

Chapter 5

Conclusions and Discussion

The National Sediment Inventory (NSI) is EPA's largest compilation of sediment chemistry data and related biological data. It includes approximately 2 million records for more than 21,000 monitoring stations across the country. EPA's evaluation of the NSI data was the most geographically extensive investigation of sediment contamination ever performed in the United States. The evaluation was based on procedures to address the probability of adverse effects to aquatic life and human health.

The characteristics of the NSI data, as well as the degree of certainty afforded by available assessment tools, allow neither an absolute determination of adverse effects on human health or the environment at any location, nor a determination of the areal extent of contamination on a national scale. However, the evaluation results strongly suggest that sediment contamination may be significant enough to pose potential risks to aquatic life and human health in some locations. The evaluation methodology was designed for the purpose of a screening-level assessment of sediment quality; further evaluation would be required to confirm that sediment contamination poses actual risks to aquatic life or human health for any given site or watershed.

Based on the number and percentage of sampling stations containing contaminated sediment within watershed boundaries, EPA identified a number of watersheds containing areas of probable concern for sediment contamination (APCs) where additional studies may be needed to draw conclusions regarding adverse effects and the need for actions to reduce risks. Although the APCs were selected by means of a screening exercise, EPA believes that they represent the highest priority for further ecotoxicological assessments, risk analysis, temporal and spatial trend assessment, contaminant source evaluation, and management action because of the preponderance of evidence in these areas. Although the procedure for classifying APCs using multiple sampling stations was intended to minimize the probability of making an erroneous classification, further evaluation of conditions in watersheds containing APCs is necessary because the same mitigating factors that might reduce the probability of associated adverse effects at one sam-

pling station may also affect neighboring sampling stations.

EPA chose the watershed as the unit of spatial analysis because many state and federal water and sediment quality management programs, as well as data acquisition efforts, are centered around this unit. This choice reflects the growing recognition that activities taking place in one part of a watershed can greatly affect other parts of the watershed, and that management efficiencies are achieved when viewing the watershed holistically. At the same time, the Agency recognizes that contamination in some reaches in a watershed does not necessarily indicate that the entire watershed is affected.

Watershed management is a vital component of community-based environmental protection. The Agency and its state and federal partners can address sediment contamination problems through watershed management approaches. Watershed management programs focus on hydrologically defined drainage basins rather than areas defined by political boundaries. These programs recognize that conditions of land areas and activities within the watershed affect the water resource. Local management, stakeholder involvement, and holistic assessments of water quality are characteristics of the watershed approach. The National Estuary Program is one example of the watershed approach that has led to specific actions to address contaminated sediment problems. Specifically, the Narragansett (RI) Bay, Long Island Sound, New York/New Jersey Harbor, and San Francisco Bay Estuary Programs have all recommended actions to reduce sources of toxic contaminants to sediment. Numerous other examples of watershed management programs are summarized in *The Watershed Approach: 1993/94 Activity Report* (USEPA, 1994g) and *A Phase I Inventory of Current EPA Efforts to Protect Ecosystems* (USEPA, 1995b).

This chapter presents some general conclusions about the extent of sediment contamination in the United States and sources of sediment contaminants. It also includes comparisons to other national studies that address the extent of sediment contamination and to a national survey of state-issued fish consumption advisories. In addition, this chapter presents the results of an analy-

sis of the sensitivity of parameters used to evaluate potential human health effects from exposure to PCBs and mercury, which was performed to show how the use of different screening values affect the results. The chapter concludes with a discussion of the strengths and limitations of the NSI data and evaluation method.

It is important to understand both the strengths and limitations of this analysis to appropriately interpret and use the information contained in this report. The limitations do not prevent intended uses, and future reports to Congress on sediment quality will contain less uncertainty. To ensure that future reports to Congress accurately reflect current knowledge concerning the conditions of the Nation's sediment as our knowledge and application of science evolves, the NSI will develop into a periodically updated, centralized assemblage of sediment quality measurements and assessment techniques.

Extent of Sediment Contamination

Based on the evaluation, sediment contamination exists at levels where associated adverse effects are probable (Tier 1) in some locations within each region and state of the country. The water bodies affected include streams, lakes, harbors, nearshore areas, and oceans. A number of specific areas in the United States had large numbers of sampling stations where associated adverse effects are probable. Puget Sound, Boston Harbor, the Detroit River, San Diego Bay, and portions of the Tennessee River were among those locations. Several U.S. harbors (e.g., Boston Harbor, Puget Sound, Los Angeles, Chicago, Detroit) appear to have some of the most severely contaminated sediments in the country. This finding is not surprising since major U.S. harbors have been affected throughout the years by large volumes of boat traffic, contaminant loadings from upstream sources, and many local point and nonpoint sources.

Thousands of other water bodies in hundreds of watersheds throughout the country contain sampling stations classified as Tier 1. Many of these sampling stations may represent isolated "hot spots" rather than widespread sediment contamination, although insufficient data were available in the NSI to make such a determination. EPA's River Reach File 1 (RF1) delineates the Nation's rivers and waterways into segments, or reaches, of approximately 1 to 10 miles in length. Based on RF1, approximately 11 percent of all river reaches in the United States contained NSI sampling stations. More than 5,000 sampling stations in approximately 2,400 river reaches across the country (4 percent of all reaches) were classified as Tier 1. Another 10,000 sampling stations

were classified as Tier 2. In total, over 5,000 river reaches in the United States—approximately 8 percent of all river reaches—include at least one Tier 1 or Tier 2 station.

EPA cannot determine the areal extent or number of river miles of contaminated sediment in the United States because the NSI does not provide complete coverage for the entire nation, sampling locations are largely based on a nonrandom sampling design, and sediment quality can vary greatly within very short distances.

