed the saltwater ERL and ERM for PCBs, as well as analogous values for 27 other chemicals or chemical groups. The authors also conducted an evaluation of the predictive ability of the SQGs for the 28 chemicals, and concluded that the prevalence of effects was relatively low at PCB concentrations greater than the ERM. Because the prevalence was only 51%, it was well below the minimum value of 75% specified by the authors as indicative of reliable ERMs. The ERMs for nearly all (i.e., 24) of the other 27 chemicals evaluated by Long et al. (1995) were found to be more reliable than the ERM for PCBs.

The authors concluded that the saltwater ERLs and ERMs should be used as informal screening tools in environmental assessments, and that they were not intended to preclude the use of toxicity tests or other measures of biological effects.

Long et al. (1998) evaluated the predictive ability of the saltwater ERM, as well as the saltwater PEL of MacDonald et al. (1996), using a database that consisted of 1,068 samples of saltwater sediments collected from throughout the United

States. They found that the predictive ability of the ERM for PCBs was only 51% (i.e., the lowest of the 27 ERMs evaluated), and the predictive ability of the PEL was only 49% (i.e., the 3rd lowest value of the 31 PELs evaluated). In addition to being among the lowest values of predictive ability observed, both values were well below the minimum value of 75% specified by the authors as indicative of reliable PELs and ERMs. Long et al. (1998) noted that these SQGs were prepared as informal (nonregulatory) benchmarks and were not based on experiments in which causality was determined, and that SQGs in general are most useful when accompanied by data from in situ biological analyses, sediment toxicity tests, or other interpretive tools.

Freshwater and saltwater AETs—The AET approach for developing SQGs is based largely on data from specific water bodies or geographic regions. The resulting AETs, therefore, have a high degree of site-specific relevance, but questionable relevance for use on a larger (e.g., national) scale. In their review of the AET approach, USEPA (1989) concluded that it contained sufficient scientific merit to establish SQGs for use at specific sites, but that it should not be used to develop general, broadly applicable SQGs. Separate sets of AETs have been developed for Puget Sound sediments (Barrick et al. 1988), saltwater sediments of California (Becker et al. 1989), and freshwater sediments of Washington State (Cubbage et al. 1997). Because the reliability of each of these sets of AETs has been determined for suites of chemicals (i.e., rather than for individual chemicals), no evaluations of the independent reliability of the AETs for PCBs have been conducted. Although the Puget Sound AETs have been promulgated as sediment quality standards for Washington State (Chapter 173–204 WAC), the state guidance specifies that sediment toxicity tests or evaluations of benthic macroinvertebrate communities can be used to confirm the presence of sediment toxicity when the standards are exceeded.

Summary of SQG limitations—In summary, the information presented above indicates that many of the original developers of SQGs for PCBs recognized the limitations of the SQGs, and recommended that they be used primarily as screening values, with their predictions being confirmed using site-specific biological evaluations. Many of the developers also recognized that the various SQGs were based on correlative, rather than cause-and-effect, information. Finally, the reliability of the freshwater and saltwater PELs and ERMs for PCBs in identifying the presence of sediment toxicity was relatively low (i.e., approximately equal to a coin toss) for studies that quantitatively evaluated predictability of those SQGs (i.e., Long et al. 1995, 1998; MacDonald et al. 1996; Smith et al. 1996).

Inconsistent grouping of SQGs

MacDonald, DiPinto, et al. (2000) stated that because the SECs for PCBs represent the central tendency of their underlying SQGs, they reconcile differences among the SQGs. In this section, the 3 groups of SQGs are evaluated to determine the degree to which the 3 SECs can be considered “consensus” values. Emphasis was placed on determining the degree to which the SQGs within each group are similar to each other, whether the SQGs were grouped correctly, and whether the groups contain redundant SQGs that provide an unwarranted enhancement of apparent agreement among the various SQGs.

Numerous authors, including MacDonald, DiPinto, et al. (2000), have used a factor of 3 to identify SQGs that are considered comparable (Long et al. 1995; MacDonald et al. 1996; Smith et al. 1996; MacDonald, Ingersoll, et al. 2000). If this same quantitative criterion is applied to the SEC categories used by MacDonald, DiPinto, et al. (2000), all of the categories span a range greater than a factor of 3, indicating that many of the SQGs within each category should not be considered comparable. For example, the minimum and maximum PCB concentrations for the TEC, MEC, and EEC differ by factors of 67, 6.1, and 6.5, respectively.

Inspection of the classification of the various freshwater and saltwater SQGs used to develop the SECs for PCBs shows that 2 freshwater SQGs were misclassified as MECs, when according to their definitions, they should have been classified as EECs: The Great Lakes NEC and the probable AET for Washington State (PAET-WA). As defined by their developers, both of these SQGs are representative of concentrations above which adverse effects are always or nearly always observed and, as such, should be classified as EECs. MacDonald, Ingersoll, et al. (2000) made this same distinction for the NEC and PAET-WA in describing why those 2 values were not used to develop the consensus-based PECs for PCBs and other chemicals. If these 2 SQGs are correctly classified as EECs, the range of those values would extend from 0.19 to 5.3 mg/kg, and the level of apparent agreement among those values would decline further, because the maximum and minimum PCB concentrations would differ by a factor of 28, rather than 6.5.

In deriving the consensus-based TECs and PECs for 28 chemicals in freshwater sediments, MacDonald, Ingersoll, et al. (2000) avoided redundancies when compiling the various underlying SQGs. For example, the authors stated that the Great Lakes ERLs and ERMs of Ingersoll et al. (1996) were not used, because they were developed from the same data that were used to derive the Great Lakes TELs and PELs (i.e., from several areas of concern in the Great Lakes). However, this criterion was not used in calculating the SECs for PCBs, where a number of the underlying SQGs were developed from the same databases and are therefore redundant, using the criteria of MacDonald, Ingersoll, et al. (2000). The inclusion of these values in the set of SQGs used to develop the SECs provides an unwarranted enhancement in the apparent degree of agreement among the various SQGs. The various redundancies are as follows:

- The Great Lakes ERL and ERM for PCBs were used along with the Great Lakes TEL and PEL, even though this redundancy was identified and avoided by MacDonald, Ingersoll, et al. (2000).
- The freshwater MET and TET and the freshwater LEL and SEL are largely redundant, because they were developed using the same SLC approach that was applied to the same database on benthic macroinvertebrate communities from the Great Lakes, with only minor modifications of the critical percentages of the benthic macroinvertebrate species they were designed to protect (i.e., 15% and 90%, as opposed to 5% and 95%, respectively).
- The freshwater PAET-WA and highest AET for Washington State are largely redundant, because they are both based on the same database collected in Washington State for the amphipod *Hyaletta azteca*, with only minor

modifications of the critical no-effect level (i.e., 95th percentile and maximum, respectively).

- The saltwater ERL and ERM are largely redundant with the saltwater TEL and PEL, because they were developed using the same biological effects database for sediments, and this redundancy is consistent with that identified by MacDonald, Ingersoll, et al. (2000) for the analogous Great Lakes ERL, ERM, TEL, and PEL.

The result of the redundancies described above is that 7 of the 30 SQGs used to develop the SECs for PCBs (i.e., 23% of the SQGs) are considered redundant with other SQGs, and therefore should not have been used in the development of the SECs. The most obvious examples of this redundancy are the saltwater TEL (0.022 mg/kg) and ERL (0.023 mg/kg), as well as the saltwater PEL (0.189 mg/kg) and ERM (0.180 mg/kg), because the 2 sets of values are nearly equal.

