EPA/600/R-08/149F
November 2009

REPORT ON THE ENVIRONMENTAL INDICATORS: AN UNCERTAINTY AND SCALING PILOT STUDY

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Preferred Citation:
U.S. EPA (Environmental Protection Agency). (2009) Report on the environmental indicators: an uncertainty and scaling pilot tudy. National Center for Environmental Assessment, Washington, DC; EPA/600/R-08/149F. Available from the National Technical Information Service, Springfield, VA., and online at http://www.epa.gov/ncea.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS AND ACRONYMS

CB	Cleanup Baseline
CDC	Centers for Disease Control and Prevention
CERCLIS	Comprehensive Environmental Response, Compensation and Liability Information System
CI	confidence interval
CWQI	Coastal Water Quality Index
CWS	community water system
DQE	Data Quality Estimate
DQO	Data Quality Objective
EDL	electrodeless discharge lamp
EMAP	Ecological Monitoring and Assessment Program
EPA	Environmental Protection Agency
eROE	online version of the ROE
FRM	Federal Reference Method
FY	fiscal year
GM EI	ground water monitoring environmental indicator
GPRA	Government Performance and Results Act
GW	ground water
HUC	hydrologic unit code
MCL	Maximum Contaminant Level
MEC	mobile examination center
NAMS	National Air Monitoring Stations
NCA	National Coastal Assessment
NCAPS	National Corrective Action Prioritization System
NCCR	National Coastal Condition Report
NCEA	National Center for Environmental Assessment
NCEH	National Center for Environmental Health
NCHS	National Center for Health Statistics
NHANES	National Health and Nutrition Examination Survey
NHIS	National Health Interview Survey
NPL	National Priority List
NRC	National Research Council
NTNCWS	non-transient non-community water systems

LIST OF ABBREVIATIONS AND ACRONYMS (continued)

ORD	Office of Research and Development
PBT	persistent bioaccumulative toxic
PDR	Public Data Release
PM	particulate matter
PSU	primary sampling units
PWS	public water systems
QA	quality assurance
QAPP	Quality Assurance Project Plan
RCRA	Resource Conservation and Recovery Act
ROE	Report on the Environment
RY	reporting year
SAB	Science Advisory Board
SDWIS/FED	Safe Drinking Water Information System/Federal Version
SLAMS	State and Local Air Monitoring Stations
TNCWS	transient non-community water systems
TPRW	total production related waste
TRI	Toxics Release Inventory
TSDF	treatment storage and disposal facilities
TT	treatment technique
USGS	United States Geological Society

PREFACE

Under the direction of the U.S. Environmental Protection Agency's Office of Research and Development's National Center for Environmental Assessment, Abt Associates Inc. prepared this assessment of uncertainty and scaling for seven of the Report on the Environment (ROE) indicators. The document is a summary of investigations into existing contextual information related to uncertainty that might help readers of the ROE interpret and better understand the accuracy of current environmental conditions and trends in environmental quality. The project focused on the principal question, "How accurate are the presented indicator values, and are changes or trends over time statistically significant?" Also related to increasing the utility of the ROE, this document presents an examination of the feasibility of presenting ROE indicators below the national scale. The information collection and research phase was undertaken between October 2006 and January 2008, and the initial draft report was prepared for internal review in June 2008. Members from EPA's Indicator Work Group conducted an internal review, and their comments were incorporated to develop an external review draft (ERD). The ERD was reviewed by an external panel of experts in February 2009. The findings presented in this final report could help inform the future development of the ROE and other indicators used by EPA.

This work was initiated in 2006 and example indicators were current as of January 2008. Several of the indicators on the electronic Report on the Environment (eROE, www.epa.gov/roe) have since been updated, and reference citations to the eROE are current as of the release date of this report.

AUTHORS, CONTRIBUTORS, AND REVIEWERS

This document was prepared by the National Center for Environmental Assessment (NCEA) within the U.S. Environmental Protection Agency's Office of Research and Development (ORD). Jay Messer was the project manager with the overall responsibility for the contents of this report.

