Methodology for Derivation of Pesticide Sediment Quality Criteria for the Protection of Aquatic Life

Phase I: Review of Existing Methodologies

Prepared for the Central Valley Regional Water Quality Control Board

Tessa L. Fojut, Ph.D.,
Martice E. Vasquez, Ph.D.
and
Ronald S. Tjeerdema, Ph.D.

Department of Environmental Toxicology
University of California, Davis

September 2011

Disclaimer

Funding for this project was provided by the California Regional Water Quality Control Board, Central Valley Region (CRWQCB-CVR). The contents of this document do not necessarily reflect the views and policies of the CRWQCB-CVR, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Methodology for Derivation of Pesticide Sediment Quality Criteria for the Protection of Aquatic Life

Phase I: Review of Existing Methodologies

Report prepared for the Central Valley Regional Water Quality Control Board

Tessa L. Fojut, Ph.D.,
Martice E. Vasquez, Ph.D.,
and
Ronald S. Tjeerdema, Ph.D.

Department of Environmental Toxicology
University of California, Davis

September 2011

Executive Summary

The goal of this project is to develop a methodology for derivation of pesticide sediment quality criteria for the protection of aquatic life in the Sacramento and San Joaquin River basins. There are three phases to this project. This is a report of the results of Phase I, which is a comparison and evaluation of existing criteria derivation methodologies from around the world. Phase II will be development of the sediment criteria derivation methodology. The new methodology may simply be one of the existing methodologies, a combination of features from existing methodologies, or an entirely new methodology based on the latest available research in aquatic ecotoxicology and environmental risk assessment. Phase III will be to apply the new methodology to derive sediment quality criteria for at least one pyrethroid pesticide of particular concern in the Sacramento and San Joaquin River basins due to listings under 303(d) of the federal Clean Water Act (CRWQCB-CVR 2006).

The approach for Phase I was to conduct an extensive literature search to find 1) criteria derivation methodologies currently in use, or proposed for use, throughout the world; 2) original studies supporting the methodologies; 3) proposed modifications of existing methodologies; and 4) relevant and recent research in ecotoxicology and risk assessment. In this report, important elements of sediment quality criteria derivation methodologies are discussed with respect to how they are, or are not, addressed by existing methodologies. Included in the discussion are methodologies from: Australia/New Zealand, Canada, the European Union/European Commission (EU/EC), France, The Netherlands, the Organisation for Economic Co-operation and Development (OECD), the United Kingdom (UK), and the United States (US), including the United States Environmental Protection Agency (USEPA) and the National Oceanic and Atmospheric Administration (NOAA), and a few individual states whose methodologies diverge somewhat from guidance from the federal agencies.

The goal of this review is to determine if there is an appropriate existing methodology that can be used to calculate sediment quality criteria for pesticides, particularly for a group of or individual pyrethroids. There are three main approaches that are currently used for development of sediment quality guidelines: empirical, mechanistic and spiked-sediment toxicity testing. In general, the empirical approaches generate concentration ranges that are very likely, likely, or not likely to cause adverse effects, while the mechanistic approaches generate single concentrations not to be exceeded that are based on the existence of a water quality criterion for the compound of interest. The third approach uses spiked-sediment toxicity data to derive criteria with statistical distributions or by applying an assessment factor (sometimes called safety factors). Several of the methodologies incorporate multiple approaches and recommend deriving criteria from spiked-sediment toxicity test data if it is available, or comparing the derived criteria to this data if it is limited.

There are two possible outcomes of this project: 1) recommend an existing methodology for adoption, or 2) develop an entirely new methodology. If a new methodology is developed, it will likely use elements from the existing methodologies, and hopefully add new techniques for more refined risk assessment than is currently available in any of the existing methodologies. The next phase of this project will be to test out and further explore the various approaches to determine which will result in the most reliable and robust methodology for developing sediment quality criteria for pesticides. The methodologies of The Netherlands (RIVM 2001) and the EU (ECB 2003) appear to be the most developed methodologies and they include both mechanistic and spiked-sediment toxicity test approaches. The methodology developed in Phase II will likely draw on these two methodologies, and will be compared to these existing methodologies.

Table of Contents

List of Tables

List of Acronyms and Abbreviations

AET	Apparent Effects Threshold
AF	Assessment Factor
ANZECC	Australia and New Zealand Environment and Conservation Council
ARMCANZ	Agriculture and Resource Management Council of Australia and New Zealand
ASTM	American Society for Testing and Materials
BAF	Bioaccumulation Factor
BCF	Bioconcentration Factor
BMF	Biomagnification Factor
BEDS	Biological Effects Database for Sediments
BSAF	Biota-sediment Accumulation Factor
CAS	Chemical Abstract Service
CCME	Canadian Council of Ministers of the Environment
CDFG	California Department of Fish and Game
CDPR	California Department of Pesticide Regulation
C_{iw}	Chemical concentration in interstitial water
CRWQCB-CVR	California Regional Water Quality Control Board, Central Valley Region
$C_{s/oc}$	Chemical concentration in whole sediment or organic carbon
CTB	Dutch Board for the Authorization of Pesticides
DCPA	Dimethyl 2,3,5,6-tetrachlorobenzene-1,4 dicarboxylate
DDT	Dichlorodiphenyltrichloroethane
DOC	Dissolved Organic Carbon
DOM	Dissolved Organic Matter
DT_x	Time for x% of a chemical to degrade
EC	European Commission
EC_x	Concentration that affects x% of exposed organisms
ECB	European Chemicals Bureau
EPA	Environmental Protection Agency
ERL	Effects Range Low
ERL(sed_{EP})	Environmental Risk Limit for the sediment compartment using EqP theory
ERL(water)	Environmental Risk Limit for aquatic species
ERM	Effects Range Median
EINECS	European Inventory of Existing Commercial Substances
EqP	Equilibrium Partitioning
ESG	Equilibrium Sediment Guideline
EU	European Union
FCV	Final Chronic Value
f_{oc}	Fraction of organic carbon
H	% Organic Matter Content
HC_p	Hazardous Concentration potentially harmful to p% of species
HOC	Hydrophobic Organic Compound
HPLC	High Pressure Liquid Chromatography
HSDB	Hazardous Substance Data Bank
IUPAC	International Union of Pure and Applied Chemistry
ISQG	Interim Sediment Quality Guideline

IWQ	Interstitial Water Quality
K_d	Solid-water partition coefficient
K_{oc}	Organic carbon-normalized solid-water partition coefficient
K_{ow}	Octanol-water partition coefficient
K_p	Solid-water partition coefficient (equivalent to K_d)
$K_{s/l}$	Sediment-liquid partition coefficient
$K_{susp\ water}$	Suspended matter-water partition coefficient
K_x	Interaction Coefficient for a synergist/antagonist at concentration x
LC_x	Concentration lethal to x% of exposed organisms
LEL	Lowest Effect Level
LOEC	Lowest Observed Effect Concentration
LOEL	Lowest Observed Effect Level
MATC	Maximum Acceptable Toxicant Concentration
MPC	Maximum Permissible Concentration
MTC	Maximum Tolerable Concentration
NC	Negligible Concentration
NEL	No Effect Level
NOAA	National Oceanic and Atmospheric Association
NOEC	No Observed Effect Concentration
NOEL	No Observed Effect Level
NSTP	National Status and Trends Program
NTIS	National Technical Information Service
OC	Organic Carbon
OCSPP	Office of Chemical Safety and Pollution Prevention
OECD	Organization for Economic Co-operation and Development
OM	Organic Matter
OMEE	Ontario Ministry of Environment and Energy
OPP	Office of Pesticide Programs
OPPTS	Office of Prevention, Pesticides and Toxic Substances
OPPT	Office of Pollution Prevention and Toxics
PAET	Probable Apparent Effects Threshold
PAH	Polycyclic Aromatic Hydrocarbons
PCB	Polychlorinated biphenyls
PEC	Probable Effect Concentration
$PEC_{oral,\ predator}$	Predicted chemical concentration a predator will receive in prey (food)
PEC_{sed}	Measured or predicted chemical concentration in sediment
PEC_{water}	Measured or predicted chemical concentration in water
PEL	Probable Effect Level
pK_a	Acid dissociation constant
PNEC	Predicted No Effect Concentration
$PNEC_{sed}$	Predicted No Effect Concentration in sediment
$PNEC_{water}$	Predicted No Effect Concentration in water
PWQO/G	Provincial Water Quality Objectives/Guidelines
QSAR	Quantitative Structure Activity Relationship
RHO_{susp}	Bulk Density of wet suspended matter

RIVM	National Institute of Public Health and the Environment, Bilthoven, The Netherlands
RL	Risk Limit
RWQCB	Regional Water Quality Control Board
SEL	Severe Effect Level
SETAC	Society of Environmental Toxicology and Chemistry
SLC	Screening Level Concentration
SPME	Solid Phase Micro-extraction
SRC_{ECO}	Ecosystem Serious Risk Concentration
SSD	Species Sensitivity Distribution
SSLC	Species Screening Level Concentration
SSTT	Spiked-Sediment Toxicity Testing
SQC	Sediment Quality Criteria
SQG	Sediment Quality Guideline
SQG_{oc}	OC-normalized sediment quality guideline
SWRCB	State Water Resources Control Board
TEC	Threshold Effect Concentration
TEL	Threshold Effect Level
TES	Threatened and Endangered Species
TU	Toxic Unit
UK	United Kingdom
US	United States
USEPA	United States Environmental Protection Agency
WQC	Water Quality Criteria

1 Introduction

The goal of this project is to develop a methodology for derivation of pesticide sediment quality criteria (SQC) for the protection of aquatic life in the Sacramento and San Joaquin River basins. The surface waters of these basins receive pesticide inputs in runoff and drainage from agriculture, silviculture, and residential and industrial storm water (CRWQCB-CVR 2009). The term pesticide is defined by the Central Valley Regional Water Quality Control Board (CRWQCB-CVR 2009) as (1) any substance, or mixture of substances which is intended to be used for defoliating plants, regulating plant growth, or for preventing, destroying, repelling, or mitigating any pest, which may infest or be detrimental to vegetation, man, animals, or households, or be present in any agricultural or nonagricultural environment whatsoever, or (2) any spray adjuvant, or (3) any breakdown products of these materials that threaten beneficial uses.

The project will be accomplished in three phases. This is a report of the results of Phase I, which is a comparison and evaluation of existing criteria derivation methodologies from around the world. Phase II will be development of the criteria derivation methodology. The new methodology may simply be one of the existing methodologies, a combination of features from existing methodologies, or an entirely new methodology based on the latest available research in aquatic ecotoxicology and environmental risk assessment. Phase III will be to apply the new methodology to derive criteria for at least one pyrethroid pesticide of particular concern in the Sacramento and San Joaquin River basins due to listings under 303(d) of the federal Clean Water Act (CRWQCB-CVR 2006).

The mission of California's nine Regional Water Quality Control Boards (RWQCBs) is "to develop and enforce water quality objectives and implementation plans which will best protect the beneficial uses of the State's waters, recognizing local differences in climate, topography, geology and hydrology" (California SWRCB 2011). Toward that mission, each RWQCB is responsible for development of a "basin plan" for its hydrologic area. The "Water Quality Control Plan (Basin Plan) for the Sacramento River and San Joaquin River Basins," (CRWQCB-CVR 2009) contains the following language regarding toxic substances in general, and pesticides in particular:

"...waters shall be maintained free of toxic substances in concentrations that produce detrimental physiological responses in human, plant, animal, or aquatic life."

"No individual pesticide or combinations of pesticides shall be present in concentrations that adversely affect beneficial uses."

"Discharges shall not result in pesticide concentrations in bottom sediments or aquatic life that adversely affect beneficial uses."

"Pesticide concentrations shall not exceed the lowest levels technically and economically achievable."

The sediment layer of aquatic ecosystems provides habitat and food sources for aquatic life. Many organic chemicals tend to accumulate in the sediment once released into the environment and these contaminants may cause toxicity even when water quality criteria are being met. The chemical and physical properties of a compound, as well as environmental factors, contribute to its tendency to accumulate in sediments and/or possibly bioaccumulate in tissues or magnify up the food chain. A measure of chemical partitioning between a solid and water is defined as the solid-water partition coefficient (K_d), which can be normalized to organic carbon (K_{oc}) to compare across different types of solids (e.g., sediment or soil). K_d is the chemical concentration in water divided by the chemical concentration in solid at equilibrium and K_{oc} is defined as K_d divided by the fraction of organic carbon (OC) in the solid. Partitioning between sediments and water is usually driven by a chemical's insolubility in water, yet the mechanism by which chemicals sorb can be influenced by environmental pH, temperature, and sorbate properties (grain size, OC quantity and makeup (e.g. black carbon), clay and mineral content, redox potential, moisture content) (Schwarzenbach et al. 2003). Bioavailability and the related toxic effects of sediment-bound chemicals are limited by the particular sorbent properties, chemical properties and organism behavior (Day et al. 1995). Many studies have demonstrated that total measured contaminant concentrations in sediments are poorly correlated to observed effects, which is most likely due to limited bioavailability of the contaminants (Conrad et al. 1999, Di Toro et al. 2002, Xu et al. 2007).

The pesticides that are most likely to cause sediment toxicity are those that are nonpolar nonionic organic compounds because they tend to sorb to solids and colloids in aqueous environments. Nonpolar pesticides have been detected in California freshwater bedded and/or suspended sediments in recent years, including herbicides (DCPA, ethalfluralin, metolachlor, oxyfluorfen, pendimethalin, prometryn, simazine, trifluralin), organochlorines (DDTs, dieldrin, endosulfan, endrin, lindane, methoxychlor), organophosphates (chlorpyrifos, diazinon, methylparathion), and pyrethroids (bifenthrin, cyfluthrin, cypermethrin, deltamethrin, esfenvalerate, fenpropathrin, lambda-cyhalothrin, permethrin) (Domalgalski et al. 2010, Hladik and Kuivila 2009, Holmes 2004, Weston et al. 2004, 2005). Most of the organochlorines that have been recently detected are no longer used; in contrast, use of pyrethroid insecticides has increased over the last decade, as they are seen as replacements for organophosphate insecticides. Pyrethroids are used in both agricultural and urban/residential settings and are characterized by extreme insolubility in water with high degrees of sorption to solids, including soils, sediment, and dissolved organic matter (Laskowski 2002). Although pyrethroids have low mammalian toxicity, they are highly toxic to aquatic invertebrates and fish, and their use has not been without adverse effects. Pyrethroids have been found at concentrations in sediments of both urban and agricultural waterways that are toxic to the freshwater invertebrate *Hyalella azteca* in the laboratory (Amweg et al. 2006, Weston et al. 2004).

There are currently no official US Environmental Protection Agency (USEPA) methods or other agreed upon approaches in the United States for generating SQC. Although water quality criteria derivation methodologies have been in place for many years, sediment proposes some unique challenges for development of single numeric concentrations below which aquatic life is protected. Sediment is a complex medium with inherent variability both spatially and temporally; to reduce this variability, concentrations of contaminants can be normalized to different sediment properties. Organic compounds are primarily thought to sorb to the OC in

sediments, and normalizing to OC content provides a way to reduce variability in sorption and toxicity measures, but it rarely reduces variation completely (Amweg et al. 2005, Xu et al. 2007). Normalization to dry weight has also been used with similar or better success compared to normalization to OC content for hydrophobic organic compounds, such as PCBs and PAHs (Ingersoll et al. 2000). Because fine particles have relatively high surface areas, they tend to be enriched in organic contaminants compared to coarser sediments, thus normalization to % fines has also been proposed (Mudroch and Azcue 1995). Bioavailability of contaminants also confounds the sediment toxicity and is an important factor to be considered when establishing any type of numerical sediment quality guidelines, and will be particularly relevant for highly hydrophobic pesticides, such as pyrethroids. Chemicals can exist in both free and bound states within sediments and many studies have suggested that the bioavailable fraction of a chemical is best predicted by the fraction freely dissolved in porewater (Bondarenko et al. 2007, Bondarenko and Gan 2009, Hunter et al. 2008, Sormunen et al. 2010, Xu et al. 2007, Yang et al. 2006a, 2006b, 2007). These predictions do not take other routes of exposure into account, such as ingestion of particulate bound chemicals by sediment dwelling organisms, which has also been suggested as an exposure route of concern (Mehler et al. 2010). These recent studies on bioavailability indicate that direct measurement of the freely dissolved concentration via passive sampling devices results in better correlations with uptake and toxicity than applying sediment-water partition coefficients to whole sediment concentrations.

The approach for Phase I was to conduct an extensive literature search to find 1) criteria derivation methodologies currently in use, or proposed for use, throughout the world; 2) original studies supporting the methodologies; 3) proposed modifications of existing methodologies; and 4) relevant and recent research in ecotoxicology and risk assessment. Several documents were found that provide a good overview of the latest scientific thinking in the field of SQC derivation. These documents include: two publications relating to a Society of Environmental Toxicology and Chemistry (SETAC) Pellston workshop entitled “Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments” (Wenning and Ingersoll 2002, Wenning et al. 2005), a report entitled “Review and recommendations of methodologies for the derivation of sediment quality guidelines” (Rowlatt et al. 2002), a review paper by Chapman (1989) titled “Current approaches to developing sediment quality criteria,” and a report prepared on behalf of the European Commission (EC) called “Towards the Derivation of Quality Standards for Priority Substances in the Context of the Water Framework Directive” (Lepper 2002). Information in these reports was used to construct Table 1, which is a list of components to consider in evaluation and development of a SQC derivation methodology.

In this report, the components listed in Table 1 are discussed with respect to how they are, or are not, addressed by existing criteria derivation methodologies. Included in the discussion are methodologies from: Australia/New Zealand, Canada, the European Union/European Commission (EU/EC), France, The Netherlands, the Organisation for Economic Co-operation and Development (OECD), the United Kingdom (UK), and the United States (US), including the USEPA and the National Oceanic and Atmospheric Administration (NOAA), and a few individual states whose methodologies diverge somewhat from guidance from the federal agencies. In some cases original documents were not available in English, but other resources containing summaries of those documents were available and were used for this report. Existing pesticides, most of the methodologies address toxicity due to metals and other inorganic

Table 1 Components to be addressed by sediment quality criteria derivation methodology.

Category	Component	Reference
Criteria types (Section 3)	Numeric criteria vs. advisory concentrations	1,2
	Multiple levels of criteria	3,4
Protection Level (Section 4)	Protect all species to protect ecosystem	2,3,5
	Protect aquatic environment	4,6,7,8,9,10,11
	Benthic communities	12
	Probability of over or under protection	3
Physical-chemical and ecotoxicity data (Section 5)	Data sources	3,13,14
	Literature search protocol	2,3,6,11
	Physical-chemical data requirements and quality	3,5,6,8,10
	Acute vs. Chronic sediment toxicity testing	2,3,15
	Laboratory vs. field data	3,11
	Traditional vs. non-traditional endpoints	2,3,11
	Ecotoxicity data quality	2,3,15
	Ecotoxicity data quantity	2,16
	Quantitative Structure Activity Relationships (QSARs)	3
Criteria calculation (Section 6)	Magnitude, duration and frequency of exposure	2,5
	Multiple exposure routes	11
	Bioavailability	3,5,6,10,11
	Equilibrium partitioning	3,4,5,6,8,10,11
	Suspended Sediments vs. bedded sediments	4,11
	Spiked-sediment toxicity testing approach	2,3,11
	Standardized sediment	3,6
	Assessment factors	2,3,4,11
	Species sensitivity distribution	3,11
	Mixtures	11
	Bioaccumulation and secondary poisoning	3,16
	Encouragement of data generation	2,3,4,11
	Utilization of available data	3,11
	Harmonization (EqP)	3,4,5,6,8,10,11

1. USEPA 2003a-c
2. CCME 1999
3. RIVM 2001
4. Lepper 2002
5. Di Toro et al. 2002
6. Persaud et al. 1993
7. MacDonald 1994
8. Rowlatt et al. 2002
9. ANZECC and ARMCANZ 2000
10. OECD 1995
11. ECB 2003
12. Diaz and Rosenburg 1996
13. TenBrook et al. 2009
14. USEPA 1985
15. Cubbage et al. 1997
16. SWRCB 2011

chemicals (e.g. ammonia), and non-pesticide organic chemicals. This project is focused on development of pesticide criteria and so the review of methodologies is likewise focused on pesticides. Some of the latest recommendations for SQC derivation methodologies are simply not technically feasible at this time due to lack of data or lack of agreement among experts on techniques. However, thorough discussions of feasibility of approaches are beyond the scope of Phase I of this project and will be reserved for Phase II (development of a methodology).

Large amounts of literature are available on contaminated sediment risk assessment, but the objective of this review is to only focus on one aspect of risk assessment, which is developing numeric SQC for which compliance can be based solely on chemistry measurements. The Central Valley Regional Water Quality Control Board will decide how to use these criteria, and they may be used as part of a risk assessment framework that includes other lines of evidence, such as sediment bioassays. Most of the methodologies refer to the values they derive as sediment quality guidelines (SQGs) instead of sediment quality criteria because they recommend using the values as part of a risk assessment framework or as triggers for further research, and not as enforceable criteria. Throughout the text we use the term SQG when referring to values derived by methodologies that do not recommend that the values be used as enforceable criteria. Much of the risk assessment literature focuses on metals and industrial and legacy chemicals, while our specific objective is to focus on pesticides. Methodologies that primarily focus on metals, dredged materials, and marine environments were excluded or not focused on in this review because freshwater environments with pesticide contamination pose significantly different issues than metal-contaminated harbors. Some of the industrial chemicals and legacy pesticides (i.e., PCBs, PAHs, DDTs, and organochlorines) have similar physical-chemical properties as hydrophobic pesticides (e.g. pyrethroids), which are the pesticides of concern for this project, so information relating to these types of compounds were included, even though these chemicals are not currently used pesticides.

2 Summary of major approaches

According to current methodologies, there are many different types of numeric criteria that may be derived, depending on how the values are to be used and how much and what types of data are available. Table 2 lists the major methodologies, the types of criteria that are derived from them, and how the criteria are used. The goal of this review is to determine if there is an appropriate existing methodology that can be used to calculate sediment quality criteria for pesticides. There are three main approaches that are currently used for development of sediment quality guidelines: empirical, mechanistic and spiked-sediment toxicity testing. In general, the empirical approaches generate concentration ranges that are very likely, likely, or not likely to cause adverse effects, while the mechanistic approaches generate single concentrations not to be exceeded that are based on the existence of water quality criteria for the compound of interest. The third approach uses spiked-sediment toxicity data to derive criteria with statistical distributions or applying an assessment factor (sometimes called safety factors). Several of the methodologies incorporate multiple approaches and recommend deriving criteria from spiked-sediment toxicity data if it is available, or comparing the derived criteria to this data if it is limited. A brief introduction to the three major approaches for SQG derivation are provided in the following subsections 2.1 through 2.3.

Table 2 Summary of major methodologies. EqP: equilibrium partitioning model, SSTT: spiked-sediment toxicity test, AF: assessment factor, SSD: species sensitivity distribution, WQC: water quality criteria.

Method Title	Source	Original Reference	Year	Jurisdiction	Criterion	Criterion derivation
Mechanistic methodologies						
Technical basis for the derivation of equilibrium partitioning sediment quality guidelines for the protection of benthic organisms: nonionic organics	USEPA	Di Toro et al.	2002	United States	Tier 1 ESGs: equilibrium sediment guidelines	EqP uses final chronic value from USEPA WQC derivation and K_{oc} to determine ESG. Compare with SSTT data; tiered approach depends on toxicity data availability
					Tier 2 ESG	Use if SSTT data are limited or unavailable to confirm the EqP prediction.
Guidance document on deriving environmental risk limits	RIVM	RIVM	2001	The Netherlands	MPC: maximum permissible concentration	Use EqP to find MPC for sediment and then compare with SSTT data
Guidance document for aquatic effects assessment	OECD	OECD	1995	OECD	MTC_{sed}: maximum tolerable concentration	Use EqP to calculate MTC_{sed}
Recommendations on the development of sediment quality guidelines.	UK	Rowlatt et al.	2002	UK	SAL: sediment action level	Use EqP to calculate SAL.
Guidelines for the protection and management of aquatic sediment quality in Ontario	OMEE	Persaud et al.	1993	Ontario	NEL: no effect level	Use EqP to calculate NEL
Spiked sediment toxicity testing methodologies						
Technical guidance document on risk assessment	EU	ECB	2003	European Union	PNEC: predicted no effect concentration	If SSTT data available, apply an AF depending on if data are acute or chronic. SSD approach possible if adequate SSTT data. If no data, use EqP.
Towards the derivation of quality standards for priority substances in the context of the water framework directive	EU	Lepper	2002	France	Threshold level 1	AF approach: Lowest NOEC/10 or $L(E)C_{50}/1000$
					Threshold level 2	AF approach: Lowest NOEC or $L(E)C_{50}/100$

Method Title	Source	Original Reference	Year	Jurisdiction	Criterion	Criterion derivation
Empirical methodologies						
Sediment quality guidelines developed for the National Status and Trends Program (NSTP)	NOAA NSTP	Long and Morgan	1990	United States	ERL/ERM: effects range low and median	Utilizes large matching sediment chemistry and biological effects database; ERL & ERM defined as concentration at 10^{th} and 50^{th} percentile of NSTP database, respectively
Australian and New Zealand guidelines for fresh and marine water quality	ANZECC & ARMCANZ	ANZECC & ARMCANZ	2000	Australia/ New Zealand	ERL/ERM	Uses North American data and then refine database using local sediment data once available
Combination of methodologies (Empirical and SSTT)						
Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life	CCME	CCME	1999	Canada	TEL/PEL: threshold and probable effects level	Empirical approach based on NSTP database for derivation of interim sediment quality guidelines
					SSTT: spiked-sediment toxicity testing	Use SSTT approach to establish definitive cause and effect relationships between chemical and response; similar to WQC but use sediment toxicity data

2.1 Mechanistic approach (equilibrium partitioning)

In the 1990s, the USEPA generated sediment quality guidelines based on total chemical concentrations using a mechanistic approach based on the equilibrium partitioning (EqP) model (Di Toro et al. 2002, USEPA 1993). In this approach, it is assumed that toxicity is only caused by the freely dissolved fraction of a contaminant. The EqP approach assumes that the freely dissolved fraction of a chemical is in equilibrium between the sediment and porewater (interstitial water between sediment particles), and that the chemical exposure is equivalent in each of these environmental compartments. In this approach, the final chronic value (FCV) from a chemical's water quality criterion (WQC) and a partition coefficient (K_d or K_{oc}) are used to derive a SQG (Di Toro et al. 2002, USEPA 1993) as follows:

$$SQG = FCV * K_d \text{ (or } K_{oc}) \quad (1)$$

where the K_d is the solid-water partition coefficient and K_{oc} is the organic carbon-normalized partition coefficient. The EqP method assumes epibenthic and benthic organisms have the same species sensitivity distribution as water column organisms; thus, it is justified to use the FCV to derive the SQG (Wenning et al. 2005).