In summary, the evaluation of the SQG groupings used by MacDonald, DiPinto, et al. (2000) showed that many SQGs within the 3 SEC groups were dissimilar, misclassified, and/or largely redundant with other SQGs in the groups. These factors considerably reduce the apparent degree of agreement of SQGs within the 3 groups, and call into question the conclusion of MacDonald, DiPinto, et al. (2000) that the SECs reconcile differences among the SQGs.

Limitations related to co-occurring chemicals

As discussed previously, many of the developers of the SQGs for PCBs recognized that most contaminated sediments contain complex mixtures of chemicals that can increase the difficulty of developing SQGs and of determining which chemicals are actually responsible for any observed toxicity (e.g., Neff et al. 1986; Long et al. 1995; Ingersoll et al. 1996; MacDonald et al. 1996). In addition, Long et al. (1998) concluded that their nationwide experience indicates that toxicants often covary with each other to a large degree. As discussed previously, MacDonald, DiPinto, et al. (2000) evaluated the predictive ability of the SECs for PCBs using 2 validation data sets, one based on freshwater studies and the other based on saltwater studies. In this section, the degree to which these predictive evaluations may have been confounded by the presence of co-occurring chemicals is evaluated, by examining the conclusions of the authors of the individual studies that make up the 2 SEC validation databases, if such conclusions were made.

Freshwater studies—

- Grand Calumet River and Indiana Harbor—Hoke et al. (1993) concluded that sediments contained a multitude of chemicals, and that toxicity appeared to be related primarily to ammonia, metals, PAHs, and petroleum hydrocarbons. Ingersoll and MacDonald (1999) also evaluated sediments in the Grand Calumet River and concluded that the chemicals causing or contributing to sediment injury include ammonia, metals, phenol, PAHs, PCBs, and pesticides. However, the authors cautioned that because all of these chemicals are present at concentrations sufficient to cause or substantially contribute to sediment toxicity, it was difficult to assign a relative priority to them.
- Indiana Harbor and Saginaw River—USEPA (1993) concluded that a broad range of contaminant concentrations were measured in the sediments, but because of the potential for contaminant interactions, the observed sediment toxicity could not be attributed to any specific

chemical. The authors concluded further that sediment toxicity was related to the sum of the concentrations of metals and PAHs.

- Potomac River—Schlekat et al. (1994) concluded that sediment toxicity was associated most strongly with ammonia, PAHs, PCBs, and chlordane.
- Waukegan Harbor—USEPA (1999) concluded that sediment toxicity was likely associated with elevated concentrations of metals, PCBs, and PAHs. For the same water body, Ingersoll and Nelson (1990) concluded that each sediment sample they evaluated contained a complex matrix of measured inorganic contaminants, organic contaminants, and numerous unidentified compounds, and that laboratory toxicity studies that test single compounds spiked into the sediment would be required to more directly determine the specific contaminants that may have caused a toxic response.

Saltwater habitats—

- Hudson–Raritan Estuary—Rice et al. (1995) found that sediment toxicity was associated with concentrations of 6 metals, PAHs, and PCBs. However, the authors stated that conclusive identification of specific causal relationships between contaminants and biological effects was not possible, primarily because of the complexity of the chemical mixtures in the sediments. Long et al. (1996a) also evaluated sediments in the Hudson–Raritan Estuary and found that sediment toxicity in various parts of the estuary correlated with concentrations of metals, PAHs, PCBs, pesticides, and dioxins. However, the authors noted that because many of these chemicals covaried with each other, the correlations alone did not provide great insight into the potential causes of the observed toxicity.
- Narragansett Bay—Munns et al. (1991) concluded that the major risks to benthic macroinvertebrate communities were related primarily to selected metals, PAHs, PCBs, and pesticides.
- San Pedro Bay—Sapudar et al. (1994) found that sediment toxicity to amphipods was correlated most strongly with concentrations of 3 metals and PAHs. However, the authors noted that because the study was not designed to determine cause–effect relationships, the data were useful only for evaluating associations between chemical concentrations and biological effects.
- Boston Harbor—Long et al. (1996b) concluded that the chemicals that most likely contributed to sediment toxicity were ammonia, several metals, PAHs, PCBs, and other chlorinated hydrocarbons. However, the authors concluded that a highly complex toxicity identification evaluation procedure would be required to specifically identify which chemicals caused the observed toxicity.
- Long Island Sound—Wolfe et al. (1994) concluded that although sediment toxicity tended to correlate with elevated chemical concentrations in sediments, the various chemicals covaried relatively strongly, so that toxicity could not be readily attributed to any particular chemical.
- Tampa Bay—Long et al. (1994) concluded that the concentrations of ammonia, numerous metals, PAHs, PCBs, and pesticides were highly correlated with sediment toxicity. However, they also noted that those correlations did not establish cause–effect relationships, and that considerably more research would be needed to

establish the causes of the observed toxicity. They noted further that the most toxic sediments frequently contained high concentrations of mixtures of chemicals, any one of which could have caused or contributed to the toxicity.

In summary, the information provided above indicates that sediments evaluated in most of the individual studies that make up the freshwater and saltwater SEC validation data sets were contaminated with numerous chemicals, in addition to PCBs. Most of the studies also recognized that conclusive determinations of which chemicals caused the observed sediment toxicity would require additional kinds of evaluations. These results indicate that use of these data sets to determine the predictive ability of the SQGs for a single chemical or group of chemicals such as PCBs is questionable, particularly if the results of this analysis are used to suggest that PCBs were the cause of any observed toxicity. In a number of cases, PCBs were not even identified by the original authors as a potential major contributor to sediment toxicity, even though they were present in the sediment samples.

Unrepresentative databases

Most of the databases used to generate the SQGs for PCBs or to validate the SECs were composed primarily of low concentrations of PCBs, with relatively few samples having elevated concentrations of PCBs (Table 3). It is likely that this pattern would result in the calculation of SQGs that are lower than the SQGs that would be calculated using databases that included greater percentages of elevated PCB concentrations. This result would likely be most apparent for those SQGs that are based on the overall distributional properties of databases, such as median values and various other percentiles of the data. Those SQGs include any based on the methodologies used to calculate an ERL, ERM, TEL, PEL, or SEC. As discussed previously, Smith et al. (1996) recognized that the database used to develop the freshwater LEL and PEL was generally biased toward reference and slightly contaminated areas, and concluded that the database should be expanded to represent a broader range of contaminated conditions.

By contrast with the SQGs based on the overall distributional characteristics of their underlying databases, AETs and NECs are usually based only on the maximum no-effect value in a database. However, a single exception is the PAET-WA, which is calculated as the 95th percentile of the no-effects data. The AETs and NECs should therefore be less affected by the overall distributions of their underlying databases. However, even the maximum no-effect values can be biased toward low values if there are few elevated PCB concentrations in a data set, particularly when concentrations of other chemicals covary with PCBs and cause or contribute to toxicity at elevated PCB concentrations, making no-effect data rare at elevated PCB concentrations.