Major portions of this report were prepared by Abt Associates Inc., Cambridge, MA, under EPA Contract EP-06-C00-0411. The following individuals are the primary authors:

Introduction:	Jay Messer[1], Andrew Stoeckle[2]
Ambient Particulate Matter Indicator:	Ellen Post[2], Victor Guzik[2]
Coastal Water Quality Index Indicator:	Paul Black[3], Paul Duffy[3], Greg McDermott[3], Mike Balshi[3]
Population Served by Community Water Systems with No Reported Violations of Health-Based Standards Indicator:	Jason Sachs[2], Victor Guzik[2]
Reported Toxic Chemicals in Wastes Released, Treated, Recycled, or Recovered for Energy Use Indicator:	Kate Ricke[2], Andrew Stoeckle[2]
High-Priority Cleanup Sites Where Ground Water is Not Continuing to Spread Above Levels of Concern Indicator:	Lauren Adams[2], Julia McGuire[2]
Blood Lead-Level Indicator:	Jason Sachs[2]
Fish Faunal Intactness Indicator:	Dennis Grossman[2], Julia McGuire[2]
Conclusions:	Jay Messer[1], Andrew Stoeckle[2]

[1]NCEA, ORD, U.S. EPA, Research Triangle Park, NC.
[2]Abt Associates Inc., Cambridge, MA.
[3]Neptune and Company, Lakewood, CO.

ACKNOWLEDGEMENTS

Content experts within and outside of the Agency generously provided their expertise and identified relevant documentation needed for the uncertainty and scaling analyses. In particular, the authors would like to thank Lawrence L. Master of NatureServe who provided methodological details underlying the Fish Faunal Intactness indicator. Dennis Mikel of EPA's Office of Air Quality Planning and Standards provided similar input for the Ambient Particulate Matter indicator.

In addition, NCEA staff, Agency staff with expert knowledge of the selected Report on the Environment (ROE) indicators, and ROE Chapter Leads provided comments as part of the Agency's review. Responsibility for the accuracy of content and presentation within this document rests with Abt Associates and the NCEA project staff.

1. INTRODUCTION

1.1. PURPOSE AND SCOPE OF ASSIGNMENT

EPA's 2008 Report on the Environment (ROE) compiles the most reliable indicators currently available to answer 23 questions of importance to the Environmental Protection Agency's (EPA's) mission and the nation's environment. The questions are divided into five topics: air, water, land, human health, and ecological condition. The report includes 85 indicators—numerical values derived from actual measurements of a state or ambient condition, exposure, or human health or ecological condition over a specified geographic area, whose trends over time represent or draw attention to underlying trends in the condition of the environment.

Each indicator underwent extensive external peer review before being incorporated into the draft ROE, and drafts of the entire report were subjected to internal EPA and interagency review. A final draft was reviewed by EPA's Science Advisory Board (SAB) and posted for public comment. Following final revisions recommended by the SAB review and updates of the indicator data, the ROE was released to the public on May 20, 2008.

Although each indicator included a list of *qualitative* descriptions regarding uncertainty, many reviewers recommended the inclusion of a more systematic, *quantitative* treatment of uncertainties associated with each indicator. The SAB review panel, in particular, noted that if the ROE was to play a role in Agency planning and decision making, then treatment of uncertainty was critical. While this could not be done in time for the 2008 ROE, EPA had anticipated this need and contracted this project to examine the feasibility of including information about uncertainty in ROE indicators in the online version of the ROE (eROE), which will be updated and revised over the coming years.

This project employed a case-study approach that focused on seven indicators from the 2008 ROE that spanned a broad range of environmental media, data source types, and sampling designs. Each case study identified the possible sources of uncertainty, quantified the resulting indicator uncertainty for the important sources wherever possible, and estimated the potential importance of sources that could not be quantified. Each case study also identified the types of data that would be needed to quantify potentially important sources of uncertainty in the future. The project examined uncertainty in the national-level indicators, as well as the effects of regional scaling of the indicators on uncertainty wherever possible (based on the available indicator data).

The purpose of the case studies was to better understand whether—and how—uncertainty could be systematically treated in the indicator presentations in the eROE. It was assumed that the specific treatment of uncertainty would vary, depending on the type of indicator data and media (e.g., biological versus chemical, air versus landscape), the nature of the database (e.g., data from an environmental monitoring network versus data reported from different sources using a variety of methods), and the sampling design (e.g., fixed site networks versus probability sampling). The project examined whether there are commonalities in these factors that might translate to the broader range of indicators in the ROE. The project focused on the principal question, “How accurate are the presented indicator values, and are changes or trends over time statistically significant?” but did not extend to examining or proposing mock-ups of how uncertainty could be displayed in future write-ups for the indicators in the eROE.