Organic carbon is assumed to be the primary factor governing the partitioning of nonionic organic chemicals between sediment and porewater and using the K_{oc} as the partition coefficient can normalize for differences in partitioning across sediments. This approach has been used for nonionic organic compounds, such as PAHs, including mixtures of PAHs, and several organochlorine pesticides. Many of the methodologies reviewed include the EqP approach, such as The Netherlands (Kalf et al. 1999, RIVM 2001), the EU (ECB 2003), Ontario (Persaud et al. 1993), France (Lepper 2002), and the OECD (1995). This mechanistic approach can also be used with other lines of evidence, such as comparison to field concentrations or spiked-sediment toxicity data, to establish SQGs.

2.2 Empirical approaches

NOAA first developed its empirical approach to deriving SQGs to interpret chemical data from their large monitoring program called the National Status and Trends Program (NSTP), and the SQGs were only intended for informal use, such as ranking areas for further study, not for regulatory purposes. Long and Morgan (1990) developed this co-occurrence approach, which is based on comparing biological effects to total chemical sediment concentrations from the NSTP database. In general, empirical approaches utilize databases containing both sediment chemistry and observed biological effects data from field collected sediments to determine numerical chemical ranges for various effects levels. This approach does not elucidate what is causing the toxicity; toxicity could be caused by the chemicals measured, other chemicals that were not measured, or by some other environmental condition. Sediment toxicity can often be a result of mixture effects and chemical and toxicity testing of field sediments does not provide the information necessary to attribute toxicity to an individual chemical. Empirical approaches only demonstrate correlations between effects and chemical concentrations, and do not demonstrate causation of toxicity by the measured contaminants, and thus are not appropriate for deriving criteria for individual chemicals. The limitations of empirical approaches are well known, and as

such, recent recommendations have stated that they must be considered with other biological effects information in a multiple lines of evidence approach (Wenning et al. 2005). The NOAA NSTP (Long and Morgan 1990) empirical methodology is widely accepted, and empirical approaches used by other jurisdictions are generally variations of this method, such as Australia/New Zealand (ANZECC and ARMCANZ 2000) Canada (CCME 1999), Ontario (Persaud et al. 1993), UK (Rowlatt et al. 2002), Washington State (Cubbage et al. 1997), California (SWRCB 2011), and Florida (MacDonald 1994).

2.3 Spiked-sediment toxicity test approach

The spiked-sediment toxicity test (SSTT) approach is similar to methods used for WQC derivation. Acute and chronic toxicity data from controlled, spiked-sediment laboratory experiments are used to set SQGs. If data are abundant, a species sensitivity distribution (SSD) can be used to derive an appropriate effect level. If data are sparse, assessment (or safety) factors may be applied to the lowest concentration lethal to 50% of exposed organisms (LC_{50}) or the lowest no observed effect concentration (NOEC); these factors can also be applied to account for uncertainty from various factors. Methodologies that include a SSTT approach are those from Canada (CCME 1999), The Netherlands (RIVM 2001), the EU (ECB 2003), and France (Lepper 2002).

3 Criteria types and uses

The USEPA is authorized to develop and implement sediment quality criteria under Section 304(a) of the Clean Water Act. The USEPA has recommended that States use numeric criteria with sediment bioassays to interpret the narrative criteria, which is typically stated as “no toxics in toxic amounts,” (USEPA 1998). The Central Valley Regional Water Quality Control Board can use numeric SQC to set sediment quality objectives or total maximum daily loads. Numeric criteria are defined as scientifically based numbers which are intended to protect aquatic life from adverse effects of pesticides, without consideration of defined water body uses, societal values, economics, or other non-scientific considerations. Each method refers to numeric criteria or guidelines by different terms, but they all fit under the definition of numeric criteria given here. For example, there are equilibrium partitioning sediment guidelines (ESGs; Di Toro et al. 2002), effects range-low or effects range-medium (ERLs or ERMs; Long and Morgan 1990), threshold effects concentrations and probable effects concentrations (TECs and PECs; MacDonald et al. 2000), no effects level, lowest effects level or severe effects level (NEL, LEL, or SEL; Persaud et al. 1993), maximum permissible concentrations (MPC; Kalf et al. 1999),and a predicted no effect concentration (PNEC; ECB 2003).

Chapman (1989) gives a detailed description of how and why SQC are developed. He lists five reasons that in addition to WQC, sediment quality criteria are also needed:

1) various toxic contaminants found in only trace amounts in the water column accumulate in sediments to elevated levels,
2) sediments serve as both a reservoir and a source of contaminants to the water column,
3) sediments integrate contaminant concentrations over time, whereas water column contaminant concentrations are much more variable,

4) both sediment contaminants and water column contaminants affect benthic and other sediment-associated organisms, and
5) sediments are an integral part of the aquatic environment, providing habitat, feeding and rearing areas for many aquatic biota.

While SQC can refer to chemical-specific concentrations in sediment expected to cause adverse biological effects, or levels of biological effects that are considered unacceptable, we will focus on SQC methods that yield chemical-specific numerical concentrations. According to Chapman (1989), the advantages to these types of criteria are that they are widely applicable and their application is straightforward, requiring no specialized biological, chemical, or other expertise. The disadvantages of individual numerical criteria are that they may overlook toxicity due to other chemicals and the values are not flexible to account for site-specific variations, particularly the bioavailability and subsequent toxicity of sediment-bound chemicals.

The USEPA (1993) states that numerical SQC can be used in a similar manner as numerical WQC, may be causing toxicity. One difference between WQC and SQC is that the regulatory basis and implementation of SQC has not been established, whereas WQC have been implemented for regulation of water contaminants since they were first published by the USEPA in 1980. The USEPA (1993) also states that the application of SQC may differ from WQC in that water column contaminants can typically be controlled by limiting the sources, whereas sediment contaminants may have been accumulating for some time, so source limitation would not necessarily eliminate toxicity. Although, it appears that source control could be effective for pyrethroids, because they typically enter water bodies via particles and have only moderately long half-lives (Gan et al. 2008, Laskowski 2002).

Many researchers have questioned the use of numeric SQC altogether, and have cautioned against their use for regulatory purposes or as pass/fail criteria. For this reason, many of the methodologies refer to the values they derive as sediment quality guidelines instead of SQC, and they recommend using the values as part of a risk assessment framework or as triggers for further research, and not as enforceable criteria. Burton (2002) stated that SQGs will probably always be used as screening tools, not enforceable regulatory values, because of the complexity of sediments and mixture interactions. Burton states that SQGs do not describe: 1) microscale variation, 2) inorganic speciation differences, 3) stressor interactions, 4) dynamics of biota, and 5) critical physicochemical parameters. Chapman (2000, 2007) is also wary of the emphasis placed on numeric chemical values, and points out that generic chemical values can help identify areas, sources and contaminants of concern, but sediment chemistry does not actually provide information on bioavailability or toxicity. Chapman and Mann (1999) identified the key limitations of numeric SQGs:

1) degree of conservatism – inaccurate and uncertain numeric results tend toward under- or overprotection,
2) bioaccumulation/biomagnification – SQGs are usually based on direct toxicity to organisms and do not address effects of long-term bioaccumulation and biomagnification,

3) bioavailability – SQGs cannot be applied to all sediment conditions because bioavailability varies greatly depending on conditions (although a mechanistic approach attempts to normalize for this variation),
4) contaminant mixtures – SQGs are often based on data from field samples that likely contain more than one chemical, which can confound SQG databases,
5) predictability – it's not clear how good SQGs are at predicting adverse ecological effects (can be a high % of false-positives or false-negatives).

3.1 Numeric criteria vs. advisory concentrations

Numeric SQC have been derived for a few compounds by the USEPA (dieldrin, endrin, PAHs), but none of these numeric criteria have been adopted as sediment quality standards. Instead, they may be used to interpret site-specific sediment chemistry data as part of an environmental risk assessment framework to predict or identify the degree and extent of contamination, or possibly to implement narrative criteria, but are not used as a pass/fail enforceable standard. Because of the way these numeric values are used, they are typically not referred to as criteria, but instead are called guidelines. These guidelines are considered advisory concentrations because the evidence thus far does not indicate that numeric SQGs are always predictive of effects. The USEPA further divides their guidelines into Tier 1 and Tier 2 categories; Tier 1 guidelines are calculated with more data and are associated with higher certainty than Tier 2 guidelines – although both tiers are still considered advisory concentrations. The Canadian method sets interim sediment quality guidelines when data are limited, yet both interim and full guidelines may be used as a basis to set enforceable, site-specific sediment quality objectives (CCME 1999).

Numeric SQGs are used in many risk assessment frameworks in the early assessment tiers to identify potentially toxic contaminant levels (NOAA, Washington State, Great Lakes; USEPA 1994a). When a contaminant exceeds a SQG in risk assessment, it is typically a trigger for further investigation, not a basis for regulatory action. Further investigation in the higher tiers of risk assessment usually include field sediment bioassays to determine if adverse effects are caused by the sediment, the results of which can lead to various management decisions (Apitz and Power 2002).

3.2 Numeric criteria of different types and levels

Many water quality criteria derivation methodologies include procedures for derivation of more than one level or type of criterion for each toxicant (ANZECC and ARMCANZ 2000, La Point et al. 2003, Lepper 2002, OECD 1995, RIVM 2001, USEPA 2003d), but this is generally not the case for sediment quality criteria derivation methodologies. Typically, there is not enough data or knowledge about sediment toxicity to enable derivation of different levels of criteria to meet different regulatory goals, such as for use in enforcement versus use in risk assessment. There are two methods that do offer more than one type of sediment criteria, those of The Netherlands and France.

Compartment-specific environmental risk limits (RLs) are derived in The Netherlands (RIVM 2001), and the same general protocols are followed, whether deriving environmental RLs

for the water or sediment compartment. The three levels of environmental RLs are the ecosystem serious risk concentration (SRC_{ECO}), the MPC and the negligible concentration (NC). The NC (concentration causing negligible effects to ecosystems) is calculated as the MPC divided by a safety factor of 100 and represents a regulatory target value. The MPC is a concentration that should protect all species in ecosystems from adverse effects. If concentrations in sediments are above the MPC, discharges can be further regulated. The SRC_{ECO} is a concentration at which ecosystem functions will be seriously affected or are threatened to be negatively affected (assumed to be when 50% of species and/or 50% of microbial and enzymatic processes are possibly affected; RIVM 2001). Sediments that exceed the SRC_{ECO} require clean-up intervention efforts.

In the French methodology (Lepper 2002), four threshold levels can be calculated for sediment and suspended matter for substances with log-normalized octanol-water partition coefficients ($\log K_{ow}$) > 3. Each threshold level corresponds to a different biological quality suitability class for water bodies. Threshold level 1, indicating negligible risk for all species, is derived from either chronic or acute toxicity data, with assessment (or safety) factors applied. The level 2 threshold indicates possible risk of adverse effects for the most sensitive species, and is derived from the same data as level 1, but with smaller assessment factors applied. Levels 3 and 4 indicate probable or significant risk of adverse ecosystem effects, respectively, and are derived solely from acute data, but Levels 3 and 4 have never been calculated for sediment because of a lack of required data. Threshold levels for sediment and suspended particles are all considered provisional because of the inherent uncertainties in the two methods used to calculate these values – equilibrium partitioning and weight-of-evidence. None of the threshold values derived by the French methodology is enforceable; values serve as references for risk assessment and actions.

By whatever name, all of the numbers discussed (including those not currently used in setting sediment quality standards or objectives) represent efforts to estimate concentrations of chemicals that, if exceeded, might lead to loss of designated uses of water bodies. When data are limited, numeric criteria of low site-specificity and high uncertainty can be derived, then as more data become available, criteria can be refined for better site-specificity and greater certainty.

4 Protection and confidence

Aquatic life sediment quality criteria are intended to protect aquatic life from exposure to toxic substances. The protection of aquatic life can be defined in various ways, from overall ecosystem protection to protection of each individual in the ecosystem. The existing methodologies specify different protection goals in terms of the level of ecosystem organization, and how to approximate the protection level by extrapolating ecosystem effects from existing data. There is also a discussion of the importance of being able to state, with a quantified level of certainty that criteria are achieving the intended level of protection.

4.1 Level of organization to protect

A full discussion defining the levels of ecosystem organization in general can be found in TenBrook and Tjeerdema (2006). Most of the methodologies reviewed designate what level of

organization is to be protected by SQC. Several SQC derivation methodologies seek to protect each species, expecting that by doing so, they will protect ecosystems. Canada's guiding principles for the development of numerical SQGs states that guidelines are "set with the intention to protect all forms of aquatic life and all aspects of the aquatic life cycle" (CCME 1999). The USEPA's goal is to be protective of benthic aquatic species, as stated in their EqP methodology (Di Toro et al. 2002). The Netherlands also has the goal of protecting all species in ecosystems from adverse effects (RIVM 2001).

Most of the reviewed methodologies specifically seek to protect aquatic ecosystems. France derives threshold levels that will maintain an ecosystem's suitability to support its biological function and other uses (Lepper 2002). The province of Ontario, Canada, states the purpose of SQGs is to protect the aquatic environment (Persaud et al. 1993). The state of Florida has the goal of protecting living resources and their habitats (MacDonald 1994). The objective for setting sediment action levels in the UK is the "maintenance of environmental quality so as to protect aquatic life and dependent non-aquatic organisms," (Rowlatt et al. 2002). In Australia/New Zealand the goal is "to maintain and enhance the 'ecological integrity' of freshwater and marine ecosystems, including biological diversity, relative abundance and ecological processes" (ANZECC and ARMCANZ 2000). The OECD guidelines provide methods for derivation of criteria "where no adverse effects on the aquatic ecosystem are expected" (OECD 1995). The predicted no-effect concentrations derived by the EU risk assessment methodology (ECB 2003) are intended to ensure "overall environmental protection." Finally, as discussed previously, the Basin Plan of the Central Valley Regional Water Quality Control Board states that "discharges shall not result in pesticide concentrations in bottom sediments or aquatic life that adversely affect beneficial uses," (CRWQCB-CVR 2009), which include preservation of wildlife and their habitats.

Diaz and Rosenberg (1996) point out that thus far, the functional component of benthic communities has been overlooked, and the concern for alteration or loss of function is greatest at the ecosystem level. Benthic organisms mediate the cycling of materials between sediments and the overlying water column via burrowing, irrigation, and other behaviors. The main effect of toxicants in sediment is to lessen the importance of key community elements associated with sediment mixing and energy flow. This happens because larger, long-lived species are eliminated and these species process large amounts of sediment and tend to stabilize annual productivity. Once community functions are stressed they tend toward 1) accumulation and storage of contaminants in sediments, which are enhanced through reduced or eliminated bioturbation and 2) shortened food chains, where energy is passed through smaller, faster, growing species. These authors posit that functional changes are more indicative of ecosystem impacts than community structure changes (Diaz and Rosenberg 1996).

4.2 Portion of species to protect

In contrast to water quality criteria derivation methodologies, most SQC derivation methodologies do not primarily rely on single-species sediment toxicity data to calculate criteria or guidelines. Not only is there a dearth of spiked-sediment toxicity data, but also a lack of consistency in the data due to variable bioavailability across different sediments. Multispecies or ecosystem data are alternative options for use in criteria derivation, but these types of data are

also relatively few and difficult to interpret. As summarized by TenBrook and Tjeerdema (2006), protection of less than 100% of species may cause unpredicted harm to an ecosystem because each species both performs a function, and takes part in complex trophic interactions in the ecosystem structure. To ensure protection of entire ecosystems, both the functions and structure must be maintained. This document presents and evaluates alternative methods for estimating ecosystem no-effect concentrations by extrapolating from available toxicity data, either from aqueous toxicity data in the equilibrium partitioning methods, from spiked-sediment toxicity data, or from co-occurrence field data. Yet, the only way to confirm that SQC are actually protective of ecosystems is to perform field or semi-field studies.

4.3 Probability of over- or underprotection

To give environmental managers some knowledge of how likely it is that a criterion will provide the intended level of protection, criteria are best expressed with associated confidence limits. Criteria that overprotect lead to unnecessary expenditures, while criteria that underprotect may lead to ecosystem damage. The effect levels derived by the French method do not have confidence limits associated with them because they are calculated by applying an assessment (or safety) factor to the single most sensitive datum (Lepper 2002). The EU also recommends applying an assessment factor to spiked-sediment toxicity test data if it is available (ECB 2003). These criteria may be protective, but there is no way to know to what degree they are likely to over- or underprotect. Uncertainty analyses can be done with the EqP approach if there are sediment toxicity data for the compound of interest. Confidence limits can be estimated as the degree to which the sediment toxicity data are predicted by the EqP model (USEPA 2003b). The EU mentions that SSDs could be used if there was data available, but since sediment data are generally lacking, they do not provide full guidance on the use of SSDs with sediment data and how confidence limits could be calculated. The Netherlands (RIVM 2001) uses a SSD technique that derives criteria at specified confidence levels, but if ample data are not available, then an assessment factor is applied to selected data, and this approach would not provide confidence limits. If a SSD were used, confidence limits provide useful information. For example, for a criterion derived at a 50% confidence level the true no-effect level may be either above or below the derived criterion with equal probability. If derived at a 95% confidence level, there is only a 5% chance that the true no-effect level lays below the derived criterion. This kind of information can provide environmental managers with some sense of the reliability of criteria.

5 Data

Quality data must be used in order to derive scientifically sound SQC, and all of the methodologies covered in this review require both physical-chemical and ecotoxicity data. The quantity of data required to derive the criteria must also be adequate to minimize uncertainty. Ideally long-term (chronic) spiked-sediment toxicity test data for benthic organisms (considering various routes of exposure) would be available for contaminants across a wide range of sediment types. This would allow criteria to be based on a known cause and effect relationship. Although these data are limited, the goal of the following section is to identify how much and what kinds of data are required to generate numerical sediment criteria using various approaches worldwide.

5.1 Data sources and literature searches

It is important to identify the data sources employed by various agencies worldwide to derive SQGs. Whether an empirical, mechanistic or spiked-sediment toxicity test approach is taken dictates the source data necessary to derive SQGs. Despite the underlying basis of the approach, all data should be evaluated for relevance and reliability for deriving a SQG. It is important for all methods to incorporate guidance on where and how to find data to ensure the most up-to-date dataset is used to calculate SQGs. The data sources and literature review protocols of agencies using the mechanistic EqP approach and SSTT approaches will be the focus of this review.

An extensive review of data sources and literature search protocols for the derivation of water quality criteria was conducted by TenBrook and Tjeerdema (2006). Since the EqP approaches for deriving SQGs use WQC, the data sources used for generating WQC are used in this approach. In the United States, pesticide registrants must submit aquatic toxicity data to the USEPA, and also to the California Department of Pesticide Regulation (CDPR) if they register for use in California. Both of these agencies maintain databases of available toxicity data for pesticides (OPP Pesticide Ecotoxicity Database and CDPR Pesticide Data Index) that may be requested from the respective agency. These databases contain both aqueous and sediment toxicity test data. In addition, the USEPA maintains the ECOTOX database that contains information about single-chemical toxicity studies for aquatic and terrestrial life published in peer-reviewed literature. The Danish WQC methodology recommends the LOGKOW (2000) database for a source of evaluated octanol-water partition coefficients, maintained by Sangster Research Laboratories. A literature search using the BIOSIS database is also recommended. In the Australia/New Zealand derivation methodology for WQC, they suggest collecting data from international criteria documents, the ECOTOX database, open literature and review papers. The online hazardous substance data bank (HSDB), as well as Verscheuren (1983; most recent version 2001 CD-ROM) and Hansch et al. (1995), are sources for physical-chemical property data, which are also needed for the derivation of SQGs. According to TenBrook et al. (2009), the WQC methodologies from France, the EU and the USEPA do not offer specific guidance about where to find data or how to complete an adequate literature search.

The Dutch protocol (RIVM 2001) for deriving SQGs gives detailed information about the different data sources that should be used to find relevant ecotoxicity data and physical and chemical property data. Data to search for include: ecotoxicity data for all aquatic species (freshwater and saltwater), soil organisms, enzymatic activities, microbial processes, sediment dwelling organisms (and birds and mammals if secondary poisoning is a concern) and partitioning coefficients. Biocide and plant protection products require environmental fate and toxicity data to be submitted to the Dutch Board for the Authorization of Pesticides (CTB) and this information can be requested for review and guidance. The Dutch method states that only primary literature is to be used for SQG derivation. It was proposed by Kalf et al. (1999) that ecotoxicity endpoints from the registration dossier and the scientific literature should be used as well as environmental fate endpoints from the registration materials. Online searches of bibliographic databases are useful sources of information: BIOSIS for ecotoxicity data and physical/chemical data, Chemical Abstracts for partitioning coefficient data, and TOXLINE for mammalian ecotoxicity data. Reliable sources for estimated or empirical physical-chemical properties can be found in the handbooks of MacKay et al. (1999) and Boethling and MacKay

(2000). Libraries recommended by the Dutch method include the Centre for Substance and Risk Assessment, the National Institute of Public Health and the Environment (RIVM) and the CTB library. The grey literature is to be searched only if there is time and a budget to do so and the secondary literature should only be used to identify primary sources (Kalf et al. 1999). The Dutch method requires the literature review to go back to at least 1970 or to the beginning of a database in order to identify all available literature, especially if there has not been an extensive review article published that identifies all primary sources of relevant data (RIVM 2001).

To derive WQC, the USEPA utilizes an aquatic acute and chronic toxicity database (OPP Pesticide Ecotoxicity Database), including both benthic and water column dwelling species of varying sensitivities. The procedure for deriving the FCV and the minimum data requirements to do so are described by USEPA (1985). The USEPA methodology states that a "complete search, retrieval and review for any applicable data must be conducted, to locate all preexisting toxicity data." It is important to re-examine the FCV of a compound to ensure that the most up-to-date toxicity data are included, or it is suggested that the FCV can be calculated if one does not exist for the chemical of concern, if the minimum data provisions set forth by USEPA (1985) are met. Literature searches are also recommended as sources of toxicity data but the USEPA methodology does not specify a procedure to evaluate the quality of a literature study.

The EU protocol is part of an overall risk assessment framework which calls for "the collection of all available information by manufacturers, importers and rapporteur" (ECB 2003). Little information is given as to where or how to find ecotoxicity data that is "complete and adequate" for use in the derivation of predicted no effect concentrations for individual chemicals. The EU protocol states that test results from peer reviewed journals are preferred, but quality review articles, summaries and abstracts may be used as supporting materials (ECB 2003). There is no mention of specific sources of data.

Ontario's EqP method does not include recommended data sources. The EqP method of the UK also lacks guidance on specific sources of data (Rowlatt et al. 2002), but published literature, commercial databases and unpublished data (e.g. manufacturer's data) must be gathered when deriving environmental quality standards in the UK (Zabel and Cole 1999). The data gathered is to be summarized and assessed for physical-chemical properties, adequacy of methodologies used, environmental fate and behavior, environmental concentrations, toxicology and bioaccumulation (Zabel and Cole 1999). The OECD (1995) does not describe where to locate data or provide a protocol for conducting an adequate literature search.

NOAA operates the National Status and Trends Program, which collects large amounts of sediment quality data. The NSTP program has operated the national mussel watch program and the bioeffects assessment program since 1986 for US coastal waters. Sediment and bivalve tissue chemistry has been collected for a suite of organics and trace metals. The bioeffects assessment program uses the sediment quality triad (sediment chemistry, sediment toxicity testing and species diversity assessment) to identify and assess contaminant exposure effects using over 40 regional studies since 1986. These monitoring efforts have yielded a large database of matching sediment chemistry and biological effects data (Long and Morgan 1990). This dataset is referred to as a biological effects database for sediments (BEDS). Because of the ongoing nature of the program, sample collection and analysis methods are standardized and well documented. Data

quality control before entry into the database is outlined in standard protocols (Long and Morgan 1990). Agencies using empirical approaches to derive SQGs require matching sediment chemistry and biological effects data for input into a BEDS. Agencies or jurisdictions without adequate local data often use the NOAA NSTP database for sediment assessments and BEDS development, and update the database as new information becomes available. Australia and New Zealand (ANZECC and ARMCANZ 2000), Florida (MacDonald 1994) and Canada (CCME 1999) all take this approach. Washington has utilized a local freshwater sediment quality database (FEDSQUAL) of matching sediment chemistry and biological effects data from both Oregon and Washington State (Cubbage et al. 1997) to derive SQGs. The California empirical method (SWRCB 2011) does not describe where to locate data or provide a protocol for conducting an adequate literature search to identify physical-chemical or ecotoxicity data.