For all but one of the SQG databases presented in Table 3, PCB concentrations less than 0.5 mg/kg accounted for more than 68% of the underlying databases. The only exception was the database used to calculate the SEC for Southern California. This pattern was particularly strong for the saltwater ERL, ERM, TEL, and PEL, the Great Lakes SQGs, and the California AETs, for which PCB concentrations in 85% or more of the underlying databases were less than 0.5 mg/kg. Taken together, those SQGs account for 12 of the 30 SQGs (40%) used by MacDonald, DiPinto, et al. (2000) to calculate the SECs for PCBs.

With respect to the freshwater and saltwater data sets used by MacDonald, DiPinto, et al. (2000) to validate the SECs for PCBs, both were composed primarily of low PCB concentrations, with 69% to 89% of those databases having concentrations less than 0.5 mg/kg (Table 3). The saltwater validation data set was particularly affected, because almost 90% of the data set was composed of PCB concentrations less than 0.5 mg/kg, and concentrations in only 2.4% of its 1151 samples exceeded the EEC of 1.7 mg/kg for PCBs.

The results of the above analyses indicate that, in general, the potential toxicity of PCBs at sediment concentrations greater than the EEC has not been well characterized in past studies, and has not been well represented in the databases used in the development and testing of available SQGs for PCBs. For example, the combined databases used to calculate the freshwater ERL, ERM, TEL, and PEL, the saltwater ERL, ERM, TEL, PEL, and the SEC for Southern California contain more than 450 samples, but PCB concentrations exceed the EEC in only 54 of those samples. This relative scarcity of information on elevated concentrations of PCBs adds uncertainty to the validity of existing SQGs. It is therefore possible that if the SQG databases contained a better representation of elevated PCB concentrations, the resulting SQGs would be considerably greater than they are at present. For example, 24 of the 54 samples (44%) with PCB concentrations exceeding the EEC in the combined SQG databases identified above are nontoxic. This fact clearly calls into question the validity of the EEC as an indicator of extreme effects.

Inferences related to predictive ability

As discussed previously, MacDonald, DiPinto, et al. (2000) determined the predictive ability of the SECs for PCBs in identifying the presence or absence of sediment toxicity using the freshwater and saltwater SEC validation data sets. The authors used criteria adapted from those of Long et al. (1998) to determine whether the SECs were reliable. Using those criteria, the MEC and EEC were considered reliable if their predictive abilities exceeded 50% and 75%, respectively. The observed predictive abilities of the MEC for the freshwater and saltwater validation data sets were 68.3% and 55.9%, respectively, and the predictive abilities of the EEC for the 2 validation data sets were 82.5% and 85.7%, respectively. MacDonald, DiPinto, et al. (2000) therefore concluded that both the MEC and EEC were accurate tools for predicting the presence or absence of toxicity in field-collected sediments.

Although the evaluations of predictive ability conducted by MacDonald, DiPinto, et al. (2000) document that the values observed for the MEC and EEC exceeded the specified performance criteria of 50% and 75%, respectively, these evaluations do not demonstrate that PCBs are causally related to the observed toxicity in the SEC validation data sets. As discussed previously, the sediments evaluated in most of the individual studies that make up the SEC validation data sets contained numerous co-occurring chemicals, such that PCBs either were not or could not be identified as the cause of any observed toxicity by the original investigators.

Aside from issues related to causality, the predictive abilities of the MEC and EEC observed by MacDonald, DiPinto, et al. (2000) cannot be placed into a context relative to most other chemicals, because comparable SECs have been developed only for total PAHs (Swartz 1999). If SECs were available for many of the other chemicals found in the studies

Table 3. Percentages of observations in various concentration intervals of PCBs in selected SQG databases and the SEC validation data sets^a

Total PCBs (mg/kg dry wt)	SQG databases											SEC validation data sets	
	Freshwater					Saltwater						Freshwater	Saltwater
	ERL/ERM	TEL/PEL	Great Lakes SQGs	Washington State AETs	ERL/ERM TEL/PEL	Southern California SEC	Puget Sound AETs	California AETs	California AETs				
<0.5	69	76	88	76	85	43	76	89				69	89
0.5–0.9	2	6	8	14	5	12	14	4				8	6
1.0–1.4	13	3	0	3	4	20	3	2				2	2
1.5–1.9	2	1	0	2	1	4	2	2				1	1
2.0–2.4	2	5	0	2	1	2	2	1				2	<1
2.5–2.9	2	1	0	<1	1	2	<1	<1				1	<1
3.0–3.4	2	3	0	1	1	2	1	0				2	<1
3.5–3.9	0	1	0	1	0	0	1	0				1	<1
4.0–4.4	2	0	0	0	0	2	0	0				2	<1
4.5–4.9	0	0	0	0	0	2	0	<1				3	<1
≥5.0	8	5	4	1	3	10	1	1				10	<1

^a AET = apparent effects threshold; ERL = effects range-low; ERM = effects range-median; PCB = polychlorinated biphenyl; PEL = probable effect level; SEC = sediment effect concentration; SQG = sediment quality guideline; TEL = threshold effect level.

that make up the SEC validation data sets, it is possible that their predictive abilities would be greater than the values found for PCBs. For example, MacDonald, Ingersoll, et al. (2000) calculated predictive abilities for the 27 chemicals that have PECs and found that the predictive abilities of all but 3 of the PECs were greater than the value of 82.3% found for the PEC for PCBs.

In summary, the predictive abilities of the MEC and EEC observed by MacDonald, DiPinto, et al. (2000) do not demonstrate that PCBs are causally related to the observed toxicity in the SEC validation data sets, or that those SECs are necessarily the best predictors of toxicity in those data sets. The meaning of the predictive abilities calculated by MacDonald, DiPinto, et al. (2000) is therefore uncertain.

Inferences related to causation

MacDonald, DiPinto, et al. (2000) compared the SECs for PCBs with 2 sets of SQGs developed using the EqP approach. The 2 sets of EqP-based SQGs included the freshwater and marine chronic SQGs for benthic aquatic life in New York State (USA; NYSDEC 1999) and the SQG developed by Bolton et al. (1985) to evaluate sediment quality conditions in freshwater and saltwater sediments throughout the United States. Assuming 1% organic carbon, the New York State Department of Environmental Conservation values for freshwater and saltwater sediments convert to 0.19 and 0.41 mg/kg on a dry-weight basis, and the value derived by Bolton et al. (1985) converts to 0.07 mg/kg on a dry-weight basis. MacDonald, DiPinto, et al. (2000) concluded that because the TEC of 0.04 mg/kg and the MEC of 0.40 mg/kg were similar to the EqP-based SQGs, confidence is increased in the 2 SECs. The authors also state that, to the extent that such chronic toxicity thresholds are causally based, the consensus-based SECs also reflect the concentrations of PCBs that are likely to cause, or substantially contribute to, sediment toxicity.

