

Recovery and validation of historical sediment quality data from coastal and estuarine areas: an integrated approach

Frank T. Manheim *, Marilyn R. Buchholtz ten Brink, Ellen L. Mecray

U.S. Geological Survey, Center for Coastal and Marine Geology, Woods Hole, MA 02543, USA

Accepted 22 June 1998

Abstract

A comprehensive database of sediment chemistry and environmental parameters has been compiled for Boston Harbor and Massachusetts Bay. This work illustrates methodologies for rescuing and validating sediment data from heterogeneous historical sources. It greatly expands spatial and temporal data coverage of estuarine and coastal sediments. The database contains about 3500 samples containing inorganic chemical, organic, texture and other environmental data dating from 1955 to 1994. Cooperation with local and federal agencies as well as universities was essential in locating and screening documents for the database. More than 80% of references utilized came from sources with limited distribution (gray literature). Task sharing was facilitated by a comprehensive and clearly defined data dictionary for sediments. It also served as a data entry template and flat file format for data processing and as a basis for interpretation and graphical illustration. Standard QA/QC protocols are usually inapplicable to historical sediment data. In this work outliers and data quality problems were identified by batch screening techniques that also provide visualizations of data relationships and geochemical affinities. No data were excluded, but qualifying comments warn users of problem data. For Boston Harbor, the proportion of irreparable or seriously questioned data was remarkably small (<5%), although concentration values for metals and organic contaminants spanned 3 orders of magnitude for many elements or compounds. Data from the historical database provide alternatives to dated cores for measuring changes in surficial sediment contamination level with time. The data indicate that spatial inhomogeneity in harbor environments can be large with respect to sediment-hosted contaminants. Boston Inner Harbor surficial sediments showed decreases in concentrations of Cu, Hg, and Zn of 40 to 60% over a 17-year period. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: geochemical databases; quality screening; contaminated harbor sediments; historical data; Boston Harbor; Massachusetts Bay

1. Introduction

This paper reports new methodologies to compile sediment quality databases from historical, heterogeneous data sources. The primary test area for this work was the Boston Harbor–Massachusetts Bay area (Fig. 1). Sediments in Boston Harbor are highly

contaminated, and some of its influence has spread into Massachusetts Bay, partly through transport by tidal and other current movements, and partly due to transport of dredged harbor sediments to offshore disposal sites. The development since 1985 of new treatment facilities and an offshore outfall, under the direction of the Massachusetts Water Resources Authority (MWRA), has been accompanied by additional sampling and monitoring (MWRA, 1997). The

* Corresponding author.

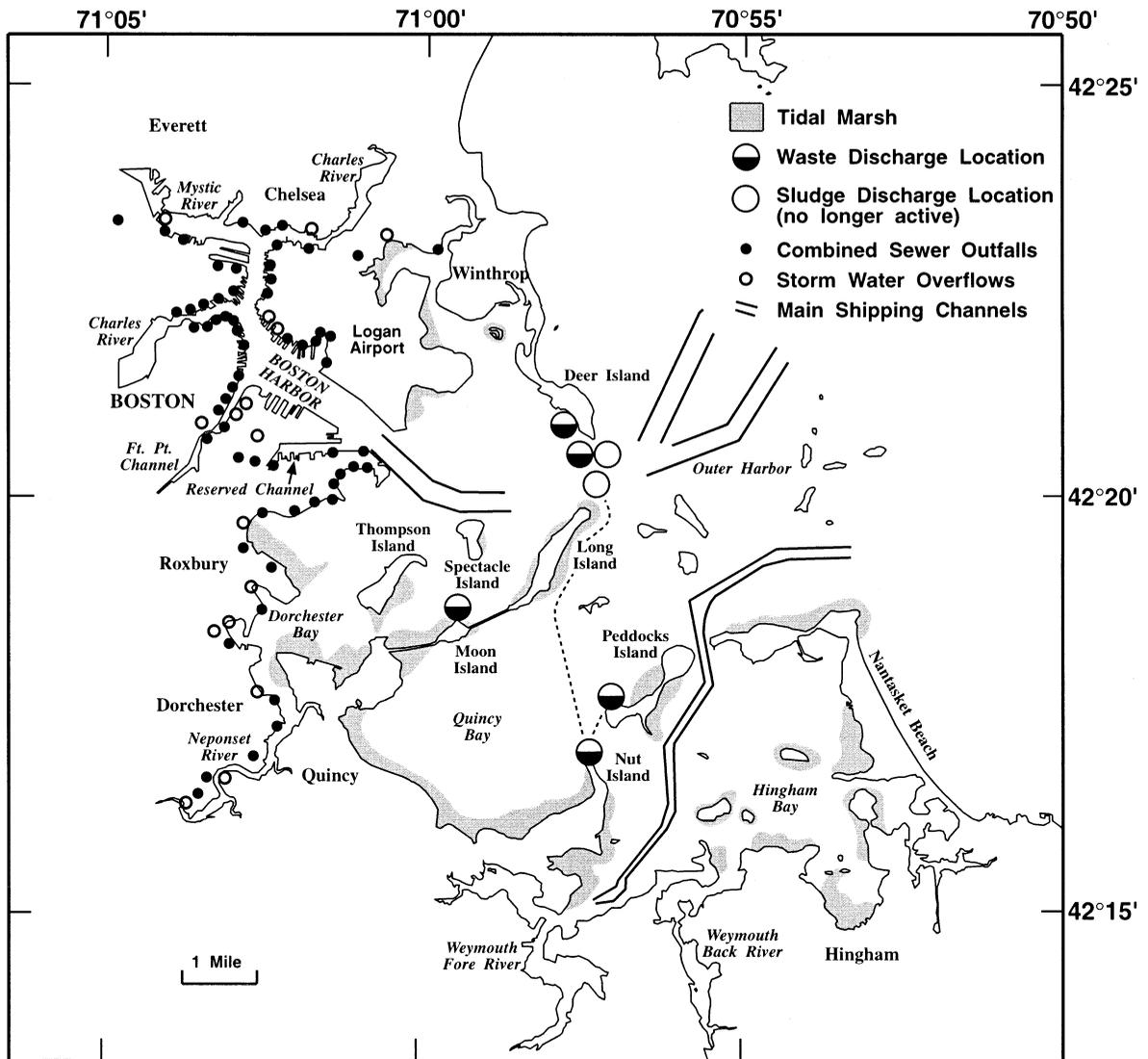


Fig. 1. Location map for Boston Harbor. Sewage effluent outfalls, combined sewer overflows (CSOs) and sediment sampling points.

current work was performed as a part of U.S. Geological Survey (USGS) cooperation with MWRA and other organizations on Boston Harbor contaminant impacts (Butman et al., 1992; Bothner et al., 1994).

In 1989 a National Research Council report (NRC, 1989) pointed out the scarcity of information about contamination in coastal and estuarine waters around the United States, but recommended against 'mapping' of contaminants in sediments. The reasons for this recommendation were not mentioned in

the report; however, the complexity of compiling heterogeneous data and uncertainties about data comparability and quality were known to be factors of concern. Over the past decades many sediment and other environmental surveys with increasingly standardized quality control protocols have been performed. However, historical data from diverse sources were rarely incorporated in local and regional databases until USEPA responded to a congressional mandate to create a nation-wide inventory of sediment quality (USEPA, 1996).

Notwithstanding the problems, justifications for compilation of quality-controlled heterogeneous data are strong. They include: (1) better definition and statistical control of the spatial distribution and heterogeneity of contaminants; (2) linkage of contaminants to data on bottom substrate type and morphology; (3) more informed decisions on when and where to use costly higher-tier measurement measurements (e.g. animal toxicity studies, acid-volatile sulfide and interstitial water analyses; USEPA, 1995 and references cited therein); (4) ability to use historical data as independent means of ascertaining changes in contaminant status with time; and (5) confirmation that concentrations of contaminants are low or within natural sediment variability (i.e. define not merely 'how dirty is dirty', but 'how clean is clean'). Definitive data provide more assurance to local populations than absence of information on presence of contamination.

At the onset of the work in Boston Harbor the total number of samples in an uncompleted database effort sponsored by USEPA for the area (Metcalf and Eddy, 1984; Mason, 1984) aggregated between 500 and 600 samples. This report cited 200 references. As a result of systematic collection, entry, and processing efforts outlined in this paper, we have been able to accumulate a total of about 3500 sediment samples from roughly 80 sources. The total number of literature citations in our reference list, which provides general environmental background data, is about 1600 (Scott, 1996; unpublished data). The techniques developed in this effort have since been applied to other coastal areas in ongoing studies, including the Gulf of Maine (Buchholtz ten Brink et al., database and manuscript in review; southern Louisiana (Manheim et al., 1997, and are in preparation for Long Island Sound and the New York Bight. Preliminary versions of results presented here were presented to workshops in preparation for a National Research Council panel report on contaminated sediments (NRC, 1997).