According to the Canadian method (CCME 1999), a comprehensive review and search of the literature is required for the compound of interest. The physical-chemical properties of the chemical are to be summarized, but there is no mention of the source of these data or guidance on how to judge data acceptable. Toxicological studies are to be found in the scientific literature and should be reviewed for quality according to the procedure outlined in the NSTP methodology. The literature review is to be used to summarize the production and uses, known environmental fate data, sources of the chemical into the aquatic environment, and to help evaluate and establish background concentrations. The Canadian method (CCME 1999) recognizes the need to find toxicological data from sediment exposures since the NOAA NSTP empirical approach does not incorporate bioavailability.

The approach taken will dictate the data required to derive SQGs. No matter the approach, guidance on where and how to find quality sources of data are integral parts of a quality methodology to set environmental limits, whether in sediment or water. The most complete and high quality dataset is desired, thus guidance for gathering relevant data should be a part of any methodology.

5.2 Physical-chemical data

Data on the physical-chemical properties of a compound can be used as a starting point for understanding how a chemical will move and persist in the environment and provide a basis for understanding which environmental compartments (air, water, soil, sediment, biota) are at risk for chemical exposure. Each SQG derivation methodology describes the kinds of physical-chemical data to be gathered in different levels of detail.

In The Netherlands, the physical-chemical data required include: the International Union of Pure and Applied Chemistry (IUPAC) name, Chemical Abstracts Service (CAS) number, EINECS number (European Inventory of Existing Commercial Substances), diagram of structural formula, empirical formula, molar mass, K_{ow}, water solubility, melting point, vapor pressure, Henry's law constant, and the acid dissociation constant (pK_a). In addition to the aforementioned parameters, K_d (referred to as K_p in RIVM 2001 or as $K_{s/l}$ in Kalf et al. 1999) and degradation rates (biotic and abiotic processes) should also be gathered.

The OECD (1995) calls for the collection of the following data: chemical structure, molecular weight, melting point, water solubility, K_{ow}, K_d, and pK_a. The bioconcentration factor (BCF) may also be required if secondary poisoning is an issue. Experimentally determined BCFs are preferred instead of those estimated from the K_{ow}, because other factors, such as metabolism of the chemical within the organism, can affect the values.

The minimum physical-chemical data requirement of the USEPA methodology (Di Toro et al. 2002) is the K_{ow} of the compound, although an experimentally determined K_{oc} is preferred. There is a similar requirement in the UK (Rowlatt et al. 2002) and Ontario (Persaud et al. 1993) methodologies. The EqP criterion calculation uses a compound's K_{oc}, which can be estimated using the K_{ow} of a compound. The Ontario method requires at least three estimates of the K_d to set a SQG using the EqP approach (Persaud et al. 1993).

Since empirical approaches use mainly matching sediment chemistry and biological effects data from field collected sediments, a compound's physical-chemical properties are not a requirement for calculations of effects range concentrations. This does not minimize the importance of this data in understanding how a chemical will move and transform in the sediment environment, although little guidance for the collection of physical-chemical property data is found in the empirical methods (NOAA NSTP, Canada, Ontario and California).

5.2.1 Physical-chemical data quality

It is extremely important to have accurate physical-chemical data, especially K_{ow} or K_{oc} values, since these are directly used in derivation of sediment quality guidelines in the EqP approach. Partition coefficients of highly insoluble chemicals, such as pyrethroids, are difficult to determine experimentally, thus, values in the literature can vary by orders of magnitudes for such compounds. As a result, specific guidance is provided by the USEPA (Di Toro et al. 2002) and Dutch (RIVM 2001) EqP methods to address data quality issues associated with partition coefficients required to calculate the SQG for an individual chemical. It is recommended in the USEPA EqP method to use K_{ow}s from Karickhoff and Long (1995, Long and Karickhoff 1996) to calculate the K_{oc}, when available. Newer experimental methodologies to determine the K_{ow}, such as the slow stir method (de Bruijn et al. 1989) and the generator column method (Woodburn et al. 1984), should be used when finding literature-based K_{ow}s. There is no mention of site-specific K_{oc}s, but discussion about the ability to calculate the K_{oc} after routine toxicity testing (generation of a sediment-water isotherm that can be normalized for organic content and allow calculation of the K_{oc}) is discussed as a possible route for data collection without additional testing (Di Toro et al. 2002). Because the EqP approach relies on WQC to calculate the SQG, other physical-chemical data quality issues would be identical to those used in WQC derivation, which have been summarized by TenBrook and Tjeerdema (2006).

The Dutch method specifies that the chemical's water solubility, Henry's law constant, log K_{ow}, and K_d should be gathered as background information for SQG derivation. The method requires that the K_d must be experimentally determined following the protocol for batch isotherm experiments for organics described by Bockting et al. (1993). All information related to the K_d is considered useful, especially the Freundlich exponent (1/n). Only K_ds with a Freundlich exponent between 0.7 and 1.1 should be used in the SQG calculation (Kalf et al. 1999). The

humus, organic matter or organic carbon content must be reported along with the K_d. Temperature is an important variable to be considered when measuring equilibrium based values, such as the K_d, water solubility, vapor pressure, and Henry's constant. It is suggested in the Dutch method that 25°C is appropriate since this is the standard temperature for laboratory toxicity tests. Additional information gathered from partition coefficient studies should include pH, cation exchange capacity and mass balance calculation ability from data (preferably both water and sediment concentrations are measured). If experimental K_{oc} data are lacking, the values are collected from the SRC database or handbooks (e.g., MacKay et al. 1999). If still not available, K_{oc} can be estimated from K_{ow}. To derive log K_{oc} from log K_{ow} using a quantitative structure activity relationship (QSAR), the QSAR regression equations of Gerstl (1990) should be used:

$$\log K_{oc} = a \log K_{ow} + b \qquad (2)$$

where a and b are constants for specific groups of chemicals, presented in the Dutch method (RIVM 2001).

In the Dutch method, chemical degradation information is used to decide whether the parent or degradation product should be tested for toxicity. If hydrolysis is a main dissipation route and the time for 50% of the chemical to degrade (DT_{50}) is less than 4 hr, the tests are to be started with the metabolites. If the DT_{50} is 24 hr or more, the aquatic toxicity test is started with the parent compound. If the DT_{50} is between 4 and 24 hr, expert judgment is used or both parent and metabolites could be tested (Kalf et al. 1999). These dissipation criteria were proposed by Mensink et al. (1995) and used in the derivation of harmonized maximum permissible concentrations in The Netherlands.

The EU method states that measured K_ds are preferred, but they may be estimated using a K_{oc} or K_{ow}. Solid-water partition coefficients may be obtained from direct measurement, simulation testing, measured by adsorption studies or the high pressure liquid chromatography (HPLC)-method, or estimated from the K_{ow} using QSARs (ECB 2003). The OECD (1995) requires the K_{ow} of a compound to be determined using the slow stir method or generator column method for compounds with a log $K_{ow} > 5$. Expert evaluation of the values is also recommended by the OECD (1995) to ensure data quality. Ontario's EqP method states both measured and calculated K_{ow}s may be used to determine the K_{oc} of a compound, but experimentally derived K_{oc} data should be used for the EqP SQG calculation whenever possible. At least three estimates of partition coefficients are required to set a SQG using the EqP approach. If less than three values are available, the SQG is considered tentative (Persaud et al. 1993).

If bioconcentration or bioaccumulation is a concern, the BCF and/or the sediment (or soil) accumulation factors can be calculated or taken from several studies. Information gathered should include: species, species properties (e.g., age, size, weight, lifestage, sex, if known), test type (semi-static, static, continuous flow, intermittent flow), water properties (hardness or salinity), exposure time and concentration, time to equilibrium and dry to wet weight ratio (RIVM 2001). The USEPA (1985) provides criteria to which studies must adhere for BCFs to be used in the derivation of final tissue residue values, as described by TenBrook et al. (2009). Briefly, BCFs must be based on measured concentrations in tissue and test solution in a flow-

through experiment conducted at steady-state conditions. The percent lipid in tissue must be reported for lipophilic compounds and chemical concentrations reported on a wet weight basis. The geometric mean of all BCFs across species is calculated for those measured under similar conditions.

5.3 Ecotoxicity data

There are many types of ecotoxicity data available in the literature, from short-term to long-term chemical exposure studies, using a variety of endpoints (lethal, sublethal, biochemical). Ecotoxicity data are generated using both single- and multiple-species testing, as well as through laboratory and field studies. A detailed discussion of the different types of ecotoxicity data found in the literature can be found in TenBrook et al. (2009). Various toxicity values, such as a lethal or effects concentration (LC_x/EC_x), are generated in different ecotoxicity tests, whether they are aqueous or sediment exposures. Internationally accepted protocols for aquatic toxicity testing have been in place for over a decade, whereas many standard methods for sediment toxicity tests are still in the process of being developed or finalized, with new research on appropriate test organisms and endpoints, and variation across sediments. While sediment toxicity test methods have been developed by the USEPA, OECD and ASTM, many of these methods are considered guidelines and are not yet fully validated and accepted. Jurisdictions using the EqP approach often cite lack of sediment toxicity data and lack of sediment toxicity testing standard protocols as reasons that an indirect approach must be used.

5.3.1 Acute vs. chronic

Sediment quality criteria aim to be protective of aquatic life during both short-term transient exposures and long-term continuous exposures. In order to assess both the long-term and short-term effects of sediment contaminants on benthic organisms, different types of toxicity tests have been developed. Acute tests are short-term and measure mortality or immobility, while chronic tests are long-term and the endpoints can include survival, growth, emergence, reproduction, and others. The USEPA Office of Research and Development (USEPA 2000a), USEPA Office of Prevention, Pesticides, and Toxic Substances (OPPTS 1996a-j), OECD (1992, 2004a-c, 2007, 2008), Environment Canada (1997a, b), and ASTM (2004, 2006a, b, 2007a, b, 2008a-d, 2010) have all developed standardized methods for sediment toxicity testing. It is important to have clear guidance in a methodology on what types of data shall be used to derive criteria, and what the resulting criteria aim to protect.

The current ASTM (2008a) method for testing toxicity of sediment-associated contaminants with freshwater invertebrates (ASTM E 1706-05) specifies that short-term (acute) tests are 10-d and include both survival and growth endpoints. Long-term (chronic) tests for *Hyalella azteca* are conducted for 42-d with endpoints at 28, 35 and 42 d, measuring survival, growth and reproduction, while a long-term test for *Chironomus dilutus* entails a 20-d life-cycle test with endpoints of growth, survival, reproduction, and emergence. Acute and chronic sediment toxicity testing protocols are also described by OPPTS 850.1735 and OPPTS 850.1735S, respectively, for the freshwater organisms, *H. azteca* and *Chironomus tentans*.

Definitions of acute and chronic test durations are not necessarily consistent across SQG methodologies. Laboratory sediment toxicity tests lasting 10-14 d were defined as acute and tests lasting 21-60 d as chronic in a USEPA report (Ingersoll and MacDonald 2002), although test durations should vary with species since life-cycle durations can vary substantially between taxa. These definitions are consistent with the ASTM sediment toxicity test guidelines for invertebrates, as described above. The definitions of acute and chronic test durations given in the Dutch method (RIVM 2001) differ somewhat from those stated by the USEPA. The Dutch guidelines give the following taxa-specific definitions for test durations: tests up to 4-d for algae and protozoa are chronic (longer if still in exponential growth phase); tests of 48- or 96-hr are acute for Crustacea and Insecta; 96-hr tests are acute for Pisces, Mollusca, and Amphibia, while 28-d early lifestage tests are considered chronic for these taxa.

The Canadian methodology does not specifically define acute and chronic toxicity, but it does state that "ideally, SQGs should be developed from detailed dose-response data that describe the acute and chronic toxicity of individual chemicals in sediment to sensitive lifestages of sensitive species of aquatic organisms" (CCME 1999). In the Canadian method, at least two of the four minimum data requirements must be chronic tests, covering partial or full life-cycles.

Washington State defines an acute test as a measurement of biological effects using surface sediment bioassays that are short in duration compared to the life-cycle of the test organism. Acute effects include mortality, larval abnormalities or other endpoints deemed appropriate. Chronic tests are defined as measurements of biological effects using surface sediment bioassays over a period not less than one complete life-cycle of the test organism. The term chronic also includes evaluations of indigenous field organisms for long-term effects as well as the effects of biomagnification and bioaccumulation. Chronic effects may include mortality, reduced growth, impaired reproduction, histopathological abnormalities, adverse effects to birds and mammals or other endpoints deemed appropriate (Cubbage et al. 1997).

Unlike exposures in the water column, benthic invertebrates are typically exposed to sediment contaminants for extended periods of time because of the accumulative nature of bedded sediments (Ingersoll and MacDonald 2002). In a study by the USEPA (2000b), longer-term (chronic) sediment toxicity tests with growth and survival endpoints tended to be more sensitive than shorter-term (acute) tests. Based on these findings, Ingersoll and MacDonald (2002) recommended that chronic toxicity tests should be used to assess effects of contaminated sediment on aquatic organisms because they are more relevant for predicting effects in aquatic ecosystems.

5.3.2 Hypothesis tests vs. regression analysis

There are two main ways to analyze ecotoxicity data, regression analysis and hypothesis testing. In regression analysis, a regression equation is calculated that relates concentration to effects (Stephan and Rogers 1985), and a concentration causing a specific effect level (LC/EC_x) can be calculated for any chosen effect level. In hypothesis testing, the effects on treatment groups are compared to a control group to determine which treatment is significantly different from the control (Stephan and Rogers 1985). The highest concentration that is not significantly different from the control is called the no observed effect concentration or level (NOEC or

NOEL), and the lowest concentration that is significantly different from the control is call the lowest observed effect concentration or level (LOEC or LOEL). It is also possible to calculate the maximum acceptable toxicant concentration (MATC) as the geometric mean of the NOEC and LOEC. Hypothesis testing is typically used for chronic tests covering full, partial, or early lifestages.

The advantages and disadvantages of these two approaches are discussed in detail by TenBrook and Tjeerdema (2006), and they will not be repeated here since these issues are identical for sediment and aqueous exposure data. These authors noted that, in general, regression methods are preferred, but they are not commonly used in chronic tests. In order to use the little chronic data available, hypothesis test values are acceptable.

In Dutch methodology, there is guidance to convert all available chronic data into NOECs, including if EC_xs are given. In this case the NOEC is defined as the EC_{10} (RIVM 2001). For sediment data, the EU method (ECB 2003) recommends that a NOEC is used to calculate the predicted no-effect concentration in sediment, but an EC_{10} can alternately be used. Similarly, in the French method, threshold levels are calculated by dividing either NOEC or LC/EC_{50} by a defined assessment factor to estimate a no-effect concentration (Lepper 2002). In the SSTT method of Canada, it is recommended that SQGs be calculated from chronic LOEC data (CCME 1999). In general, the methods that provide guidance on using SSTT data to derive criteria recommend the use of chronic data, which are typically reported as hypothesis test results.

5.3.3 Single-species (laboratory) vs. multispecies (field/semi-field) data

Single-species laboratory tests are used in criteria derivation by the SSTT approach, for verification in the EqP approach, and are added to co-occurrence datasets because the tests are standardized and relatively easy to interpret. While field studies, mesocosm/microcosm tests, and multispecies laboratory tests better approximate natural ecosystems, these types of tests are criticized for lack of standardization, lack of replication, and difficulty of interpretation, as noted by TenBrook and Tjeerdema (2006). These authors give a thorough discussion of the advantages and disadvantages of these types of studies for aqueous exposures that also apply to sediment exposures, and conclude that examining these types of studies can be useful, but that it is not likely that criteria will be derived solely based on multi-species studies because of issues relating to cost-effectiveness, reproducibility, and reliability.

Several methodologies mention the use of multispecies field studies in the final stages of criteria derivation. Both The Netherlands (RIVM 2001) and the EU (ECB 2003) methods recommend comparing the results of multispecies field studies to the criteria derived from single-species laboratory toxicity tests. They note that field studies are difficult to interpret because they are so varied in their test parameters and exposures, unlike laboratory tests, and the Dutch method offers a set of criteria by which to evaluate field studies. The Dutch method states that NOECs from multispecies tests should be compared to the derived criteria to establish if the derived criteria could be underprotective of ecosystems (RIVM 2001).

5.3.4 Traditional vs. non-traditional endpoints

Traditional endpoints are those effects assessed in standard test methods that are clearly linked to population-level effects, and include survival, growth and reproduction. Traditional endpoints are recommended for use in all of the methodologies that were reviewed, but some methods utilize non-traditional endpoints on a case-by-case basis. Non-traditional endpoints can include endocrine disruption, enzyme induction or inhibition, behavioral effects, histological effects, stress protein induction, changes in RNA or DNA levels, mutagenicity, and carcinogenicity (TenBrook and Tjeerdema 2006).

The Canadian guidelines indicate that toxicity tests should follow standard methods that assess ecologically relevant endpoints, which include survival, growth, reproduction, and developmental effects (CCME 1999). Similarly, The Netherlands methodology states that only endpoints that affect the species at the population level are included (RIVM 2001), which are those related to survival, growth, and reproduction. The reproductive effects can include histopathological effects on reproductive organs, spermatogenesis, fertility, pregnancy rate, number of eggs produced, egg fertility, and hatchability. The Dutch method also states that data on other effects should be collected for comparison to the derived criteria to ensure the criteria are protective, especially for chemicals with specific modes of action (e.g., phthalates are suspected to be endocrine disruptors).

The EU method (ECB 2003) states studies that do not test endpoints used in standard methods (related to survival, growth, or reproduction) may possibly be used in effects assessment, but only if an expert judges such data can be included with other standard endpoints. Emergence, sediment avoidance, and burrowing activity are also considered as relevant endpoints, in addition to survival, growth, and reproduction. Non-traditional endpoints might include other behavior effects, photosynthesis, or cellular or subcellular effects.

5.3.5 Data estimated from interspecies relationships

Toxicity data are not available for many species that are likely to be present in ecosystems, and one proposed way to increase the number of species represented without performing toxicity tests is to estimate toxicity values using interspecies relationships. For aquatic toxicity data, the USEPA has developed a program to estimate acute toxicity based on data for more common test species that have larger datasets (Raimondo et al. 2010). Unfortunately, this program does not include data for sediment exposures and thus, cannot be used in sediment quality criteria derivation, and no other interspecies correlation approaches were identified that included sediment exposures.

5.3.6 Ecotoxicity data quality

Data quality is usually assured by using standardized methods of toxicity testing. In the development of SQGs, different types of sediment toxicity testing have been used, depending on the experimental question being addressed. In this section, sediment toxicity tests are reviewed, including those used for empirical approaches, as well as spiked-sediment toxicity tests applicable to derivation of SQC using a species sensitivity distribution or assessment factor approach. Ways to evaluate whether individual studies have properly followed a standard method are also reviewed.

5.3.6.1 Standard methods

A selected list of current standard sediment toxicity testing methods and related protocols is provided in Table 3. The Dutch method suggested that there was a lack of internationally accepted protocol for toxicity studies in sediment-water systems (RIVM 2001), but this appears to be outdated. There are now standard methods available from several jurisdictions that were developed since 2000, which are further described in this section. The ASTM E 1706-05 (2008a) method, entitled "Standard test methods for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates," describes short-term 10-d testing protocols for *Hyalella azteca* and *Chironomus dilutus* (formerly *C. tentans*) using whole sediments from field-collected or laboratory-spiked sediments. This method also provides guidance on conducting short-term sediment toxicity tests using *Chironomus riparius*, *Daphnia magna*, *Ceriodaphnia dubia*, *Hexagenia* spp., *Tubifex tubifex* and *Diporeia* spp. Long-term sediment toxicity testing guidance is also provided in ASTM 1706-05(2008a) for *H. azteca* and *C. dilutus*, and can be applied to the other organisms for which there is short-term testing guidance. Bioaccumulation of sediment-associated contaminants is addressed in a separate method (ASTM 1688), which details a 28-d study with the oligochaete *Lumbriculus variegatus*. According to ASTM 1706-05 (2008a), future method updates will include results of research into the use of formulated sediment, refinement of sediment spiking procedures and evaluation of endpoint sensitivities.

Species included in the ASTM 1706-05 (2008a) method, representing freshwater organisms with different feeding and habitat requirements, are presented in Table 4. The datasets for *Hexagenia* spp., *T. tubifex*, and *Diporeia* spp. are not as robust as the data available using *H. azteca* and *C. dilutus*, and thus their sediment toxicity tests are currently considered guidelines and not standard methods (ASTM 1706-05 (2008a)). In addition to the invertebrate method, there is also ASTM guidance on conducting sediment toxicity tests with amphibians (ASTM E 2591-07), but this protocol is also considered a guideline, and not an official test method. The lack of standard protocols for testing sediment toxicity toward a wider range of benthic community members has been a critique of SSTT derivation methodologies, but it should be noted that some guidance is currently available, and it appears that more standard methods will be available in the near future.

Acute and chronic sediment toxicity testing protocols are also described by OPPTS 850.1735 and OPPTS 850.1735S, respectively, for the freshwater organisms, *H. azteca* and *C. tentans*. The OPPTS 850.1735 method is based on the USEPA (1994b) protocol entitled "Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates" (EPA 600-R24-024), and represents the harmonized version of the USEPA's ecological effects test methods. The USEPA's Office of Prevention, Pesticides, and Toxic Substances, newly named the Office of Chemical Safety and Pollution Prevention (OCSPP) has developed harmonized test guidelines for pesticide and toxic substances registration. Not all of the guidelines are considered final, including the ecological effects test guidelines (OPPTS series 850), but can still be used for study protocol development. The goal of the harmonized guidelines is to minimize the variation among the testing procedures used to fulfill data requirements for the Toxic Substance Control Act and the Federal Insecticide, Fungicide, and Rodenticide Act. This includes guidance documents prepared by the OECD Office of Pollution Prevention and Toxics (OPPT) and the USEPA Office of Pesticides Programs

Table 3 Selected list of sediment toxicity testing methods and related protocols.

Method Source	Method Number	Title
OPPTS (1996a)	850. 1735 (S)	Whole sediment acute (or chronic) toxicity: Invertebrates, freshwater.
OPPTS (1996b)	850. 1740	Whole sediment acute toxicity: Invertebrates, marine.
OPPTS (1996c)	850. 1790	Chironomid sediment toxicity test.
OPPTS (1996d)	850. 1800	Tadpole/sediment subchronic toxicity test.
OPPTS (1996e)	850. 1850	Aquatic food chain transfer.
OPPTS (1996f)	850. 1900	Generic freshwater microcosm test, laboratory.
OPPTS (1996g)	850. 1925	Site-specific aquatic microcosm test, laboratory.
OPPTS (1996h)	850. 1950	Field testing for aquatic organisms.
OPPTS (1996i)	850. 1010	Aquatic invertebrate acute toxicity test, freshwater, daphnids.
OPPTS (1996j)	850. 1075	Fish acute toxicity test, freshwater and marine.
ASTM (2008a)	E 1706-05 (2008)	Standard tests method for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates.
ASTM (2008b)	E 1367-03 (2008)	Standard test method for measuring the toxicity of sediment-associated contaminants with estuarine and marine invertebrates.
ASTM (2008c)	E 1391-03 (2008)	Standard guide for collection, storage, characterization, and manipulation of sediments for toxicological testing and for selection of samplers used to collect benthic invertebrates.
ASTM (2008d)	E 1525-02 (2008)	Standard guide for designing biological tests with sediments.
ASTM (2010)	E 1688-10	Standard guide for determination of the bioaccumulation of sediment-associated contaminants by benthic invertebrates.
ASTM (2006a)	E 2455-06	Standard guide for conducting laboratory toxicity tests with freshwater mussels.
ASTM (2007b)	E 2591-07	Standard guide for conducting whole sediment toxicity tests with amphibians.
ASTM (2006b)	E 1295-01 (2006)	Standard guide for conducting three-brood, renewal toxicity tests with *Ceriodaphnia dubia*.
ASTM (2004)	E 1193-97 (2004)	Standard guide for conducting *Daphnia magna* life-cycle toxicity tests.
USEPA (1994b)	EPA 600-R24-024	Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates (1994).
USEPA (1994a)	EPA 905-R94-002	Assessment guidance document, Great Lakes Program (EPA 600-R94-025; EPA 600-R99-064).
OECD (2004a)	218	OECD No. 218: Sediment-water chironomid toxicity using spiked sediment.
OECD (2004b)	219	OECD No. 219: Sediment-water chironomid toxicity using spiked water.
OECD (2007)	225	OECD No. 225: Sediment-water *Lumbriculus* toxicity test using spiked sediment.
OECD (1992)	210	OECD No. 210: Fish, early-life stage toxicity test.
OECD (2004c)	202	OECD No. 202: *Daphnia* sp. acute immobilization test.
OECD (2008)	211	OECD No. 211: *Daphnia magna* reproduction test.

Table 4 Species included in the ASTM 1706-05 (2008a) method, representing freshwater organisms with different feeding and habitat requirements.

Species	**Phylum (Sub-Phylum or Class)/Family**	**Habitat**	**Eating Characteristics**	**Other important characteristics**
Hyalella azteca	Arthropoda(Crustacea)/ Hyalellidae (amphipod)	E	Some subsurface deposit feeding	Wide tolerance of sediment grain size
Daphnia magna	Arthropoda(Crustacea)/ Daphniidae (water flea)	WC	Filter feeder	
Chironomus riparius	Arthropoda(Insecta)/ Chironomidae (midge)	E	Filter feeder/surface deposit feeder	Larvae burrow into sediment (direct contact); wide tolerance of sediment grain size
Chironomus dilutus (formerly C. tentans)	Arthropoda(Insecta)/ Chironomidae (midge)	E	Filter feeder/surface deposit feeder	Same as *Chironomus riparius*
Ceriodaphnia dubia	Arthropoda(Crustacea)/ Daphniidae (water flea)	WC	Filter feeder	
Tubifex tubifex	Annelida/Tubificidae (oligochaete)	E, I	Subsurface deposit feeder	Tolerant of variation in sediment particle size and proportion OM; important ecological link in aquatic food chain and active in bioturbation
Hexagenia spp.	Arthropoda(Insecta)/ Ephemeridae (mayfly)	E, I	Surface particle collector	Nymphs burrow into sediment (direct contact); prefers fine/organically enriched sediments
Diporeia spp.	Arthropoda(Crustacea)/ Pontoporeiidae (amphipod)	E, I	Deposit feeder	Relatively insensitive to grain size
Lumbriculus variegatus	Annelida/Lumbriculidae (oligochaete)	E, I	Subsurface deposit feeder	Inhabits a wide variety of sediment types

Habitats: I – Infaunal, E – Epibenthic, WC – Water Column

(OPP). OPPT methods were published in the Federal Register and OPP methods were published by the National Technical Information Service (NTIS). The test guidelines in the OPPTS 850series for aquatic fauna ecological assessment are listed in Table 3.