A review of the basis for each of the EqP-based SQGs described above showed that none of them was based on sediment toxicity to benthic macroinvertebrates. This fact was also noted by Fuchsman et al. (2006). The freshwater SQG of 0.19 mg/kg developed by NYSDEC (1999) and the SQG of 0.07 developed by Bolton et al. (1985), and based on JRB (1984), were derived from the final residue value (FRV) identified by USEPA (1980) for protection of wildlife from consumption of contaminated fish. That FRV was developed using the lowest permissible tissue concentration for protection of mink (*Mustela vison*; Platonow and Karstad 1973) and a bioconcentration factor based on salmonid fishes. The saltwater SQG of 0.41 mg/kg developed by NYSDEC (1999) was derived from the FRV identified by USEPA (1980) for protection of the marketability of fish for human consumption. That FRV was developed using the US Food and Drug Administration's historical action level of 5.0 mg/kg for PCBs and a bioconcentration factor based on fish species consumed by humans. Given that all 3 of the EqP-based SQGs are based on bioaccumulation of PCBs in fishes, they have no relevance for assessing the validity of the SECs for use as indicators of sediment toxicity to benthic macroinvertebrates. As noted by MacDonald, DiPinto, et al. (2000), the SECs do not consider the potential for bioaccumulation of PCBs in aquatic organisms. The conclusion of MacDonald, DiPinto, et al. (2000) that the EqP-based SQGs increase confidence in the SECs is therefore invalid.

MacDonald, DiPinto, et al. (2000) also describe spiked-sediment toxicity tests conducted by Swartz et al. (1988), where the marine amphipod *Rhepoxynius abronius* was exposed for 10 d to various concentrations of Aroclor 1254, and an LC50 of 8.8 mg/kg was determined. To convert the 10-d LC50 value to a chronic value, MacDonald, DiPinto, et al. (2000) applied an acute to chronic ratio of 11 that was reported for the freshwater amphipod *Gammarus pseudolimnaeus* by USEPA (1980), and derived an estimated chronic toxicity value of 0.8 mg/kg. MacDonald, DiPinto, et al. (2000) concluded that, because the MEC and EEC for PCBs were comparable to the estimated chronic toxicity value, confidence is increased that the 2 SECs are representative of the PCB concentrations that are likely to cause or substantially contribute to sediment toxicity.

A review of the spiked-sediment study conducted by Swartz et al. (1988) showed that the sediments used in the toxicity tests had an unusually low organic content, compared to conditions commonly found in contaminated aquatic sediments in the field. The sediments were collected from Yaquina Bay (OR, USA), a relatively uncontaminated marine embayment. The organic content of the sediments was estimated by Swartz et al. (1988) as percent total volatile solids (TVS), which was measured as percent weight loss after drying the sediments at 550 °C for 1 h. The authors stated that total organic carbon (TOC) typically represents 17.3% of the TVS content of sediments at the Yaquina Bay sediment collection site. The value of TVS measured in the sediments used for the toxicity study ranged from 1.28% to 1.49%, with a mean value of 1.39%. The mean TOC content of the sediments can therefore be estimated as 0.24%. This TOC value is unusually low relative to the values typically found in contaminated aquatic sediments. For example, the overall mean TOC content in the water bodies included in the freshwater SEC validation data set used by MacDonald, DiPinto, et al. (2000) was approximately 5%, with a range of 1.9% to 14% for specific water bodies (Table 2). For the saltwater SEC validation data set, overall mean TOC content was approximately 2%, with a range of 1.5% to 2.5% for specific water bodies (Table 2).

Swartz et al. (1988) acknowledged that because of the low amount of TVS (and therefore low TOC content) in the test sediment, the concentrations of test chemicals in porewater during the toxicity testing may have been high relative to contaminated sites. The implication of this experimental limitation is that the LC50 value for Aroclor 1254 determined by Swartz et al. (1988) may underestimate the LC50 values that would be found using sediments with TOC contents comparable to those found at most contaminated sites. A 2nd experiment conducted by Swartz et al. (1988) supports this conclusion.

In the 2nd experiment, Swartz et al. (1988) evaluated amphipod toxicity at 2 TVS concentrations: 1.30% and 1.72%. The estimated TOC concentrations associated with the TVS values would be 0.22% and 0.30%, respectively. The authors evaluated concentrations of Aroclor 1254 of 4.8 and 2.1 mg/kg with the lower TVS, and found amphipod survival of 89% and 95%, respectively. They also evaluated concentrations of Aroclor 1254 that were approximately twice as high (i.e., 9.4 and 4.6 mg/kg) with the higher TVS level, and found that amphipod survival increased to 91% and 97%, respectively, despite the higher concentrations of Aroclor 1254. These results show that the toxicity of Aroclor 1254

was reduced with increased TVS levels, and presumably TOC concentrations as well. It is noteworthy that although the higher concentration of 9.4 mg/kg tested with a TVS level of 1.72% approximated the LC50 value of 8.8 mg/kg determined in the 1st experiment with a TVS level of 1.49%, amphipod survival (i.e., 91%) was substantially greater than 50%. These results provide additional evidence that the toxicity of Aroclor 1254 declined with increasing levels of TVS, and presumably TOC.

If the EqP approach (DiToro et al. 1991) is used to estimate the LC50 value that would have been observed using sediments with a TOC content of 2% (i.e., the approximate mean and minimum value found in the saltwater and freshwater SEC validation data sets, respectively), the estimated LC50 value would be approximately 73 mg/kg. Using the acute to chronic ratio of 11 employed by MacDonald, DiPinto, et al. (2000), the estimated chronic toxicity value would be approximately 6.6 mg/kg, which is considerably greater than the EEC of 1.7 mg/kg. These results, and the information presented above, indicate that the estimated chronic toxicity value of 0.8 mg/kg derived by MacDonald, DiPinto, et al. (2000) is unusually low as a result of the low TOC content of the test sediments used by Swartz et al. (1988). Therefore, the conclusions of MacDonald, DiPinto, et al. (2000) with respect to the implications of the agreement between the estimated chronic value of 0.8 mg/kg and the SECs are invalid with respect to their applicability to most contaminated sediments.

In summary, the comparisons made by MacDonald, DiPinto, et al. (2000) between the SECs and SQGs based on the EqP approach do not indicate that the SECs for PCBs are causally related to sediment toxicity, because the SQGs are based on bioaccumulation in fishes rather than toxicity to benthic macroinvertebrates. In addition, the LC50 for PCBs determined by Swartz et al. (1988) has limited applicability to sediments in most contaminated aquatic environments, because it was determined using sediments with an unusually low TOC content, estimated as 0.24%.

Site-specific contradictions

As discussed previously, most contaminated sediments contain elevated concentrations of multiple chemicals, making it difficult to reach definitive conclusions regarding the toxicity of single chemicals. However, 2 of the freshwater validation data sets used by MacDonald, DiPinto, et al. (2000) were collected in water bodies where PCBs were the primary chemicals found at elevated concentrations: Fox River–Green Bay (WI, USA; Call et al. 1991) and Waukegan Harbor (IL, USA; USEPA 1999). In this section, the results of those 2 studies are used to demonstrate how application of the SECs for PCBs to site-specific conditions can result in erroneous predictions of sediment toxicity. In addition, the results of those studies are used to evaluate potential concentration–response relationships between sediment toxicity and concentrations of PCBs.

Concentrations of PCBs at the 13 stations sampled in Fox River–Green Bay by Call et al. (1991) exhibited a relatively wide range (i.e., 0.09–6.8 mg/kg), with concentrations at 11 stations exceeding the MEC and concentrations at 7 of those stations also exceeding the EEC. Sediment toxicity (i.e., based on percent survival) was evaluated using the 10-d chironomid test with *Chironomus riparius* and the 10-d mayfly test with *Hexagenia limbata*. Although comparison of PCB concen-

trations with the MEC and EEC predicts that 11 and 7 stations should be toxic, respectively, only 1 station for each sediment toxicity test was identified as toxic in the SEC validation data set (MESL 1999). Therefore, the site-specific predictive ability of the MEC was only 9% for each toxicity test, and the site-specific predictive ability of the EEC was only 14% for each test. These values are well below the target accuracies for those 2 SECs of 50% and 75%, respectively, as specified by MacDonald, DiPinto, et al. (2000), and demonstrate that use of the SECs to predict the presence of sediment toxicity in Fox River–Green Bay would result in the erroneous identification of numerous sediment samples as toxic.