Most of the NSI data were compiled from nonrandom monitoring programs. Such monitoring programs focus sampling efforts on areas where contamination is known or suspected to occur. As a result, assuming all other factors are the same, the frequency of Tier 1 or Tier 2 classification based on the NSI data evaluation is probably greater than that which would result from purely random sampling. Swartz et al. (1995) demonstrated the effects of nonrandom sampling design on the frequency of detecting contaminated sampling stations. They compared the percent of sediment sampling stations that exceeded PAH screening effects levels (ERL, SQC, AET) based on random sampling station selection (Virginian Province EMAP stations) to the percent of sampling stations that exceeded those levels based on sampling station selection on the basis of known PAH contamination (such as creosote-contaminated Eagle Harbor, Washington). They found that the frequency of exceeding a sediment chemistry screening value in sampling stations known to be contaminated was 5 to 10 times greater than that for randomly selected sampling stations.

The percentage of all NSI sampling stations where associated adverse effects are "probable" or "possible but expected infrequently" (i.e., 26 percent in Tier 1 and 49 percent in Tier 2) does not represent the overall condition of sediment across the country: the overall extent of contaminated sediment is much less, as is the percentage of sampling stations where contamination is expected to actually exert adverse effects. For example, a reasonable estimate of the national extent of contamination leading to adverse effects to aquatic life is between 6 and 12 percent of sediment underlying surface waters. This is primarily because the majority of sampling stations in the NSI are located in known or suspected areas of sediment contamination (i.e., sampling stations were not randomly selected). However, some individual data sets that are included in the NSI, as well as the results of independent investigations conducted by other researchers, can be applied to represent the areal extent of sediment contamination in their respective study areas. EPA's EMAP data collection effort featured a probabilistic, or random, sampling design. In the Virginian and Louisianian EMAP

Provinces, located on the Mid-Atlantic and Gulf coasts respectively, 104 of 678 (15.3 percent) of sediment samples were toxic to amphipods. With a 5 percent false positive rate (statistical $\alpha=0.05$), EMAP toxicity data suggest that about 10 percent of marine and estuarine sites are sufficiently contaminated to cause lethality to benthic organisms (Richard Swartz, personal communication, December 27, 1996). In another recent study, Long et al. (1996) examined amphipod survival in test sediment collected from 1,176 locations in 22 estuarine areas throughout the nation. These authors concluded that the areal extent of toxic sediment comprised approximately 11 percent of the combined study area.

To apply the NSI evaluation to estimate the areal extent of toxic sediment in the United States, three factors must be accounted for: (1) most of the NSI data were generated from sampling targeted toward areas of known or suspected contamination, (2) sediment chemistry screening values only identify sediment associated with a *probability* of toxicity, and (3) toxicity is demonstrated at some sampling stations where sediment chemistry screening values are not exceeded. The latter condition could be a result of false positives (i.e., laboratory toxicity that would not be present in the field), toxic chemicals present in the field but not measured or evaluated, or toxicity that correlative screening values do not predict (e.g., by definition 10 percent of toxic samples in the "effects distribution" lie below the ERL).

Using information from available data and published studies, the effects of each of the above factors can be quantified. Swartz et al. (1995) suggest that exceeding a sediment chemistry screening value at sites of known or suspected contamination is 5 to 10 times more likely than at sites where sediment is randomly sampled. However, comparison of Tier 1 classification for Virginian and Louisianan EMAP data to the entire NSI data base suggests that the mix of sampling strategies in the NSI data base as a whole results in screening value exceedance at 2 to 4 times as many sampling stations than purely random sampling. Long et al., (in press), as well as a comparison of matched sediment chemistry and toxicity data within the NSI, suggest that approximately 40 percent of Tier 1 sampling stations, and 20 percent of Tier 2 sampling stations, would exhibit significant lethality to bottom dwelling aquatic organism. Both data sets also suggest that significant lethality occurs at approximately 10 percent of Tier 3 stations, where no screening value is exceeded. Alternatively, one could assume that significant laboratory toxicity at randomly sampled locations classified as Tier 3 only represents "false positives", and therefore that no toxicity occurs at Tier 3 sampling stations classified from random sampling.

In the NSI evaluation, 3,283 and 9,688 of the 17,884 sampling stations with sediment chemistry data available were classified as Tier 1 and Tier 2, respectively, for risk to bottom dwelling aquatic organisms. Using a 40 percent probability of lethality at Tier 1 and a 20 percent probability of lethality at Tier 2, and further assuming 10 times less frequent Tier 1 and Tier 2 classification (upper end of range from Swartz et al., 1995) in a random sample and no lethality at Tier 3 sampling stations, the estimated extent of sediment contamination in the United States associated with lethality to bottom dwelling aquatic organisms is 2 percent. At the other extreme, assuming 2 times less frequent Tier 1 and Tier 2 classification (lower end of range from EMAP/NSI comparisons) in a random sample and a 10 percent probability of lethality at all resulting Tier 3 sampling stations (11,399; including the additional sampling stations previously classified as Tier 1 and Tier 2 before adjusting for random sampling), the estimated extent of sediment contamination associated with lethality to bottom dwelling aquatic organisms is 15 percent. Avoiding either extreme, assuming 2 to 5 times less frequent Tier 1 and Tier 2 classification in a random sample and a 10 percent probability of lethality for only the original Tier 3 sampling stations (4,913; prior to adjusting for random sampling), the range narrows to 6 to 12 percent—about 1,000 to 2,000 toxic sampling stations out of approximately 18,000. This range encompasses the areal extent point estimates from EMAP toxicity data and Long et al. (1996). EPA believes these are reasonable estimates of the extent of sediment contamination across the United States.

The results of the NSI data evaluation must be interpreted in the context of data availability. Many states and EPA Regions appear to have a much greater incidence of sediment contamination than others. To some degree, this appearance reflects the relative abundance of readily available electronic data, not necessarily the relative incidence of sediment contamination. For example, 182 of the 920 river reaches in Illinois contain a Tier 1 sampling station, whereas only 9 of the 5,490 reaches in Montana contain a Tier 1 sampling station. However, the NSI includes sampling station data for over 50 percent of the river reaches in Illinois but less than 1 percent of the river reaches in Montana. Therefore, although the absolute number of Tier 1 and Tier 2 stations in each state is important, relative comparisons of the incidence of sediment contamination between states is not possible because the extent of sampling and data availability vary widely.