1.1. Principles

The main principles underlying the current work include: (1) cooperation among agencies and organizations to identify and compile data from diverse sources; (2) a comprehensive data dictionary; (3) use of widely accessible software and hardware; (4) data

entry and batch screening techniques; and (5) data validation and quality control procedures.

2. Methods and examples

2.1. Cooperation and sources of data

Cooperation through shared expertise and resources is essential to insure both a thorough exploitation of dispersed existing data and knowledgeable interpretation and effective use of the compiled data. Earlier stages of the sediment database work were reported by Manheim and Hathaway (1991) and resulted in data exchanges with or transfer to the U.S. Environmental Protection Agency (USEPA), Region I, USEP's National sediment Inventory (NSI), and the U.S. Army Corps of Engineers (New England Division) (Buchholtz ten Brink et al., 1992; Manheim et al., 1992; Manheim, 1993).

Our search for references found that about 80% of the environmental research literature compiled for Boston Harbor–Massachusetts Bay (Scott, 1996; database in revision) consists of 'gray literature'. The proportion of samples supplying raw data for our chemical database in this category is still higher. Gray literature is here defined as research data that are available in limited-distribution documents such as non-serial, in-house and consultant reports, and unpublished papers and electronic data files. Citations to such work are often not included in abstracting services or library catalogs or accessions. Theses and dissertations are an intermediate publication type. They are valuable because they may report extensive raw data and describe methodologies in detail. Limiting the search to electronic media would have excluded more than three quarters of our present data set.

2.2. Data dictionary

A comprehensive and clearly defined data dictionary for sediments and operational parameters provided the data structure and data entry template for most data processing. The data dictionary incorporates station data and sediment parameters from our and other large national sediment databases. In our analytical (secondary-user) database version, the

key referencing unit is the sediment sample, which is identified by a unique Reference ID number. These Reference ID numbers link data stored in six or more flat-file data tables containing: station data, inorganic data, organic chemical data (three files), and texture (grain size) data. Text fields provide descriptions or comments that define methods and quality-control parameters. Finally, supplementary tables contain information such as glossaries and organizational addresses. These data are meant to be self-explanatory insofar as possible. Comments relating to specific data are provided in fields adjacent to the data whenever possible to ensure proper use and transfer of the data.

The data dictionary in spreadsheet format, is suited for convenient hard-copy output. In 'vertical' format it lists field names vertically, with filenames in various lengths and full definitions provided in rows. For normal data entry and processing, the data dictionary is transposed into 'horizontal' mode. It can now serve as a data entry template and format for data processing and analysis. When data entry is completed, the flat file tables may be directly used for creating graphical and interpretive output. Examples of part of the station data and inorganic chemical data fields are provided in Tables 1 and 2. The data dictionary can be continuously updated (supplemented with new fields) at any time. Releases of the database are designed to always be accompanied by a current data dictionary.

2.3. Database structure

In computer database terms, operational structures (relational) are distinguished from analytical structures (flat file). Our data entry and processing, as well as output to graphical interpretation, are performed on flat files (spreadsheet software). The fields reflected in the data dictionary are designed to interface with the flat-file format. They contain redundancies, like alternative formats for location as latitude/longitude or state plane coordinates, and text. Redundancy is avoided in standard relational dbms's. It is deliberately used here in order that users can find fields in which to enter raw data in its original reported form and units. Qualifying fields are retained in the master tables to associate comments closely with data. These features expand

the total number of fields (which can be increased as needed) to nearly 1000, although only a small fraction of these is normally used in any data entry or analytical task. Given normal software field limits of 256 (Table 3), we have subdivided the total parameter (field) list into six linkable tables. A full bibliographic reference database is maintained in separate bibliographic software for searchability and transferability to desired editorial style.

The data structures used are similar to the flat file structures used in the NOAA National Standards and Trends program COSED data base (NST, 1994).

Working strategy has resemblances to the recently developed 'data mining' paradigm (Edelstein, 1997) for databases. In addition to permitting verification of existing hypotheses by querying databases, our databases are used to explore data relationships leading to new hypotheses, e.g. what internal relationships among concentrations of organic congeners are most consistent? Why? Unlike industrial 'data mining' systems, the operations proceed by manual rather than automated search operations. Data structures and procedures are therefore designed to make these operations as simple and convenient as possible.

2.3.1. Database management software utilization

Database management software (dbms), PARADOX™ and ACCESS™, was used to accommodate large databases (e.g. bathymetric data and large sediment sample files), and data queries and perform transformations that cannot be accomplished using spreadsheets. This software was not utilized in true relational mode. True relational databases isolate semantically related multiple relationships and eliminate redundancies. They are required for very large data sets, and provide great flexibility in data manipulation and querying. The cost is that data structures become complex and non-intuitive for most persons and the systems require professionally trained data management specialists.

Thus, for the present systems greater aggregation of relationships (grouping of spreadsheet fields, as in the case of data qualifiers) is utilized than in standard dbms operations. Although this involves limitations, including increasing the number of data tables, it keeps software simple, and maintains the coherence of data groups in ways suited to work tasks. It facilitates flexible use of text and comment

Table 1
Data dictionary, station data table

USGS row #	Short_field_name_10_characters_long	Medium length field name 25 characters long	Full length field name	Description of parameter and its fields
<i>Table of inorganic data</i>				
1	LOCAL_ID	Local Row or ID Number	Local Row or ID Number	Same as in Sample Header section = ID for use by user in maintaining sample order. This number can be changed by users.
2	UNIQUE_ID	Unique Sample ID (US#)	Cover-Id Unique Sample Identifier (US#)	Same as in Sample Header section = Database ID unique to this specific sample; assigned by USGS, cannot be changed, is used to link data between tables.
31	SRCE_OR_RF	Into Source or Reference	Source of Information, Reference, or Reference ID No.	Same as in Sample Header section = library reference or repository for hardcopy, or digital data.
85	LAB_INORG	Inorganics Testing Lab	Inorganics Testing Lab	Name or code for laboratory that performed the analysis for metals, see listing for abbreviations.
86	LAB_ID_NO	Laboratory's sample ID	Laboratory sample ID number	Laboratory's ID number indicating specific sample (for metals analysis).
87	LAB_JOB	Laboratory's JOB number	Laboratory's JOB number	Laboratory's ID number indicating job number or sample-tracking info. (for metals analysis).
88	ANAL_TECH	Analyt techn	Analytical technique	Method used for analysis of each metal; e.g., AA (flame, furnace, etc.), ICP, ICP-MS, MS; include code to methods' reference when given.
89	COMMENT1	Comments 1 (metals)	Analytical comments 1 (metals)	Any further information about analysis for all or specific metals.
90	COMMENT2	Comments 2 (other)	Analytical comments 2 (other inorganics)	Questions needing further investigations and any information about analysis that did not fit into previous comment field for all or specific metals.
91	REPL_NO	Rep no_of n (metals)	Replicate no_of n (metals)	Number in set of replicate analysis of one sample. Leave blank when no replicates.
92	TOT_REPL	Total # reps (metals)	Total replicates n (metals)	Total number of analysis in set of replicate analysis.
93	TEST_DATE	Testing Date	Testing date	Date of metals analysis by testing lab in 'mo/dy/yr'
94	TEST_MO	Test month	Test month	Month of metals analysis by testing lab.
95	TEST_DAY	Test day	Test day	Day of the month of metals analysis by testing lab.
96	TEST_YR	Test year	Test year	Year of metals analysis by testing lab.
108	AG_UG_G	Ag (silver) µg/g	Ag (silver) µg/g	Concentration of Silver (Ag) in the sample in units of µg/g. Do not enter detection limit values here.
109	AG_Q	Ag q	Ag qualifier	Any qualifier information or comments about the silver concentration; e.g. 'less than' (< or l); analytical trouble with this sample; corrections made during VALIDIS; indications of poor quality data; etc. original value and units if not µg/g
110	AG_DL	Ag det. lim.	Ag detection limit	The lowest detectable concentration of silver for this lab and this methodology.
111	AL_UG_G	Al (aluminum) µg/g	Al (aluminum) µg/g	Concentration of aluminum (Al) in the sample in units of µg/g (micrograms per gram). Do not enter detection limit values here.
112	AL_Q	Al q	Al qualifier	Any qualifier information or comments about the aluminum concentration; e.g. 'less than' (< or l); analytical trouble with this sample; corrections made during VALIDIS; indications of poor quality data; etc. Original value and units if not µg/g.
113	AL_DL	Al det. lim.	Al detection limit	The lowest detectable concentration of aluminum for this lab and this methodology.
114	AL_OU	Al original units	Al original units	The type of unit that the original aluminum concentration is recorded in.