Fleming et al. (1996) performed intra- and inter-laboratory comparisons of sediment toxicity tests and they found that most of the variability between test results could be attributed to sediment spiking procedures. They determined that standardized spiking methods would need to address sediment heterogeneity, appropriate characterization of the variables controlling sorption and bioavailability, equilibration and aging times. Fuschman and Barber (2000) offer a simpler solution to address sediment spiking issues – measurement of sediment concentrations pre- and post-test to confirm the actual test exposures and check for chemical losses. Sediment spiking guidance provided by the ASTM (E 1367-03 (2008b)) includes confirmation of sediment concentrations before toxicity test initiation.

5.3.6.2 Relevance and reliability of studies

A detailed description of the processes by which aquatic ecotoxicity data are judged for quality in The Netherlands, UK, Canada, and Australia/New Zealand is presented by TenBrook et al. (2009). EU protocols specify adequate and complete ecotoxicity data generated using standardized, internationally accepted protocols (Table 3) should be used in criteria derivation. Test procedures and designs which deviate from the standard are usually reviewed using best professional judgment (ECB 2003). The adequacy of a study applies to the reliability of the study: the quality of test method used (ASTM, OECD, USEPA test methods with good laboratory practice) and the description of the methods and protocols in the study. The relevance of the study includes the use of appropriate endpoints under relevant conditions and that the substance tested is representative of the substance being assessed.

In The Netherlands, a reliability index is used as a ranking system for ecotoxicity data quality (RIVM 2001). A score of 1 indicates the methodology used in the study is in accordance with accepted international test guidelines and/or Mensink et al. (1995); a score of 2 indicates less accordance with accepted test methods and a score of 3 does not fit the quality of data to be used to calculate the maximum permissible concentration. The Dutch method also states that the purity of the test substance has to be at least 80%. Data generated from a less pure substance cannot be included in the dataset used for calculation but can be used as supporting information. An exception is made for granulates and wettable powders where purity is between 20% and 80% if absences of carrier toxicity has been established. Studies using polluted animals are rejected, and aquatic studies must have at least an 80% recovery of the substance and may not test concentrations that exceed 10 times the aqueous solubility of the substance (with a maximum solvent concentration of 1 mL/L). Sediment toxicity test requirements are not as detailed as they are for aquatic studies in the Dutch methodology, but they suggest the consideration of the following test parameters: sediment characteristics (% organic carbon, particle size distribution, field or standard sediment), amounts of sediment and water, test method (static or flow-through), spiking method, measured concentrations in sediment and/or water and if the concentrations are at equilibrium, description of system including if sediment is suspended or bedded, and exposure route of organisms.

In Canada, “accurate and precise” data generated using standard sampling techniques and appropriate test methods are important to maintain data integrity. Sediment characteristics also need to be determined (e.g., grain size, total organic carbon) for interpretation of biological effects. The CCME (1999) specifically states test methods should include light and dark cycles and verification of the condition of test organisms throughout the duration of exposure. Chemical concentrations in the water should be measured at the beginning and end of the test in both the overlying water and sediment compartments. Data on the health and survival of the test organism before exposure should be documented for at least one week prior to start of test. Test organisms should not be used if significant mortality has occurred during this time frame.

The NSTP empirical approach used in Canada also requires quality ecotoxicity data for incorporation into a BEDS. A detailed description of the evaluation of ecotoxicity data is provided by CCME (1999). To ensure sediment ecotoxicity data quality, sampling, storage and handling of sediments should be consistent with standard protocols (e.g., ASTM 2008a-d, Environment Canada 1994a, Loring and Rantala 1992). In terms of holding time and storage, sediments should tested within 2 wk of collection and not be frozen. Toxicity tests following standard protocols are considered acceptable (e.g., ASTM 1990a, b, Environment Canada 1992a-c, 1995). Non-standard toxicity testing methods should be evaluated on a case by case basis. Sediment chemical concentrations must be measured (nominal not acceptable) with the number of measurements dependent on the chemical and test duration, using appropriate analytical techniques (CCME 1999). The sediment should be characterized for total organic carbon, particle size distribution, acid volatile sulfides, pH, redox conditions and sediment type. The overlying water should be characterized for pH, dissolved oxygen, total suspended solids, suspended and dissolved organic carbon and water hardness (and/or alkalinity) or salinity. Embryonic development, early lifestage survival, growth, reproduction and adult survival are preferred endpoints, although other endpoints related to organism pathology or behavior (avoidance, burrowing) may be considered. Control survival and response must be measured and within acceptable limits and should be appropriate for the lifestage of the organism tested. Aquatic ecotoxicity tests may be included in the dataset for sediment assessment using the Canadian NSTP methodology. These aquatic tests can be static, static renewal or flow-through. Maintenance of adequate environmental conditions must be demonstrated for the test duration. Unacceptable data are those lacking sufficient information to assess the adequacy of the test design, procedures and/or results, and are not included in the dataset for the SQG derivation.

The California State Water Resources Control Board requires that all test methods must adhere to USEPA or ASTM methodologies or otherwise must be approved by the State and Regional Water Boards (SWRCB 2011). In Washington, a quality assurance grade has been applied to each entered piece of data into the database and made available for reference. The grade (A-F) is assigned to each investigation based on protocols conducted and presented in the final report. The grade does not necessarily reflect data quality, but the quality of the amount and types of quality assurance procedures completed in the investigation (Cubbage et al. 1997).

5.3.7 Ecotoxicity data quantity

A full review of the data quantity required for calculation of WQC for use in EqP derivation methodologies is presented by TenBrook et al. (2009). The focus of this section is to

present direct methods of SQG derivation that use sediment toxicity tests as input data for criteria calculation, although there is very little guidance on this matter in any of the methodologies. The Canadian SSTT approach outlines the minimum dataset for deriving freshwater SQGs: at least four studies are required on at least two or more sediment-resident invertebrate species found in North American waters (CCME 1999). One benthic arthropod and benthic crustacean species must be included. At least two of these studies must be partial or full life-cycle tests that consider ecologically relevant endpoints (e.g., growth, reproduction, developmental effects). They also state that "ecologically relevant species" should be the focus of the data review. Alternatively, in the Canadian NSTP based approach, the minimum toxicological dataset required for interim sediment quality guideline derivation is at least 20 entries into the effects and no effects dataset. The Dutch method states that when sediment toxicity data are lacking, aquatic toxicity data can be used to indirectly calculate the sediment maximum permissible concentration via the EqP approach (RIVM 2001). Ontario's use of the screening level concentration (SLC) empirical approach provides guidance on effects and no effects database construction, similar to the Canadian NSTP approach. The range of concentrations entered into the database should span 2 orders of magnitude and include both heavily contaminated and relatively clean sites. At least 75% of the database entries must be benthic infaunal species, with proper taxonomic identification to at least the genus level. A minimum of 10 observations are required to calculate a species SLC (SSLC), and a minimum of 20 different SSLCs are required to calculate the SLC (Persaud et al. 1993).

For California, the SWRCB (2011) method states that a minimum of one short-term survival test and one sublethal test are required for each sediment sample collected from a particular station. Acceptable test methods for the short-term tests include whole sediment exposures for a 10-d duration, with survival as the endpoint using acceptable test organisms (*Eohaustorius estuarius, Hyalella azteca*, *Leptocheirus plumulosus*, *Rhepoxynius abronius*) tolerant of the sample salinity and grain size characteristics. Acceptable sublethal testing methods are whole sediment exposures for 28-d, using growth as an endpoint using *Neanthes arenaceodentata* as the test organism. The other acceptable test is a 48-hr sediment-water interface exposure using embryo development as an endpoint with *Mytilus galloprovincialis* as test organism. Sediment toxicity results are compared and categorized as nontoxic, low toxicity, moderate toxicity and high toxicity relative to control survival, as described by the SWRCB (2011). The average of all test responses determines the final line of evidence category and if the average falls in between, the average is rounded up to the next response category.

Burton et al. (1996) suggested that for field bioassays to adequately detect sediment toxicity, test design should consist of 2-3 assays of various grouping of the following species: *Hyalella azteca, Chironomus tentans, Chironomus riparius, Ceriodaphnia dubia, Daphnia magna, Pimephales promelas, Hexagenia bilineata, Diporeia sp, Hydrilla verticillata, and Lemna minor.*

5.4 Quantitative structure activity relationships (QSARs)

A review of the role of QSARs in filling data gaps for WQC derivation is presented in TenBrook et al. (2009). QSARs are a mathematical relationship between a compound's structure and its toxicity. In terms of sediment quality, QSARs could be used to predict K_{oc} from K_{ow}, as

discussed in the Dutch method. QSAR toxicity predictions for pesticides with specific modes of action, such as pyrethroids and organophosphates, are still in the early stages of development, but may be of use in the future (Zvinavashe et al. 2009).

5.5 Data combination and exclusion

Once all relevant data are collected and evaluated for reliability, it is possible that there could be multiple data for a single species, and guidance is needed on how to combine data or exclude some data. While many methods give guidance on how to deal with this for aquatic data, only the Dutch method specifically includes sediment data in their data reduction guidelines. Most other methods assume that there will be so few data that multiple data for a species will not be an issue, or if EqP is used for sediment quality criteria calculation then aquatic data are used instead of sediment data. TenBrook and Tjeerdema (2006) describe how most WQC guidelines recommend calculating the geometric mean of multiple data, and conclude that the geometric mean is preferable to the arithmetic mean for toxicity data.

The Dutch guidelines (RIVM 2001) select toxicity data to obtain one single reliable value for each species, for a given type of toxicity value (LC_{50} vs. NOEC). To exclude or combine data, the following guidance is given: 1) the geometric mean of multiple data based on the same endpoint should be calculated for each species, 2) if there are multiple endpoints for a given species, the most sensitive one is selected, 3) if there are tests with different lifestages for a given species, the most sensitive one is selected. The Dutch also provide guidance on how to convert all acceptable chronic data into NOECs, which is detailed in TenBrook and Tjeerdema (2006).

The EU method (ECB 2003) does not give specific guidance on data reduction for the sediment compartment, as it does for the aquatic compartment, mainly because it is expected that multiple data for one species will not be available. If the EqP approach is used to calculate a criterion for the sediment, then aqueous data are used so the data reduction procedures for aquatic data would be utilized. If there were multiple data for one species for the sediment compartment it would seem reasonable to follow the data reduction procedures outlined for the aquatic compartment, which are described by TenBrook and Tjeerdema (2006). Similarly, the OECD (1995) method does not offer guidance for reducing sediment data as it does for aquatic toxicity data because the EqP approach is used, which uses aquatic toxicity data.

Although SSTT data are likely to be few, there may be multiple data for the most common test species (i.e., *Hyalella azteca, Chironomus dilutus*), so guidance should be given on how to select which endpoint or duration is most appropriate for use in criteria derivation. Similar guidance as seen in WQC methods can be used to combine multiple data for a given endpoint/species combination.

6 Criteria calculation

The goal of this section is to present the various approaches and their procedures for the derivation of sediment quality criteria. The EqP and SSTT approaches are described, including the assessment factor (AF) and species sensitivity distribution approaches. Empirical approaches are also presented, although their limitations for deriving single numeric criteria have been

discussed previously. Other factors to consider in criteria derivation are discussed, including mixtures, bioaccumulation, secondary poisoning, threatened and endangered species, harmonization of criteria across environmental compartments, and data utilization.

6.1 Exposure considerations

It is important to consider the exposure factors that affect sediment toxicity, such as magnitude, duration and frequency of the exposure as well as exposure routes and sediment or particulate characteristics which contribute to bioavailability. This section provides a summary of how the different methodologies consider an organism's exposure to a sediment contaminant, as well as the most recent research on these topics.

6.1.1 Magnitude, duration, and frequency

Exposures to sediment contaminants will vary in magnitude, duration and frequency depending on the particular environmental conditions. It was suggested by TenBrook et al. (2009) that water quality criteria be defined in terms of magnitude, duration, and frequency in order to determine exceedances; this review will also discuss these aspects with regard to sediment quality criteria. TenBrook and Tjeerdema (2006) give the example that a criterion designed to protect against ongoing, chronic toxicant exposures that is stated only in terms of magnitude will be overprotective in a case of a brief, mild excursion above the criterion, but will be underprotective in the case of a brief, large excursion. In order to provide appropriate criteria, there are two approaches that derivation methodologies can take with regards to magnitude, duration and frequency: 1) incorporate some combination of magnitude, duration and frequency in each criterion statement, or 2) derive the magnitude only and leave duration and frequency determinations to site-specific management decisions (TenBrook and Tjeerdema 2006). Among the methodologies reviewed, derivation of the numeric magnitude of a SQG is addressed by all of them, while the duration and frequency components are only considered indirectly through guidance on monitoring for compliance.

As discussed by TenBrook et al. (2009), exposure duration is an important consideration in the Sacramento-San Joaquin River basins where short-term toxic pulses occur regularly as a result of runoff from rain events or agricultural discharges. While brief pulses of sediment-bound pesticides may cause short-term adverse effects (Balthis et al. 2010, Hose et al. 2002), the resulting deposited sediment contaminants may not cause long-term adverse effects (Forbes and Cold 2005), although repeated pulses may not allow for full recovery after exposure (Wallace et al. 1989). The magnitude and duration aspects together address the differences between pulse emissions and on-going chronic exposure as a result of residue accumulation in the sediments. In WQC, the duration aspect has often been addressed by deriving two separate criteria with magnitudes that vary depending on a short-term (acute criterion) or long-term duration (chronic criterion), while all of the SQG methods only recommend deriving a single criterion, typically with chronic data. The EqP approaches utilize WQC, which are described in terms of exposure magnitude, frequency and duration by the USEPA WQC method (USEPA 1985). Yet unlike in WQC, for which separate acute and chronic criteria are derived, the USEPA EqP methodology uses only the final chronic value, without the duration and frequency components, to determine a single chronic SQG. The other EqP approaches (The Netherlands, EU, Ontario, France, OECD,

and UK) also use a chronic water quality value to derive a single SQG. The use of chronic exposure data is one way to incorporate duration because the magnitude derived using this data should be protective of long-term exposures. Acute SQGs are not considered representative of contaminated sediment exposures because sediment contaminants are not as transient as aqueous contaminants and often accumulate in the sediments, leading to long-term exposures. The USEPA (2003a, 2003b) states that the duration and frequency components are not given in SQGs because it is expected that the concentration of sediment contaminants will be relatively stable over time, leading to chronic, relatively constant exposures of benthic species.

The frequency of exceedance component is designed to ensure that an ecosystem impacted by an excursion of the criterion has time to fully recover before another excursion might occur, because it is assumed that adverse effects may be compounded if multiple excursions occur before an ecosystem is recovered. Studies of toxicity due to pesticide-contaminated sediment suggest that recovery times may vary from as little as 1 mon up to 3 yr, and is dependent on many variables (Balthis et al. 2010, Caquet et al. 2007, Hatakeyama and Yokoyama 1997, Woin 1998, Yasuno et al. 1982). Frequency of exceedances may be the most difficult aspect to address for sediment because unlike water, sediment residues can accumulate over time. Sampling in the same location over time may not reveal new inputs, but instead may confound results with residual pesticides, for which exceedances were previously recorded. All of the methodologies included in this review only give instructions to derive the magnitude, and leave the duration and frequency components up to site-specific judgment by environmental managers.

Sediment quality guidelines are generally used to complement existing sediment assessment tools, assess sediment contamination and serve as targets for maximum contaminant loading in a water body (Di Toro et al. 2002). SQGs are currently not used in regulatory contexts alone, and as a result there is little to no discussion regarding the allowable duration and frequency of a SQG exceedance in the current methodologies. In most SQG methods, magnitude exceedances of SQGs are used as triggers for further study of the contaminant and development of management practices designed to reduce chemical loadings. If environmental managers want to consider the duration or frequency of sediment contaminant exposures, those aspects may be incorporated in the design of monitoring programs. For example, the Dutch and German methods use the 90^{th} percentile of annual monitoring data for compliance based on the concentration in suspended particles (German and Dutch) or the concentration in bedded sediments (Dutch only); in these cases, the duration and frequency of exceedances are dependent on the sampling design (Lepper 2002). The EU is debating whether to base compliance of guidelines on either the arithmetic mean or the 90^{th} percentile of the levels monitored in suspended particulate matter on an annual basis (Lepper 2002). Using the suspended particulate matter to check for compliance avoids the issue of sampling accumulated contaminants, versus newly deposited contaminants, although resuspension of bedded sediments could still confound monitoring if sampling sites have high resuspension fluxes.

In summary, the magnitude of SQGs is clearly addressed by existing methodologies, while the duration and frequency components are only indirectly addressed. The duration component may be addressed by the use of chronic data to derive SQGs that are likely to be protective of long-term exposures, which are more representative of contaminated sediment

exposure scenarios. Frequency of exceedance has not been addressed by any of the existing methodologies, and determination of an appropriate frequency for compliance testing of bedded sediments could be problematic because sediment contaminants can accumulate over time, unlike water column contaminants. One way to address the problem of accumulation would be to test suspended particles in compliance monitoring, so that only new inputs would be monitored, rather than accumulated residues, although this ignores the possibility that particle properties change when they become bedded. Another possibility is to use the freely dissolved concentration in porewater for compliance and setting criteria because the bioavailable or bioaccessible fractions would be expected to degrade or dissipate more quickly than the bound fraction.

6.1.2 Multipathway exposure

Benthic organisms may be exposed to sediment contaminants via: 1) absorption of freely dissolved chemical from sediment porewater (interstitial water between sediment particles) or overlying water through gills or onto their body surface, 2) ingestion of contaminated sediment or food particles, and 3) through direct contact with contaminated sediments. The predominant uptake route will be both species- and compound-dependent, and there are several studies that demonstrate different exposure routes dominating uptake in different conditions (Boese et al. 1990, Ingersoll et al. 2000, Lohmann et al. 2004, Lu et al. 2004, Savage et al. 2002, Selck et al. 2003). All of the possible exposure routes should be considered in SQC derivation, because if an exposure route is overlooked the resulting criterion may be underprotective.

The EqP approach only considers exposure to freely dissolved residues in sediment porewater because toxicity is predicted using water-only exposures in which organisms were fed uncontaminated food. There is evidence that accumulation of hydrophobic organic compounds (HOCs) varied little across organisms with various feeding habits, indicating that ingestion was not the main route of exposure (Tracey and Hansen 1996). In order to include the ingestion exposure route, the EU EqP method is modified for chemicals that have a log $K_{ow} > 5$ by applying an additional assessment factor of 10 to the predicted no effects concentration for sediment (ECB 2003). The Dutch method also references the modified EU EqP method to account for sediment ingestion (RIVM 2001). SSTTs offer a more realistic view of possible exposure routes because benthic organisms are tested in direct contact with contaminated sediments. If the tested organism ingests sediment particles as a feeding behavior, this would also be accounted for by performing a SSTT. Yet ingestion of contaminated food is overlooked in standard test methods, so the ingestion exposure route may be underestimated compared to what occurs in the field where algae, bacteria, detritus and other food sources could also take up contaminants.

6.1.3 Bioavailability

In an aquatic environment, nonionic organic compounds will either be freely dissolved in the water column or porewater or sorbed to particulates or colloidal matter, generally known as dissolved organic matter (DOM). Sediment-bound contaminants, the subject of this review, are those that are sorbed to bedded or suspended particles or DOM. It is widely accepted that organisms are generally not exposed to chemicals in the bound state, but rather are primarily

exposed to the fraction that has desorbed and is in the freely dissolved state (You et al. 2011). The exceptions to this are that organisms can be exposed to HOCs that are sorbed to food particles (e.g., algae, bacteria, detritus) or sediments that are ingested, since contaminants could desorb in the digestive tract of the organism (Mayer et al. 2001, Mehler et al. 2011), or can be exposed through direct contact with the contaminated sediment (Savage et al. 2002). For the purpose of this review, the bioavailable fraction of a contaminant refers to the fraction of a chemical that is available for uptake by organisms, which encompasses all exposure routes (uptake of freely dissolved chemical, ingestion, and direct contact). Bioavailability is dependent on many factors, including sediment characteristics (e.g., particle size, source of organic matter), organism characteristics (e.g., behavior, feeding), chemical properties, contact time, environmental conditions (e.g., temperature, pH), and biological activity in the ecosystem (e.g., biotic transformation, cycling, and burial) (Diaz and Rosenberg 1996, You et al. 2011). The questions we seek to answer in this section are: what fraction of contaminants are organisms exposed to, and what is the best technique to accurately predict the bioavailable concentration for criteria derivation and compliance?

6.1.3.1 Bioavailable fraction

The bioavailable fraction can be defined at the fraction of a compound that is available to an organism for uptake, which primarily consists of the fraction in the freely dissolved state (You et al. 2011). Many studies have shown that there are good correlations between freely dissolved concentrations and organism uptake or toxicity (Amweg et al. 2005, 2006, Trimble et al. 2008, Weston 2004, 2005, 2008). In a study by Xu et al. (2007), 10-d acute sediment toxicity tests were performed with *Chironomus tentans,* and the concentrations of three pyrethroids were expressed on the basis of five different phases: whole sediment, OC in sediment, whole porewater, dissolved organic carbon (DOC) in porewater, and freely dissolved in porewater. When the LC_{50}s were expressed as the concentration freely dissolved in porewater the variation across sediments was significantly reduced, and appeared to be matrix-independent, in contrast to the LC_{50}s expressed on a whole sediment or whole porewater basis, which varied widely. Normalizing to the OC or DOC content did reduce variation in the LC_{50}s, but some variation remained, unlike when expressed as the freely dissolved concentration. This study illustrates that organisms are exposed primarily to the freely dissolved fraction of a compound.

Other researchers have also defined the “bioaccessible” fraction, which is the fraction of chemical that has the potential to become available (by desorbing from sediment or DOM to become freely dissolved), although it is currently in the bound state (Semple et al. 2004, You et al. 2011). This suggests that just measuring the concentration that is currently freely dissolved may underestimate the total risk from the contaminant because as the freely dissolved fraction is taken up by organisms or transported downstream, additional contaminant would desorb.

6.1.3.2 Prediction or measurement of the bioavailable fraction

It has been demonstrated that nonionic compounds primarily sorb to organic matter (OM) contained in sediments or DOM, and the abundance of OM is usually expressed as the OC content of a sorbent (Schwarzenbach et al. 2003). Because OM primarily controls sorption and desorption of these compounds, solid-water partition coefficients are often normalized to the OC

content to reduce variability of the partition coefficients across sorbents with good success. Normalization of sediment concentrations of various hydrophobic organic compounds to the OC content has similarly been used to predict the bioavailable fraction in sediment, and many studies have demonstrated good correlations between biological effects and the OC-normalized sediment concentration (Amweg et al. 2005, 2006, Trimble et al. 2008, Weston 2004, 2005, 2008). There are limitations to this approach for several reasons, including: OC is quite variable, sorption can also be dependent on particle size, and highly hydrophobic compounds can also sorb to mineral domains. For example, toxicity of *Hyalella azteca* in sandy sediments was lower than predicted based on measured pyrethroid concentrations in the sediments, indicating that the mineral phase is also a factor controlling pyrethroid bioavailability (You et al. 2008a). Also, pyrethroid sorption to glassware has been extensively documented (Zhou et al. 1995, Oudou and Hansen 2002, Wheelock et al. 2005), implicating the mineral phase as an important pyrethroid sorbent.

Two alternative techniques have been developed to predict or measure the freely dissolved fraction: matrix-solid-phase microextraction (matrix-SPME) and Tenax® extraction. These two techniques were thoroughly described and evaluated in a recent review by You et al. (2011), the conclusions of which are summarized here. Both are matrix-independent, so the issues regarding varied sediment characteristics are eliminated. These techniques have resulted in concentrations that correlate to biological effects or uptake for various pesticides and other HOCs, although they measure different things (Parsons et al. 2007, Trimble et al. 2008, You et al. 2006, 2007, 2008a, b). Matrix-SPME is based on equilibrium partitioning and typically uses a fiber coated with polydimethylsiloxane as the sorbent to which freely dissolved HOCs in sediment porewater partition. The placement of the fiber in the matrix does not disturb equilibrium because the fiber only has the capacity to sorb less than 5% of the analytes. The residues on the fiber can be solvent extracted or thermally desorbed directly into an analytical instrument for detection and quantification of the freely dissolved chemical concentration. Tenax is a sorbent that can be used to measure desorption of hydrophobic organic compounds. Tenax powder is added to a sediment-water system, which alters the equilibrium of the system because it has a strong sorption affinity for HOCs. Any analytes that are freely dissolved sorb onto the Tenax, which disturbs equilibrium and leads to more analytes to desorb from the sediment. The Tenax is subsequently removed, solvent extracted, and analyzed at different time points to measure a desorption rate. The fraction of the compound that rapidly desorbs from the sediment is the main source of the freely dissolved compound in porewater, which can be correlated to organism uptake. An advantage of SPME is that it can be used *in situ* or in a laboratory setting. If the goal is to reach true equilibrium, SPME sampling can be very time- and labor-intensive because it can takes weeks to months for some systems to reach equilibrium, although one method reports equilibrium can be reached in <5 d for pyrethroids if a samples are agitated (Hunter et al. 2009). Laboratory protocols do not require equilibrium conditions during SPME porewater extractions, only that the SPME fiber is exposed the same length of time in each sample, typically 20-40 min (Bondarenko et al. 2007, Bondarenko and Gan 2009, Xu et al. 2007). The advantages of Tenax are that is has lower detection limits than SPME and can be processed more quickly than SPME.