For a number of reasons, some potentially contaminated sediment sites were missed in this evaluation. The most obvious reason is that the NSI does not include all

sediment quality data that have ever been collected. For example, the NSI does not include many EPA Superfund Program data and therefore sampling stations in the vicinity of hazardous waste sites might not have been included in the NSI evaluation. Additional data sets will be added to the NSI for future evaluations to provide better national coverage. In addition, some data in the NSI were not evaluated because of questions concerning data quality or because no locational information (latitude and longitude) was available.

Sources of Sediment Contamination

Some of the most significant sources of persistent and toxic chemicals have been eliminated or reduced as the result of environmental controls put into place during the past 10 to 20 years. For example, the commercial use of PCBs and the pesticides DDT and chlordane has been restricted or banned in the United States. In addition, effluent controls on industrial and municipal point source discharges and best management practices for the control of nonpoint sources have greatly reduced contaminant loadings to many of our rivers and streams.

The results of better controls over releases of sediment contaminants are evident from studies such as that conducted by Swartz et al. (1991) on the Palos Verdes Shelf. These researchers examined sediment cores collected at two sites on the Palos Verdes Shelf near the Los Angeles County Sanitation District's municipal wastewater outfalls, and at two reference sites in Santa Monica. They found that the vertical distribution of sediment toxicity near the outfalls was significantly correlated with profiles of total organic carbon and sediment chemical contamination. Dating of core horizons showed that sediment toxicity also was significantly correlated with historical records of the mass emission rate of suspended solids from the outfalls. The vertical profiles showed that the toxicity of surficial sediments increased after the initiation of the discharge in the 1950s, remained relatively high until the early 1970s, and then decreased after the implementation of source controls and improved effluent treatment (Swartz et al., 1991).

Based on the NSI data evaluation, metals and persistent organic chemicals are the contaminants most often associated with sediment contamination. Despite recent progress in controlling sediment contaminant releases to the environment, active sources of these contaminants still exist. These include nonpoint source loadings such as surface water runoff and atmospheric deposition, point source loadings, and resuspension of in-place sediment contaminants from historical sources.

Some correlations between land use and sediment contamination caused by specific classes of chemicals were identified in Chapter 4. Agricultural land use was correlated with the extent of sediment contaminated with organochlorine pesticides in APC watersheds, especially those with more than 75 percent of land area devoted to crop production or rangeland. In contrast, the extent of sediment contaminated with PAHs, mercury, and other metals in APC watersheds correlated with the extent of urban land use. Land use did not appear to be associated with the extent of PCB contamination.

Comparison of NSI Evaluation Results to Results of Previous Sediment Contamination Studies

The results of this study are consistent with the findings of other national assessments of sediment contamination. For example, in EPA's 1992 *National Water Quality Inventory* report, 27 states identified 770 known contaminated sediment sites (USEPA, 1994e). The identified "sites" probably best correlate to river reaches from this analysis in terms of areal extent. The NSI evaluation identified approximately 2,400 river reaches in 50 states that contain a Tier 1 sampling station. In the *National Water Quality Inventory* report, the states frequently listed metals (e.g., mercury, cadmium, and zinc), PCBs, DDT (and its by-products), chlordane, and priority organic chemicals as the cause of sediment contamination. They identified industrial and municipal discharges (past and present), landfills, resource extraction, abandoned hazardous waste disposal sites, and combined sewer overflows as the most important sources of sediment contamination.

In a 1987 overview of sediment contamination (which was based on a limited amount of national data), EPA estimated that hundreds of sites located in all regions of the United States have in-place sediment contaminants at concentrations of concern (USEPA, 1987). The study identified harbor areas, both freshwater and marine, as some of the most severely impacted areas in the country. The study identified municipal and industrial point source discharges, urban and agricultural runoff, combined sewer overflows, spills, mine drainage, and atmospheric deposition as frequently cited sources of sediment contamination.

In 1994, the National Oceanic and Atmospheric Administration (NOAA) released its *Inventory of Chemical Concentrations in Coastal and Estuarine Sediments* (NOAA, 1994). This study categorized 2,800 coastal sites as either "high" or "hot" based on the contaminant concentrations found at the sampling locations. NOAA did not

use risk-based screening values for its analysis. Using the National Status and Trends Mussel Watch data set, “high” values were defined as the mean concentration for a specific chemical plus one standard deviation. High values corresponded to about the 85th percentile of contaminant concentration. “Hot” concentrations were defined as those exceeding five times the “high” values. Most of the “hot” sites were in locations with high ship traffic, industrial activity, and relatively poor flushing, such as harbors, canals, and intracoastal waterways (NOAA, 1994). Mercury and cadmium exceeded the NOAA “hot” thresholds at a greater percentage of sites where they were measured (about 7 percent each) than other sediment contaminants.

Comparison of NSI Evaluation Results to Fish Consumption Advisories

EPA recently published a National Listing of Fish Consumption Advisories issued by state governments. As of 1994, 1,532 fish consumption advisories were in place in 46 states. (Each advisory might apply to several water body segments, or reaches, as defined in this study.) Mercury was the contaminant most often associated with fish consumption advisories; 1,119 water bodies had advisories that included mercury. States also issued a large number of advisories because of high levels of chlordane, PCBs, and dioxins in fish tissue.

A direct comparison of the fish advisory contaminants and NSI contaminants is not possible because states often issue advisories for groups of chemicals. Nevertheless, five of the top six contaminants associated with fish advisories (PCBs, DDT, dieldrin, chlordane, and dioxins) are also among the contaminants most often responsible for the Tier 1 classification of water bodies based on potential human health effects (Table 5-1). As illustrated in Figure 5-1, many sampling stations categorized as Tier 1 or Tier 2 for

human health effects are located in water bodies for which fish consumption advisories have been issued for the chemical(s) responsible for the Tier 1 or Tier 2 categorization. Tier 1 and Tier 2 stations are located predominantly where data have been collected and compiled for the NSI, whereas fish consumption advisories are located in states with active fish advisory programs. Unlike the NSI data evaluation, which is applied consistently to available data, risk assessment methods used by states may vary.