Table 2
Data dictionary, inorganic chemistry table

USGS row #	Short_field_name_10_characters_long	Medium length field name 25 characters long	Full length field name	Description of parameter and its fields
<i>Table of station data</i>				
1	LOCAL_ID	Local row or ID No	Local Row or local ID No.	ID for use by user in maintaining sample order. This number can be changed by users.
2	UNIQUE_ID	Unique Sample ID (US#)	Unique Sample ID Identifier (US#)	Database ID unique to this specific sample; assigned by USGS, this is the common parameter for joining all tables in this database.
3	ORIGNL_ID	Preceding database ID	Preceding database sample ID	Database ID unique to this specific sample in the preceding database.
3.1	ORIGL_FILE	Preceding file name	Preceding database file name	Database file name from which this specific sample came.
4	REPLCT_NO	Repl. #	Replicate number	A number for a replicate where, for any parameter(s), more than one analysis have been made of the same sample. For the purpose of unambiguous joining of multiple tables, this number is incorporated in the unique sample identification number.
5	ST_PLANE_E	State Plane E	State Plane E	Easting coordinate in the State Plane System for the datum shown in Datum, Position.
6	ST_PLANE_N	State Plane N	State Plane N	Northing coordinate in the State Plane System for the datum shown in Datum, Position.
7	LATITUDE	Latitude (Decimal)	Latitude (decimal)	Latitude in decimal-degrees (south latitudes are negative).
8	LAT_DEG_N	Lat(Deg)N	Latitude (degrees) N	Degrees of latitude in whole degrees.
9	LAT_MIN_N	Lat(Min)N	Latitude (minutes) N	Minutes of latitude in whole or decimal units.
10	LAT_SEC_N	Lat(Sec)N	Latitude (seconds) N	Seconds of latitude in whole or decimal units.
11	LAT_ORIG	Lat Orig. Position Format	Latitude, original position format and precision	Format and precision of originally recorded latitude (indicate format and significant digits, e.g. ddimms.s; dd.dddd; ffff).
12	LONGITUDE	Longitude (Dec.) (- = W)	Longitude (decimal) (- = West long.)	Longitude in decimal-degrees (West longitude is depicted by negative values).
13	LON_DEG_W	Lon(Deg)W	Longitude (degrees) W	Degrees of longitude (West) in whole degrees.
14	LON_MIN_W	Lon(Min)W	Longitude (minutes) W	Minutes of longitude (West) in whole or decimal minutes.
15	LON_SEC_W	Lon(Sec)W	Longitude (seconds) W	Seconds of longitude (West) in whole or decimal units.
16	LON_ORIG	Lon Orig. Position Format	Longitude, original position format and precision	Format and precision of originally recorded longitude (west), e.g. ddimms.s; dd.dddd; fffff.
17	ORIG_LOC	Orig. Locn. If Unusual	Original location in original format, if unusual	Original location if given in units other than latitude and longitude (e.g., State Plane or distance from a point).
18	DATUM_POS	Datum, Pos.	Datum, position	North American Datum (NAD) for the State Plane position coordinates given.
19	DATUM_ELEV	Datum, Elev.	Datum, elevation	Datum standard for determining lat./lon. from location map.
20	NAV_MODE	Navigationl Mode	Navigationl mode	Navigationl system used, e.g. LORAN C, GPS, triangulation, read from a sketch map).
21	TIME_DELAY1	Time Delay 1	Time Delay 1	Navigationl time delay (first value) used to calculate position from signals.
22	TIME_DELAY2	Time Delay 2	Time Delay 2	Navigationl time delay (second value) used to calculate position from signals.
23	SOUNDING_M	Sounding (M)	Sounding (meters)	Measured depth of water overlying sediment at sample time, in meters.
24	SNDNG_ORIG	Sounding (in Orig. Units)	Sounding (in original units, if not meters)	Measured depth of water overlying sediment at sample time, in original units.
25	SNDG_Units	Sounding Original Units	Sounding original units	Depth units (meters, feet, fathoms, etc.).
26	AGNCL_SPON	Agency1 (Sponsoring)	Agency1 (Sponsoring)	Agency or researcher sponsoring or publishing the work, see listing for abbreviations.

Table 3
Software limitations for desktop data management software

Type	Number of records/table	Characters per record	No. fields	Characters per fieldname	Characters per field
Dbms	2 billion	65,000	256	10 ^a –54	256 ^b
Spreadsheet	8,000–65,550 ^c	65,000–130,000	256–512 ^d	256	256

^a Older versions of dBase™. ^b Except memo fields, which have limited interactive capacity. ^c EXCEL™ has increased from 16,000 to current 65,550 records. ^d LOTUS 123™ has a 512 field capacity. Alpha Five has a 128 field capacity.

fields, meaningful field headings, and minimizes the risk that data quality flags (critical for heterogeneous data sets) become separated from the analytical data.

2.4. Data entry and screening

References having data to be entered were pre-screened by a scientist or environmental specialist, who noted the nature and quality of data to be entered and resolved problems about data definitions, gaps in coverage, units, sampling or analytical issues. After checking compatibility of field definitions and possible sample redundancy, data were encoded as fully as the completeness of original data sources allowed. When electronic data were involved, comparability of data fields and software data transfer formats was confirmed. The data dictionary mentioned previously provides comment fields for detection limit and data quality, which includes notes on nonstandard units. Fields are also provided for extended comments on data quality or documentation for each table and for special parameters (e.g. location methodology). References are entered in bibliographic software format, which allows output to be made in a variety of editorial styles, and makes electronic searches on annotations possible.

Original (raw) data are preserved wherever available. Transformations (e.g. location to decimal degrees) are provided in separate fields. This minimizes artifacts of conversion, and often provides useful information on the way data were prepared (e.g. rounding conventions).

Data processing typically goes on concurrently with further data entry. Checks on data consistency, sample designation, location and other station data may reveal clerical problems that can be corrected. Missing or incorrect locations are one of the more frequent problems or gaps in data. The most useful format for plotting locations is decimal latitudes

and longitudes. Often, however, locations are given in degrees, minutes and seconds, state plane coordinates, or only in the form of location plots. The latter are converted to decimal degrees using standard conversion software. Location sketches are converted by optical scale matching equipment and standard base maps like USGS topographic maps or NOAA harbor charts, computer digitization, or manual equivalents. Data lacking specific locations are entered because these may be found in other documents, or more generally identified data may have value for characterizing larger data sets in spite of the absence of specific locations (e.g. frequency plots such as Figs. 2 and 3). Once data are entered from the reference they undergo quality/validation checks as summarized below.

2.5. Data validation and quality control

Data quality problems were identified by batch screening techniques since standard quality-assurance (QA/QC) protocols currently used for new sampling and laboratory analyses by federal and state agencies and firms that consult for them (Baker and Kravitz, 1992; USEPA, 1995) are usually inapplicable to historical data. Techniques included screening for outliers and comparison of data from diverse origins.

Early work in Boston Harbor demonstrated that common problems like sample identification, errors in unit designations, and gaps in location data were repairable (Manheim and Hathaway, 1991). Moreover, problems of analytical origin tended to occur in batches. Our quality control effort has therefore been directed toward establishing practical methods for identifying data that fall outside normal ranges of sample heterogeneity and analytical variability, i.e. deviate from internal relationships characteristic for well-controlled data. We use batch screening

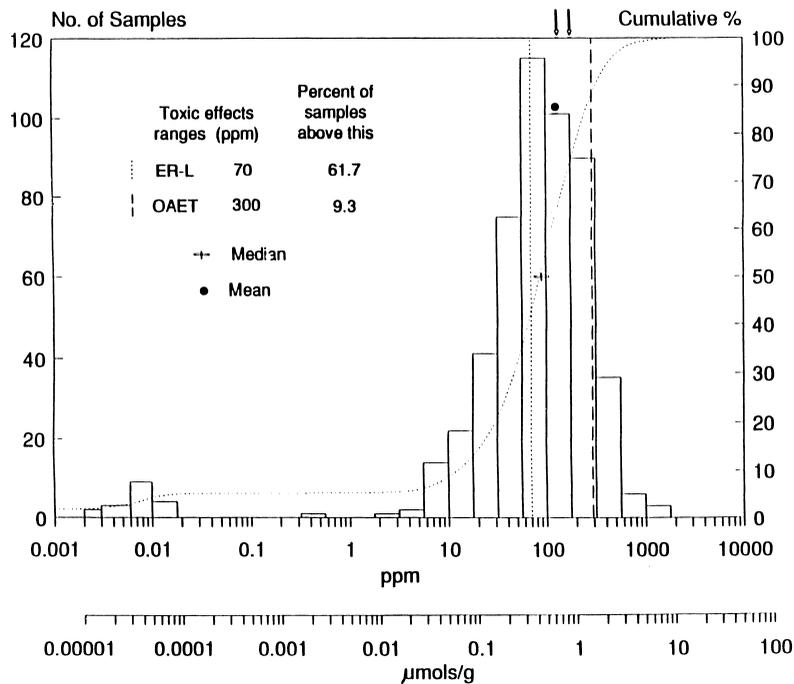


Fig. 2. Frequency plot of raw copper values in 'Boston Harbor Data Management File' (Manheim and Hathaway, 1991). Arrows along the top of the diagram show the values for the element at NS&T monitoring sites in Boston Harbor. Curved dotted line is cumulative distribution. Dashed and dotted vertical lines refer to NS&T apparent toxicity thresholds utilized by Long and Morgan (1990).

techniques referred to generally as a VALIDS procedure (Validation by Interactive Data Screening) to accomplish this task.