For Waukegan Harbor, concentrations of PCBs at the 18 stations sampled by USEPA (1999) also exhibited a relatively wide range (i.e., 0.9–8.9 mg/kg), with concentrations exceeding the MEC and EEC at 18 and 17 stations, respectively. Sediment toxicity was evaluated using the 28-d amphipod test with *Hyalella azteca*. According to USEPA (1999), significant reductions ($p \leq 0.05$) in survival and growth were found at 6 and 17 stations, respectively, based on statistical comparisons with results for the negative controls (i.e., a formulated sediment). However, USEPA (1999) noted that mean survival in 4 of the 6 samples that were significantly different ($p \leq 0.05$) from the mean value of 100% found in the formulated sediment ranged from 84% to 88%, and were greater than the minimum negative control value of 80% for test acceptability. Mean amphipod survival in the 2 remaining samples was 79% and 66%. USEPA (1999) concluded that the sediments from Waukegan Harbor were generally not lethal to the amphipods.

Although significant reductions ($p \leq 0.05$) in amphipod growth were found at 17 of the 18 stations sampled in Waukegan Harbor, the fact that the statistical comparisons were made with results from negative controls composed of formulated sediments, rather than sediment from a reference area, may have unduly influenced the results. Because the formulated sediment had few similarities to the sediments from Waukegan Harbor, the meaning of the apparent reductions in growth, and survival as well, is uncertain. For example, concentrations of TOC (1.6%) and fine-grained material (i.e., silt and clay; 27%) in the formulated sediment were considerably lower than the values found in the test sediments (i.e., 2.2%–7.8% and 47%–98%, respectively). In addition, at the initiation of toxicity testing, the concentration of dissolved oxygen in porewater of the formulated sediment (9.6 mg/L) was substantially greater than concentrations in the test sediments (1.6–5.2 mg/L), and concentrations of total ammonia (0.39 mg/L) and total sulfide (<1 µg/L) in porewater of the formulated sediment were considerably lower than concentrations in the test sediments (4.3–63 mg/L and 4–330 µg/L, respectively). Given these relatively large dissimilarities between many of the characteristics of the formulated and test sediments, it is possible that those differences alone resulted in subtle differences in amphipod growth between the 2 kinds of sediment. Furthermore, although the range of observed mean amphipod lengths in the test sediments (3.5–4.3 mm) was lower than the mean value of 4.6 mm found in the formulated sediment, they were all greater than the performance level for negative control sediments (3.2 mm) used by USEPA (2000) to evaluate laboratory performance of the 28-d *H. azteca* test.

Based on the information on Waukegan Harbor sediments presented above, it is clear that lethal toxicity was not

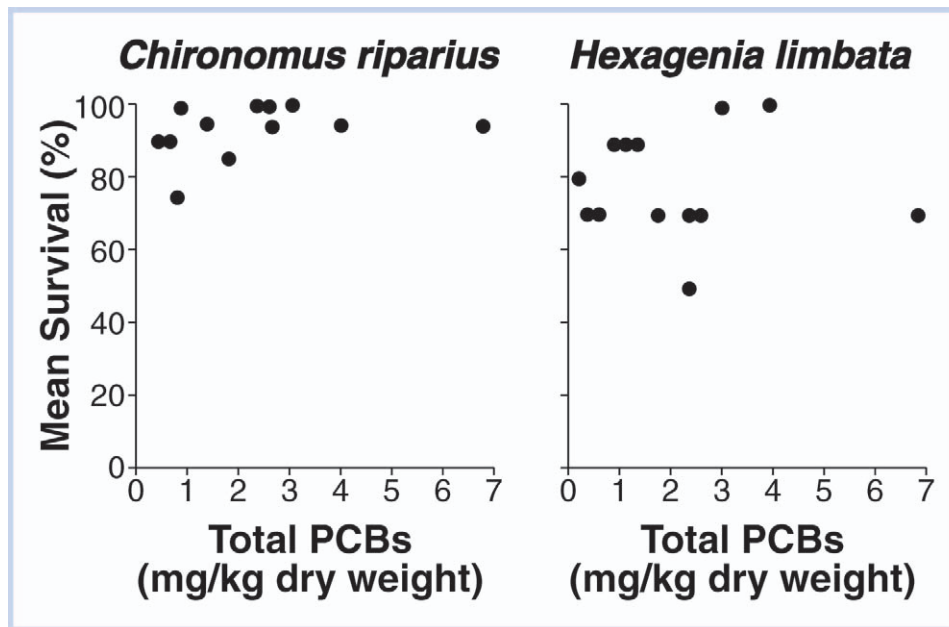


Figure 1. Relationship between sediment toxicity endpoints for the 10-d *Chironomus riparius* and *Hexagenia limbata* tests and concentrations of total PCBs from Fox River–Green Bay (WI, USA).

widespread in the samples. Although statistical differences in amphipod length were found for all but one of the sediment samples, it is questionable whether those differences were related to chemical toxicity or to differences in other sediment characteristics between the formulated and test sediments. The determination of whether the growth results are indicative of chemical toxicity is important, because 17 of those values account for 52% of the toxic samples with PCB concentrations greater than the EEC in the freshwater SEC validation data set (MESL 1999). If those 17 values were reclassified as nontoxic, the predicative ability of the EEC would drop from the value of 82.5% presented in MacDonald, DiPinto, et al. (2000) to only 55%. Furthermore, if the 4 survival values of 84% to 88% were reclassified as nontoxic because survival exceeded the minimum allowable negative control value, the predicative ability of the EEC would decline further to 45%. These examples demonstrate how sensitive the predicative ability of the EEC is to the single data set for Waukegan Harbor, and therefore call into question the robustness of the calculated predicative ability.

Given the uncertainties described above regarding the designation of stations in Waukegan Harbor as toxic or nontoxic, an evaluation of potential concentration–response relationships between the toxicity endpoints and concentrations of PCBs is particularly valuable, because the interpretations of the results do not depend on comparisons with negative controls or reference sediments. In addition, evaluations of concentration–response relationships can sometimes help reduce interpretive difficulties associated with multiple co-occurring chemicals. That is, if a chemical is the primary cause of the observed sediment toxicity, one would expect to find relatively monotonic toxic responses in relation to increasing concentrations of the chemical, regardless of the presence of co-occurring chemicals. If such a response is not found, it is likely that the chemical is not a major contributor to toxicity.

For Fox River–Green Bay, concentration–response relationships were not found for either the 10-d *Chironomus riparius*

test or the 10-d *Hexagenia limbata* test, because percent survival did not exhibit a consistent negative relationship with increasing concentrations of PCBs (Figure 1). In addition, survival of *C. riparius* exceeded the minimum allowable negative control value of 80% for all but 1 sample, including the value of 95% found for the highest PCB concentration of 6.8 mg/kg. Survival of *H. limbata* exhibited a wider range along the gradient of PCB concentrations. Although survival was less than the minimum acceptable control value of 80% at approximately half the stations, values as high as 100% were found at PCB concentrations as high as 3.1 and 4.0 mg/kg.