Although there is good agreement for other chemicals, mercury is notably absent from the Tier 1 category in Table 5-1. Using the NSI evaluation methodology, mercury cannot place a sampling stations in Tier 1 for potential human health effects. For chemicals other than PCBs and dioxins, sediment chemistry and fish tissue

Table 5-1. Comparison of Contaminants Most Often Associated With Fish Consumption Advisories and Those Which Most Often Cause Stations to Be Placed in Tier 1 or Tier 2 Based on the NSI Data Evaluation

Chemical ^a	# of Water Bodies with Fish Advisories ^c	Number of River Reaches That Include at Least One Tier 1 or Tier 2 Station Based on the NSI Data Evaluation of Human Health Fish Consumption Advisories Parameters ^d		
		Tier 1	Tier 2 ^e	Total
Mercury	1,119	0	89	89
PCBs	387	1,498	732	2,230
Chlordane	114	11	1,026	1,037
Dioxins	53	242	8	250
DDT and metabolites	28	19	656	675
Dieldrin	15	9	1,296	1,305
Selenium	12	0	4	4
Mirex	10	0	15	15
PAHs	5	0	529	529
Toxaphene	4	0	183	183
Hexachlorobenzene	3	0	53	53
Lead	2	0	259	259
Hexachlorobutadiene	2	0	6	6
Creosote ^b	2	-	-	-
Chromium	1	0	6	6
Copper	1	0	4	4
Zinc	1	0	14	14

^aOther chemical groups responsible for fish consumption advisories (i.e., pesticides [24 water bodies], “multiple” [4 water bodies], “not specified” [4 water bodies], and metals [6 water bodies]) could not be directly compared to NSI chemicals.

^bNo reference values were available for creosote; therefore, it was not evaluated in the NSI data evaluation.

^cDoes not include statewide advisories

Mercury: New York, New Jersey, Maine, Massachusetts, Michigan, coastal Florida

Chlordane: Missouri

PCBs: New York

Dioxin: coastal Maine

^dA water body can be composed of numerous river reaches.

^eRiver reaches that include at least one Tier 2 sampling station but no Tier 1 sampling stations.



Figure 5-1. Tier 1 and Tier 2 Sampling Stations for Potential Risk to Human Health Located Within Water Bodies with Fish Consumption Advisories in Place for the Same Chemical Responsible for the Tier 1 or Tier 2 Classification.

data must both indicate human health risk for Tier 1 assignment. Unfortunately, the bioaccumulation potential of mercury based on concentrations in sediment cannot be assessed because the biota sediment accumulation factors (BSAFs) used for this study apply only to nonionic organic compounds. In addition, available fish tissue data for mercury did not place a large number of sampling stations in Tier 2 for potential human health effects, compared to the number of fish consumption advisories issued.

There are three possible explanations for the relatively small number of sampling stations categorized as Tier 2 for mercury in comparison to the number of fish consumption advisories in place for mercury. The first explanation is that the NSI evaluation was limited to data from resident demersal species, whereas data used in support of issuing state fish advisories probably included pelagic and migratory species. The second possible explanation is that the evaluation parameters used in the analysis were not as stringent as the ones used to support fish consumption advisory issuance. The third explanation is that the NSI does not include all of the data used by the states to issue fish advisories.

To examine these possible explanations, EPA performed additional analyses of mercury fish tissue data included in the NSI. The current evaluation, using a fish tissue screening value of 1 part per million (ppm), yields 103 Tier 2 sampling stations (4 percent of all stations with detectable levels). If data from all edible pelagic and migratory species are included in the analysis, there are 374 Tier 2 sampling stations (9 percent of all stations with detectable levels). A fish tissue threshold of 0.6 ppm, derived using the more stringent reference dose (0.00006 mg/kg-day) recommended to states for issuing fishing advisories to protect against developmental effects among infants (USEPA, 1994f), yields 821 Tier 2 sampling stations (20 percent of all stations with detectable levels) when applied to all edible species using the consumption rate for an average consumer of 6.5 grams per day. However, fish consumption advisories are often issued for more highly exposed populations, such as recreational or subsistence fishers. The 0.2 ppm Canadian guideline limit for mercury in fish that are part of a subsistence diet yields 2,308 Tier 2 sampling stations (56 percent of all stations with detectable levels) when applied to all edible species in the NSI database. Further details of the additional mercury analyses are provided in Appendix H.

The conclusion resulting from these additional analyses is that all three explanations for the discrepancy in numbers of fish advisories and Tier 1 and Tier 2 sampling stations for mercury probably have an effect. Most fish consumption advisories are issued to protect

infants from developmental effects for populations where exposure is greater than 6.5 grams of fish per day. It is also likely that many of the data used to develop state fish consumption advisories are not included in the NSI, or are not evaluated for sediment contamination because they are measurements in pelagic or migratory fish.

Sensitivity of Selected PCB Evaluation Parameters

Because PCBs and dioxin are extremely hydrophobic chemicals commonly associated with sediment, and because of their toxicity to humans, EPA believes that elevated levels of PCBs and dioxins in fish tissue of resident, demersal species are sufficient evidence to indicate a higher probability of adverse human health effects and to place a sampling station in Tier 1. Based on the NSI data evaluation, PCBs were responsible for the Tier 1 classification of more sampling stations than any other chemical. Therefore, EPA conducted a sensitivity analysis of some PCB evaluation parameters to determine the effect on the number of sampling stations classified as Tier 1 or Tier 2.

In the NSI evaluation, EPA selected a precautionary approach for the analysis of PCBs. The approach is precautionary because it does not require matching sediment chemistry and tissue residue data for PCB, and it is based on the risk of cancer for all PCBs congeners or total PCB measurements. However, some PCB congeners are considered a greater threat for noncancer effects than for cancer. The evaluation currently places 2,256 tissue sampling stations in Tier 1 based on human health cancer risk. Only 542 of these sampling stations included matching sediment and tissue data for PCBs. Therefore, the number of sampling stations classified as Tier 1 would have decreased significantly if this match had been required.