1.2. METHODOLOGY

ROE Indicators Selected for Investigation

- Ambient Particulate Matter (PM) Concentrations
- Coastal Water Quality Index (CWQI)
- Populations Served by Community Water Systems (CWS) with No Reported Health-Based Violations
- Reported Toxic Chemicals in Wastes Released, Treated, Recycled, or Recovered for Energy Use
- High-Priority Cleanup Sites where Contaminated Ground Water (GW) is Not Continuing to Spread above Levels of Concern
- Children’s Blood Lead-Levels
- Fish Faunal Intactness

Each case study followed a similar—but not identical—approach to the investigation and presentation of uncertainty. The first step was to construct a root-cause analysis diagram for the possible sources of uncertainty in the indicator, based on the data flows from data collection in the field all the way to the presentation of the indicator (see examples in Chapters 2–7). Each case study looked for three sources of uncertainty (including both random error and bias):

- Uncertainty caused by the failure of the monitoring design to capture the true value of the indicator in time and space (e.g., sampling error in probability designs, failure of the design to detect rare events, sensitivity of the indicator to a few large sources or events).
- Uncertainty caused by the measurement itself (e.g., instrument error, inadequate quality control).
- Uncertainty caused by data reporting, management (e.g., entry errors, database corruption, computation errors), and analysis (e.g., methods used to aggregate data).

Note that an indicator may cover only part of a target population (e.g., only counties likely to exceed an air quality standard or facilities of a minimum size), or only part of the year (only the ozone season or only the summer for some types of surface water). The ROE identifies any such limitations of the indicator, but such limitations do not introduce uncertainty into the indicator value itself. However, in some cases, the project did examine the uncertainty about assumptions that other locations, or sources, or seasons in the above examples would be of concern, and, thus, should, if possible, be included in the indicator.

Following identification of the sources of error, a judgment was made (1) regarding the level of sensitivity of the indicator to each possible source of uncertainty; and (2) whether uncertainty in the reported indicator values could be quantified from readily available information. Where such information was available, uncertainty was calculated. The resulting uncertainty is shown graphically or in table form, as appropriate. For sources whose effects on uncertainty could not be quantified, their potential to introduce significant uncertainty into the indicator was estimated, and the data that would be needed to quantify the uncertainty are identified.

Finally, the case studies for all indicators were compared to determine the likely implications of the seven pilot indicator case studies for quantification of the remaining indicators in the ROE. This comparison is included in the conclusion of this paper.

2. AMBIENT PARTICULATE MATTER (PM) CONCENTRATIONS INDICATORS

2.1. DESCRIPTION OF AMBIENT PM INDICATOR

The ROE introduces the Ambient PM indicator with the following description (U.S. EPA, 2008d):

> PM is the general term used for a mixture of solid particles and liquid droplets found in the air. Airborne PM comes from many different sources. Primary particles are released directly into the atmosphere from sources such as cars, trucks, heavy equipment, forest fires, and burning waste. Primary particles also consist of crustal material from sources such as unpaved roads, stone crushing, construction sites, and metallurgical operations. Secondary particles are formed in the air from reactions involving precursor chemicals such as sulfates (which are formed from sulfur dioxide emissions from power plants and industrial facilities), nitrates (which are formed from nitrogen dioxide emissions from cars, trucks, and power plants), and carbon (which are formed from reactive organic gas emissions from cars, trucks, industrial facilities, forest fires, and biogenic sources such as trees).
>
> Ambient air monitoring stations throughout the country measure air concentrations of two size ranges of particles: $PM_{2.5}$ and PM_{10}. $PM_{2.5}$ consists of fine particles with aerodynamic diameters less than or equal to 2.5 µm. PM_{10} includes both fine particles ($PM_{2.5}$) and coarse particles, which are particles with aerodynamic diameters less than 10 µm. The heavier PM_{10} particles tend to exhibit more localized effects, whereas $PM_{2.5}$ tends to exhibit a more regional effect as the primary and secondary particles that form it are more easily transported.
>
> This indicator presents ambient concentrations of PM_{10} from 1990–2004 based on the annual average and the annual second maximum 24-hour average. The indicator also presents ambient concentrations of $PM_{2.5}$ from 1999–2004 based on the annual average and the 98th percentile of the 24-hour average. Regional data are presented from 1990–2004. Both annual average and 24-hour average concentrations are presented to capture trends related to chronic as well as acute exposure. Trend data are based on measurements from two nationwide networks of monitoring stations: the National Air Monitoring Stations (NAMS) and State and Local Air Monitoring Stations (SLAMS)[1] (U.S. EPA, 2008d).