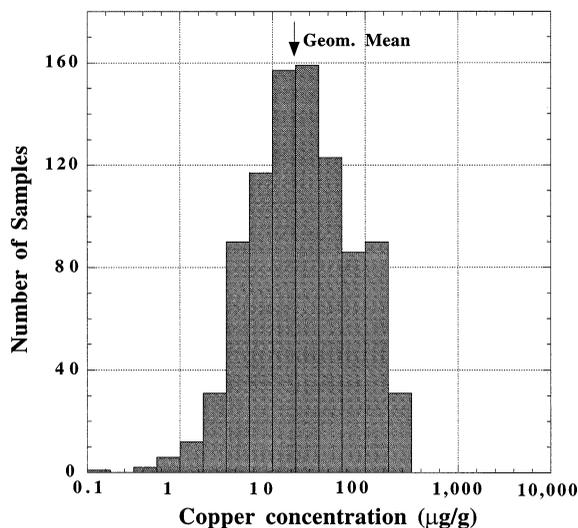


Fig. 3. Frequency plot of copper values in NOAA NS&T benthic surveillance stations (NST, 1996).

We have chosen to retain all raw data in databases, regardless of problems associated with given data sets. Exclusion would involve arbitrary choices about data acceptability and uses. It might create unwarranted assumptions about the quality of data retained. Exclusion of seriously flawed data already available in the public record might result in other workers utilizing them without realizing that problems had already been encountered with them. We make an effort to retain data as originally reported and provide additional comments for the user when the data are repaired or identified as having questionable quality.

3. Examples of display of geochemical parameters for quality validation

3.1. Frequency plots

Trace elements in naturally occurring sediments and rocks have long been known to follow log-normal distributions of occurrence vs. concentration (Ahrens, 1954). Anthropogenic contaminant concen-

trations in larger national data sets for sediments have likewise been observed to follow lognormal distributions (e.g. O'Connor, 1990). Histogram plots provide helpful delineation of the frequency of analytical data and reveal outlier groupings, and in some cases, problems in reporting of units (organic constituents). Points within these populations can be inspected to find explanations for the behaviors. Fig. 2 shows a plot of raw data for copper in sediments, processed from the earliest electronic data file for Boston Harbor (Manheim and Hathaway, 1991). This demonstrates the typical lognormal distribution, along with small subsidiary modes.

The small aberrant population on the tail toward lower values was found to mainly represent elutriate or biological samples mislabeled as sediments. The NOAA National Standards and Trends (NS&T) data set (NST, 1996) plotted here for Cu (Fig. 3) also shows lognormal character, but has a much lower mean than the Boston Harbor data set. This data set was obtained from the Benthic Surveillance project, covering estuaries, harbors and coastal areas around the United States, and was chosen as a representative data set to demonstrate the general distribution of contaminant ratios.

3.2. Concentration range screening

Many problems with data quality can be found by inspection of samples having contaminant concentrations in the high and low ends of the range. Table 4 gives an example of such reference levels. Others could be used as appropriate to the area and constituents in question. The expected maximum and minimum data are mainly provided to provide guidelines to detect possible 'unreasonable' values due to correctable errors in unit designations. Table 5 shows a set of Boston Harbor sediments sorted on Hg in descending order, along with other metals. Values exceeding five times the 'ERM' toxicity screening values for metal toxicity in sediments (Long et al., 1995) are shaded. For historical data, supplementary laboratory or other tests for toxicity relationships pertinent to benthic organisms are not possible. However, bulk-sediment screening criteria or quality guidelines associated with concentration levels of contaminants have been found to offer conservative indexes of toxicity level.

Table 4

Concentration range guidelines for metals in bulk dry sediments

Element	Soil	SQAG (ERL)	SQAG (ERM)	Minimum range	Maximum range	Ref.
Ag	0.05	1	3.7	0.02		
Al	71000			10000	100000	N
As	6	8.2	70	1		N
B	20			5	150	N
Ba	500			50	1000	N
Be	0.3			0.1	10	N
Ca	15000			2000	100000	N
Cd	0.35	1.2	9.6	0.04		
Co	8			0.5	100	N
Cr	70	81	370	4		
Cu	30	34	270	1		
Fe	40000			2000		N
Hg	0.06	0.15	0.71	0.01	0.71	
K	14000			1000	40000	N
Mg	5000			2000	50000	N
Mo	1.2			0.5	25	N
Na	5000			3000	50000	N
Ni	50	20.9	51.6	3		
P	800			200	30000	N
Pb	35	46.7	218	2		
S	700			300	30000	N
Sb	1			0.1	25	NST
Se	0.4			0.05	5	NST
Si	33000			10000	480000	N
	0					
Tl	0.2			0.05	20	NST
U	2			0.5	10	NST
V	90			5	350	NST
Zn	90	150	410	5		
Corg	20000			1000	50000	N
Norg (Kj)	2000			200	5000	N

All values in are in $\mu\text{g/g}$.

Average for soil data (natural) is taken as most closely resembling uncontaminated estuarine sediments (from Bowen, 1979).

SQAG refers to Sediment Quality Assessment Guidelines (Long et al., 1995).

ERL refers to Toxic Effects-Range Low, from Long et al., 1995.

ERM refers to Toxic Effects-Range Medium, from Long et al., 1995.

N refers to references for maximum–minimum ranges range for normally non-toxic elements in natural sediments, based on standard geochemical summaries (Wedepohl, 1978; Bowen, 1979).

NST refers to a 'high' range for potentially toxic elements for which no SQAG or equivalent values are available, estimated from data in NOAA NST sediment monitoring sample database (NST, 1996).

The five highest (extreme) Hg values have concentration levels more than twenty times higher than the next highest data point. These values came from samples having other metal values that were *below*

Table 5

Selected metal concentrations in the 1993 Boston Harbor sediment database, sorted on Hg in descending order. Shading denotes samples whose concentrations are more than five times the ERM toxicity level (column 4 in Table 4). Heavy borders delineate a Hg value believed to be erroneously carried down by the citing source from a Cu value in the original data resource. Light shading refers to values below the expected minimum range shown in Table 4

UNIQUE_ID	LATITUDE	LONGITUDE	Source Code	CR_UG_G	CU_UG_G	HG_UG_G	PB_UG_G	ZN_UG_G
US00008	42.385	-71.04611	1	0.83	80	5760	1200	465
US00006	42.385	-71.04611	1	1.9	260	3160	400	29.5
US00003	42.385	-71.04611	1	1.2	200	2350	505	7.5
US00004	42.385	-71.04611	1	2.4	250	1960	1100	86250
US00009	42.385	-71.04611	1	1.4	47	1140	450	175
US02736	42.38416	-71.04722	81	57.8	68.8	68.8	151.2	264
US01992	42.3675	-71.04361	57	250	460	26	400	800
US02008	42.38666	-71.04027	57	97	1650	23	2000	7200
US01983	42.35833	-71.05	57	120	480	20	230	1100
US00306	42.35958	-71.04805	26	68	408	14	805	59.6
US07693				90	64	14	85	190
US00318	42.36072	-71.049	26	28	75	12	244	283
US02705	42.34472	-70.96527	81	136	75	12	77	210
US02149	42.30833	-70.97638	67	302	192	9.4	234	242
US02714	42.35277	-71.05138	81	195	218.2	9.4	500	595
US00308	42.3593	-71.04905	26	44	627	8	1244	2156
US02687	42.31805	-71.02166		240	110	8	142.5	190
US00307	42.35941	-71.04894	26	64	173	7	524	380
US02069				451	338	6.8	222	410
US02159	42.34305	-71.01388	67	133	81	6.8	84	160
US02162	42.35833	-70.98472	67	134	292	6.8	84	160
US01507	42.30722	-71.03694	48	128	108	6.7	117	183
US02068				653	649	6.6	337	583
US01982	42.3575	-71.05	57	250	390	6.5	310	950
US02163	42.36527	-71.04861	67	418	339	6.3	402	506
US02153-1	42.32361	-71.00555	67	292	239	6.1	215	317
US02157	42.33472	-70.95972	67	255	239	6.1	215	317

levels expected for naturally occurring (pristine) sediments in this area. One sample from this set also reported grossly anomalous Zn concentrations, without supporting evidence. It was clear from these and other indications that all metal data from the laboratory that had reported out these data should be flagged with warnings.