EPA performed additional evaluations to determine the number of sampling stations that exceed other screening values which are less precautionary than those selected for the PCB evaluation in this study. The complete results are presented in Appendix H, which includes a comparison of the number of sediment and fish tissue sampling stations with detectable levels of PCBs that exceed various evaluation parameters for both aquatic life and human health.

Sampling station evaluation based on PCB contamination is quite sensitive to the selection of evaluation parameters. For protection of fish consumers, there are essentially three distinct levels of protection. Using an EPA cancer risk of 10^{-5} (i.e., a 1 in 100,000 extra chance

of cancer over a lifetime of 70 years) or greater, 85 percent or more of the sampling stations with detectable PCB levels are classified as Tier 1. About one-half to two-thirds of the sampling stations are classified as Tier 1 for exceedances of PCB levels protective of noncancer health effects, cancer risk at a 10^{-4} risk level, or levels exceeding the wildlife criterion. Less than one-third of the stations are classified as Tier 1 using the FDA level of protection. As documented in Appendix H, these percentages vary depending on use of a BSAF safety factor, and whether one is examining the set of fish tissue data or sediment chemistry data. These three levels of protection vary within two orders of magnitude, a range that covers most of the distribution of PCB measurements.

Although sampling station classification for PCB contamination is quite sensitive to selection of evaluation parameters, overall station classification using the complete NSI evaluation for all chemicals is more robust. Using the selected PCB evaluation parameters, there are 15,922 total Tier 1 and Tier 2 sampling stations. If PCBs are dropped from the analysis entirely, the total number of Tier 1 and Tier 2 sampling stations remains about the same (less than a 5 percent decrease), but the number of Tier 1 sampling stations decreases by approximately 40 percent. If PCBs are evaluated using a noncancer human health threshold, the total number of Tier 1 and Tier 2 sampling stations decreases by less than 2 percent and the number of Tier 1 sampling stations decreases by approximately 12 percent. Figure 5-2 shows the location of Tier 1 and Tier 2 sampling stations that exhibit potential human health risks for all chemicals other than PCBs for comparison to Figure 3-6 in the results section. Approximately 78 percent (6,670 of 8,523) of the total number of Tier 1 and Tier 2 sampling stations indicating human health risk remain after excluding PCBs from the evaluation.

Strengths of the NSI Data Evaluation

For this report to Congress, EPA has compiled the most extensive data base of sediment quality information currently available in electronic format. To evaluate these data, EPA has applied sediment assessment techniques in a weight-of-evidence approach recommended by national experts. The process to produce this report to Congress has engaged a broad array of government, industry, academic, and professional experts and stakeholders in development and review stages. The evaluation approach utilizes sediment chemistry, tissue residue, and toxicity test results. The assessment tools employed in this analysis have been applied in North America with results published in peer reviewed literature. Toxicity test data were generated using established standard meth-

ods employed by multiple Federal agencies. The evaluation approach addresses potential impacts to both aquatic life and human health.

Because of the complex nature of the reactions among different chemicals in different sediment types, in water, and in tissues, no single sediment assessment technique can be used to adequately evaluate potential adverse effects from exposure to all contaminants. Uncertainties and limitations are associated with all sediment quality evaluation techniques. To compensate for those limitations, EPA has used multiple assessment techniques, alone and in combination, to evaluate the NSI data. For example, EPA developed draft SQCs based on the best scientific data available and extensive peer review. Therefore, EPA believes that the draft SQCs are reliable benchmarks for protecting sediment quality, and with measured TOC can indicate a higher probability for adverse effects to aquatic life. In addition, EPA believes that other sediment chemistry screening values (ERMs/ERLs, PELs/TELs, AETs, and SQALs) are also useful indicators of probability for aquatic life impacts. The Agency applied a weight-of-evidence approach for evaluating contaminant levels using these screening values, requiring the exceedance of multiple upper sediment chemistry screening values (i.e., ERM, PEL, AET-high, or SQAL) for classification of Tier 1 sampling stations.

The screening values used to evaluate the NSI data include both theoretical and correlative approaches. The theoretical approaches (e.g., draft SQCs, SQALs, and TBPs) are based on the best information available concerning how chemicals react in sediments and organisms and how organisms react to those chemicals. The correlative approaches (i.e., ERM/ERLs, PEL/TELs, and AETs) are based on matched sediment and biological data gathered in the field and in the laboratory, and they provide substantial evidence of actual biological effects from sediments contaminated with specific concentrations of the chemicals.

The NSI evaluation approach includes assessments of potential impacts to both human health and aquatic life. Some chemicals pose a greater risk to human health than to aquatic life; for others, the reverse is true. By evaluating both potential human health and aquatic life impacts, EPA has ensured that the most sensitive endpoint is used to assess environmental impacts.

Because sediment chemistry data are not the only indicators of potential environmental degradation due to sediment contamination, the NSI data evaluation approach also includes evaluations of fish tissue residue and toxicity data. If high levels of PCBs or dioxins (which are highly



Figure 5-2. Sampling Stations Classified as Tier 1 or Tier 2 for Potential Risk to Human Health Excluding PCBs.

hydrophobic organic chemicals commonly found associated with sediments) were measured in fish tissue at a given sampling station, the station could be categorized as Tier 1 with no corroborating sediment chemistry data. For other chemicals, high concentrations in tissues alone were not sufficient to categorize a sampling station as Tier 1; corroborating sediment chemistry data were also required. For a sampling stations to be categorized as Tier 1 based on toxicity data alone, multiple toxicity tests with positive results using two different test species were required. One of the tests had to be a solid-phase test.

Although EPA has developed draft SQCs for only five nonionic organic chemicals, the Agency has developed similar values, the SQALs, for an additional 35 chemicals as part of the NSI data evaluation. The SQALs have allowed EPA to evaluate more chemicals using multiple assessment techniques, thereby adding more weight of evidence to the results of this evaluation.