[1]While the 2008 ROE indicator was in development, EPA's Ambient Air Monitoring Regulations were amended to eliminate the distinction between NAMS and SLAMS.

The ROE also notes the following limitations of the indicator:

- Because there are far more PM_{10} and $PM_{2.5}$ monitors in urban areas than in rural areas, the trends might not accurately reflect conditions outside the immediate urban monitoring areas.
- Potential biases may exist for some ambient concentration measurements due to losses from volatilization of nitrates and other semivolatile materials and retention of particle-bound water associated with hygroscopic species.
- Due to the relatively small number of monitoring sites in some EPA Regions, the regional trends are subject to greater uncertainty than the national trends. Some EPA Regions with low average concentrations may include areas with high local concentrations, and vice versa.
- To ensure that long-term trends are based on a consistent set of monitoring sites, selection criteria were applied to identify the subset of PM monitoring sites with sufficient data to assess trends over the time frames covered by this indicator. Monitoring sites without sufficient data are not included in the trend analysis. Some excluded monitoring sites reported PM concentrations above the level of the PM standard during the years covered by this indicator. In 2007, for example, 39 monitoring sites nationwide (in addition to the trend sites shown in Exhibit 2-20, panel B) recorded PM_{10} concentrations above the level of the NAAQS but did not have sufficient long-term data to be included in this indicator.

The $PM_{2.5}$ annual average for a year (hereafter referred to as simply “the PM indicator”) is actually an average of each of the monitor-specific annual averages for that year. In order to calculate national average ambient PM concentrations, an appropriate set of monitors in the NAMS/SLAMS network is first identified by narrowing the universe of monitors to those that have at least 5 years of valid data during the 6-year period from 1999 to 2004. A “valid” year of data is a year with at least 11 daily measurements recorded in each calendar quarter (U.S. EPA, 2008e). For each monitor that satisfies the inclusion criteria, a seasonally weighted annual average is calculated for each year in which there are sufficient daily measurements.[2] For any year with insufficient daily measurements (a “missing” year), the annual average PM concentration at the monitor is estimated through linear interpolation, or, if the missing year is at either the beginning or end of the 6-year period, it is then replaced with the concentration for the nearest valid year. The $PM_{2.5}$ annual average for a year is then calculated as the average of the (actual or linearly interpolated) monitor-specific annual averages for that year across the set of monitors with valid data.

[2]We assume that “seasonally weighted” means that the four quarterly averages are first calculated and then the average of these four quarterly averages is calculated. This mitigates any bias from missing values that are nonrandom or nonuniformly distributed throughout the year.

According to the ROE (see Figure 2-1), 707 of the 781 PM monitoring sites had sufficient information to be used for the $PM_{2.5}$ national averages for the 6-year time period. The national point estimate of average concentrations in 2004 represents an 11% decrease since 1999.