6.1.3.3 Bioavailability in current methodologies

Some of the current methodologies address the issue of bioavailability. Di Toro et al. (2002) state in the USEPA EqP method that one of the principal reasons they chose the EqP approach instead of other approaches is because it addresses the issue of varying bioavailability of chemicals across sediments. The EqP approach is founded on two assumptions: 1) aquatic environments are approximately at equilibrium and 2) nonionic organic compounds sorb primarily to sediment OC or DOM. Given these assumptions, the freely dissolved concentration in porewater can be predicted from sediment concentrations using K_{oc}s, which are determined in laboratory experiments at equilibrium. Thus, the prediction of the bioavailable concentration in the EqP approach is highly dependent on which K_{oc} is used. Of the agencies worldwide that use the EqP approach (USEPA, The Netherlands, EU, Ontario, OECD), there is little to no mention of determination of site-specific sediment K_{oc}s for the chemical of concern. Due to the variability in measured K_{oc}s for the same compound, site-specific K_{oc}s would seem to be most representative of the site sediment, especially for the pyrethroid class of compounds, which are analytically challenging to handle as a result of their extreme insolubility in water. These EqP methodologies also allow for the estimation of K_{oc} from K_{ow} data, which may not be as accurate as experimentally determined K_{oc}s.

All EqP approaches account for bioavailability by using the freely dissolved fraction of a contaminant determined using a partitioning coefficient, and most agencies have chosen to focus on bedded sediments with the exception of the EU methodology, which is based on suspended sediment. Contaminants on suspended sediments are thought to reflect recent inputs of contaminants while bedded sediments are considered to be repositories for sediment-associated contaminants. The EU method rationalizes the use of suspended sediment concentrations for compliance because this fraction will settle and be the main food source for detritivorous benthic organisms (Crane 2003). Suspended solid concentrations in tributaries of the Sacramento River (1-330 mg/L) Basin and the San Joaquin River Basin (1-5280 mg/L) fluctuate over a wide concentration range, and may affect residue transport more than partitioning between the overlying water column and the sediment (TenBrook et al. 2009), so using suspended sediment concentrations may be a good way to measure the bioavailable fraction in this system. The USEPA EqP method also recognizes that bioavailability is dependent on the particle size distribution of sediments, and they recommend that large particles should be removed from sediment before chemistry analysis so as not to over-represent the contribution of large particles (and their small surface areas) to sequestration of residues (Di Toro et al. 2002).

The Dutch SSTT method (RIVM 2001) normalizes all SSTT data to a standardized sediment to reduce the variation in SSTT data across different sediments used in toxicity testing. For organic chemicals, SSTT data (e.g., NOEC, LC_x, or EC_x) are normalized to the organic matter content of the sediment by the following equation:

$$EC_{x(\text{std sed})} = EC_{x(\text{exp})} * H_{\text{std sed}} / H_{\text{exp}} \qquad (3)$$

where $EC_{x(\text{std sed})}$ is the NOEC or EC/LC_{50} normalized for standard sediment (mg/kg dry weight effect concentration), $EC_{x(\text{exp})}$ is the NOEC or EC/LC_{50} calculated for the experimental sediment (mg/kg dry weight effect concentration), $H_{\text{std sed}}$ is the OM content (%) of the standard sediment, and H_{exp} is the OM content (%) of the experimental sediment. The standard sediment is defined as having 10% OM content and 25% clay content (both on a dry weight basis). According to the

Dutch method, the OC content can be estimated from the OM content (dry weight basis) by dividing it by a factor of 1.7 (OC = OM/1.7), thus the OC content for standard sediment with 10% OM would have 5.88% OC.

In the Canadian SSTT method, it is stated that bioavailability is addressed by using the SSTT approach itself, since the organisms are only exposed to the bioavailable fraction. The Canadian method also addresses bioavailability by stating that when there is sufficient information available to define the influence of a particular factor on the toxicity of a substance, then the SQGs should be developed to reflect this relationship and those factors should be measured in samples. The example they give in the Canadian method is the relationship between OC content and nonpolar organic substances, indicating that the concentration of these substances should be normalized to the sediment OC content in SSTT data and compliance testing.

A bioavailability line of evidence has been proposed as part of the weight-of-evidence approach in California (Maruya et al. 2010). This line of evidence would relate porewater concentrations, passive sampling device measurements, or measurements of the rapidly desorbing fraction of a contaminant to known toxic effects concentrations for hydrophobic organic compounds.

6.2 Summary of methodologies

The basic methodologies are described in more detail here than in section 2, describing the differences between jurisdictions in the three main approaches: 1) mechanistic (EqP), 2) spiked-sediment toxicity tests, and 3) empirical. Brief descriptions are given of the many empirical approaches, but because most of these approaches do not yield single numerical values, but rather concentration ranges, most details are left out. Included in each section is a discussion on the advantages and disadvantages of each basic approach framed in terms of the ability of the approach to derive reliable numeric criteria for sediment-associated pesticides.

6.2.1 Equilibrium partitioning (mechanistic approach)

6.2.1.1 USEPA (Di Toro et al. 2002, USEPA 1993)

The EqP model as a mechanistic approach to SQG derivation was first developed by the USEPA specifically for nonionic organic compounds (Di Toro et al. 2002, USEPA 1993). The USEPA has taken this approach because it accounts for chemical bioavailability differences in varying sediment types and defines a biological effects based concentration in sediments using aqueous exposure data. The supporting rationale suggests that this approach is most likely to produce a protective concentration from biological effects in the field, and the method is robust enough to use in a regulatory setting. This method utilizes the data and knowledge base built in the derivation of WQC and can serve as a tool to protect uncontaminated sites and assist in the restoration of impaired sites.

The central assumption of the EqP model is that the chemical is in equilibrium between the sediment and porewater. This leads to the secondary assumption that the overall exposure is

equivalent, no matter the exposure route, if the system is at equilibrium conditions between phases. This assumption is supported by the finding that biota-sediment accumulation factors were similar within and among habitat groups for pesticides, PCBs and PAHs, indicating that the exposure route had little effect on accumulation (Tracey and Hansen 1996). The fate of highly hydrophobic organic compounds (log K_{ow} > 5.5) is primarily driven by sorption to the organic carbon in sediment. Therefore, the organic carbon-normalized sediment concentration is used as the bioavailable fraction in sediments with >0.2% OC content.

Another assumption is that an observed effects concentration can be predicted within the uncertainty of the model, and this uncertainty needs to be quantified and accompany the SQG (Di Toro et al. 2002). The aspect of uncertainty is the theory that aqueous exposure toxicity data can be used to predict sediment toxicity. This assumption has been evaluated and scrutinized over the previous decade in order to answer the question: Do benthic species have similar chemical exposure sensitivities as the aquatic species used to derive WQC? The USEPA has demonstrated good correlations between observed mortality and the mortality predicted using the EqP model for three sediments and seven chemicals, as well as good correlations between acute WQC of benthic and water column organisms, indicating that benthic species (epibenthic and infaunal) do have similar sensitivities to the species used in WQC derivation (Di Toro et al. 2002).

The FCV of the water quality criterion is used in the USEPA EqP method as the chemical concentration protective of benthic life. As a result, a SQG can be derived for a chemical if the FCV and the K_{oc} are known, and the uncertainty in the SQG can be quantified. The equations used to derive the USEPA EqP SQGs are as follows:

$$SQG = K_d * FCV \quad (4)$$

$$K_d = C_s / C_w = K_{oc} * f_{oc} \quad (5)$$

$$SQG = f_{oc} * K_{oc} * FCV \quad (6)$$

$$SQG_{oc} = K_{oc} * FCV \quad (7)$$

where K_d is the sediment-water partition coefficient, C_s is the equilibrium chemical concentration in the sediment, C_w is the freely dissolved porewater concentration, f_{oc} is the fraction of organic carbon in sediment, and K_{oc} is the organic carbon-porewater partition coefficient (Di Toro et al. 2002).

The USEPA takes a two-tiered approach for deriving SQGs, depending on the availability of data. The minimum data requirement to compute a Tier 1 SQG is 1) K_{ow} measured with current techniques, 2) the FCV using most current toxicity data, and 3) sediment toxicity testing to validate predictions of the EqP calculation. There are Tier 1 SQGs currently available for endrin, dieldrin and PAH mixtures from the USEPA. If there is inadequate data for a Tier 1 SQG, a Tier 2 SQG may be calculated if the following are available: 1) K_{ow} measured with current techniques, 2) the FCV or secondary chronic value is available, and 3) sediment toxicity testing is recommended but not required.

6.2.1.2 EU (ECB 2003)

The European Union (EU) methodology outlined in the Technical Guidance Document (ECB 2003), uses either an AF SSTT approach or the EqP approach to generate the predicted no effects concentration for sediment ($PNEC_{sed}$), depending on data availability. If there is inadequate toxicity data for sediment dwelling organisms, the EqP approach is taken, and the predicted no effects concentration for water ($PNEC_{water}$) and the suspended matter-water partitioning coefficient are used to calculate $PNEC_{sed}$. If only acute toxicity results are available (at least one) for benthic organisms, risk assessment is conducted on the basis of the result of the most sensitive species (using an AF of 1000) and the results of the EqP prediction. If there are long-term toxicity data for benthic organisms, the $PNEC_{sed}$ can be derived using AFs and this result should prevail over other methods. The equation for calculation of the $PNEC_{sed}$ (mg/kg) using the EqP approach is as follows:

$$\mathrm{PNEC_{sed}} = \frac{K_{susp-water}}{\mathrm{RHO_{susp}}} * \mathrm{PNEC_{water}} * 1000 \qquad (8)$$

where $PNEC_{water}$ (mg/L; similar to the FCV used by the USEPA), RHO_{susp} (1150 kg/m^3) is the bulk density of wet suspended matter, and $K_{susp\text{-}water}$ is the partition coefficient between suspended matter and water (m^3/m^3). Chemical uptake is only considered for the water phase when using the EqP methodology and as a result, uptake can be underestimated if the log K_{ow} is between 3 and 5. If the log K_{ow} is greater than 5, as is the case with many pyrethroids, the EU EqP method is modified. The calculated $PNEC_{sed}$ should be divided by 10 to account for sediment ingestion (ECB 2003). Also, if the measured (or predicted) chemical sediment concentration (PEC_{sed}) at a site is greater than the $PNEC_{sed}$ ($PEC_{sed}/PNEC_{sed} > 1$), spiked-sediment toxicity tests with benthic organisms need to be conducted for a refined sediment compartment assessment (ECB 2003). The EU EqP method uses suspended solids instead of bedded sediments to account for bioavailability in the SQG calculation. The rationale for using suspended particulate matter is that these are the materials that will eventually settle down and be the most important fraction for supporting benthic life. For compliance monitoring, the sediment quality standard is compared to the chemical concentration in the suspended matter.

6.2.1.3 The Netherlands (RIVM 2001)

The Netherlands also uses EqP as the basis for setting SQGs. Whereas the USEPA uses the FCV to compute a SQG, the Dutch derive the environmental risk limit for sediment. The equation to calculate the environmental RL for the sediment compartment using EqP theory ($ERL(sed_{EP})$) on a dry weight basis is as follows:

$$\mathrm{ERL(sed)_{EP}} = \mathrm{ERL(water)} * K_d \text{ (or } K_{oc}) \qquad (9)$$

where ERL(water) is the environmental risk limit for aquatic species and K_d is the sediment-water partition coefficient (L/kg). The Dutch method states that if there is experimental sediment toxicity data available, it is used to evaluate whether the derived $ERL(sed)_{EP}$ is protective enough.

6.2.1.4 OECD (1995)

The OECD guidance document of aquatic effects assessment recommends the use of the EqP, the interstitial water quality (IWQ) or the spiked-sediment toxicity test approach for the development of SQGs (OECD 1995). The OECD EqP method is described using the following equation, similar to the Dutch EqP method, to calculate the maximum tolerable concentration in sediment (MTC_{sed}):

$$MTC_{sed} = MTC_{water} * K_d \quad (10)$$

where MTC_{water} is the maximum tolerable concentration in water. Insufficient sediment toxicity data and lack of accepted sediment toxicity testing methods are provided as rationale for using the EqP approach instead of the IWQ or SSTT approaches. The IWQ approach is similar to EqP but *measured* interstitial water (porewater) concentrations are used instead of predicted concentrations.

6.2.1.5 Ontario (Persaud et al. 1993)

The Ontario guidelines suggest that the EqP approach be taken to determine the no effects concentration for nonpolar organic compounds. It is suggested that this criterion is the most stringent and is a level at which sediment contamination presents no threat to water quality and uses, benthic biota, wildlife or human health (Persaud et al. 1993). The equation provided in the Ontario EqP method is as follows:

$$SQG = K_{oc} * PWQO/G \quad (11)$$

where PWQO/G is the Provincial water quality objective/guideline. The Ontario SQG is then converted to bulk sediment concentration by assuming 1% total organic carbon sediment content.

6.2.1.6 France (Lepper 2002)

The French method to setting threshold effect levels for sediment and suspended matter under SEQ-Eau includes the EqP approach, but EqP is only used if toxicity data are lacking and the SSTT assessment factor approach cannot be used (Lepper 2002). The co-occurrence empirical approach (TEL/PEL derivation) was previously used in France (Wenning et al 2005). If the EqP approach is used, an additional safety factor of 10 is applied to set the level 1 threshold effect, which corresponds to very suitable aquatic ecosystems with negligible risk of adverse effects for all species. The EqP value with no additional safety factor applied is considered the level 2 threshold effect, which corresponds to a suitable aquatic ecosystem with possible risk of adverse chronic sublethal effects for the most sensitive species. For suspended matter, the threshold levels 1 and 2 are extrapolated from the sediment values by multiplying sediment threshold levels by a factor of 2 for organic substances.

6.2.1.7 UK (Rowlatt et al. 2002)

The UK does not have a formal policy for setting SQGs, but the EqP approach has been used since 1989 to set values for several metals, following the method described by Pavlou and Weston (1984). The sediment action level is calculated as follows:

$$C_{s/oc} = K_{oc} * C_{iw} \quad (12)$$

where $C_{s/oc}$ is the concentration in whole sediment or organic carbon, and C_{iw} is the concentration in interstitial water (porewater). The $C_{s/oc}$ is equal to the sediment action level when C_{iw} is equal to the aquatic action level.

6.2.1.8 Evaluation of EqP approach

Various reports and researchers have described the advantages and disadvantages of the equilibrium partitioning model approach, and they are compiled here, but listed with respect to the goal of developing single-value numeric SQC.

Advantages:
1) Results in a single numeric criterion.
2) There are WQC available for several sediment contaminants of concern, including chlorpyrifos, diazinon, and five pyrethroids.
3) Makes use of well-established toxicological database so new field data are not needed (Chapman 1989, Rowlatt et al. 2002).
4) Makes use of organic carbon (Rowlatt et al. 2002).
5) Makes use of chemical equilibria, which are often well-known (Rowlatt et al. 2002).
6) Efficient for determining which chemicals are likely contributors to toxicity (Rowlatt et al. 2002).
7) EqP is biologically based to the extent that existing WQC are biologically based, and therefore, provide more defensible guidelines than the "background approach" (Persaud et al. 1993).
8) SQC can be considered no-effect levels for the protection of the end uses the WQC were designed to achieve (Persaud et al. 1993).
9) Relies on an existing toxicological rationale (from WQC), thus a new toxicological evaluation is not required as long as WQC will protect benthic organisms (Chapman 1989).

Disadvantages:
1) Some pesticides have log $K_{ow} > 5.5$, which may confound bioavailability considerations.
2) Does not account for uptake through food chain and may underestimate uptake via ingestion (Persaud et al. 1993, Rowlatt et al. 2002).
3) Some partition coefficients are uncertain – including those for pyrethroids and other highly hydrophobic compounds (Chapman 1989, Persaud et al. 1993, Rowlatt et al. 2002).
4) Assumes that sediment infauna have the same sensitivity as other aquatic life (Persaud et al. 1993, Rowlatt et al. 2002).
5) Does not account for the presence of mixtures (Chapman 1989, Rowlatt et al. 2002), except for PAH mixtures (USEPA 2003c)
6) Relies on the existence of a WQC for the chemical of concern, and that WQC are accurate (Chapman 1989, Persaud et al. 1993, Rowlatt et al. 2002).

7) Assumption of equilibrium between the various phases may be unrealistic. Uptake kinetics can be multi-phasic and very slow for HOCs, and turbation of natural sediments can prevent complete equilibrium. Due to seasonal variations in particle composition, size and association with HOCs, equilibrium cannot be obtained between HOCs and growing particles (e.g., algae, plankton) (Crane et al. 1996, Naf et al. 1996, O'Connor 2000, Parsons et al. 2007, Rowlatt et al. 2002).
8) Does not use toxicity data derived from the sediment of interest (Chapman 1989).

The prime advantage of the EqP approach is that criteria can be derived without large co-occurrence field datasets or SSTT data, both of which are limiting factors for many pesticides of interest, including pyrethroids. WQC and partitioning coefficients are currently available for several pesticides of interest, so it would be possible to calculate sediment quality criteria with this approach.

The prime disadvantage is that some criteria derived using EqP have been shown to be unrealistic because assumptions about equilibrium and partitioning to OC may not be valid. The EqP approach was pursued by the USEPA beginning in the late 1970s to the early 1980s. In 1989, the Science Advisory Board (SAB) Sediment Quality Subcommittee reviewed the EqP approach and brought to light uncertainties in the technical basis of the approach. The Committee concluded that more information was required and uncertainties must be addressed before it could be used for SQC derivation. In response, the USEPA reduced the uncertainty in the EqP approach, although not all issues were resolved. The following is a summary of the Committee's 1992 evaluation (USEPA 1992) of the progress made toward minimizing uncertainty associated with EqP predictions and the outstanding issues remaining to be resolved.

1. Limited field data prevents assessment of uncertainty of EqP derived values to natural environment.
2. Assumptions of EqP may not be valid at site in question (listed below); these factors limit reliability and applicability of the approach but do not negate it.
    a. OC is not the only factor controlling bioavailability
    b. Different sensitivities of water column and benthic organisms
    c. Sediment and porewater may not be in equilibrium
    d. K_{ow} may not always be a good predictor of K_{oc}
    e. Kinetic limitations of chemicals partitioning to and from sediments
    f. Short-term bioassays may underestimate long-term effects
3. Supporting experiments showing OC-normalized sediment concentrations correlate with bioavailability are limited. Field exposure and method validation needs to be pursued further. Data for acenaphthene, dieldrin, endrin, fluoranthene and phenanthrene used to evaluate the method initially by the Committee supports the hypothesis, but they are limited to short-term studies representing only a few species. Bioaccumulation, biomarker and/or population/community level studies should be conducted in the field as well as toxicity testing.
4. EqP does not account for ingestion by uptake of food (although this point has been disputed by the authors of the EqP approach).
5. Uncertainty of EqP has improved through more accurate determinations of K_{ow}, but uncertainties should still be investigated and quantified.

Criticisms of the EqP approach in the literature tend to reflect the SAB review. Iannuzzi et al. (1995) argued in an editorial letter that there is consensus among scientists that a biological effects based approach to sediment quality criteria should be used. The authors reflect concern for the use of EqP as a regulatory tool because EqP is not directly based on biological effects, does not account for mixtures, or elucidate cause of toxicity. Instead, they suggest that a co-occurrence approach should be used since it is related to biological effects. A rebuttal to this letter (Ankley et al. 1996) highlights frequent misconceptions with the intended use of EqP for regulatory purposes, as well as misunderstandings with the technical assumptions of the EqP model. The rebuttal states that the EqP model is based on valid science, the limitations are known, and uncertainty in the model can be quantified, but they agree that it should not be used as a stand-alone tool for evaluating sediment quality on a pass/fail basis. The USEPA's use of porewater concentrations from laboratory spiked-sediment toxicity tests to validate the EqP model has also been criticized because these laboratory tests may not be reflective of the environment (O'Connor 2000). Environmental porewater concentration data is limited and spiked chemicals may be more available than more aged residues found in the field.

6.2.2 Spiked-sediment toxicity test approaches

Methodologies can use SSTT data in several ways. There are two ways to calculate SQC using SSTT data: 1) by applying an assessment factor to the lowest datum, or 2) by plotting all of the available data and fitting a statistical distribution to them, and then choosing a percentile to use as the SQC.

6.2.2.1 Assessment Factor (AF)

The EU only uses the EqP approach as a screening tool. If there is long-term toxicity data for benthic organisms, the $PNEC_{sed}$ can be derived using AFs and this is the required approach if the data requirements are met. If only acute whole sediment toxicity results are available (at least one) for benthic organisms, EU sediment risk assessment is conducted on the basis of the most sensitive species using an AF of 1000, and the results of the EqP prediction. Exposure routes during toxicity testing should be characterized and long-term exposure data are preferred to acute studies. Long-term tests with sublethal endpoints (reproduction, growth, emergence, sediment burrowing activity and avoidance) are regarded as the most relevant exposure scenarios of sediment contaminants (ECB 2003). Thus, the $PNEC_{sed}$ calculated from long-term tests are derived from the lowest NOEC/EC_{10} and application of an AF of 100 if one long-term test (NOEC or EC_{10}) is available, an AF of 50 if two long-term tests are available, and an AF of 10 if three long-term tests with species representing different living and feeding conditions are available.

The Netherlands method requires that an AF approach is used if there are not enough data available to fulfill the minimum dataset for use of a statistical distribution, for what they call "Preliminary effect assessment." The size of the AF depends on the number and kind of data available (acute vs. chronic, taxa), and the AFs range from 1 to 1000. The Netherlands follows the AF guidance in the EU method (ECB 2003), unless the EU minimum dataset is not available, in which case they use modified AFs from the USEPA Great Lakes Program (USEPA 2003d).

The Canadian method includes a SSTT approach, to be used in combination with an empirical approach (CCME 1999). If the minimum dataset requirements are met, an assessment (safety) factor is applied to account for uncertainty in the dataset due to inter- and intra-species variation, different endpoint measured, factors that control bioavailability, and laboratory to field extrapolations (including synergistic and antagonistic effects). Individual studies were examined to quantify the margin of safety due to various sources of uncertainty in the dataset. An assessment factor of 20 is recommended for use when deriving a sediment guideline from the median lethal concentration of acute studies. An assessment factor of 5 is recommended when developing a guideline based on chronic studies. It should be noted that the assessment factors were derived based on a limited dataset for only three compounds – zinc, cadmium, and fluoranthene.

In the French methodology, AFs ranging from 1 to 1000 are applied to single-species toxicity data. The size of the AF depends on what type of data are available, a lower AF is applied to chronic than to acute data (Lepper 2002).

6.2.2.1.1 Derivation and justification of AFs

Before the appropriate AF is chosen according to the EU AF method (ECB 2003), careful evaluation of studies is recommended to determine which exposure routes have been considered, and whether any possible exposure routes were underestimated due to study design (e.g., feeding organisms with unspiked food could reduce exposure via sediment ingestion). Additional AFs could be considered if important exposure routes were overlooked in a toxicity study. The magnitude of the AFs used by the EU decrease as uncertainty in the data decreases. AFs are designed to account for uncertainties in extrapolating from laboratory data to benthic ecosystems, including intra- and inter-laboratory variation in toxicity data, intra- and inter-species biological variance, short-term to long-term toxicity extrapolation, and laboratory data to field impact extrapolation (Lepper 2002). A factor of 1 to 10 is added for each layer of extrapolation, so that AFs range from 10-1000 depending on what type of and how many data are available. The Dutch AF method follows the EU method exactly. In the EU method, the $PNEC_{sed}$ is derived from the lowest available $NOEC/EC_{10}$ obtained in long-term tests by application of the following AFs:

Available test results	AF
At least 1 short-term test with a sediment-dwelling organism	1000
1 long-term test ($NOEC/EC_{10}$)	100
2 long-term tests ($NOEC/EC_{10}$) with species representing different living and feeding conditions	50
3 long-term tests ($NOEC/EC_{10}$) with species representing different living and feeding conditions	10

The Canadian AF methodology (CCME 1999) uses AFs of varying magnitude to compensate for uncertainties such as intra- and inter-species variations, use of various endpoints, acute to chronic extrapolation, bioavailability issues, and extrapolation from laboratory to field conditions (including mixtures). In this method, AFs are determined on a case-by-case basis, depending on what information is available. While most of the uncertainty factors are not

inherently addressed in the SSTT approach, inter-species variation is accounted for by inclusion of relatively sensitive species in the minimum data requirements, and bioavailability issues are inherently accounted for in SSTTs. To account for each of the remaining uncertainty factors, SSTT results from various lifestages of an organism, various endpoints, and acute and chronic tests are compared. Using these data, ratios are calculated to convert from the less sensitive test result to the more sensitive test result. All ratios are then combined to yield final AFs for extrapolating from acute (LC_{50}) data or from chronic (NOEC) data. According to the Canadian methodology, different AFs could be calculated for individual chemicals or classes of chemicals, or generic AFs could be calculated for all relevant chemicals.

In general, the summary of the French method by Lepper (2002) does not describe how the magnitude of the AFs were determined or justified, except that the factors are lower when chronic data are used than when acute data are used. To calculate a threshold level 1 for sediments, the lowest reliable NOEC or EC_{10} is divided by an AF of 10 or the lowest reliable LC/EC_{50} is divided by an AF of 1000. Threshold 1 corresponds to the concentration at which there is negligible risk of adverse effects for all species, while threshold level 2 corresponds to the concentration at which there is possible risk of adverse chronic sublethal effects for the most sensitive species of the ecosystem. Threshold level 2 is equal to the lowest reliable chronic NOEC (AF of 1) or the lowest reliable LC/EC_{50} divided by an AF of 100. To derive threshold levels for suspended sediment, the threshold levels for sediment are multiplied by a factor of 2.