Limitations of the NSI Data Evaluation

This methodology was designed for the purpose of a screening-level assessment of sediment quality. A considerable amount of uncertainty is associated with the site-specific measures, assessment techniques, exposure scenarios, and default parameter selections. Therefore, the results of evaluating particular sampling stations based on this methodology should be followed up with more intensive assessment efforts, when appropriate (e.g., for water bodies with multiple Tier 1 sampling stations located in APCs). Two types of limitations are associated with the evaluation of the NSI data: limitations associated with the data themselves and limitations associated with the evaluation of the data.

Limitations of Data

The NSI is a multimedia compilation of environmental monitoring data obtained from a variety of sources, including state and federal government offices. Inherent in the diversity of data sources are contrasting monitoring objectives and scopes, which make comparison of data from different data sets difficult. For example, several of the databases contain only information from marine environments or other geographically focused areas. The potential for inconsistencies in measured concentrations of contaminants at different stations exists for samples taken from different monitoring programs. For example, sampling different age profiles in sediments, applying different sampling and analysis methods, and sampling for different objectives can affect the results of

the NSI evaluation. Although numerous data sets identified sampling and laboratory methods, most data did not have this information. In addition, some data sets included in the NSI were not peer-reviewed (i.e., Region 4's Sediment Quality Inventory, the Gulf of Mexico Program's Contaminated Sediment Inventory, and some data sets from EPA's STORET). Furthermore, each monitoring program used unique sampling and analysis protocols. For example, PCBs, the chemical group most often responsible for placing sites in Tier 1, were measured by nearly all of the programs but were analyzed and reported as aroclor-specific data, congener-specific data, total PCBs, or a combination of these.

The only quality assurance/quality control (QA/QC) information required for data to be included in the NSI was information on the source of the data and the location of the sampling station. Available information on several types of QA/QC procedures that can influence the quality of the data and can be used to check the quality of data was included in the NSI. None of this information, however, was required before a data set could be included in the NSI. Evaluation of such information can provide an indication of the quality of the data used to target a specific site. Table 5-2 presents a summary of the known QA/QC information associated with each of the data sets included in the NSI.

Data reporting was also inconsistent among the different data sources. Inconsistencies that required resolution included the lack or inconsistent use of Chemical Abstract Service (CAS) numbers, analyte names, species names, and other coding conventions, as well as the lack of detection limits and associated data qualifiers (remark codes). The evaluation of toxicity data required the presence of control data. Control data were not often initially reported with the data, and significant follow-up work was required to acquire such data. In addition, 4 of the 11 sources of toxicity test data used in the NSI evaluation did not report the use of laboratory replicates.

Some of the data included in the NSI were compiled as early as 1980 (the data cover the period of 1980-93) and might not reflect current conditions. The analysis did not include a temporal assessment of trends in sediment contaminant levels. Emissions of many prominent contaminants declined during the 1980s, and significant remediation efforts have taken place at many locations since that time. In addition, dredging, burial, and scouring might have removed contaminants from some sampling stations. The lack of a trend analysis in sediment contamination over time is an important limitation of this study and will be investigated in future NSI evaluations.

Table 5-2. National Sediment Inventory Database: Summary of QA/QC Information

Database	Are There QA/QC Reports to Accompany the Data?	Were the Data Peer-Reviewed?	Are the Sampling and Analytical Methods Identified in the Database?	Are the Detection Limits for the Analytes Included in the Database?	Comments
ODES	Yes	Yes, 301(h) data	Yes	Yes	Data Qualifiers
EMAP (VA and LA Provinces)	Yes	Yes	Yes	Yes	Data Qualifiers
Seattle; U.S. Army Corps of Engineers	Yes	Yes	Yes	Yes	Data Qualifiers
Region 4	Some	No	Some	Yes	Data Qualifiers
Gulf of Mexico	Some	No	Some	Yes	Data Qualifiers
COSED	Yes	Yes	Yes	Some	
Great Lakes	Yes	Yes	Yes	Yes	
DMATS	Some	Yes	Yes	Yes	Data Qualifiers
STORET	Unknown	Unknown	No	Yes	Data Qualifiers
Massachusetts Bay (USGS)	Some	Yes	Yes	Yes	

Some data parameters are consistently absent throughout the NSI database. (Refer to Appendix A, Tables A-1 and A-2, for information on the number of NSI stations at which the various types of data were collected.) For example, very few site-specific TOC or AVS data are available, and toxicity data or matched sediment chemistry and biological data were available at relatively few sampling stations. For many of the fish tissue data included in the NSI, the species was not identified.

The lack of AVS data in the NSI was a significant limitation for the evaluation of metals data. The NSI includes a relatively large amount of metals data, and the data indicate that metals concentrations in sediment are elevated in many areas. At some stations the elevated metals concentrations might indicate a potential problem; however, no sampling stations in the NSI could be placed in Tier 1 solely from measured concentrations of cadmium, copper, nickel, lead, or zinc. This reflects in large part the absence of AVS data, which are required to place sampling stations contaminated with those metals in Tier 1.

The unavailability of matching sediment chemistry and tissue residue data also limited the NSI data evaluation. In several instances, fish tissue was not analyzed for the same suite of chemicals for which sediment was analyzed. Spatial and temporal limitations of the data might have directly affected the analysis. Although some sediment chemistry and tissue residue data might have

been collected in the same or very similar sampling stations, if the station names were not identical, the data could not be treated as if they were collected from the same location. This very likely resulted in an underestimate of the number of Tier 1 stations identified based on potential human health effects. The underestimate occurred because exceedances of sediment TBP and tissue levels (EPA risk levels and FDA levels) at the same sampling station were required to categorize stations as Tier 1.