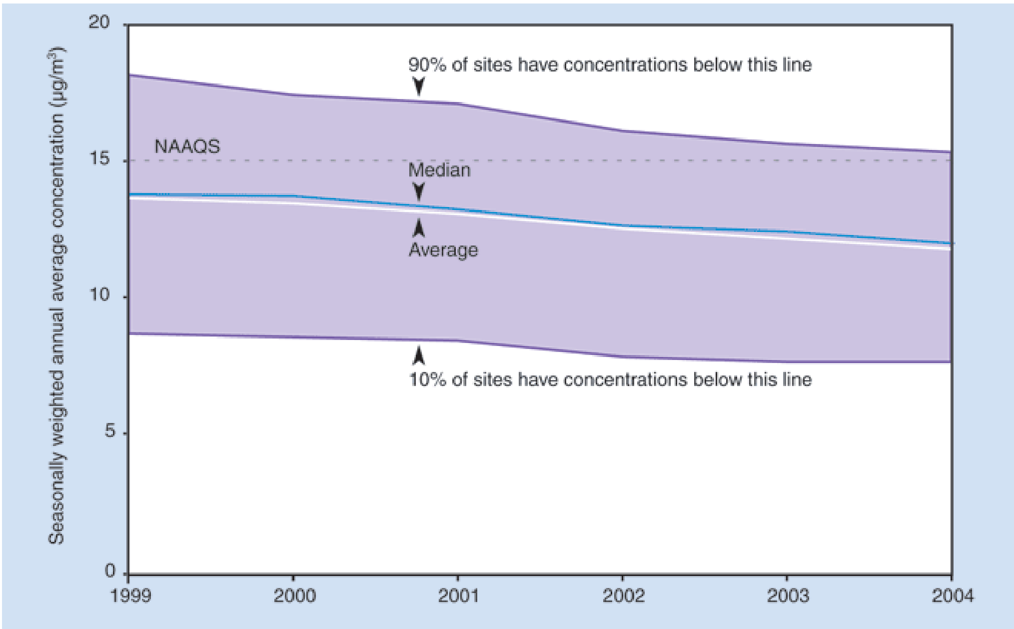


Figure 2-1. Ambient $PM_{2.5}$ concentrations from 1999–2004 (U.S. EPA, 2007b).

This case study characterizes the uncertainty related to the Ambient PM indicators using available information. The case study focuses on the $PM_{2.5}$ annual averages only (rather than on PM_{10} or the 10 or 90 percentiles). However, we assumed that similar sources and trend effects of uncertainty exist for the other PM indicators presented in the ROE. Figure 2-2 depicts the general sequence of data capture and processing involved in preparing ROE indicators. For the Ambient PM indicator, we identified the "Data Capture Plans and Methods," "Equipment," and "Information Processing Rules and Procedures" steps as the major sources of uncertainty pertaining to the ROE end use. Because the analysis under this case study is limited to examination of available data, quantitative analysis of missing daily values of monitor-specific annual averages and missing monitor-specific annual averages is used to characterize the uncertainty introduced to the ROE indicator. Data processing errors and measurement errors are the other major source of uncertainty and are characterized qualitatively. Thus, uncertainty in the national-level indicator is not assessed quantitatively.

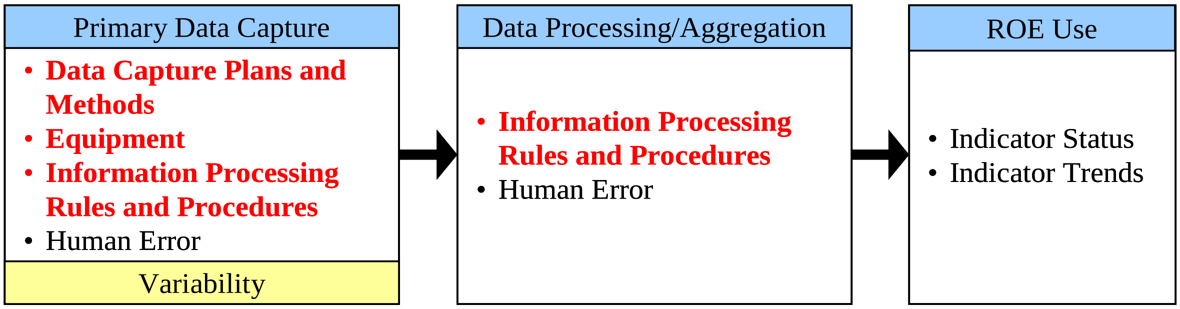


Figure 2-2. General sequence of data capture and processing involved in preparing the ambient particulate matter indicator.

2.2. OVERVIEW OF UNCERTAINTY

Figure 2-3 presents the data flow that culminates in the ROE indicator presentation, along with elements of uncertainty associated with each step in this data flow. As shown in the chart, the PM indicator has several potential sources of uncertainty:

> **Overview: Uncertainty**
>
> - Measurement errors, estimations of missing PM concentrations, and data processing errors are not expected to have any significant impact on the national or regional mean annual $PM_{2.5}$ estimates.
> - Biases from measurement error and nearest-year replacement shift the current national trendline down (lower mean annual averages), and reduce the slope, respectively.

- Measurement error
- Missing daily values in the calculation of monitor-specific annual averages
- Missing monitor-specific annual averages
- Data processing errors

Each of these sources of uncertainty may result in over- or underestimation of PM concentration to some degree. However, unless the source of uncertainty is also a source of systematic bias, the effect of errors in one direction (e.g., underestimation) is likely to be dampened by the effect of errors in the opposite direction (e.g., overestimation) when data are averaged for the final indicator. The listed uncertainty elements and their likely effect on the final ROE indicator are discussed in further detail in Sections 2.2.1 through 2.2.5.