The sixth-highest Hg value (68.8 $\mu\text{g/g}$) suffered from another type of problem that occasionally occurs in secondary sources. This anomalous Hg value was carried down by mistake in a secondary compilation from a Cu value identified as an entry in the original data set. The originating laboratory reported a value of 1.2 $\mu\text{g/g}$ in its primary records, in agreement with other well-controlled data for the area.

3.2.1. Outliers

Some comment is needed on range data with respect to outlier removal. Range data may be useful to remove outliers for determination of means and other statistical distributions. However, they are not sufficient to identify outliers related to data quality problems. A common range method to delineate as outliers utilizes the average less three times the standard deviation, or the average plus three times the standard

deviation. The spread of the Boston Harbor data set is so large that the lower limit for all metals would extend to negative numbers. On the high side the $3 \times \text{SD}$ approach would fail to detect some problem data discussed in the next section but would discriminate against high-concentration data points deemed to be valid. Approaches using the geometric mean and deviation would be better but still yield unacceptable results. Multivariate statistics that take into account anomalies in more than one constituent are sometimes used to delineate outlier *samples*, rather than data. However, the spread of the data again tend to defeat such approaches for Boston Harbor sediments, and those in other impacted harbor environments. In short, establishing ranges is an important first step in evaluating the distribution of data, especially those in high (potentially toxic) or anomalously low concentration levels. But statistical measurements relating to ranges are insufficient guides to problem data.

3.3. Ratio plotting and use of standard data sets: zinc normalization

Contaminants often occur together because of common sources or chemical behaviors. Deviations

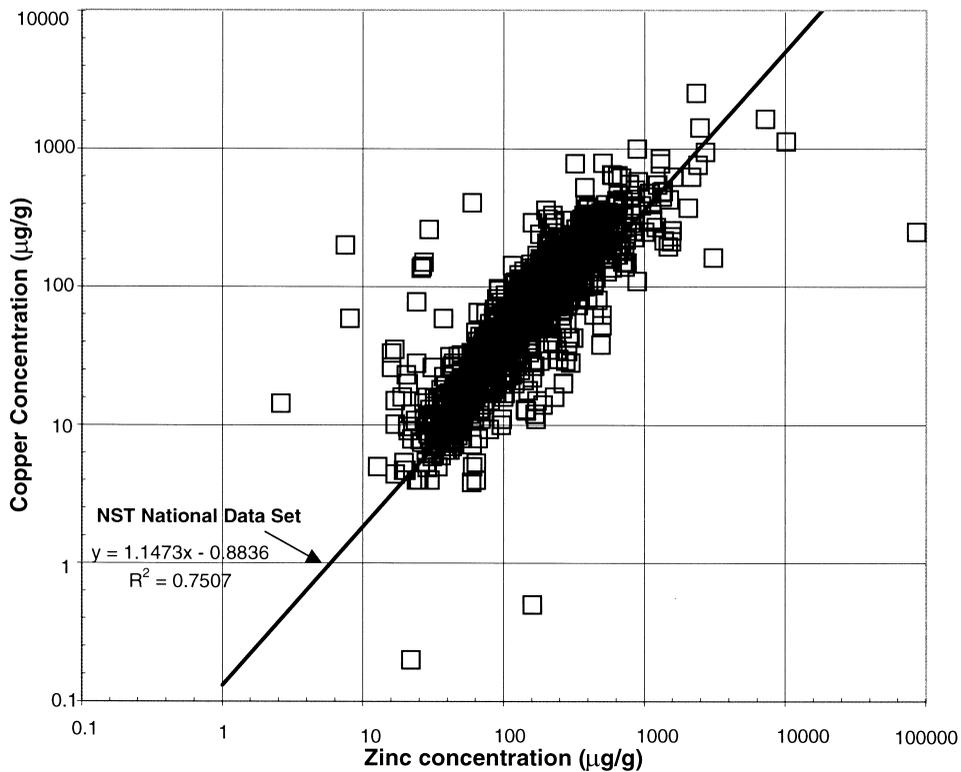


Fig. 4. Plot of Cu against zinc for the Boston Harbor–Massachusetts Bay data set. The regression line is for the NST data set (NST, 1994).

from typical associations can be used to further identify data quality. Plotting metals against Zn concentrations takes advantage of the fact that interelement ratios in contaminated sediments show much greater coherence than do the absolute concentrations of the components. As further elaborated in work in preparation, zinc is a suitable element for normalization, because it is almost always included in contaminant data sets, often displays the largest concentrations among the metals, and is a nearly ubiquitous marker wherever anthropogenic influences are present in waterways. As may be seen in Fig. 4, plotting Cu against Zn reveals a consistent relationship on a log/log plot. The utility of such plots is increased by adding the regression line for large standardized data sets as a normalizing tool. Regressions for the NS&T sediment data set (NST, 1996), compiled from hundreds of benthic monitoring stations, are incorporated in Figs. 4 and 5. Close agreement with the NS&T regression line is not necessarily expected, but samples and groups of samples hav-

ing a common source, which show large departures from local trends (e.g. more than a factor of about two) call for closer examination. If no factors or sources explaining the aberrant behavior are found, in conjunction with well-qualified data from similar environments, samples may be flagged as requiring further confirmation or potentially questionable.

Outlier points on such plots may be examined to determine if they are due to regional variations in contaminants, or due to systematic errors in analysis. Fig. 5, for Hg and Zn, shows much greater scatter than does Cu/Zn. This is attributable, among other reasons, to the lower concentrations of Hg, and greater diversity of sources than for Cu and Zn. The aberrant data Hg data depicted in Table 5 may be seen scattered near the top of the foregoing figure.

3.4. Organic contaminants

Log/log plots, as employed for metals, above, are critical to perform validation controls on or-

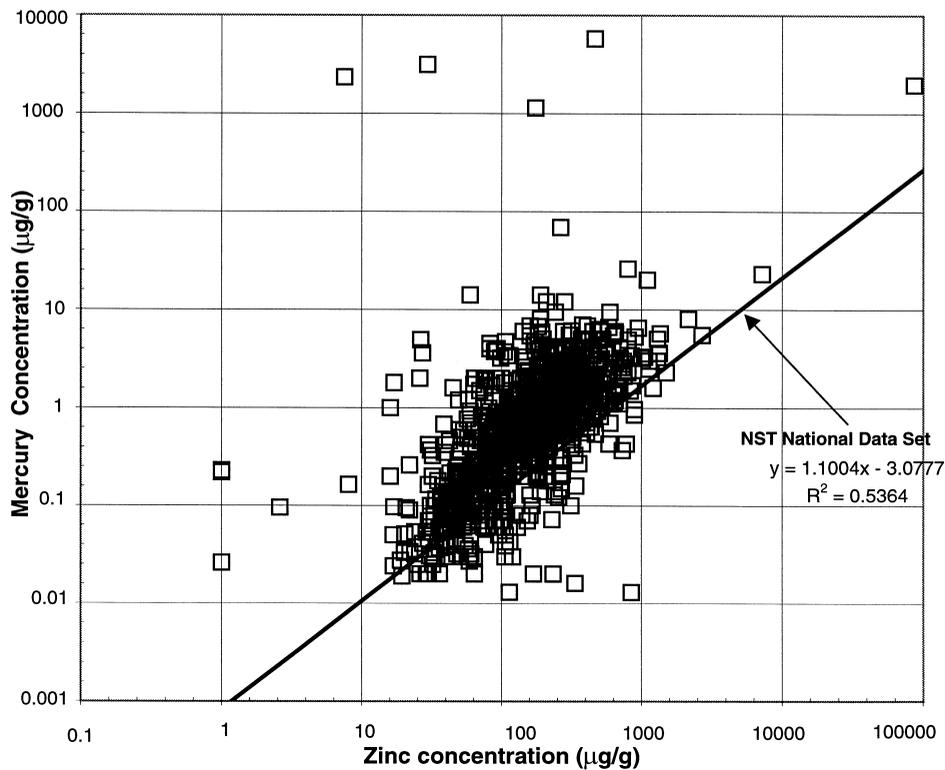


Fig. 5. Plot of Hg against zinc for the Boston Harbor–Massachusetts Bay database. The regression line is for the NST data set.

ganic contaminants. Fig. 6 shows a plot of polychlorinated biphenyl congeners for Boston Harbor–Massachusetts Bay samples (1994 cutoff date) for all data sources (656 samples). Of the 47 sources 25 were U.S. Army Corps of Engineers permit files with only a few data each. This plot of raw data reveals several types of information: (1) there is predictable increase in scatter toward lower values as data approach detection limits; (2) linear rows of data points in the vertical and horizontal directions for certain data sets signify detection limits that were reported as measured values; (3) ‘Total PCB’ (as reported) values tend to show more scatter than congeners for given geographical areas. Some Total PCB data plotted three orders of magnitude below the present regression lines for the three congeners because of incorrect unit designations. Points identified visually on a spreadsheet plot can be automatically identified with respect to data series and record number by clicking with the mouse on recent editions of spreadsheet software.