The lack of consistency among the different monitoring programs in the suite of chemicals analyzed also represents an area of uncertainty in the NSI data evaluation. Certain databases contain primarily information describing concentrations of metals or pesticides, whereas others (e.g., STORET and ODES) contain data describing concentrations of nearly every chemical monitored in all of the NSI data. Many monitoring programs use a screening list of chemicals that are indicator pollutants for contaminated sediments. Thus, many of the specific chemicals assessed in the NSI data evaluation are not always measured in samples. In addition, certain classes of in-place sediment contaminants might not be recognized as causing significant impacts and thus are not routinely measured.

Information describing local background levels of sediment contaminants was usually not presented with the data included in the NSI and thus was not considered when the significance of elevated contaminant concen-

trations in sediment was evaluated. Background conditions can be important in an evaluation of potential adverse effects on aquatic life because ecosystems can adapt to their ambient environmental conditions. For example, high metals concentrations in samples collected from a particular station might occur from natural geological conditions at that location, as opposed to the effects of human activities.

Most data are associated with a specific location. As a result, establishing the extent of contaminated sediment within a water body is not possible because it is difficult to assess the extent to which a monitoring station represents a larger segment of a water body. Furthermore, the NSI data are geographically biased. More than 50 percent of all sampling stations evaluated in the NSI are located in 8 states (Washington, Florida, Illinois, California, Virginia, Ohio, Massachusetts, and Wisconsin), which have more than 700 monitoring stations each. Finally, EPA did not verify reported latitude and longitude coordinates for each sampling station.

Limitations of Approach

Sediment Chemistry Screening Values

There are significant gaps in our knowledge concerning sediment-pollutant chemistry (especially bioavailability) and direct and indirect effects on aquatic biota. The certainty with which sediment toxicity can be predicted for each chemical using the various screening values included in the NSI evaluation can vary significantly based on the quality of the available data and the appropriateness of exposure assumptions. For example, draft SQCs and SQALs are not equivalent, even though they were developed using the same methodology. EPA has proposed SQCs for five chemicals based on the highest quality toxicity and octanol/water partitioning data, which have been reviewed extensively. The draft SQCs have also undergone extensive field validation experiments. However, SQALs for additional chemicals are in many cases based on a less extensive toxicity data set and have not been field validated. The AET values used in this evaluation were based on empirical data from Puget Sound. Direct application of values from Puget Sound to a specific location or region in another part of the country might be overprotective or underprotective of the resources in that area. Extensive collection of data and additional analyses would be required to develop AETs for other locations.

The bioavailability of metals in sediment is addressed by the comparison of the molar concentration of sulfide anions (i.e., acid-volatile sulfide [AVS]) to the molar con-

centration of metals (i.e., simultaneously extracted metals [SEM]). The [SEM]-[AVS] difference is most applicable as an indicator of when metals are not bioavailable. If [AVS] exceeds [SEM], there is sufficient binding capacity in the sediment to preclude metal bioavailability. However, if [SEM] exceeds [AVS], metals might be bioavailable or other nonmeasured phases might bind up the excess metals. To apply the [SEM]-[AVS] difference to indicate positive bioavailability and toxicity for this evaluation, EPA used laboratory data that indicated the probability of observed toxic effects at various [SEM]-[AVS] levels. Based on these data, EPA defined the Tier 1 level as [SEM]-[AVS]>5. Thus, this use of [SEM]-[AVS] represents a hybrid of a theoretical approach and a correlative approach.

Only those chemicals for which sediment chemistry screening values (i.e., draft SQCs, SQALs, ERLs/ERMs, PELs/TELs, and AETs) are available were evaluated in the analysis of NSI data. Therefore, the methodology could not identify contamination associated with chemical classes such as ionic organic compounds (e.g., alkyl phenols) and organometallic complexes (e.g., tributyl tin).

Biological effects correlation approaches such as ERLs or PELs are based on the evaluation of paired field and laboratory data to relate incidence of adverse biological effects to the dry-weight sediment concentration of a specific chemical at a particular sampling station. Researchers use these data sets to identify level-of-concern chemical concentrations based on the probability of observing adverse effects. Exceedance of the identified level-of-concern concentration is associated with a likelihood of adverse organism response, but it does not demonstrate that a particular chemical is solely responsible. In fact, a given sample typically contains a mixture of chemicals that contribute to observed adverse effects to some degree. Therefore, these correlative approaches tend to result in screening values that are lower than the theoretical draft SQCs and SQALs, which address the effects of a single contaminant. However, these correlative approaches are better at predicting toxicity in complex mixtures of contaminants in sediment. The effects range approaches to assessing sediment quality also do not account for such factors as organic matter content and AVS, which can mitigate the bioavailability and, therefore, the toxicity of contaminants in sediment.

Another concern is the application of screening values based on freshwater data (draft SQCs and SQALs) and those based on saltwater data alone (ERLs/ERMs, PELs/TELs, and AETs) to evaluate sediment contaminant concentrations in the NSI from both freshwater and saltwater habitats. Freshwater organisms exhibit tolerance to

toxic chemicals similar to that of saltwater species when tested in their respective water; however, estuarine organisms might be less tolerant if osmotically stressed (Rand and Petrocelli, 1985). Thus, the relative toxicity of a chemical in water (i.e., its chronic threshold water concentration) is usually within an order of magnitude for saltwater and freshwater species, although final chronic values and proposed sediment quality criteria values are usually slightly higher for saltwater species. Ingersoll et al., (1996) reported similar reliability and predictive ability between marine and freshwater guidelines. In addition Long et al., (1995) compared the ERLs and ERM with comparable values derived for freshwater by the Ontario Ministry of the Environment and the agreement was extremely good. Because of limitations of time and resources, sampling stations in the NSI were not classified by salinity regime, and further site-specific evaluations are required to more definitively assess the toxicity at the stations. However, the application of several different screening values should provide a reasonable estimate of probability of risk to aquatic life in freshwater, estuarine, and marine habitats.