6.2.2.1.2 Aggregation of taxa

As stated by TenBrook and Tjeerdema (2006), a criterion is calculated based on the most sensitive species using the AF methodology. Factors such as taxon or toxicant mode of action are not taken into account, and all of the AF methods considered include both plant and animal data together in the dataset. None of the AF methodologies mention pooling taxa. However, separate freshwater and saltwater criteria are typically derived.

6.2.2.2 Species sensitivity distribution (SSD)

Statistical extrapolation methods are used to derive water quality criteria in many jurisdictions and it has been suggested that this technique could also be applicable to derivation of sediment quality criteria. In this approach, the cumulative probabilities of the toxicity values for a given chemical are plotted and then a statistical distribution is fit to the data. Next, the criterion is derived by choosing a percentile cutoff of the statistical distribution, typically the 5^{th} percentile. A thorough review of the statistical distribution approach to WQC derivation can be found in TenBrook et al. (2009), and it will not be repeated here for the sake of brevity. In the majority of the methodologies reviewed, concerns about low data availability have limited discussion on using statistical distributions for the derivation of SQC. Only The Netherlands and the EU methodologies specifically discuss the use of SSDs for derivation of SQGs, and they are described in more detail below. Both of these methods use only chronic data in the distributions and use the 5^{th} percentile from the distribution as the cutoff.

6.2.2.2.1 The Netherlands (RIVM 2001, 2004)

The Dutch method states that an environmental risk limit, such as the maximum permissible concentration, can be derived for sediment, air, water, soil or groundwater using a statistical approach if chronic data from four or more species of at least four different taxonomic groups are available for a particular environmental compartment (RIVM 2001). The Dutch method uses a log-normal distribution, and calculates the hazard concentrations (HC_p) for the 5th and 50th percentiles (HC_5 and HC_{50}) as follows:

$$\log HC_p = \overline{x} - k * s \quad (13)$$

where HC_p is the hazardous concentration for p% of species, $\overline{x}$ is the mean of log-transformed NOEC data, *k* is the extrapolation constant depending on percentile level of certainty and sample size (Table 1 in Aldenberg and Jaworska 2000), and *s* is the standard deviation of log-transformed data. HC_ps and their associated 90% confidence intervals are easily obtained using a computer program designed for making these calculations (RIVM 2004). The HC_5 is used as the MPC, which in turn can be used to derive environmental quality standards. The negligible concentration is calculated as the MPC/100, which is the target value. The HC_{50}s are used as intervention values, indicating that the ecosystem is seriously threatened because 50% of species are adversely affected.

As of 2001, Dutch Environmental Quality Standards had been derived for 147 organic substances and pesticides in sediment, but the data requirement of four chronic data (4 NOECs) were not available for any of these substances, so the standards for these compounds were derived using equilibrium partitioning (Sijm et al. 2001).

6.2.2.2.2 EU (ECB 2003)

The EU also suggests a statistical distribution approach is preferable to the assessment factor approach and if there is adequate sediment toxicity data to fit to a distribution, this method should be employed (ECB 2003). The EU uses the Dutch SSD procedure, but they do not specify that the log-normal distribution should be used, rather they recommend that whichever distribution best fits the data should be used. Two statistical tests are proscribed to check goodness of fit for this determination: the Anderson-Darling test or the Kolmogorov-Smirnov test. Once a distribution is chosen, the PNEC may be calculated as follows:

$$PNEC = 5\%\ SSD\ (50\%c.i.)/AF \quad (14)$$

where PNEC is the predicted no effect concentration, 5% SSD is the 5th percentile from species sensitivity distribution, 50% c.i. is the 50% confidence interval, and AF is the assessment factor, which ranges from 1-5.

The AF is determined based on professional judgment, designed to reflect additional uncertainties in the calculation, although they suggest starting with an AF of 5, and only decreasing the AF after considering the following points:

1. the overall quality of the database and endpoints covered, e.g. if all data are from chronic studies covering *all* sensitive lifestages;

2. the diversity and representativity of the taxonomic groups covered by the database, and the extent to which differences in the life forms, feeding strategies and trophic levels of the organisms are represented;
3. knowledge on presumed mode of action of the chemical (covering also long-term exposure);
4. statistical uncertainties around the 5^{th} percentile estimate, e.g., reflected in the goodness of fit or the size of confidence interval around the 5^{th} percentile, and consideration of different levels of confidence;
5. comparisons between field and mesocosm studies and the 5^{th} percentile to evaluate the laboratory to field extrapolation.

The EU also recommends that any NOECs below the 5^{th} percentile should be discussed in the final report, and checking to see if they are all from one trophic group that may have increased sensitivity compared to other taxa. The PNEC should also be calculated according to the assessment factor approach and compared to the PNEC derived using a SSD, and based on all considerations, one of these PNECs should be chosen as the final value.

6.2.2.3 Evaluation of SSTT approaches

Various reports and researchers have described the advantages and disadvantages of the SSTT approach, and they are compiled here, but listed with respect to the goal of developing single-value numeric SQC.

Advantages:
1) Similar to WQC: technically acceptable and legally defensible (Chapman 1989, Ingersoll and MacDonald 2002, Rowlatt et al. 2002).
2) Does not require prior knowledge of mechanisms of uptake (Chapman 1989, Rowlatt et al. 2002).
3) Direct cause-effect relationship can be determined (Ingersoll and MacDonald 2002, Persaud et al. 1993).
4) Laboratory SSTTs can be performed with any chemical and do not require *a priori* assumptions about the specific mechanism of interaction (exposure route) (Chapman 1989, Ingersoll and MacDonald 2002).
5) SSTTs measure bioavailable fraction of contaminant (Ingersoll and MacDonald 2002).

Disadvantages:
1) Few data are available, particularly chronic data, and available data are only for a few standard test species (Chapman 1989, Rowlatt et al. 2002).
2) Little information currently available relating acute and chronic effects to use when chronic data are not available (Chapman 1989, Rowlatt et al. 2002).
3) It is not practical to test all possible mixtures that could occur in the environment (Chapman 1989, Rowlatt et al. 2002).
4) No basis at present for extrapolating to no-effect-concentrations in sedimentary communities (Rowlatt et al. 2002).
5) Sediment-spiking techniques are not standardized, and differences in methods can strongly influence the results (Chapman 1989, Ingersoll and MacDonald 2002, Persaud et al. 1993,

Rowlatt et al. 2002).
6) Laboratory tests may not be representative of field-contaminated sediment, where conditions may vary considerably from those in the lab, particularly bioavailability effects (Ingersoll and MacDonald 2002, Persaud et al. 1993, Rowlatt et al. 2002).

The primary advantages of the SSTT approach are that it is technically acceptable, shows a direct cause-effect relationship, and it addresses the issue of bioavailability, while the primary disadvantages of using the SSTT approach are the dearth of available data and that the data may not be comparable because different spiking methods can strongly influence the results.

6.2.3 Empirical approaches

Many countries that use biological effects based approaches are utilizing empirically based, rather than mechanistically based methodologies. Empirically based methods utilize large datasets containing matching sediment chemistry and toxicity data from field collected sediments. The SQGs generated under the empirical approaches do not allow causality to be attributed to any one chemical, and in most cases were not intended for regulatory use, but rather for risk assessment and ranking of sites for further study. Mixture effects are indirectly taken into account in the toxicity tests, but again, the mixture or one chemical cannot be directly linked to the biological effects seen in the field samples. With that said, there are large amounts of sediment chemistry and toxicity data generated through both national and state-wide programs. In North America, most of the jurisdictions use biological effects data, while in many European countries, a reference site approach has historically been used to evaluate sediment contamination (Wenning et al. 2005). For example, the Belgian province of Flanders bases sediment quality guidelines on the geometric mean of the average reference site concentrations, and they also use biological assessments and toxicity testing in their evaluations. Italy also uses a similar reference site approach (Wenning et al 2005).

6.2.3.1 Effects range approach (US – NOAA NSTP, Canada, Australia/New Zealand)

The effects range approach (Long and Morgan 1990, Long and MacDonald 1992, Long et al. 1995) was developed as a means to interpret monitoring data collected through the National Oceanic and Atmospheric Association National Status and Trends Program (NOAA NSTP; Wenning et al. 2005). A large database (sorted in ascending order of dry weight chemical concentrations) of matching sediment chemistry and biological effects data was used to calculate the lower 10^{th} percentile of the data, defined as the effects range-low (ERL) concentration level, and the 50^{th} percentile, defined as the effects range-median (ERM) concentration level. The dataset was modified over time (Long et al. 1995), excluding freshwater data and adding saltwater data from coastal and estuarine waters of the Atlantic, Pacific and Gulf of Mexico. Existing sediment guideline values and results from spiked-sediment toxicity tests were also included in the database (Wenning et al. 2005). The SQGs calculated by NOAA were derived on a sediment dry weight basis, not normalized to organic carbon, and bioavailability is not directly accounted for. This methodology has been taken as part of a weight-of-evidence approach to SQGs and has a different intention than the derivation of numeric criteria for individual contaminants based on biological effects.

Other regulatory agencies worldwide use the NSTP database as a major data source for the derivation of empirically based SQGs. The Australia/New Zealand (ANZECC and ARMCANZ 2000) and Canadian (CCME 1999) methodologies rely on North American databases for ecotoxicity information to derive ERL/ERM values, and then use local data to refine the values when more information is available. The Canadian guidelines calculate separate freshwater and saltwater SQGs, and specify that sediment concentrations are reported on a dry weight basis. If data are limited, an interim sediment quality guideline (ISQG) can be calculated but data gaps must be described to suggest limitations of ISQG. The Canadian guidelines also suggest incorporating SSTTs in the future to validate effect ranges calculated using the NSTP database.

These SQGs have not been adopted by NOAA, USEPA, or the Army Corps of Engineers because they were not intended to be used as standards and have not been proven reliable for predicting sediment toxicity (O'Connor 2000). Some researchers have used the ERL as the chemical concentration above which there is observed toxicity and this usage is not supported by Thomas O'Connor of NOAA (O'Connor 2004). He states that the ERL is "a low point on a continuum of bulk chemical concentrations in sediment that roughly relate to sediment toxicity," and criticized other researchers for using the exceedance of an ERL to exaggerate sediment contamination (O'Connor 2004). Further limitations of the effects range approach include inability to determine cause of toxicity and lack of consideration for chemical and environmental factors that may contribute to or be responsible for the observed biological effects.

6.2.3.2 Effects level approach (Florida)

Florida also uses the NOAA NSTP method, called the effects level approach, as part of a weight-of-evidence approach to SQGs and localized the approach by including data collected from Florida coastal waters and the Gulf of Mexico to develop a relevant biological effects database for sediment (BEDS). The effects level approach is similar to the effects range approach except the effect level approach separates the biological effect and no effect data within the database. Reference site data or a site in which the average chemical concentration was less than twice the level at sites considered nontoxic, were considered "no effect" data (MacDonald et al. 1996). The threshold effects levels (TELs) were calculated as the geometric mean of the lower 15^{th} percentile concentration of the effects dataset and the 50^{th} percentile of the of the no-effect dataset. The probable effects levels (PELs) were calculated as the geometric mean of the 50^{th} percentile concentration of the effects dataset and the 85^{th} percentile of the no-effects dataset. The BEDS developed for the effects range approach was widely expanded to include separate marine (MacDonald et al. 1996) and freshwater (Ingersoll et al. 1996, Smith et al. 1996) sediments for determination of the TEL/PEL. Like the effects range approach, the effects level approach also does not account for bioavailability of contaminants, which would be necessary for a valid technical approach.

6.2.3.3 Apparent effects thresholds approach (Washington/Oregon/Puget Sound)

A database of matching sediment chemistry and biological effects from the Puget Sound, WA was used to derive the apparent effects threshold (AET) approach to generate SQGs. This is also very similar to the NSTP approach, but they use a different database called the freshwater

sediment quality database (FEDSQUAL), which was developed from 33 freshwater studies and 245 stations in Washington and Oregon (Willamette River) using sediment chemistry and bioassay data from Microtox, *Hyalella azteca, Chironomus tentans, Daphnia magna, Ceriodaphnia dubia* and *Hexagenia limbata*. The database is a subset of the marine sediment quality database (SEDQUAL), and different saltwater (Barrick et al. 1988) and freshwater (Cubbage et al. 1997) databases have been developed for AET calculations. The AET is defined as the chemical concentration above which biological effects are always expected to occur (Barrick et al. 1988). The probable apparent effects threshold (PAET) was defined as the 95th percentile of the sample concentration with no significant biological effects above the lowest concentration associated with an effect (Cubbage et al. 1997). The AET approach includes benthic community endpoints in the database, whereas the effects range and levels approaches predominantly use lethality as endpoints (Wenning et al. 2005). It has been suggested that freshwater sediment quality values derived based on Microtox PAETs for organic chemicals were best able to predict biological effects (OECD 1995).

The USEPA (1989) Science Advisory Board (SAB) reviewed the AET approach and suggested it should not be used to develop "broadly applicable sediment quality criteria" due to the site specific nature of the approach, lack of validation, inability to establish causality and inability to account for differences in bioavailability across sediment types. The SAB did acknowledge that the AET approach was a technically valid approach for estimation of sediment quality at a particular location, if properly validated (USEPA 1989).

6.2.3.4 Screening level concentration approach (USEPA, Ontario)

The screening level approach (Neff at al. 1986, 1987) uses matching sediment chemistry and the presence or absence of benthic species in field samples to derive a species screening level concentration. A frequency distribution of a particular species present at a sampling site is plotted against the organic carbon-normalized sediment concentration of a chemical to derive the SSLC (90th percentile concentration of the frequency distribution). A frequency distribution of the SSLCs are plotted to determine the screening level concentration, which is the 95th percentile of the SSLC distribution. The SLC is an estimate of the highest concentration that can be tolerated by 95% of the benthic species (Neff et al. 1986).

6.2.3.5 Logistic regression model approach (California)

Empirically based narrative sediment objectives are currently used in California for the protection of benthic communities in bays and estuaries. Sediment toxicity, benthic community condition, and sediment chemistry are used to support a multiple lines of evidence approach to meet the narrative sediment quality objectives. These three metrics are combined in order to assess a sampling station as unimpacted, likely unimpacted, possibly impacted, likely impacted, clearly impacted or inconclusive. Stations having a status of unimpacted and likely unimpacted are considered to meet the protection objectives. An individual line of evidence is not considered reliable when used alone. Sediment toxicity and benthic community condition categories are used in a decision matrix to determine the severity of biological effects at a sampling site. Sediment toxicity and chemistry categories are used in a decision matrix to determine the potential for chemically-mediated effects. A station level assessment is then made using a

decision matrix based on the potential for chemically-mediated effects and the severity of effects using the Chemical Score Index and the California Logistic Regression Model. Limitations of this approach include over- or underestimation of risk to benthic communities, causality cannot be indicated for specific chemicals, and the lack of consideration of particle size, physical disturbance, or organic enrichment as possible contributors to biological effects (SWRCB 2011).

A logistic regression model can be used to relate a chemical's concentration in sediment to the probability of toxic effects (Field et al. 1999, 2002). Matching sediment chemistry and toxicity data from marine amphipods were initially combined in a database for derivation of the logistic regression equation used for SQG formulation (Field et al. 1999, 2002, Wenning et al. 2005). In California, many empirical approaches have been evaluated for SQG derivation. A study by Ritter et al. (2007) compared the performance of an effects range median, a logistic regression model, a sediment quality guideline quotient 1 (SQGQ1) and a consensus approach to deriving SQGs. A new chemical score index approach was also compared to these existing empirical methods. Results indicated that all methods performed similarly (Ritter et al. 2007).

6.2.3.6 Probable effects approach (Great Lakes)

Sediment quality in the Great Lakes region has been assessed by deriving a probable effects concentration from 92 reports and over 1500 samples with matching sediment chemistry and toxicity data from field collected sediments throughout North America (Ingersoll et al. 2001). The database includes 10-14 d and 28-42 d sediment toxicity data for the amphipod, *Hyalella azteca*, and 10-14 d toxicity tests with the midge, *Chironomus riparius* or *C. tentans*. The PECs derived were shown to predict sediment toxicity on both a regional and national basis (Ingersoll et al. 2001). Mean PEC quotients were calculated to provide an overall measure of chemical contamination and to support an evaluation of the combined effects of multiple contaminants in sediments. This is considered a consensus based approach to sediment quality guidelines.

6.2.3.7 Evaluation of empirical approaches

The co-occurrence empirical approaches will generally not work for the purpose of deriving single-value numeric criteria for pesticides for several reasons. The primary reason is that, while co-occurrence data incorporates environmental mixtures, it does not provide a way to tease out which contaminant is causing the effect, or what degree of effect is caused by a single contaminant when multiple contaminants are present. Reliable individual numeric criteria cannot be obtained with the empirical approaches because of the possibility that some observed effects were not caused by the contaminant of concern (caused by other measured or unmeasured contaminants, or caused by other environmental factors). Bioavailability variation across sediment types is not accounted for in empirical approaches which limit their usefulness in deriving reliable and predictive guidelines.

Another reason an empirical approach will not meet our project goals is that all of these approaches require datasets of co-occurring biological and chemical data for the contaminants of interest, that are preferably site-specific; these types of datasets are not available for many of the pesticides of interest for the Sacramento-San Joaquin River watersheds. Some of the methods

also require or suggest performing additional field bioassays and benthic community surveys to include in the determination of SQGs, but these types of additional information will not likely be available for this project. While these approaches do have advantages in other situations of sediment contamination assessment, particularly if the databases already exist, the lack of co-occurrence data for pyrethroids and other pesticides of interest makes these approaches impractical. Even if the necessary database did exist, empirical approaches have not been shown to reliably predict sediment toxicity, and do not have a valid technical foundation because they overlook bioavailability.

A recent evaluation of the predictive abilities of the probable effects, effects range and logistic range model approaches in the Calcasieu Estuary, Louisiana demonstrated that 10-d sediment toxicity was underestimated using all three approaches (MacDonald et al. 2011). The probable effects approach was more reliable when predicting toxicity using the estuary dataset compared to the national dataset; therefore, the authors recommend using a site-specific approach for benthic organisms risk assessment (MacDonald et al. 2011). In another assessment of empirical approaches, Fuchsman and Barber (2000) reported that concentrations greatly exceeding those associated with toxicity from co-occurrence empirical approaches caused no adverse effects in SSTTs for two compounds, indicating that it is possible to attribute toxicity to the incorrect contaminant.

6.3 Other considerations

6.3.1 Mixtures

In the field, contaminant mixtures in sediments are ubiquitous, as many years of field sampling data have demonstrated (Gilliom 2007). Because of the high likelihood that organisms will encounter chemical mixtures in the field, it would be ideal to have a way to quantify the effects of mixtures so that they may be incorporated into criteria derivation and compliance. As noted above, a disadvantage of the EqP and SSTT approaches is that neither of them easily incorporates mixture toxicity into criteria calculation. In contrast, the empirical methods inherently take mixtures into account because samples are analyzed for multiple compounds and field sediment bioassays are performed, which would detect toxicity of all contaminants present, analyzed or not. Yet the toxicity cannot be directly related to any of the known or unknown chemicals using an empirical approach, which may make it difficult to move forward with a management plan.

When an organism is exposed to more than one chemical, the toxic effects can be additive, synergistic, or antagonistic. The current state of knowledge on pesticide interactions is not developed to be able to easily predict what the mixture effects will be when there are compounds in a mixture from different classes with different modes of action (Lydy et al. 2004). Prediction or modeling of additive mixture toxicity has been further developed than for non-additive interactions. The concentration addition model, in which chemicals have the same mode of action, but do not interact with each other (Plackett and Hewlett 1952), is widely used, including by the Central Valley Regional Water Quality Control Board and is mentioned in the California SQG methodology (Maruya et al. 2010). The concentration addition model is applied in the Water Quality Control Plan (Basin Plan) for the Sacramento River and San Joaquin River

Basins in cases where multiple chemicals with similar modes of action are present in a water body (CRWQCB-CVR 2006).

Toxic unit (TU) analysis has been used to predict additive toxicity for chemical mixtures. A toxic unit is calculated by dividing the measured concentration in sediment (C_{sed}) by the laboratory-determined LC_{50} of a particular organism as follows:

$$TU = C_{sed}/LC_{50} \quad (15)$$

Typically, both the sediment concentration and LC_{50} will be normalized to the organic carbon content from the respective sediments. Weston et al. (2004) used OC-normalized TUs for pyrethroids and organochlorines to determine which contaminants were present in sediments at high enough concentrations to cause the observed toxicity. TU analysis has also been used by CDPR to predict potential toxicity to the amphipod *Hyalella azteca* from pyrethroids detected in sediments (Starner and Kelley 2005). To calculate the number of TUs for each compound, the OC-normalized sediment concentration was divided by the OC-normalized LC_{50}. For the CDPR analysis, pyrethroids were assumed to be additive, so if multiple pyrethroids were present in a sample, the TUs were simply added together. If the sum of TUs > 1, then the concentration exceeds the LC_{50}, and the proportional toxicity would be expected. The toxic unit approach is a simple way to determine additive toxicity, and can be used in compliance determination by replacing the LC_{50} with the SQC of the respective compound (TenBrook et al. 2009). The use of TUs to determine mixture toxicity is discussed in relation to measurement of bioavailable fractions for the bioavailability line of evidence in the California SQG methodology (Maruya et al. 2010).

Non-additive mixture effects were addressed with regard to criteria compliance by TenBrook et al. (2009). They concluded that at this time, non-additive mixture effects can only be considered in criteria compliance when valid interaction coefficients (*K*) are available over a range of concentrations. When either a synergist or antagonist is present in a mixture with the pesticide of interest, an interaction coefficient (*K*) can be used to describe the joint mixture effect, as follows (Finney 1942):

$$K_x = EC_{50(0)} / EC_{50(x)} \quad (16)$$

where K_x is the interaction coefficient when the synergist/antagonist is at concentration *x*, $EC_{50(0)}$ is the EC_{50} determined for a chemical in absence of the synergist/antagonist, and $EC_{50(x)}$ is the EC_{50} determined for a chemical in the presence of a synergist/antagonist at concentration *x*.

Once a K_x is known, then the measured concentration of a pesticide can be multiplied by K_x to yield an adjusted concentration of the pesticide:

$$C_a = C_m * K_x \quad (17)$$

where C_a is the adjusted concentration of a pesticide, C_m is the measured concentration of the pesticide, and K_x is the interaction coefficient for a synergist/antagonist at concentration *x*.

To use *K*s in criteria compliance, K_xs must be known for a wide range of synergist/antagonist concentrations in order to determine a relationship between synergist/antagonist concentrations and mixture toxicity effects to use for predicting the adjusted pesticide concentration at any given synergist/antagonist concentration. These relationships would have to be established for multiple species, or for one particularly sensitive species to compare the adjusted concentrations to the criteria for compliance determination (TenBrook et al. 2009).

Of the mechanistic and SSTT approaches, the Dutch, OECD, EU, UK, French, and Canadian SSTT methodologies do not directly address mixtures or multiple stressors. The EU method states that the possibility of multiple contaminants in the field, and their possible additive, synergistic, or antagonistic effects, are one of the uncertainties that the assessment factors are designed to address (ECB 2003). The co-occurrence methodologies incorporate mixture toxicity in that the field samples are typically tested for a suite of compounds, so there is information about multiple contaminants in a sample, but there could also be unknowns that were not analyzed. These methodologies also conduct field sediment bioassays that expose organisms to whatever mixture might be present in the sample, so mixture toxicity is measured, even if all of the possible contaminants were not quantified.

The USEPA derived SQGs for a group of PAHs, and this ∑PAH approach incorporated EqP, quantitative structure activity relationships, toxic units, additivity and concentration-response models to address the mixture effect of PAH toxicity in sediments. Data used to generate this model was from both published and unpublished sources of spiked-sediment toxicity tests and field sediment bioassays (Swartz et al. 1995) that followed standard methods. A more general approach to PAH mixtures and narcotic chemical mixtures was developed by Di Toro and McGrath (2000) and has been officially issued by the USEPA as a method to derive equilibrium sediment benchmarks for PAH mixtures (USEPA 2003c).

In general, additive mixture effects can be incorporated in criteria compliance using the concentration addition model for either the EqP or SSTT approaches when it has been established that it is reasonable to assume additivity. When it is demonstrated or can be assumed that mixture effects will be additive, toxic unit analysis is a simple way to check for compliance, as long as there is reliable single compound LC_{50} data available for each compound in the mixture. For non-additive mixture effects, interaction coefficients can be used if ample data are available. More complex mixtures, involving both synergists and antagonists cannot be incorporated into compliance determination at this time, although some complex models do exist to predict effects in these situations (Rider and LeBlanc 2005, TenBrook et al. 2009). The SQG methodologies following empirical approaches inherently incorporate mixture toxicity, but are not necessarily able to attribute toxic effects to particular chemicals or mixtures.

6.3.2 Threatened and endangered species

Threatened and endangered species (TES) are particularly sensitive to stressors in the environment, thus it should be ensured that sediment quality criteria will be protective of these species. TES typically have a limited range, so in setting SQC for the Sacramento and San Joaquin River basins, only local species should be considered. There are various benthic

organisms currently on the California Department of Fish and Game list of TES, including 16 amphibians and eight crustaceans (CDFG 2011), which could potentially be affected by sediment contaminants. None of the SQG methodologies specifically address protection of TES, perhaps because it is assumed that if the most sensitive species in an ecosystem are protected, TES should also be protected. There are few SSTT data available in general, and toxicity data for TES are likely to be even scarcer, so being able to predict toxicity values for TES would be very valuable. Unfortunately, interspecies correlations based on surrogate species have not been developed for sediment toxicity, as they have for aqueous toxicity, and it is unlikely that quantitative structure activity relationships will be appropriate to use with pesticides of interest because they are limited for use with compounds with a narcotic mode of action. It may be difficult to incorporate specific procedures to ensure protection of benthic TES at this time because of the few toxicity data likely to be available, but any data that is available should be used to evaluate the derived criteria to assess protection of listed species.