Mix-ups can occur because total PCB is frequently listed as $\mu\text{g/g}$ (ppm), whereas congeners are normally listed as ng/g (ppb), a 1000-fold concentration difference. Plots thus provide quick checks on unit consistency. Total PCB values reported from earlier dates are expected to have poorer correlation because of uncertainties relating to the compositions measured and the way these were computed.

Fig. 7 plots the PCB 101 and PCB 153 pair, using different symbols for five randomly chosen, well-qualified sources of data. The consistent agreement, except at very low values, can be partly attributed to the fact that extended organic congener analysis has become more widely available in the last ten years. Laboratories undertaking analysis of complex mixtures of specific congeners generally have considerable expertise and tend to follow consistent QA/QC procedures. See, however, caution regarding regionally variable source factors, below.

Similar plots of homologues for the polyaromatic hydrocarbons are shown for the high molecular

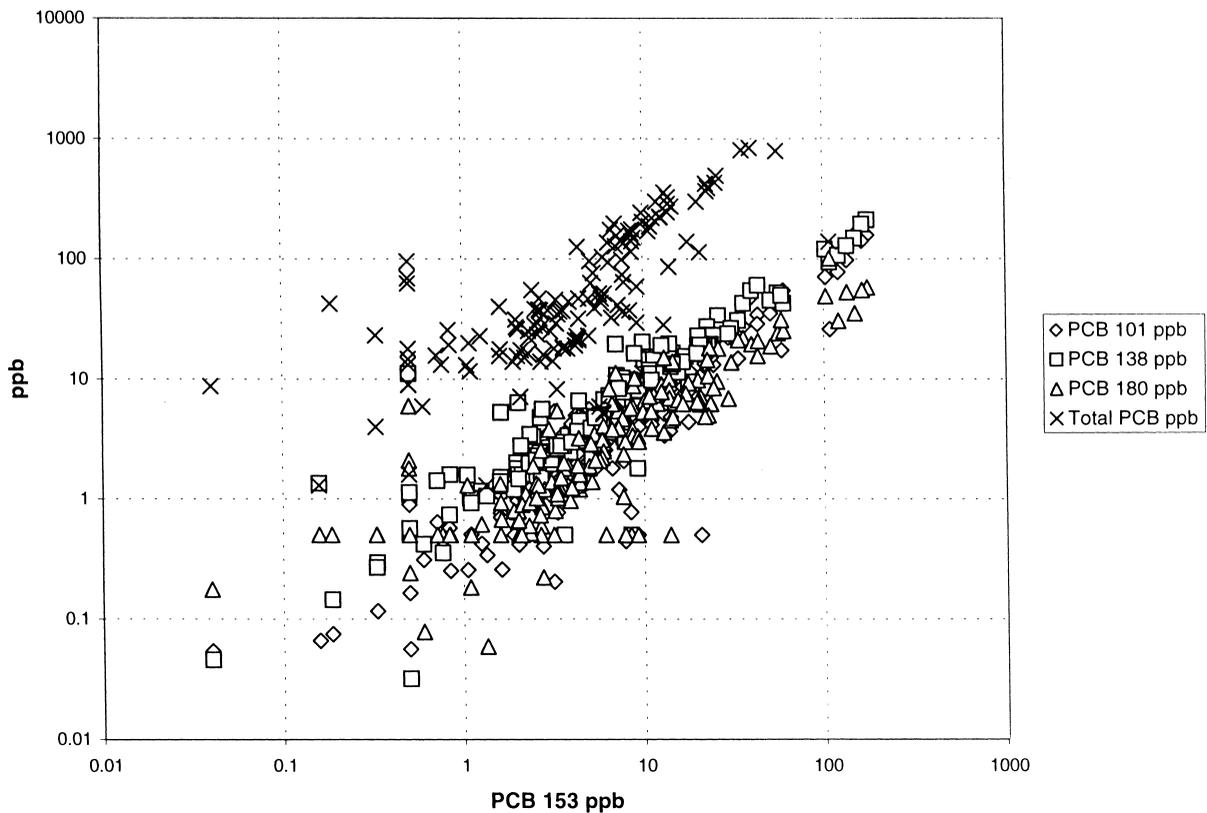


Fig. 6. Plot of PCB congeners for Boston Harbor–Massachusetts Bay data, all sources.

weight hydrocarbons (Fig. 8). Although phenanthrene is a 3-ring compound that nominally belongs to the light hydrocarbon group according to the NST (1994) delineation, it is closer to the heavier hydrocarbons in its behavior during uptake and weathering on sediments. It is particularly suitable as a normalizing component because it is often the most frequently observed PAH homologue.

Plotting for analytical consistency is ideally accomplished by using samples from similar areas, including samples from well-controlled laboratories. Once questionable data are identified, plotting may be useful for other diagnostic purposes, such as exploring questions of contaminant sources.

Plots of pesticide derivatives follow similar patterns. For example DDE 4,4' and DDD 4,4' are breakdown products of DDT, and much more widely detected than the latter in general sediments. They tend to show similar transport and weathering behaviors. These derivatives normally show good coher-

ence in data from well-controlled data sources, and are therefore suited to assess analytical consistency among different sources of data. On the other hand, the chlordane group of pesticides remained in production well after DDT was prohibited in 1972. It may have a different time-distribution pattern as well as chemical affinities from DDT. In like fashion, one would first seek closely related pesticide species for data quality checking before utilizing ratios or other display of data that excludes questioned samples to explore for environmentally significant distributions of various pesticides.

3.5. Texture normalization

Normalization of data based on sediment grain size, carbon content, Al or Fe content or surface area has been frequently employed to reduce the effect of variable dilution by detrital minerals (Horowitz, 1991). We have not applied texture normalization

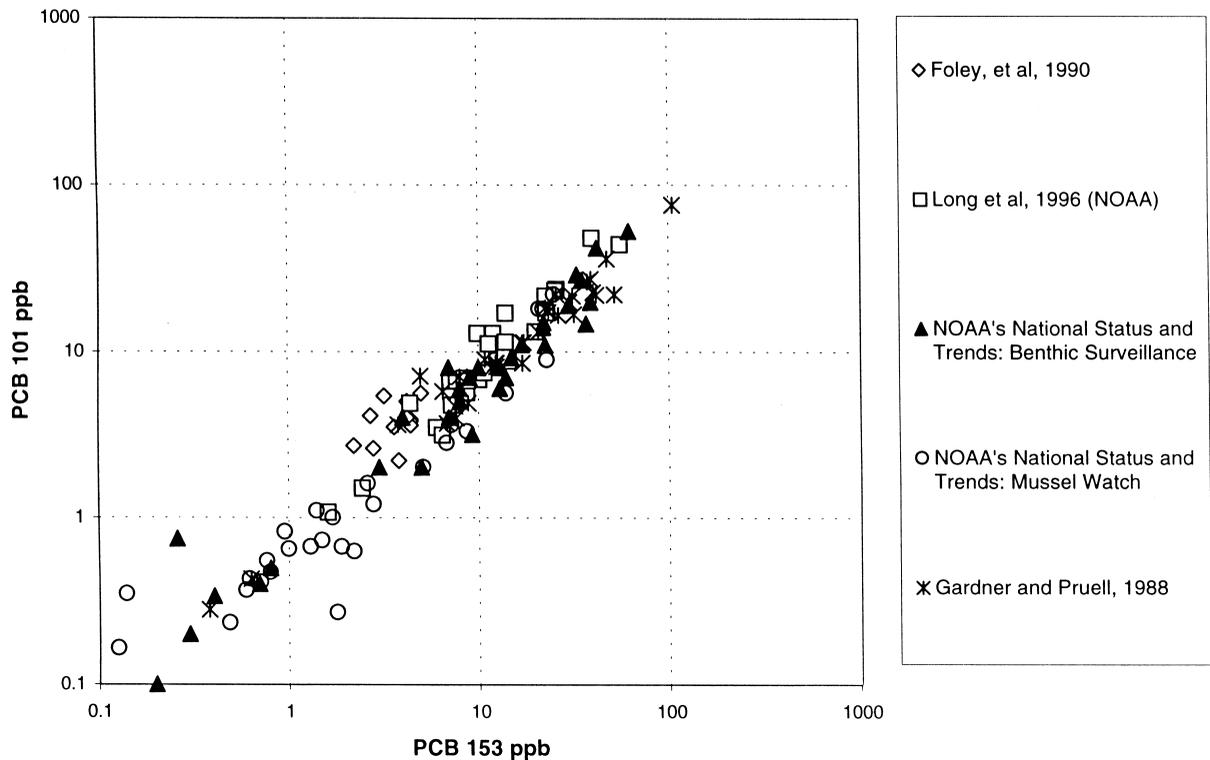


Fig. 7. Plot of PCB congeners 101 and 153 analyzed by five arbitrarily selected, well-controlled data sources: Foley et al., 1990 (diamonds); Long et al., 1996 (squares); NST, 1996 (triangles and circles); Gardner and Pruell, 1988 (asterisks).