Additional false positive and false negative classifications of risk to aquatic life from sediment contaminant concentrations could occur when a default value for organic carbon content is applied. Draft SQCs and SQALs are based on the partitioning of a chemical between organic carbon in the sediment and pore water at equilibrium. Because the organic carbon content of most sediment samples in the NSI is unknown, these sediment samples were assumed to contain 1 percent organic carbon. Total organic carbon (TOC) can range from 0.1 percent in sandy sediments to 1 to 4 percent in silty harbor sediments and 10 to 20 percent in navigation channel sediments (Clarke and McFarland, 1991). Long et al. (1995) reported an overall mean TOC concentration of 1.2 percent from data compiled from 350 publications for their biological effects database for sediments. Ingersoll et al. (1996) reported a mean TOC concentration of 2.7 percent with a 95 percent confidence interval of only 0.65 percent. In contrast, the concentration ranges of contaminants normalized to dry weight typically varied by several orders of magnitude. Therefore, normalizing dry-weight concentrations to a relatively narrow range of TOC concentrations had little influence on relative concentrations of contaminants among samples. Similar findings were reported by Barrick et al., (1988) for AETs and Long et al. (1995) for ERM calculated using sediment concentrations normalized to TOC concentrations.

Uncertainty associated with the equilibrium partitioning theory for developing draft SQCs and SQALs includes the degree to which the equilibrium partitioning model

explains the available sediment toxicity data (USEPA, 1993d). An analysis of variance using freshwater and saltwater organisms in water-only and sediment toxicity tests (using different sediments) was conducted to support development of the proposed sediment criteria. This analysis indicated that varying the exposure medium (i.e., water or sediment) resulted in an estimate of variability that should be used for computing confidence limits for the draft SQCs. The methodology used to derive the octanol/water partitioning coefficient and the final chronic value can also influence the degree of uncertainty associated with the draft SQCs. Differences in the response of water column and benthic organisms, and limitations in understanding the relationship of individual and population effects to community-level effects, have also been noted (Mancini and Plummer, 1994). Site-specific modifications to screening values derived using the equilibrium partitioning model have been recommended to better address chemical bioavailability and species sensitivities (USEPA, 1993b). Sediment chemistry screening values developed using the equilibrium partitioning approach also do not address possible synergistic, antagonistic, or additive effects of contaminants.

Based on the theoretical calculations used to compute SQAL values, it is possible that SQALs might be orders of magnitude larger or smaller than other screening values used for the analysis (ERLs/ERMs, PELs/TELs, and AETs). This might be a result of the limited aquatic toxicity data used to develop SQAL values for some of the contaminants for which water quality criteria are unavailable. EPA did not develop SQALs for this analysis in those cases where toxicity data were considered inadequate. The approach used to develop SQALs, and to choose chemicals for which SQALs could not be developed, is presented in Appendix B.

Fish Tissue Screening Values

The approach used to assess sediment chemistry data for the potential to accumulate in fish tissue also represents a theoretical approach with field-measured components. In addition to applying a site-specific or default organic carbon content, the TBP calculation includes a field-measured biota sediment accumulation factor (BSAF) to account for the relative affinity of a chemical for fish tissue lipids or sediment organic carbon. The BSAF will account for the effects of metabolism and biomagnification in the organism in which it is measured. The primary limitation of this approach is the applicability of a field-measured BSAF, or a percentile from a distribution of values, at a variety of sites where the conditions may vary.

TBPs were assumed to be equivalent to levels detectable in fish tissue. However, this approach might not completely account for biomagnification in the food chain, especially when using a BSAF derived from a benthic organism. In addition, it is assumed that sediment does not move, that contaminant sources other than sediment are negligible, that fish migration does not occur, and that exposure is consistent. The TBP calculation assumes that various lipids in different organisms and organic carbon in different sediments are similar and have distributional properties similar to the field-measured values used to derive BSAFs. Other simplifying assumptions are that chemicals are similarly exchanged between the sediments and tissues and that compounds behave alike, independent of site conditions other than organic carbon content. In reality, physical-chemical processes (e.g., diffusion through porous media and sediment mixing) can vary and limit the rate at which chemicals can exchange with bottom sediments. Uptake of contaminants by aquatic organisms is also a kinetic (rate-controlled) process that can vary and be slowed, for example, by awkward passage of a bulky molecule across biological membranes. Also, a BSAF of 1 (thermodynamic equilibrium) was used to estimate TBPs for many nonpolar organics. This BSAF might overestimate or underestimate the bioaccumulative potential for certain nonpolar organic chemicals because it is assumed that there is no metabolic degradation or biotransformation of such chemicals. Site-specific organic carbon content was often not available, which leads to additional uncertainty concerning the comparability of BSAFs among different locations. In addition, development of the BSAFs used in the TBP evaluation relied on a large amount of data that have not been published or peer-reviewed. Because of these factors, actual residue levels in fish resulting from direct and/or indirect exposure to contaminated sediment might be higher or lower. There is therefore uncertainty regarding sampling stations classifications based on comparison of estimated

TBPs with FDA tolerance/action and guideline levels and EPA risk levels.

TBPs could not be calculated for polar organic compounds or heavy metals. Therefore, sampling stations could not be classified using FDA levels or EPA risk levels for those chemicals using a TBP approach (although fish tissue monitoring data are often available for many stations).

Uncertainties and numerous assumptions are associated with exposure parameters and toxicity data used to derive EPA risk levels and FDA tolerance/action and guideline levels. For example, the derivation of EPA risk levels is based on the assumption that an individual consumes on average 6.5 g/day of fish caught from the same site over a 70-year period. Also, the TBP calculation for human health assessments assumes fish tissue contains 3 percent lipid. This value is intended to be indicative of the fillet rather than the whole body. Generally, the exposure assumptions and safety factors incorporated into toxicity assessments might overestimate risks to the general population associated with sediment contamination, but might underestimate risks to populations of subsistence or recreational fishers.

Other Limitations

Because a numerical score was not assigned to each sampling station to indicate the level of contamination associated with that station, it is not possible to determine which of the stations in Tier 1 should be considered the “most” contaminated. Such a numerical ranking system was intentionally not used for the NSI data evaluation because EPA does not believe that such ranking is appropriate for a screening-level analysis such as this, given the level of uncertainty.