6.3.3 Bioaccumulation/secondary poisoning

Chemicals that accumulate in the sediments often also have a propensity to accumulate in organisms; this accumulation may lead to secondary poisoning effects as contaminants magnify up the food chain. According to the OECD (1995), substances with a log $K_{ow} > 3$ and molecular weight < 1000 may bioaccumulate, and these are similar properties as described for chemicals likely to accumulate in sediments. Several of the SQG methodologies address bioaccumulation by calculating what the expected level of the compound would be in a predator that consumes aquatic organisms if the sediment concentration were equal to the SQG. When the EqP approach is used (USEPA, OECD, France, UK, The Netherlands, EU), it would be most appropriate to check for bioaccumulation using the water quality criteria, on which the SQG is based, to remove one layer of extrapolation. Guidance for extrapolating secondary poisoning effects from aqueous concentrations is covered by TenBrook and Tjeerdema (2006).

The Canadian methodology states that bioaccumulation in the context of SQGs is defined as “the process by which substances are accumulated by aquatic organisms from all routes of exposure.” To address bioaccumulation up the food chain, bioaccumulation factors (BAFs) and tissue residue guidelines for the protection of wildlife consumers of aquatic life can be used to calculate a SQG that would be protective of higher trophic levels. If this tissue-based SQG is lower than the ecotoxicity-based SQG, then the final SQG can be altered to be protective of all trophic levels (CCME 1999).

The Dutch methodology includes consideration of secondary poisoning of predators caused by uptake of substances by prey organisms (RIVM 2001). According to this method, biota-sediment accumulation factors (BSAFs) and chronic toxicity studies for birds and mammals are used to assess possible secondary poisoning effects. The individual bird or mammal NOECs are divided by a BSAF to obtain a concentration in sediment, and then the data are added to the ecotoxicity dataset and used for derivation of the environmental risk limit. A MPC is also calculated separately for predators and is compared to the MPC calculated for the sediment compartment.

The EU method (ECB 2003) states that a chemical should be assessed for bioaccumulation if it has the following characteristics: a log $K_{ow} \geq 3$, is highly adsorptive, belongs to a class of substances known to have a potential to accumulate in living organisms, there are indications from structural features, and there is no mitigating property such as hydrolysis (half-life less than 12 hr). For compounds with log $K_{ow} \geq 4.5$, direct uptake of freely dissolved chemicals from the water may not be the most important uptake route, and exposure routes such as ingestion of contaminated food or sediment may predominate. No direct guidance is given to predict the concentration in food of predators when this is the case, but extensive guidance is given on how to assess risk to predators from dietary uptake, with the assumption that the prey is living in a water body with the contaminant at the level of the predicted environmental concentration in water (PEC_{water}), as discussed by TenBrook and Tjeerdema (2006):

$$PEC_{oral,\ predator} = PEC_{water} * BCF_{fish} * BMF_{fish} \qquad (18)$$

where $PEC_{oral,\ predator}$ is the predicted environmental concentration a predator will receive in prey (food), BCF_{fish} is the bioconcentration factor for fish on a wet weight basis (concentration in fish/concentration in water), and BMF_{fish} is a biomagnification factor in fish (concentration in fish/concentration in food).

In order to assess the risk of secondary poisoning from sediment contaminants, the PEC_{water} can be replaced with the PEC_{sed} as follows:

$$PEC_{oral,\ predator} = PEC_{sed} * BCF_{fish} * BMF_{fish} \qquad (19)$$

By substituting NOECs for the PECs in the above equation, one can solve for $NOEC_{sediment}$ given a $NOEC_{oral,\ predator}$, BCF_{fish}, and BMF_{fish}:

$$NOEC_{sediment} = NOEC_{oral,\ predator} / BCF_{fish} * BMF_{fish} \qquad (20)$$

The BCF_{fish} accounts for any uptake from the water column, while the BMF_{fish} accounts for ingestion of contaminants from eating organisms from lower trophic levels, which may have acquired contaminants from any possible exposure route. In case measured BCF or BMF values are not available, the EU methodology provides default values for organic substances based on the log K_{ow} of the substance. Measured BAFs could also be used for this assessment in place of (BCF_{fish} * BMF_{fish}), because a BAF encompasses both dietary and aqueous uptake. The $NOEC_{sediment}$ can then be compared to the SQC to check if secondary poisoning would be expected if the sediment concentration was equal to the SQC.

The Australia/New Zealand methodology gives specific guidance on how to check that the SQC do not lead to concentrations that exceed human health standards for consumption of aquatic organisms (ANZECC and ARMCANZ 2000). A sediment criterion based on a bioconcentration factor can be calculated with the following equations:

$$C_{sed} = C_{org} * K_d / r * BCF \qquad (21)$$

$$K_d = f_{oc} * K_{oc} \quad (22)$$

where C_{sed} is the sediment criterion, C_{org} is the human health standard for fish, r is an empirical concentration ratio of suspended matter to sediment (r = 2 for organics), BCF is the bioconcentration factor for a sediment-ingesting aquatic organism, and f_{oc} is the fraction of organic carbon of the sediment. This criterion can be compared to the derived SQC to ensure that it will be protective of human health standards.

Because the physical-chemical properties that make a chemical likely to be associated with sediments are similar to those that make a chemical likely to bioaccumulate, secondary poisoning should be considered in sediment quality criteria derivation. There are several ways to predict toxicity due to bioaccumulation and biomagnification, and many methodologies already incorporate these procedures into the criteria derivation process. By using BCFs, BAFs and BMFs, along with dietary exposure toxicity data or human health standards, it can be ensured that a proposed SQC will not lead to secondary poisoning.

6.3.4 Harmonization across media

To harmonize a sediment quality criterion across various media is to ensure that the derived SQC will not cause excursions of criteria in other compartments. Sediment contaminants may desorb into the porewater or overlying water, and contaminants in water may volatilize into the air, which could then deposit on soil or water surfaces. Because sediment and overlying water are such closely associated environmental media, harmonization across these two compartments is the focus in most SQG methodologies that address this topic.

Cross-media harmonization is addressed by the EqP approach because SQGs are derived based on the chronic water quality criterion (FCV) and the equilibrium partitioning coefficient between sediment OC and water (Di Toro et al. 2002). Thus, the SQG is calculated so that if the concentration in sediment is equal or below the SQG, then concentration in water will be equal or below the WQC, assuming the system is at equilibrium. All of the EqP methodologies (USEPA, OECD, The Netherlands, EU, France, UK, and Ontario) are inherently harmonized between the sediment and water compartments, but other compartments are not addressed (e.g., air, soil).

The Dutch method (RIVM 2001) states the maximum permissible concentration derived for the primary (emission) compartment of concern should be used to predict concentrations in receiving compartments at steady-state. The concentrations in the receiving compartments should then be compared to the MPCs derived for these compartments, and if the MPC of the emission compartment leads to MPC exceedances in the receiving compartments, then the MPC of the emission compartment should be adjusted. When the emission compartment is sediment, the harmonization procedure depends on how the MPC_{sed} was derived. If EqP was used, then it is already harmonized with the water compartment and no further assessment is needed. If there was enough data to use a statistical extrapolation, then no further harmonization is required because this is considered a refined effects assessment. If the sediment environmental risk limit was derived using assessment factors, then this value should be compared to the value derived

using the EqP approach, and the lower of the two is taken as the harmonized environmental risk limit.

Harmonization between the sediment and water compartments should be assessed because it is likely that there will be existing WQC for compounds for which SQC are being developed. If EqP is used to derive SQC, then compartments are already harmonized, but if assessment factors or statistical distributions are used to calculate SQC, then these criteria should be harmonized with the aqueous compartment.

6.3.5 Utilization of all available data and encouragement to generate data

In their review of WQC derivation methodologies, TenBrook and Tjeerdema (2006) state, "A recurring theme throughout this review is that ecotoxicity data are generally too scarce to allow for derivation of criteria with a high level of certainty that they will neither over- nor underprotect aquatic ecosystems," and the same statement applies to this review of SQG methodologies. These authors conclude that a chosen method should encourage the generation of additional ecotoxicity data by all stakeholders, because having a large, broad, high quality toxicity dataset is the best way to produce criteria with a high level of certainty. Of the three main approaches discussed in this review, only the empirical approach utilizes all available data to derive SQGs. The EqP approach essentially ignores any existing sediment toxicity data in favor of (likely) larger aquatic toxicity datasets, and the assessment factor SSTT approach only utilizes the single lowest value in the dataset. The SSD SSTT approach uses all available laboratory toxicity data, but excludes field bioassays. Yet, the empirical approaches are not able to result in single compound numeric criteria based on a causal relationship between sediment chemistry and observed biological effects, so we are left with the EqP and SSTT approaches.

If the EqP approach is used, stakeholders have no incentive to generate sediment toxicity data because this approach only uses aquatic toxicity data, which tends to be more plentiful than sediment toxicity data. On the other hand, adoption of the SSTT approaches would encourage data generation by either the AF or SSD procedures because the magnitude of AFs decrease and SSDs become less uncertain as the number of data increase. Thus, to utilize all available data and encourage generation of data, a SQC methodology could have different derivation procedures depending on the size of the dataset, for which uncertainty decreases as the number of data increases. For example, if 5 or more data are available, an SSD would be used, if 1-5 data were available an AF would be used, of which the magnitude would decrease as the number of data increased, and if no sediment data were available the EqP approach would be used. Uncertainty can be quantified when an SSD is used, which provides information regarding the degree to which resulting criteria are likely to be under- or overprotective, so this approach would be favored if enough data were available.

7 Conclusions

We evaluated three main approaches that are currently used for deriving sediment quality guidelines: empirical, mechanistic equilibrium partitioning, and spiked-sediment toxicity testing approaches. Empirical approaches make use of large datasets that include both field bioassay and laboratory toxicity data, with matching chemistry data, to determine ranges of sediment

concentrations that are likely or not likely to cause toxicity. The empirical approaches would not be useful for determining SQC for pesticides in the Sacramento-San Joaquin River basins because a direct cause-effect relationship between a single sediment contaminant and toxicity cannot be discerned and thus far chemistry measurements have not accounted for bioavailability, leading to numeric values with high uncertainty and low reliability; in addition, little if any matching chemistry and toxicity data are available for pesticides. The EqP approach uses aquatic toxicity data with the equilibrium partitioning model to derive SQC. This approach could be used in the Sacramento-San Joaquin River basins for any pesticides for which there are existing water quality criteria, but the assumption of equilibrium in aquatic ecosystems may not be valid, and this approach neglects any available sediment toxicity data. The SSTT approaches utilize sediment toxicity data, creating a scientifically defensible foundation for SQC, but there are still experimental uncertainties regarding spiking technique and equilibration times. The species sensitivity distribution method allows for the calculation of confidence intervals so that uncertainty of the criteria may be quantified, but requires a relatively large number of data, while the assessment factor method allows for the calculation of conservative criteria based on few data, but without the quantification of uncertainty. Several of the existing methodologies incorporate more than one approach, depending on the available data and properties of the compound of interest.

A summary of differences and similarities between key elements of seven methodologies that were the focus of this review are displayed in Table 5. One element regarding sediment contamination that will need to be addressed to have a technically defensible methodology is how to incorporate bioavailability, including various exposure routes. Bioavailability will need to be incorporated in both criteria derivation and compliance determination (sampling) to ensure that the numbers are comparable. Recent research on bioavailability of sediment contaminants indicates that the freely dissolved fraction in the porewater corresponds to the bioavailable fraction, but for some species exposure via ingestion of contaminated food and/or sediments or direct contact will also be significant, and exposure could be underpredicted if these exposure routes are overlooked.

There are two possible outcomes of this project: 1) recommend an existing methodology for adoption, or 2) develop an entirely new methodology. If a new methodology is developed, it will likely use elements from the existing methodologies, and hopefully add new techniques for more refined risk assessment than is currently available in any of the existing methodologies. The next phase of this project will be to test out and further explore the various approaches to determine which will result in the most reliable and robust methodology for developing sediment quality criteria. The Netherlands (RIVM 2001) and EU (ECB 2003) methodologies, which include both SSTT and EqP approaches, appear to be the most developed SQC derivation methodologies for deriving single numeric criteria, and will be compared to whatever methodology results from Phase II of this project.

Table 5 Summary of the similarities and differences between main methodologies capable of generating numeric sediment quality criteria.

Method Considerations			USEPA (Di Toro et al. 2002)	RIVM (2001)	ECB (2003)	CCME (1999)	Ontario (Persaud et al. 1993)	France (Lepper 2002)	OECD (1995)
Ecotoxicity data used for derivation		Directly		x	x	x		x	
		Indirectly	x	x	x		x	x	x
		Non-traditional endpoints			x	x			
EqP	K_{oc}	Measured	x	x	x		x	x	x
		Estimated	x	x	x		x	x	x
		Site-specific							
	FCV	Species sensitivity distribution	x	x	x				x
		Assessment factor	x	x	x	x	x	x	x
SSTT	SSD	Distribution stated		x	x				
		Minimum number of data required		4	10				
		Minimum number of taxa required		4	8				
	AF	Minimum number of data required		1	1	4		1	
		Minimum number of taxa required		1	1	2		1	
Criteria considerations		Acute		x		x		x	
		Chronic preferred	x	x	x	x	x	x	x
		Magnitude	x	x	x	x	x	x	x
		Duration							
		Frequency							
		Bioavailability	x	x	x	x	x	x	x
		Bioaccumulation		x	x	x			
		Additivity	x		x				
		Threatened and endangered species							
Sediment considerations	Normalization of sediments	Standardized sediment		x					
		Organic carbon content	x	x	x	x	x	x	x
		Suspended sediments			x			x	
		Bedded sediments	x	x		x	x	x	x
		Ingestion		x	x				

8 References

Aldenberg T, Jaworska JS (2000) Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. Ecotox Environ Safe 46:1-18.

Amweg EL, Weston DP, Ureda N (2005) Use and toxicity of pyrethroid pesticides in the Central Valley, CA, USA. Environ Sci Technol 24:966-972.

Amweg EL, Weston DP, You J, Lydy MJ (2006) Pyrethroid insecticides and sediment toxicity in urban creeks from California and Tennessee. Environ Sci Technol 40:1700-1706.

Ankley GT, Berry WJ, Di Toro DM, Hansen DJ, Hoke RA, Mount DR, Reiley MC, Swartz RC, Zarba CS (1996) Use of equilibrium partitioning to establish sediment quality criteria for nonionic chemicals: a reply to Iannuzzi et al. Environ Toxicol Chem 15:1019-1024.

ANZECC and ARMCANZ (2000) Australian and New Zealand guidelines for fresh and marine water quality. Australian and New Zealand Environment and Conservation Council and Agriculture and Resource management Council of Australia and New Zealand, Canberra, Australia.

Apitz SE, Power EA (2002) From risk to sediment management: an international perspective. J Soil Sediment 2:61-66.

ASTM (1990a) Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. American Society of Testing and Materials. ASTM Designation: E 1367-90. 24 pp.

ASTM (1990b) Standard guide for conducting sediment toxicity tests with freshwater invertebrates. American Society of Testing and Materials. ASTM Designation: E 1383-90.

ASTM (2004) Standard guide for conducting *Daphnia magna* life-cycle toxicity tests. American Society of Testing and Materials. ASTM designation: E 1193-97(2004).

ASTM (2006a) Standard guide for conducting laboratory toxicity tests with freshwater mussels. American Society of Testing and Materials. ASTM designation: E 2455-06.

ASTM (2006b) Standard guide for conducting three-brood, renewal toxicity tests with *Ceriodaphnia dubia*. American Society of Testing and Materials. ASTM designation: E 1295-01(2006).

ASTM (2007a) Standard guide for conducting sediment toxicity tests with polychaetous annelids. American Society of Testing and Materials. ASTM designation: E 1611-00 (reapproved 2007).

ASTM (2007b) Standard guide for conducting whole sediment toxicity tests with amphibians. American Society of Testing and Materials. ASTM designation: E 2591-07.

ASTM (2008a) Standard tests method for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates. American Society of Testing and Materials. ASTM designation: E 1706-05 (2008).

ASTM (2008b) Standard tests method for measuring the toxicity of sediment-associated contaminants with estuarine and marine invertebrates. American Society of Testing and Materials. ASTM designation: E 1367-03 (2008).

ASTM (2008c) Standard Guide for Collection, Storage, Characterization, and Manipulation of Sediments for Toxicological Testing and for Selection of Samplers Used to Collect Benthic Invertebrates. American Society of Testing and Materials. Report E1391 - 03.

ASTM (2008d) Standard Guide for designing biological tests with sediments. American Society of Testing and Materials. ASTM designation: E 1525-02(2008).

ASTM (2010) Standard guide for determination of the bioaccumulation of sediment-associated contaminants by benthic invertebrates. American Society of Testing and Materials. ASTM designation: 1688-10.

Balthis WL, Hyland JL, Fulton MH, Pennington PL, Cooksey C, Key PB, DeLorenzo ME, Wirth EF (2010) Effects of chemically spiked sediments on estuarine benthic communities: a controlled mesocosm study. Environ Monit Assess 161:191-203.

Barrick R, Becker S, Brown L, Beller H, Pastorok R (1988) Sediment quality values refinement: 1988 Update and evaluation of Puget Sound AET, Volume 1. PTI Environmental Services, Bellevue, WA 74 p.

Bockting GJM, Van de Plassche EJ, Struijs J, Canton JH (1993) Soil-water partition coefficients for organic compounds. Report 679101 013, RIVM, Bilthoven, The Netherlands.

Boese BL, Lee II H, Specht DT, Randall RC (1990) Comparison of aqueous and solid-phase uptake for hexachlorobenzene in the tellinid clam *Macoma nasuta* (Conrad): A mass balance approach. Environ Toxicol Chem 9:221-231.

Boethling RS, MacKay D (2000) Handbook of property estimation methods for chemicals. Environmental and health sciences. Lewis Publishers, Boca Raton, Florida, USA.

Bondarenko S, Gan J (2009) Simultaneous measurement of free and total concentrations of hydrophobic compounds. Environ Sci Technol 43:3772-3777.

Bondarenko S, Spurlock F, Gan J (2007) Analysis of pyrethroids in sediment pore water by solid-phase microextraction. Environ Toxicol Chem 26:2587-2593.

Burton GA, Jr (2002) Sediment quality criteria in use around the world. Limnology 3: 65-75.

Burton GA, Ingersoll CG, Burnett LC, Henry M, Hinman ML, Klaine SJ, Landrum PF, Ross P, Tuchman M (1996) A comparison of sediment toxicity test methods at three Great Lake areas of concern. International Association of Great Lakes Research 22:495-511.

California SWRCB (2011) State Water Resources Control Board web site. [Accessed January 24, 2011]. Available from: http://www.waterboards.ca.gov/about_us/water_boards_structure/mission.shtml.

CCME (1999) Protocol for the derivation of Canadian sediment quality guidelines for the protection of aquatic life. Canadian Council of Ministers of the Environment. 1995 CCME EPC-98E.

Caquet T, Hanson ML, Roucaute M, Graham DW, Lagadic L (2007) Influence of isolation on the recovery of pond mesocosms from the application of an insecticide. II. Benthic macroinvertebrate responses. Environ Toxicol Chem 26:1280-1290.

CDFG (2011) State and federally listed endangered and threatened animals of California. State of California, The Natural Resources Agency, Department of Fish and Game. [Accessed May 24, 2011] Available at: http://www.dfg.ca.gov/biogeodata/cnddb/pdfs/TEAnimals.pdf.

CDPR Pesticide Data Index [Accessed 24 May 2011]. Available at: http://apps.cdpr.ca.gov/ereglib/main.cfm

Chapman PM (1989) Current approaches to developing sediment quality criteria. Environ Toxicol Chem 8:589-599.

Chapman PM (2000) Editorial: Why are we still emphasizing chemical screening-level numbers? Mar Pollut Bull 40:465-466.

Chapman PM (2007) Editorial: Do not disregard the benthos in sediment quality assessments! Mar Pollut Bull 54:633-635.

Chapman PM, Mann GS (1999) Sediment quality values (SQVs) and ecological risk assessment (ERA). Mar Pollut Bull 38:339-344.

Conrad AU, Fleming RJ, Crane M (1999) Laboratory and field response of *Chironomus riparius* to a pyrethroid insecticide. Water Res 33:1603-1610.

Crane M (2003) Proposed development of sediment quality guidelines under the European Water Framework Directive: a critique. Toxicol Lett 142:195-206.

Crane M, Everts J, van de Guchte F, Heimbach F, Hill I, Matthiessen P, Stronkhorst J (1996) Research needs in sediment bioassay and toxicity testing. In: Munawar M, Dave G (eds) Development and progress in sediment quality assessment: Rationale, challenges, techniques and strategies. SPB Academic Publishing, Amsterdam, The Netherlands.

CRWQCB-CVR (2006) 2006 CWA Section 303(d) List of Water Quality Limited Segments. Central Valley Regional Water Quality Control Board web site. [Accessed January 24, 2011]. Available from: http://www.waterboards.ca.gov/water_issues/programs/tmdl/docs/303dlists2006/swrcb/r5_final303dlist.pdf.

CRWQCB-CVR (2009) The Water Quality Control Plan (Basin Plan) for the California Regional Water Quality Control Board Central Valley Region, fourth edition, the Sacramento River Basin and the San Joaquin River Basin. [Accessed January 24, 2011]. Available from: http://www.swrcb.ca.gov/rwqcb5/water_issues/basin_plans/sacsjr.pdf

Cubbage J, Batts D, Breindenbach S (1997) Freshwater sediment quality values in Washington State. Washington State Department of Ecology, Environmental Investigations and Laboratory Service Program, Olympia, WA.

Day KE, Clements WH, DeWitt T, Landis WG, Landrum P, Morrisey DJ, Riley M, Rosenberg DM, Suter GW (1995) Chapter 12, Workgroup summary report on critical issues of ecological relevance in sediment risk assessment. In: Ingersoll CG, Dillon T, Biddinger GR (eds) Ecological Risk Assessment of Contaminated Sediments. SETAC Press, Pensacola, FL.

de Bruijn J, Busser F, Seinen W, Hermens J (1989) Determination of octanol/water partition coefficients for hydrophobic organic chemicals with the slow stirring method. Environ Toxicol Chem 8: 499-412.

Diaz RJ, Rosenberg R (1996) The influence of sediment quality on functional aspects of marine benthic communities. In: Munawar M, Dave G (eds). Development and progress in sediment quality assessment: Rationale, challenges, techniques and strategies. SPB Academic Publishing, Amsterdam, The Netherlands.

Di Toro DM, McGrath JA (2000) Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. II. Mixtures and sediments. Environ Toxicol Chem 9:1487-1502.

Di Toro DM, Hansen DJ, DeRosa LD, Berry WJ, Bell HE, Reiley MC, Zarba CS (2002) Technical basis for the derivation of equilibrium partitioning sediment quality guidelines (ESGs) for the protection of benthic organisms: Nonionic organics. Draft report. 822-R-02-041. USEPA. Office of Science and Technology and Office of Research and Development, Washington, D.C.

Domalgalski JL, Weston DP, Zhang M, Hladik M (2010) Pyrethroid insecticide concentrations and toxicity in streambed sediments and loads in surface water of the San Joaquin Valley, California, USA. Environ Toxicol Chem 29:813-823.

ECB (2003) Technical guidance document on risk assessment in support of commission directive 93/67/EEC on risk assessment for new notified substances, commission regulation (EC)

no. 1488/94 on risk assessment for existing substances, directive 98/8 EC of the European Parliament and of the council concerning the placement of biocidal products on the market. Part II. Environmental Risk Assessment. European Chemicals Bureau, European Commission Joint Research Center, European Communities.

Environment Canada (1992a) Biological test method: Acute test for sediment toxicity using marine and estuarine amphipods. Report EPS 1/RM/26. Environmental Protection, Conservation and Protection, Ottawa, Canada.

Environment Canada (1992b) Toxicity tests using luminescent bacteria (*Photobacterium phosphoreum*). Report EPS 1/RM/24. Environmental Protection, Conservation and Protection, Ottawa, Canada.

Environment Canada (1992c) Fertilization assay with echinoids (sea urchins and sand dollars). Report EPS 1/RM/27. Environmental Protection, Conservation and Protection, Ottawa, Canada.

Environment Canada (1994a) Guidance for the collection, handling, transport, storage and manipulation of sediments for chemical characterization and toxicity testing. Draft report. Environmental Protection, Conservation and Protection, Ottawa, Canada.

Environment Canada (1995) Guidance document on measurement of toxicity test precision using control sediments spiked with a reference toxicant. Report EPS 1/RM/30. Ottawa, Canada.

Environment Canada (1997a) Biological test method: Test for growth and survival in sediment using larvae of freshwater midges (*Chironomus tentans* or *Chironomus riparius*). Technical report EPS 1/RM/32. Ottawa, Canada.

Environment Canada (1997b) Biological test method: Test for growth and survival in sediment using the freshwater amphipod *Hyalella azteca.* Technical report EPS 1/RM/33. Ottawa, Canada.

Field LJ, MacDonald DD, Norton SB, Ingersoll CG, Severn CG, Smorong D, Lindskoog R (2002) Predicting amphipod toxicity from sediment chemistry using logistic regression models. Environ Toxicol Chem 21:1993-2005.

Field LJ, MacDonald DD, Norton SB, Severn CG, Ingersoll CG (1999) Evaluating sediment chemistry and toxicity data using logistic regression modeling. Environ Toxicol Chem 18:1311-1322.