to the sediments for several reasons. First, many of the heterogeneous data sets contained the only key pollutants and lacked normalizing metals. Variable analytical methodology also biases use of the more refractory metals (Al, Fe). Second, though in some areas contaminant levels correlate with texture or organic carbon, this relationship is not consistent. Normalization could potentially introduce significant error. Recent data (Bothner et al., 1997; L. Torres, unpublished information) suggest that part of the reason for this behavior is the past presence of localized point sources of contaminants, along with spatially heterogeneous natural sediment distribution in this glaciated area. One might point out in this context that the hard copy reports of the National Standards and Trends program have in the past reported 'mud' (silt plus clay) normalized data, whereas the data obtained by web site or floppy disc are unmodified.

3.6. Spatial variability in composition

Recent studies confirm that Boston Inner Harbor sediments have a high degrees of spatial variability in contaminant concentration, often within short distances, e.g. 200 m. This effect is interpreted to be most pronounced in the center of the urban harbor complex. It is not due to unreliable data, as has sometimes been assumed. Rather, it reflects real heterogeneity, often associated with point contaminant sources and presence of CSOs. Such spatial inhomogeneity complicates the tracing of changes in bottom sediment quality with time through use of samples taken at different dates (Fig. 9). However, when sufficient samples are available to provide robust statistics, the aggregate effects are consistent with outer harbor studies of more homogeneously distributed sediments using dated cores (Bothner et al., 1997). They show decreases of 40 to 60% in Cu, Hg, and Zn concentrations in surficial sediments collected over a 17-year period. A corollary conclusion

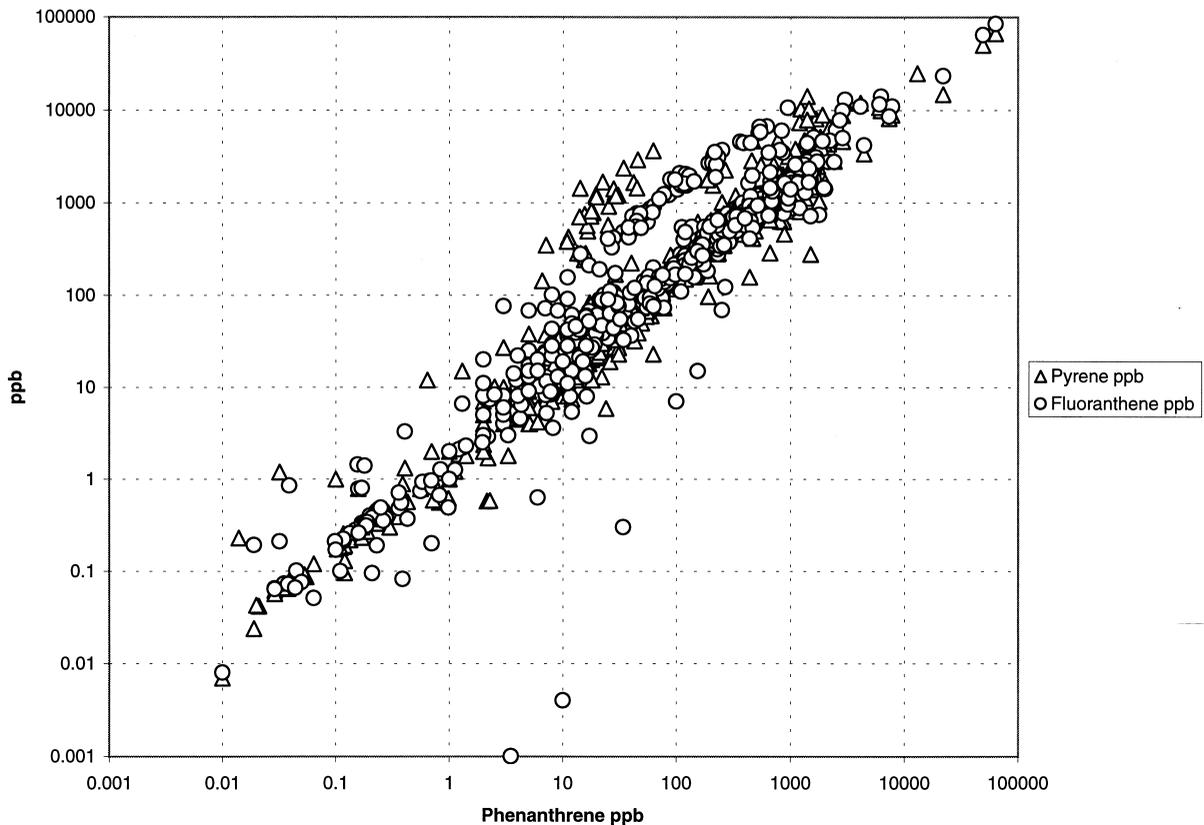


Fig. 8. Plot of two high-molecular weight PAH hydrocarbons against phenanthrene.

from the above observations would be that isolated cores within inner harbor environments may be expected to be much less representative of larger areas than areas of fine-grained sediment accumulations more distant from point sources.

4. Summary and conclusions

Assessment of a large, heterogeneous sediment data set from the Boston Harbor–Massachusetts Bay area has shown the following.

(1) Heterogeneous historical data can be successfully compiled even though data sources may lack specific information on the quality of analytical procedures. Comprehensive data rescue in Boston Harbor–Massachusetts Bay has greatly enhanced spatial resolution of contaminant distribution, helped identify sources, and helped document changes in the

contaminant status of local environments with time. This kind of data rescue opens up large quantities of existing information to cost-effective use without new field studies.

(2) The study suggests that a relatively small fraction of the data, <5% for zinc and copper in Boston Harbor–Massachusetts Bay, was associated with irremediable problems affecting their use for environmental assessment. A combination of editing for clerical errors, concentration range testing in spreadsheet format, and use of key geochemical relationships repair some data and reveal the presence of systematic errors consistent ways. Univariate or multivariate statistical methods by themselves are not reliable measures of outliers related to questionable analytical data. Ratio plots, regional groupings, and comparison with standard data sets serve as supplementary discriminant functions. Problematic data most often occur in batches from common sources or subsets.

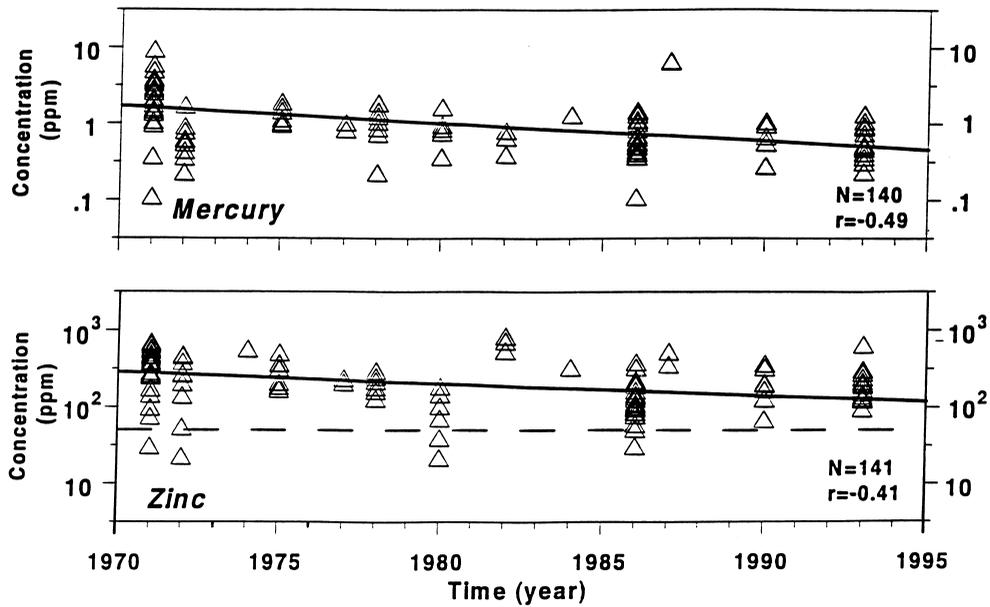


Fig. 9. Plot of surficial sediment samples from Boston Inner Harbor taken over time (modified from Bothner et al., 1997).