For Waukegan Harbor, no consistent concentration–response relationships were found for the *Hyalella azteca* test across the range of PCB concentrations evaluated, based on either the 28-d or 42-d exposure periods evaluated by USEPA (1999; Figure 2). However, percent survival began to exhibit a decline at PCB concentrations of 6 to 8 mg/kg for both exposure periods. For the 28-d exposure period, amphipod survival exceeded the minimum allowable negative control value of 80% for all but 1 sample, including the value of 90% found at a PCB concentration as high as 7.7 mg/kg. Survival was substantially reduced only at the highest PCB concentration of 8.9 mg/kg, for which a value of 66% was found. For the 42-d exposure period, amphipod survival exceeded 80% for all but 3 samples, including the value of 95% found at a PCB concentration as high as 7.4 mg/kg. Survival was substantially reduced only at the highest PCB concentration of 8.9 mg/kg, for which a value of 63% was found. For the growth endpoints in both the 28- and 42-d exposure periods, amphipod length showed no evidence of a negative relationship with increasing PCB concentrations, and the highest length values for the 28- and 42-d tests (i.e., 4.3 and 4.6 mm, respectively) were found at PCB concentrations of 7.4 and 7.3 mg/kg, respectively.

In summary, the evaluations described in this section for Fox River–Green Bay and Waukegan Harbor indicate that use of the SECs for PCBs to identify the presence of sediment toxicity on a site-specific basis can result in highly erroneous

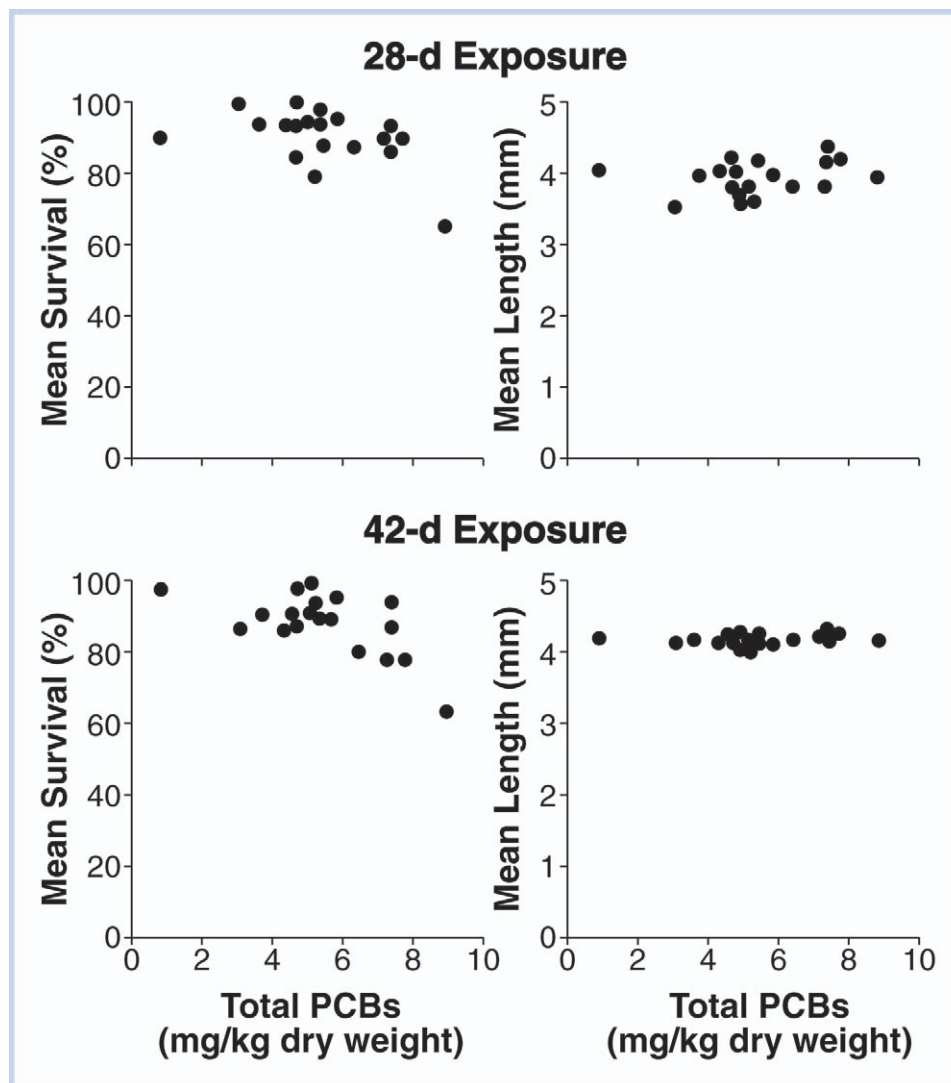


Figure 2. Relationship between sediment toxicity endpoints for the 42-d *Hyalella azteca* test and concentrations of total PCBs in Waukegan Harbor (IL, USA).

conclusions, because the predicative ability of the MEC and EEC were both less than 15% when applied to sediments from Fox River–Green Bay. In addition, although all but 1 station sampled in Waukegan Harbor were identified as toxic by USEPA (1999), most of those designations may be largely an artifact of the use of a negative control sediment that had little similarity to the test sediments, because survival in 4 of the 6 toxic stations based on that endpoint exceeded the minimum allowable negative control value, and all of the length values for the toxic stations exceeded the performance level for negative control sediments used by USEPA (2000). Aside from statistical determinations of the presence or absence of significant ($p \leq 0.05$) toxicity, evaluations of concentration–response relationships between the various sediment toxicity endpoints and concentrations of PCBs in both Fox River–Green Bay and Waukegan Harbor showed that there was no consistent negative relationship between toxicity and PCB concentrations, and that some of the highest values of survival and growth in the toxicity tests were found at PCB concentrations as high as 7.3 to 7.7 mg/kg, casting considerable doubt on the validity of the 1.7 mg/kg EEC as an indicator of extreme effects.

CONCLUSIONS

Results of this critical evaluation of the SECs for PCBs indicate that the SECs are simple mathematical constructs that share the same limitations as their underlying SQGs. The SECs are questionable “consensus” values, because many of their underlying SQGs are dissimilar, misclassified, or redundant with other SQGs. Because nearly all of the data sets included in the databases used to calculate the underlying SQGs, or in the 2 data sets used to validate the SECs, are affected by elevated concentrations of multiple co-occurring chemicals, it is not possible to conclusively identify PCBs as the cause of any of the observed sediment toxicity. The SECs, and most of their underlying SQGs, are likely biased by the fact that their underlying databases are composed primarily of PCB concentrations <0.5 mg/kg, and contain relatively few samples with PCB concentrations greater than the EEC. Comparisons between the SECs and bioaccumulation-based SQGs calculated using the EqP approach provide no information on whether the SECs are causally related to sediment toxicity. The primary available LC50 value for PCBs conducted using spiked-sediment toxicity tests has limited applicability to most contaminated aquatic environments,

because it was determined using sediment with an unusually low TOC content. Finally, site-specific application of the SECs indicated that the predictive ability of both the MEC and EEC was very low, that concentration–response relationships were not found between PCBs and a variety of test species and toxicity endpoints, and that some of the highest values of survival and growth in the toxicity tests were found at PCB concentrations more than 4 times as great as the EEC, casting considerable doubt on the validity of that SEC as an indicator of extreme effects.