Finney DJ (1942) The analysis of toxicity tests on mixtures of poisons. Ann Appl Biol 29:82-94.

Fleming RJ, van de Guchte C, Grootelaar L, Ciarelli S, Borchert J, Looise B, Guerra MT (1996) Sediment tests for poorly water soluble substances: European intra- and inter-laboratory comparisons. In: Munawar M, Dave G (eds) Development and progress in sediment quality assessment: Rationale, challenges, techniques and strategies. SPB Academic Publishing, Amsterdam, The Netherlands.

Forbes VE, Cold A (2005) Effects of the pyrethroid esfenvalerate on life-cycle traits and population dynamics of *Chironomus riparius* - Importance of exposure scenario. Environ Toxicol Chem 24:78-86.

Fuchsman PC, Barber TR (2000) Spiked sediment toxicity testing of hydrophobic organic chemicals: bioavailability, technical considerations, and applications. Soil Sediment Contam 9:197-218.

Gan J, Bondarenko S, Spurlock F (2008) Persistence and phase distribution in sediment. In: Gan J, Spurlock F, Hendley P, Weston DP. Synthetic Pyrethroids: Occurrence and Behavior in Aquatic Environments. Oxford University Press, New York, NY.

Gerstl Z (1990) Estimation of organic chemical sorption by soils. J Contam Hydrol 6:357–375.
Gilliom RJ (2007) Pesticides in U.S. streams and groundwater. Environ Sci Technol 41:3409-3414.
Hansch C, Leo A, Hoekman D (1995) Exploring QSAR. Hydrophobic, electronic, and steric constants. American Chemical Society, Washington, DC.
Hatakeyama S, Yokoyama N (1997) Correlation between overall pesticide effects monitored by shrimp mortality test and change in macrobenthic fauna in a river. Ecotoxicol Environ Saf 36:148-161.
Hladik ML, Kuivila KM (2009) Assessing the occurrence and distribution of pyrethroids in water and suspended sediments. J Agric Food Chem 57:9079-9085.
Holmes RW (2004) Monitoring of sediment-bound contaminants in the lower Sacramento River watershed. Surface water ambient monitoring program (SWAMP). Lower Sacramento River watershed. Final report. California Environmental Protection Agency. Regional Water Quality Control Board Central Valley Region. Available at: http://www.waterboards.ca.gov/centralvalley/water_issues/water_quality_studies/sedimentchem.pdf.
Hose GC, Lim RP, Hyne RV, Pablo F (2002) A pulse of endosulfan-contaminated sediment affects macroinvertebrates in artificial streams. Ecotoxicol Environ Saf 51:44-52.
Hunter W, Xu YP, Spurlock F, Gan J (2008) Using disposable polydimethylsiloxane fibers to assess the bioavailability of permethrin in sediment. Environ Toxicol Chem 27:568-575.
Hunter W, Yang Y, Reichenberg F, Mayer P, Gan J (2009) Measuring pyrethroids in sediment pore water using matrix-solid phase microextraction. Environ Toxicol Chem 28:36-43.
Ingersoll CG, Haverland PS, Brunson EL, Canfield TJ, Dwyer FJ, Henke CE, Kemble NE (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, Ivey CD, Brunson EL, Hardesty DK, Kemble NE (2000) An evaluation of the toxicity: Whole sediment vs. overlying-water exposures with the amphipod *Hyalella azteca*. Environ Toxicol Chem 19:2906-2910.
Ingersoll CG, MacDonald DD (2002) A Guidance Manual to Support the Assessment of Contaminated Sediments in Freshwater Ecosystems. Volume III- Interpretation of the results of sediment quality investigations. Great Lakes National Program Office: Chicago, IL. EPA-905-B02-001-C.
Iannuzzi TJ, Bonnevie NL, Huntley SL, Wenning RJ, Truchon SP, Tull JD, Sheehan PJ (1995) Comments on the use of equilibrium partitioning to establish sediment quality criteria for nonionic chemicals. Environ Toxicol Chem 14:1257-1259.
Kalf DF, Mensink BJWG, Montforts MHMM (1999) Protocol for the derivation of harmonized maximum permissible concentrations (MPCs). RIVM report 601506 001. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.
Karickhoff SW, Long, JM (1995) Internal report on summary of measured, calculated and recommended LogKow values. Internal Report. Environmental Research Laboratory, U.S. Environmental Protection Agency, Athens, GA.
La Point TW, Belanger SE, Crommentuijn T, Goodrich-Mahoney J, Kent RA, Mount DI, Spry DJ, Vigerstad T, Di Toro DM, Keating FJ Jr, Reiley MC (2003) Problem Formulation. In: Reevaluation of the State of the Science for Water-Quality Criteria Development. SETAC Press, Pensacola, FL. pp. 1-14.

Laskowski DA (2002) Physical and chemical properties of pyrethroids. Rev Environ Contam Toxicol 174:49-170.
Lepper P (2002) Towards the derivation of quality standards for priority substances in the context of the water framework directive. Final report of the study contract No. B4-3040/2000/30637/MAR/E1: Identification of quality standards for priority substances in the field of water policy. Fraunhofer-Institute Molecular Biology and Applied Ecology, Munich, Germany.
LOGKOW (2000) LOGKOW octanol-water partition coefficient program. Now called KowWin. Syracuse Research Corporation, New York, NY. Available at http://www.syrres.com/esc/est_soft.htm.
Lohmann R, Burgess RM, Cantwell MG, Ryba SA, MacFarlane JK, Gschwend PM (2004) Dependency of polychlorinated biphenyl and polycyclic aromatic hydrocarbon bioaccumulation in *Mya arenaria* on both water column and sediment bed chemical activities. Environ Toxicol Chem 23:2551-2562.
Long JM, Karickhoff SW (1996) Protocol for setting Kow values. Draft, September 1996. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Ecosystems Research Division, Athens, GA. 19 pp.
Long ER, Morgan LG (1990) The potential for biological effects of sediment-sorbed contaminants tested in the national status and trends program. NOAA technical memorandum NOS OMA 52. National Oceanic and Atmospheric Administration. Seattle, Washington.
Long ER, MacDonald DD (1992) National status and trends program approach. In: Sediment classification methods compendium, Chapter 14. US Environmental Protection Agency, Washington, DC. EPA 823-R-92-006.
Long ER, MacDonald DD, Smith SL, Calder FD (1995) Incidence of adverse effects within ranges of chemical concentrations in marine and estuarine sediments. Environ Manage 19: 81-97.
Loring DH, Rantala RTT (1992) Manual for the geochemical analysis of marine sediments and suspended particulate matter. Earth Sci Rev 32:235.
Lu X, Reible DD, Fleeger JW (2004) Relative importance of ingested sediment versus pore water as uptake routes for PAHs to the deposit-feeding oligochaete *Ilyodrilus templetoni*. Arch Environ Contam Toxicol 47:207-214.
Lydy M, Belden J, Wheelock C, Hammock B, Denton D (2004) Challenges in regulating pesticide mixtures. Ecol Soc 9(6):1.
MacDonald DD (1994) Approach to the assessment of sediment quality in Florida coastal waters. Volume 1 - Development and evaluation of sediment quality assessment guidelines. Florida Department of Environmental Protection, Tallahassee, FL.
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, Ingersoll CG, Berger TA (2000) Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Arch Environ Contam Toxicol 39:20-31.
MacDonald DD, Ingersoll CG, Smorong DE, Sinclair JA, Lindskoog R, Wang N, Severn C, Gouguet R, Meyer J, Field J (2011) Baseline ecological risk assessment of the Calcasieu

Estuary, Louisiana: Part 2. An evaluation of the predictive ability of effects-based sediment-quality guidelines. Arch Environ Contam Toxicol 61:14-28.
MacKay D, Shiu W-Y, Ma K-C (1999) Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals. CRC-LLC netbase, CD-rom version.
Maruya KA, Landrum PF, Burgess RM, Shine JP (2010) Incorporating contaminant bioavailability into sediment quality assessment frameworks. pp. 153-175. In: Stephen B. Weisberg SB, Miller K (eds), Southern California Coastal Water Research Project 2010 Annual Report. Southern California Coastal Water Research Project, Costa Mesa, CA.
Mayer LM, Weston DP, Bock MJ (2001) Benzo[a]pyrene and zinc solubilization by digestive fluids of benthic invertebrates - A cross-phyletic study. Environ Toxicol Chem 20:1890-1900.
Mehler WT, You J, Maul JD, Lydy MJ (2010) Comparative analysis of whole sediment and porewater toxicity identification evaluation techniques for ammonia and non-polar organic contaminants. Chemosphere 78:814-821.
Mehler WT, Li H, Pang J, Sun B, Lydy MJ, You J (2011) Bioavailability of hydrophobic organic contaminants in sediment with different particle-size distributions. Arch Environ Contam Toxicol 61:74-82.
Mensink BJWG, Montforts M, Wijkhuizen Maslankiewiez L,Tibosh H, Linders JBHJ (1995) Manual for summarizing and evaluating the environmental aspects of pesticides. Report 679101-022 RIVM, Bilthoven, The Netherlands.
Mudroch A, Azcue JM (1995) Manual of Aquatic Sediment Sampling. Lewis Publishers, CRC Press, Boca Raton, FL.
Naf C, Axelman J, Broman D (1996) Organic contaminants in sediments of the Baltic Sea: Distribution, behaviour and fate. In: Munawar M, Dave G (eds). Development and progress in sediment quality assessment: Rationale, challenges, techniques and strategies. SPB Academic Publishing, Amsterdam, The Netherlands.
Neff JM, Bean DJ, Cornaby BW, Vaga RM, Gulbransen TC, Scanlon JA (1986) Sediment quality criteria methodology validation: Calculation of screening level concentrations from field data. Report to US Environmental Protection Agency, Office of Water Regulations and Standards, Washington, DC. 60p.
Neff JM, Word JQ, Gulbransen TC (1987) Recalculation of screening level concentrations for nonionic organic contaminants in marine sediment. Final report. US Environmental Protection Agency, Office of Water.
O'Connor TP (2000) Misfit between sediment toxicity and chemistry. Mar Pollut Bull 40:59-64.
O'Connor TP (2004) The sediment quality guideline, ERL, is not a chemical concentration at the threshold of toxicity. Mar Pollut Bull 49:383-385.
OECD (1992) Guidelines for testing chemicals, Section 2: Effects on biotic systems. Test No. 210: Fish, early-life stage toxicity test. Organisation for Economic Co-operation and Development, Paris, France.
OECD (1995) Guidance Document for Aquatic Effects Assessment. Organisation for Economic Co-operation and Development, Paris, France.
OECD (2004a) Guidelines for testing chemicals, Section 2: Effects on Biotic Systems. Test No. 218: Sediment-Water Chironomid Toxicity Using Spiked Sediment. Organisation for Economic Co-operation and Development, Paris, France.

OECD (2004b) Guidelines for testing chemicals, Section 2: Effects on Biotic Systems. Test No. 219: Sediment-Water Chironomid Toxicity Using Spiked Water. Organisation for Economic Co-operation and Development, Paris, France.
OECD (2004c) Guidelines for testing chemicals, Section 2: Effects on Biotic Systems. Test No. 202: Daphnia sp. Acute Immobilization Test. Organisation for Economic Co-operation and Development, Paris, France.
OECD (2007) Guidelines for testing chemicals, Section 2: Effects on Biotic Systems. Test No. 225: Sediment-Water Lumbriculus Toxicity Test Using Spiked Sediment. Organisation for Economic Co-operation and Development, Paris, France.
OECD (2008) Guidelines for testing chemicals, Section 2: Effects on Biotic Systems. Test No. 211: Daphnia magna Reproduction Test. Organisation for Economic Co-operation and Development, Paris, France.
OPP Pesticide Ecotoxicity Database. [Accessed 24 May 2011]. Available at: http://www.ipmcenters.org/Ecotox/DataAccess.cfm
OPPTS 850. 1735 (1996a) Whole sediment acute toxicity: Invertebrates, freshwater. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-354.
OPPTS 850. 1740 (1996b) Whole sediment acute toxicity: Invertebrates, marine. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-355.
OPPTS 850. 1790 (1996c) Chironomid sediment toxicity test. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-313.
OPPTS 850. 1800 (1996d) Tadpole/sediment subchronic toxicity test. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-132.
OPPTS 850. 1850 (1996e) Aquatic food chain transfer. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-133.
OPPTS 850. 1900 (1996f) Generic freshwater microcosm test, laboratory. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-134.
OPPTS 850. 1925 (1996g) Site-specific aquatic microcosm test, laboratory. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-173.
OPPTS 850. 1950 (1996h) Field testing for aquatic organisms. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-135.
OPPTS 850. 1010 (1996i) Aquatic invertebrate acute toxicity test, freshwater, daphnids. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-114.
OPPTS 850. 1075 (1996j) Fish acute toxicity test, freshwater and marine. USEPA Office of Prevention, Pesticides and Toxic Substances, EPA 712-C-96-118; OECD 203.
Oudou HC, Hansen HCB (2002) Sorption of lambda-cyhalothrin, cypermethrin, deltamethrin, and fenvalerate to quartz, corundum, kaolinite, and montmorillonite. Chemosphere 49: 1285-1294.
Parsons J, Jesus Belzunce Segarra M, Cornelissen G, Gustafsson O, Grotenhuis T, Harms H, Janssen CR, Kukkonen J, van Noort P, Ortega Calvo JJ, Solaun Etxeberria O (2007) Characterization of contaminants in sediments – effects of bioavailability on impact. In: Barcelo D, Petrovic M (eds) Sustainable Management of Sediment Resources: Sediment Quality and Impact Assessment of Pollutants. Elsevier, Oxford, UK.
Pavlou SP, Weston DP (1984) Initial evaluation of alternatives for development of sediment related criteria for toxic contaminants in marine waters (Phase II). Report prepared for US Environmental Protection Agency, Washington, DC.

Persaud D, Jaagumagi R, Hayton A (1993) Guidelines for the protection of and management of aquatic sediment quality in Ontario. Ontario Ministry of the Environment, Water Resources Branch.

Plackett RL, Hewlett PS (1952) Quantal responses to mixtures of poisons. J Roy Stat Soc B 14:141-163.

Raimondo S, Vivian DN, Barron MG (2010) Web-based Interspecies Correlation Estimation (Web-ICE) for Acute Toxicity: User Manual. Version 3.1. Office of Research and Development, U.S. Environmental Protection Agency, Gulf Breeze, FL. EPA/600/R-10/004.

Rider CV, LeBlanc GA (2005) An integrated addition and interaction model for assessing toxicity of chemical mixtures. Toxicol Sci 87:520-528.

Ritter KJ, Bay SM, Smith RW, Vidal-Dorsch DE, Field LJ (2007) Development and evaluation of sediment quality guidelines based on benthic macrofauna responses. Technical Report 544. Southern California Coastal Water Research Project, Costa Mesa, CA.

RIVM (2001) Guidance document on deriving environmental risk limits. Traas TP (ed) RIVM report 601501 012. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

RIVM (2004) ETX 2.0. Normal distribution based hazardous concentration and fraction affected. Designed by Van Vlaardingen P, Traas T, Aldenberg T, Wintersen A. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

Rowlatt S, Matthiessen P, Reed J, Law R, Mason C (2002) Recommendations on the development of sediment quality guidelines. Environment Agency, Bristol, UK.

Savage WK, Quimby FW, DeCaprioi AP (2002) Lethal and sublethal effects on polychlorinated biphenyls on *Rana sylvatica* tadpoles. Environ Toxicol Chem 21:168-174.

Schwarzenbach RP, Gschwend PM, Imboden DM (2003) Environmental Organic Chemistry, 2nd edition. John Wiley & Sons, Inc., Hoboken, NJ.

Selck H, Palmqvist A, Forbes VE (2003) Uptake, depuration, and toxicity of dissolved and sediment-bound fluoranthene in the polychaete, *Capitella* sp. I. Environ Toxicol Chem 22:2354-2363.

Semple KT, Doick KJ, Jones KC, Burauel P, Craven A, Harms H (2004) Defining bioavailability and bioaccessibility of contaminated soil and sediment is complicated. Environ Sci Technol 38:228A-231A.

Sijm D, De Bruijn J, Crommentuijn T, van Leeuwen K (2001) Environmental quality standards: endpoints or triggers for a tiered ecological effect assessment approach? Environ Toxicol Chem 20:2644-2648.

Smith SL, MacDonald DD, Keenleyside KA, Ingersoll CG, Field LJ (1996) A preliminary evaluation of sediment quality assessment values for freshwater ecosystems. J Great Lakes Res 22:624-638.

Sormunen AJ, Tuikka AI, Akkanen J, Leppanen MT, Kukkonen JVK (2010) Predicting the bioavailability of sediment-associated spiked compounds by using the polyoxymethylene passive sampling and Tenax ® extraction methods in sediments from three river basins in Europe. Arch Environ Contam Toxicol 58:80-90.

Starner K, Kelley K (2005) Pyrethroid concentrations in surface water and bed sediment in high agricultural use regions of California. Poster presented at 26th Annual SETAC Conference, November 2005 at Baltimore, Maryland.

Stephan CE, Rogers JW (1985) Advantages of using regression analysis to calculate results of chronic toxicity tests. In: Bahner RC, Hansen DJ (eds) Aquatic Toxicology and Hazard Assessment Eighth Symposium. ASTM STP 891. American Society for Testing and Materials, Philadelphia, PA, pp. 328-338.

Swartz RC, Schults DW, Ozretich RJ, Lamberson JO, Cole FA, DeWitt TH, Redmond MS, Ferraro SP (1995) ∑PAH: A model to predict the toxicity of field collected marine sediment contamination by polycyclic aromatic hydrocarbons. Environ Toxicol Chem 14:1977-1987.

SWRCB (2011) Appendix A. Draft proposed amendments to the water quality control plan for enclosed bays and estuaries plan part 1: Sediment quality. Accessed Feb 15, 2011: http://www.swrcb.ca.gov/water_issues/programs/bptcp/sediment.shtml.

TenBrook PL, Tjeerdema RS (2006) Methodology for derivation of pesticide water quality criteria for the protection of aquatic life in the Sacramento and San Joaquin Rivers basins. Phase I: Review of existing methodologies. Report prepared for the Central Valley Regional Water Quality Control Board.

TenBrook PL, Tjeerdema RS**,** Hann P, Karkoski J (2009) Methods for deriving pesticide aquatic life criteria. Rev Environ Contam Toxicol 199:19–110.

Tracey GA, Hansen DJ (1996) Use of biota-sediment accumulation factors to assess similarity of nonionic organic chemical exposure to benthically-coupled organisms of differing tropic mode. Arch Environ Contam Toxicol 30:467-475.

Trimble TA, You J, Lydy MJ (2008) Bioavailability of PCBs from field-collected sediments: application of Tenax extraction and matrix-SPME techniques. Chemosphere 71:337-344.

USEPA (1985) Guidelines for developing numerical national water quality criteria for the protection of aquatic organisms and their uses. PB-85-227049. US Environmental Protection Agency National Technical Information Service, Springfield, VA.

USEPA (1989) An SAB report: Evaluation of the Apparent Effects Threshold (AET) approach for assessing sediment quality. Sediment Criteria Subcommittee. EPA-SAB-EETFC-89-027.

USEPA (1992) An SAB report: Review of sediment criteria development. Sediment Quality Subcommittee of the Ecological Processes and Effects Committee. EPA-SAB-EPEC-93-002.

USEPA (1993) Technical basis for deriving sediment quality criteria for nonionic organic contaminants for the protection of benthic organisms by using equilibrium partitioning. Report. EPA-822-R-93-011. USEPA. Office of Science and Technology and Office of Research and Development, Washington, D.C.

USEPA (1994a) Assessment Guidance Document: Assessment and Remediation of Contaminated Sediments Program, Great Lakes National Program Office, Chicago, IL. EPA 905-R94-002.

USEPA (1994b) Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. EPA 600-R24-024.

USEPA (1998) Contaminated Sediment Management Strategy (EPA 823-R-98-001). http://www.epa.gov/waterscience/cs/stratndx.html

USEPA (2000a) Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Second edition. EPA 600/R-99/064. Office of Research and Development. Washington D. C.

USEPA (2000b) Prediction of sediment toxicity using consensus-based freshwater sediment quality guidelines. EPA 905/R-00/007. Great Lakes National Program Office, Chicago, IL.
USEPA (2003a) Procedures for the derivation of equilibrium partitioning sediment benchmarks (ESBs) for the protection of benthic organisms: Dieldrin. EPA 600 R 02 010. Office of Research and Development, Washington D.C.
USEPA (2003b) Procedures for the derivation of equilibrium partitioning sediment benchmarks (ESBs) for the protection of benthic organisms: Endrin. EPA 600 R 02 009. Office of Research and Development, Washington D.C.
USEPA (2003c) Procedures for the derivation of equilibrium partitioning sediment benchmarks (ESBs) for the protection of benthic organisms: PAH mixtures. EPA 600 R 02 013. Office of Research and Development, Washington D.C.
USEPA (2003d) Water quality guidance for the Great Lakes system. Federal Register, 40 CFR Part 132. US Environmental Protection Agency, Washington, DC.
Verscheuren K. 1983. Handbook of environmental data on organic chemicals, 2nd edition, Van Nostrand Reinhold Co., New York, NY.
Verscheuren K (2001) Handbook of environmental data on organic chemicals, 4th edition, CD-ROM, Wiley Interscience, New York, NY.
Wallace JB, Lugthart GJ, Cuffney TF, Schurr GA (1989) The impact of repeated insecticidal treatments on drift and benthos of a headwater stream. Hydrobiologia 79:135-147.
Wenning RJ, Batley GE, Ingersoll CG, Moore DW (eds) (2005) Use of sediment quality guidelines and related tools for the assessment of contaminated sediments. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL. 815 p.
Wenning RJ, Ingersoll CG (2002) Summary of the SETAC Pellston Workshop on Use of Sediment Quality Guidelines and Related Tools for the Assessment of Contaminated Sediments; 17-22 August 2002; Fairmont, Montana, USA. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola FL.
Weston DP, Holmes RW, You J, Lydy MJ (2005) Aquatic toxicity due to residential use of pyrethroid insecticides. Environ Sci Technol 39:9778-9784.
Weston DP, You J, Lydy MJ (2004) Distribution and toxicity of sediment-associated pesticides in agriculture dominated water bodies of California's central valley. Environ Sci Technol 38:2752-2759.
Weston DP, Zhang M, Lydy MJ (2008) Identifying the cause and source of sediment toxicity in an agriculture-influenced creek. Environ Toxicol Chem 27:953-962.
Wheelock CE, Miller JL, Miller MJ, Phillips BM, Gee SJ, Tjeerdema RS, Hammock BD (2005) Influence of container adsorption upon observed pyrethroid toxicity to *Ceriodaphnia dubia* and *Hyalella azteca.* Aquat Toxicol 74:47-52.
Woin P (1998) Short- and long-term effects of the pyrethroid insecticide fenvalerate on an invertebrate pond community. Ecotoxicol Environ Saf 41:137-156.
Woodburn KB, Doucette WJ, Andren AW (1984) Generator column determination of octanol-water partition coefficients for selected polychlorinated biphenyl congeners. Environ Sci Technol 18:457-459.
Xu Y, Spurlock F, Wang Z, Gan J (2007) Comparison of five methods for measuring sediment toxicity of hydrophobic contaminants. Environ Sci Technol 41:8394-8399.
Yang WC, Gan JY, Hunter W, Spurlock F (2006a) Effect of suspended solids on bioavailability of pyrethroid insecticides. Environ Toxicol Chem 25:1585-1591.

Yang WC, Hunter W, Spurlock F, Gan J (2007) Bioavailability of permethrin and cyfluthrin in surface waters with low levels of dissolved organic matter. J Environ Qual 36:1678-1685.
Yang WC, Spurlock F, Liu WP, Gan. JY (2006b) Inhibition of aquatic toxicity of pyrethroid insecticides by suspended sediment. Environ Toxicol Chem 25:1913-1919.
Yasuno M, Fukushima S, Hasegawa J, Shioyama F, Hatakeyama S (1982) Changes in the benthic fauna and flora after application of temephos to a stream on Mt. Tsukuba. Hydrobiologia 89:205-214.
You J, Harwood, AD, Li H, Lydy MJ (2011) Chemical techniques for assessing bioavailability of sediment-associated contaminants: SPME versus Tenax extraction. J Environ Monit 13:792-800.
You J, Landrum F, Lydy MJ (2006) Comparison of chemical approaches for assessing bioavailability of sediment-associated contaminants. Environ Sci Technol 40: 6348-6353.
You J, Pehkonen S, Weston DP, Lydy MJ (2008a) Chemical availability and sediment toxicity of pyrethroid insecticides to Hyalella azteca: application to field sediment with unexpectedly low toxicity. Environ Toxicol Chem 27:2124-2130.
You J, Weston DP, Lydy MJ (2008b) Quantification of pyrethroid insecticides at sub-ppb levels in sediment using matrix-dispersive accelerated solvent extraction with tandem SPE cleanup. In: Gan J, Spurlock F, Hendley P, Weston DP (eds) Synthetic pyrethroids: Occurrence and Behavior in Aquatic Environments. American Chemical Society, Washington, D.C.
You J, Pehkonen S, Landrum PF, Lydy MJ (2007) Desorption of hydrophobic compound from laboratory-spiked sediment measured by Tenax absorbent and matrix solid-phase microextraction. Environ Sci Technol 41:5672-5678.
Zabel TF, Cole S (1999) The derivation of environmental quality standards for the protection of aquatic life in the UK. J CIWEM 13:436-440.
Zhou, JL, Rowland S, Mantoura, RF (1995) Partition of synthetic pyrethroid insecticides between dissolved and particulate phases. Water Res 29:1023-1031.
Zvinavashe E, Murk AJ, Rietjens IMCM (2009) On the number of EINECS compounds that can be covered by (Q)SAR models for acute toxicity. Toxicol Lett 184:67-72.