(3) All raw data are retained in the compiled database. The rationale for this policy includes the desirability of flagging potentially compromised data (which might otherwise be inadvertently or independently used for statistical or other characterization). Rejecting some data would require the setting of arbitrary criteria that might limit access of users to given data. It might also set up potentially unwarranted assumptions about the quality of retained data. Users have the ability and responsibility to set their own ultimate quality criteria. The authors and USGS take no responsibility for data other than those produced by USGS, except for exercising reasonable care in compiling information from their respective sources.

(4) Data rescue has revealed in more detail the presence of erratic distributions of contaminants in inner harbor environments. This potential must be considered in attempting modeling, averaging of data, or interpretations of cores from inner harbor areas.

Acknowledgements

We acknowledge the partial financial support of the Massachusetts Water Resources Authority toward studies reported here, and thank collaborators

at the MWRA, U.S. EPA, U.S. Army Corps of Engineers and other local institutions for their cooperation and assistance. The analyses and cooperation of Michael Bothner have been of critical importance to this work. Finally, we acknowledge the assistance of Polly Hastings and Jamey Reid, who prepared many of the data sets and illustrations reported here.

References

- Ahrens, L.H., 1954. The lognormal distribution of the elements (a fundamental law of geochemistry and its subsidiary). *Geochim. Cosmochim. Acta* 5, 49–73.
- Baker, B., Kravitz, M. (Eds.), 1992. *Sediment Classification Methods Compendium*, Sediment Oversight Technical Committee. U.S. Environmental Protection Agency Report, EPA 823-R-92-006, chapter paginated, 14 chapters.
- Bothner, M.H., Buchholtz ten Brink, M., Butman, B., Knebel, H.J., Manheim, F.T., Signell, R.P., 1994. *Circulation and contaminant transport in Massachusetts Coastal Waters: a summary of achievements and future plans*. U.S. Geol. Surv. Open-File Rep. 94-649, 30 pp.
- Bothner, M.H., Buchholtz ten Brink, M., Manheim, F.T., 1997. Metal concentrations in surface sediments of Boston Harbor — changes with time. *Mar. Environ. Res.* 45, 127–155.
- Bowen, H.J.M., 1979. *Environmental Chemistry of the Elements*. Academic Press, London, New York, 333 pp.
- Buchholtz ten Brink, M.R., Manheim, F.T., Hathaway, J.C., 1992.

- Results of Boston Harbor Contaminated Sediment Database Development. Report on Cooperative database development with U.S. Army Corps of Engineers, New England District, 11 chapters and electronic media.
- Butman, B., Bothner, M.H., Hathaway, J.C., Jenter, H.L., Knebel, H.J., Manheim, F.T., Signell, R.P., 1992. Contaminant transport and accumulation in Massachusetts Bay and Boston Harbor: a summary of United States Geological Survey studies in Massachusetts Bay. U.S. Geol. Surv. Open-File Rep. 92-202.
- Edelstein, H., 1997. Data mining: exploiting the hidden trends in your data. *DB2 Mag.* 2: 19–23.
- Foley, E., Blecinski, C., Eustace, H., Howell, C., Kronick, A., O'Neill, S., Palmer, K., Wong, W., Rigatti, M., 1990. Final Data Report for evaluation of the reference site suitability for the Massachusetts Bay Disposal Site. Arthur D. Little, Inc., Cambridge, Massachusetts; prepared for the U.S. Environmental Protection Agency, Region I, Water Management Div., 25 pp.
- Gardner, G.R., Pruell, R.J., 1988. A histopathological and chemical assessment of winter flounder, lobster and soft-shelled clam indigenous to Quincy Bay, Boston Harbor and an in situ evaluation of oysters including sediment (surface and cores) chemistry. Task II and III Report by U.S. Environmental Protection Agency, Environmental Research Laboratory, Narragansett, Rhode Island, 80 pp. with tables and appendices.
- Horowitz, A.J. 1991. A Primer on Sediment-Trace Elements Chemistry. Lewis Publishers, Chelsea, MI, 2nd ed., 136 pp.
- Long, E.R., Morgan, L.G., 1990. The potential for biological effects of sediment-sorbed contaminants tested in the National Status and Trends Program, Seattle, WA. NOAA Tech. Mem., NOS OMA 62, 175 pp.
- Long, E.R., MacDonald, D.D., Smith, S.L., Calder, F.D., 1995. Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environ. Manage.* 19, 81–97.
- Long, E.R., Sloane, G.M., Carr, R.S.Á., Scott, K.J., Thursby, G.B., Wade, T.L., 1996. Sediment toxicity in Boston Harbor: magnitude, extent, and relationships with chemical toxicants. NOAA Tech. Mem., NOS ORCA 96, 133 pp.
- Manheim, F.T. (Ed.), 1993. Composition of contaminated sediments along the Atlantic Margin. Report by the USGS-Branch of Atlantic Marine Geology, Contaminated Sediments Project, to the U.S. Environmental Protection Agency National Contaminated Sediment Data Inventory, and Tetratex, Inc., Baltimore MD (Contractor), 7 chapters, 5 appendices and 8 floppy disks (mainly compressed files).
- Manheim, F.T., Hathaway, J.C., 1991. Polluted sediments in Boston Harbor, Massachusetts Bay: progress report on the Boston Harbor Data Management File. U.S. Geol. Surv. Open-File Rep. 91-331, 27 pp.
- Manheim, F.T., Hathaway, J.C., Buchholtz ten Brink, M.R., 1992. Compiling sediment and pollutant data bases from the historical record: results of pilot studies from the Boston Harbor–Massachusetts Bay Program. *Proc. EPA's Contaminated Sediment Management Forum*, v. EPA 823-R-92-007, pp. 28–33.
- Manheim, F.T., Flowers, G.C., McIntire, A.G., Marot, M., Holmes, C., 1997. Environmental geochemistry and sediment quality in Lake Pontchartrain: database development and review. *Gulf Coast Assoc. Geol. Soc.* 47, 337–349.
- Mason, C., 1984. Unpublished memorandum and correspondence with Office of Program Support, U.S. Environmental Protection Agency, Region I; supplied by S. Curran, Massachusetts Coastal Zone Management Program, Boston, MA.
- Metcalf and Eddy, 1984. Boston Harbor Data Management File, Electronic data file in SAS format, with supplementary hard copy bibliography. Prepared for the U.S. Environmental Protection Agency, Region I; coding incomplete; available from U.S. Environmental Protection Agency, Harbor Studies, JFK Building, Boston, MA, 02128.
- MWRA, 1997. Massachusetts Water Authority Technical and Miscellaneous Reports. Unpublished list, 7 pp. (available from B. McCarthy, Environmental Quality Dept., MWRA, 100 First Ave., Boston MA, 02129).
- NRC, 1989. Contaminated Marine Sediments: Assessment and Remediation. National Academy Press, Washington DC, 493 pp.
- NRC, 1997. Contaminated Sediments in Ports and Waterways: Cleanup Strategies and Technologies. National Research Council, National Academy Press, Washington DC, 295 pp.
- NST, 1994. Inventory of chemical concentrations in coastal and estuarine sediments. NOAA Tech. Mem., NOS ORCA 76, 66 pp.
- NST, 1996. Analyses of sediment samples for the NOAA National Standards and Trends (NS&T) monitoring program, described by: Lauenstein, G., Cantillo, A.Y., Sampling and analytical methods of the National Status and Trends Program, National Benthic Surveillance and Mussel Watch Projects, 1984–1992; web site <http://www.orca.nos.noaa.gov/projects/nsandt/nsandt.html>.
- O'Connor, T.P., 1990. Coastal environmental quality in the United States. A Special NOAA 20th Anniversary Report, 34 pp.
- Scott, D. (Compiler), 1996. Massachusetts Bays Program Bibliography, Floppy disk with User's Guide. Massachusetts Bays Program, Boston, MA, 24 pp.; Cooperative venture of Massachusetts Coastal Zone Management Office and U.S. Environmental Protection Agency.
- USEPA, 1995. QA/QC guidance for sampling and analysis of sediments, waters, and tissues for dredged material evaluations, prepared by Tetratex Inc. for the U.S. Environmental Protection Agency and U.S. Army Corps of Engineers. USEPA Report 823-B-95-001, chapter paginated, about 250 pp. in loose-leaf format.
- USEPA, 1996. The National Sediment Quality Survey. A Report to Congress on the Extent and Severity of Sediment Contamination in Surface Waters of the United States. Draft, July 1996, United States Environmental Protection Agency, Office of Science and Technology, Standards and Applied Science Division, Washington DC, chapter paginated, 9 chapters.
- Wedepohl, K.H. (Ed.), 1978. Handbook of Geochemistry. Springer Verlag, Berlin, Heidelberg, New York, Volumes I–V.