Based on the results of this study, we conclude that the SECs for PCBs should be used only in the screening-level evaluations that typically precede more direct assessments of sediment toxicity at individual study sites, and should not be used to predict the presence of sediment toxicity. In such screening-level assessments, measured sediment concentrations that are less than the TEC can likely be assumed to be nontoxic with respect to PCBs, because of the very low concentration of that SEC. However, conclusions concerning the existence of significant sediment toxicity related to PCBs should be based primarily on site-specific assessments of biological effects that address such factors as site-specific bioavailability and the presence or absence of concentration–response relationships. Laboratory studies that address potential causality, such as toxicity identification evaluations and spiked sediment studies, may also be useful.

Most contaminated sediment sites, especially those in urban areas, contain complex mixtures of chemicals in the sediments. At such sites, both the kinds of co-occurring chemicals and their relative concentrations can vary considerably, reducing the utility of any kind of SQG expressed as the concentration of a single chemical or chemical group. The inherent complexity introduced by co-occurring chemicals does not lend itself to simplified predictive relationships based on a single chemical or chemical group, such as PCBs.

Given the limitations of the SECs for PCBs identified in this paper, we conclude that these values cannot be used alone to reliably identify PCBs as the cause of any toxicity observed in the kinds of sediments typically found at contaminated sites. Contrary to the conclusions of MacDonald, DiPinto, et al. (2000), the SECs for PCBs do not reconcile existing SQGs, do not reflect causal effects, and should not be used to determine the spatial extent of injury to sediment-dwelling organisms.

Acknowledgment—Funding for the preparation of this manuscript was provided by the General Electric Company. The views expressed are solely those of the authors.

REFERENCES

- Barrick R, Becker S, Pastorok R, Brown L, Beller H. 1988. Sediment quality values refinement: 1988 update and evaluation of Puget Sound AET. Final report. Seattle (WA): US Environmental Protection Agency. EPA 68-01-4341.
- Becker DS, Barrick RC, Read LB. 1989. Evaluation of the AET approach for assessing contamination of marine sediments in California. Bellevue (WA): PTI Environmental Services. Report 90-3WQ.
- Bolton HS, Breteler RJ, Vigon BW, Scanlon JA, Clark SL. 1985. National perspective on sediment quality. Washington DC: US Environmental Protection Agency. EPA 68-01-6986.
- Call DJ, Balder MD, Brooke LT, Lozano SJ, Vaishnav DD. 1991. Sediment quality evaluation in the Lower Fox River and southern Green Bay of Lake Michigan. Final report. Superior (WI): University of Wisconsin-Superior. USEPA CR-815232.
- Chapman PM, Dexter RN, Long ER. 1987. Synoptic measures of sediment contamination, toxicity and infaunal community composition (the sediment quality triad) in San Francisco Bay. *Mar Ecol Prog Ser* 37:75–96.
- Cabbage J, Batts D, Briedenbach S. 1997. Creation and analysis of freshwater sediment quality values in Washington State. Technical report. Olympia (WA): Washington Department of Ecology. Publication 97-323.
- Dickson KL, Waller WT, Kennedy JH, Arnold WR, Desmond WP, Dyer SD, Hall JF, Knight JT, Malas D, Martinez ML, Matzner SL. 1989. A water quality and ecological survey of the Trinity River, Vols 1 and 2. Denton (TX): Institute of Applied Sciences, University of North Texas.
- Di Toro DM, Zarba CS, Hansen DJ, Berry WJ, Swartz RC, Cowan CE, Pavlou SP, Allen HE, Thomas NA, Paquin PR. 1991. Technical basis for establishing sediment quality criteria for non-ionic organic chemicals using equilibrium partitioning. *Environ Toxicol Chem* 10:1541–1583.
- Environment Canada and Ministère de l'Environnement du Québec. 1992. Interim criteria for quality assessment of St. Lawrence River sediment. Final report. Ottawa (ON): Environment Canada.
- Fairey R, Bretz C, Lamerdin S, Hunt J, Anderson B, Tudor S, Wilson CJ, LaCaro F, Stephenson M, Puckett M, Long ER. 1996. Chemistry, toxicity and benthic community conditions in sediments of the San Diego Bay region. Final report. Sacramento (CA): State Water Resources Control Board.
- Fuchsman PC, Barber TR, Lawton JC, Leigh KB. 2006. An evaluation of cause–effect relationships between polychlorinated biphenyl concentrations and sediment toxicity to benthic invertebrates. *Environ Toxicol Chem* 25:2601–2612.
- Hoke RA, Giesy JP, Zabik M, Unger M. 1993. Toxicity of sediments and sediment pore waters from the Grand Calumet River-Indiana Harbor, Indiana, area of concern. *Ecotoxicol Environ Saf* 26:86–112.
- Ingersoll CG, Haverland PS, Bunson EL, Canfield TJ, Dwyer FJ, Henke CE, Kemble NE, Mount DR, Fox RG. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. *J Great Lakes Res* 22:602–623.
- Ingersoll CG, MacDonald DD. 1999. An assessment of sediment injury in the west branch of the Grand Calumet River. Columbia (MO): US Geological Survey; Ladysmith (BC): MacDonald Environmental Sciences Ltd.
- Ingersoll CG, Nelson MK. 1990. Testing sediment toxicity with *Hyalella azteca* (Amphipoda) and *Chironomus riparius* (Diptera). In: Landis WG, van der Schalie WH, editors. Aquatic toxicology and risk assessment: Thirteenth volume. Philadelphia (PA): American Society for Testing and Materials. STP 1096. p. 93–109.
- JRB. 1984. Initial evaluation of alternatives for development of sediment related criteria for toxic contaminants in marine waters (Puget Sound). Phase II: Development and testing of the sediment-water equilibrium partitioning approach. Bellevue (WA): JRB Associates. USEPA report 910/9-83-117.
- Kemble NE, Brunson TJ, Canfield FJ, Dwyer FJ, Ingersoll CG. 1998. Assessing sediment toxicity from navigational pools of the Upper Mississippi River using a 28-day *Hyalella azteca* test. *Arch Environ Contam Toxicol* 35:181–190.
- Long ER, Field LJ, MacDonald DD. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. *Environ Toxicol Chem* 17:714–727.
- Long ER, MacDonald DD, Smith SL, Calder FD. 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environ Manag* 19:81–97.
- Long ER, Morgan LG. 1991. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program. Seattle (WA): National Oceanic and Atmospheric Administration. NOS OMA 52.
- Long ER, Wolfe DA, Carr RS, Scott KJ, Thursby GB, Windome HL, Lee R, Calder FD, Sloane GM, Seal T. 1994. Magnitude and extent of sediment toxicity in Tampa, Florida. Silver Spring (MD): National Oceanic and Atmospheric Administration. NOS ORCA 78.
- Long ER, Wolfe DA, Scott KJ, Thursby GB, Stern EA, Peven C, Schwartz T. 1996a. Magnitude and extent of sediment toxicity in the Hudson-Raritan Estuary. Silver Spring (MD): National Oceanic and Atmospheric Administration. NOS ORCA 88.
- Long ER, Wolfe DA, Scott KJ, Thursby GB, Stern EA, Peven C, Schwartz T. 1996b. Sediment toxicity in Boston Harbor: magnitude, extent, and relationships with chemical toxicants. Silver Spring (MD): National Oceanic and Atmospheric Administration. NOS ORCA 96.

- MacDonald DD. 1997. Sediment injury in the Southern California Bight: Review of the toxic effects of DDTs and PCBs in sediments. Technical report. Long Beach (CA): National Oceanic and Atmospheric Administration.
- MacDonald DD, Carr RS, Calder FD, Long ER, Ingersoll CG. 1996. Development and evaluation of sediment quality guidelines for Florida coastal waters. *Ecotoxicology* 5:253–278.
- MacDonald DD, DiPinto LM, Field J, Ingersoll CG, Long ER, Swartz RC. 2000. Development and evaluation of consensus-based sediment effect concentrations for polychlorinated biphenyls. *Environ Toxicol Chem* 19:1403–1413.
- MacDonald DD, Ingersoll CG, Berger TA. 2000. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. *Arch Environ Contam Toxicol* 39:20–31.
- MESL. 1999. Development and evaluation of consensus-based sediment effect concentrations for PCBs in the Hudson River. Final report. Prepared for Damage Assessment Center, National Oceanic and Atmospheric Administration. Ladysmith (BC): MacDonald Environmental Sciences Ltd.
- Munns WR Jr, Mueller C, Cobb DJ, Gleason TR, Pesch GG, Johnson RK. 1991. Risk assessment pilot study, phase I, Naval Construction Battalion Center. Final Report. Narragansett (RI): US Environmental Protection Agency.
- Neff JM, Bean DJ, Cornaby BW, Vaga RM, Gulbransen TC, Scalon JA. 1986. Sediment quality criteria methodology validation: Calculation of screening level concentrations from field data. Final report. Washington DC: US Environmental Protection Agency.
- Neff JM, Word JQ, Gulbransen TC. 1987. Recalculation of screening level concentrations for nonpolar organic contaminants in marine sediments. Final report. Washington DC: Washington Environmental Program Office.
- NYSDEC. 1999. Technical guidance for screening contaminated sediments. Technical report. Albany (NY): Division of Fish and Wildlife, Division of Marine Resources.
- Pastorok RA, Becker DS. 1990. Comparative sensitivity of sediment toxicity bioassays at three Superfund sites in Puget Sound. In: Landis WG, van der Schalie WH, editors. Aquatic toxicology and risk assessment: Thirteenth volume. Philadelphia (PA): American Society for Testing and Materials. STP 1096. p 123–139.
- Persaud D, Jaagumagi R, Hayton A. 1993. Guidelines for the protection and management of aquatic sediment quality in Ontario. Final report. Toronto (ON): Ontario Ministry of Environment and Energy.
- Platonow NS, Karstad LH. 1973. Dietary effects of polychlorinated biphenyls on mink. *Can J Comp Med* 37:391–400.
- Rice CA, Plesha PD, Casillas E, Misitano DA, Meador JP. 1995. Growth and survival of three marine invertebrate species in sediments from the Hudson-Raritan Estuary, New York. *Environ Toxicol Chem* 14:1931–1940.
- Sapudar RA, Wilson CJ, Reid ML, Long ER, Stephenson M, Puckett M, Fairey R, Hunt J, Anderson B, Holstad D, Newman J, Birosik S, Smyth H. 1994. Sediment chemistry and toxicity in the vicinity of the Los Angeles and Long Beach Harbors. Draft final report. Sacramento (CA): State Water Resources Control Board.
- Schimmel SC, Melzian BD, Campbell DE, Strobel CJ, Benyl SJ, Rosen JS, Buffum HW. 1994. Statistical summary. EMAP-Estuaries. Virginian Province—1991. Washington DC: US Environmental Protection Agency. EPA 620/R-94/005.
- Schlekat C, McGee BL, Boward DM, Reinharz E, Velinsky DJ, Wade TL. 1994. Biological effects associated with sediment contamination in the Potomac and Anacostia rivers in the Washington, DC, area. *Estuaries* 17:334–344.
- Smith SL, MacDonald DD, Keenleyside KA, Ingersoll CG, Field J. 1996. A preliminary evaluation of sediment quality assessment values for freshwater ecosystems. *J Great Lakes Res* 22:624–638.
- Strobel CJ, Buffum HW, Benyl SJ, Petrocelli EA, Reifsteck DR, Keith DJ. 1995. Statistical summary. EMAP-Estuaries. Virginian Province—1990–1993. Narragansett (RI): US Environmental Protection Agency. EPA 620/R-94/-26.
- Swartz RC. 1999. Consensus sediment quality guidelines for polycyclic aromatic hydrocarbon mixtures. *Environ Toxicol Chem* 18:780–787.
- Swartz RC, Kemp PF, Schults DW, Lamberson JO. 1988. Effects of mixtures of sediment contaminants on the marine infaunal amphipod *Rhepoxynius abronius*. *Environ Toxicol Chem* 7:1013–1020.
- [USEPA] US Environmental Protection Agency. 1980. Ambient water quality criteria for polychlorinated biphenyls. Washington DC: Office of Water Regulations and Standards. EPA 440/5-80-068.
- [USEPA] US Environmental Protection Agency. 1989. Evaluation of the apparent effects threshold (AET) approach for assessing sediment quality. Washington DC: USEPA, Science Advisory Board. SAB-EETFC-89.
- [USEPA] US Environmental Protection Agency. 1993. Biological and chemical assessment of contaminated Great Lakes sediment. Chicago (IL): USEPA. EPA-905-R93-006.
- [USEPA] US Environmental Protection Agency. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. Chicago (IL): USEPA. EPA 905-R96-008.
- [USEPA] US Environmental Protection Agency. 1997. An assessment of sediments from the Upper Mississippi River. Washington DC: USEPA, Office of Science and Technology. EPA 823-R-97-005.
- [USEPA] US Environmental Protection Agency. 1999. Evaluation of toxicity and bioaccumulation of contaminants in sediment samples from Waukegan Harbor, Illinois. Chicago (IL): USEPA. EPA-905-R-99-009.
- [USEPA] US Environmental Protection Agency. 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates, second edition. Duluth (MN); Washington DC: USEPA. EPA/600/R-99/064.
- Velinsky DJ, Wade TL, Schlekat CE, McGee BL, Presley BJ. 1994. Tidal river sediments in the Washington, DC, area. I. Distribution and sources of trace metals. *Estuaries* 17:305–320.
- Von Stackelberg K, Menzie CA. 2002. A cautionary note on the use of species presence and absence data in deriving sediment criteria. *Environ Toxicol Chem* 21:466–472.
- Wade TL, Velinsky DJ, Reinharz E, Schlekat CE. 1994. Tidal river sediments in the Washington, DC, area. II. Distribution and sources of organic contaminants. *Estuaries* 17:321–333.
- Wenning RJ, Batley GE, Ingersoll CG, Moore DW, editors. 2005. Use of sediment quality guidelines and related tools for the assessment of contaminated sediments. Pensacola (FL): SETAC.
- Wolfe DA, Bricker SB, Long ER, Scott KJ, Thursby GB. 1994. Biological effects of toxic contaminants in sediments from Long Island Sound and Environs. Silver Spring (MD): National Oceanic and Atmospheric Administration. NOS ORCA 80.