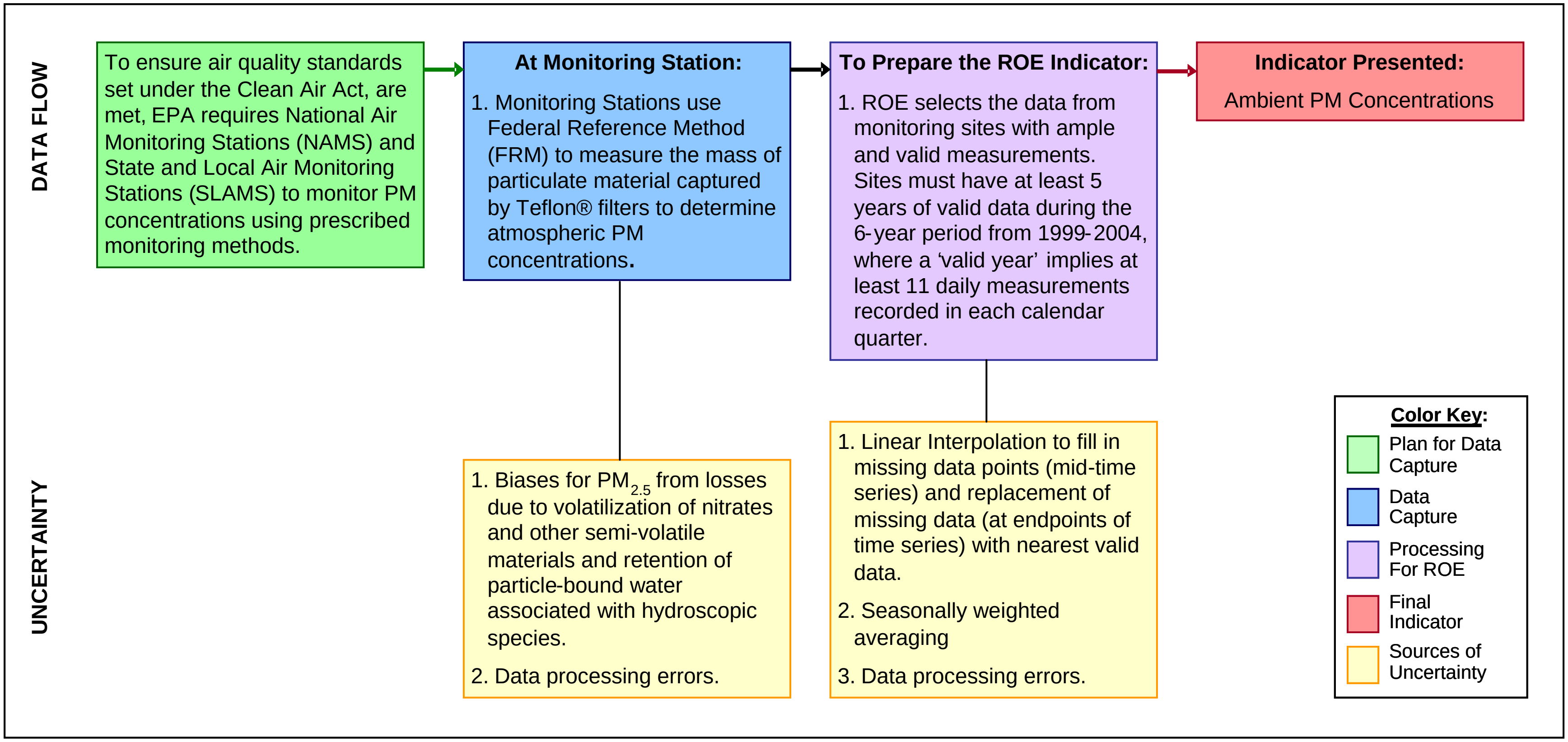


Figure 2-3. Uncertainty and data flow: ambient $PM_{2.5}$.

2.2.1. Measurement Error

Measurements of ambient $PM_{2.5}$ concentrations are made at NAMS/SLAMS monitors nationwide. EPA requires state/local monitoring agencies to use a U.S. Federal Reference Method (FRM) or a Federal Equivalent Method sampler to measure the mass of material captured by Teflon® filters to determine atmospheric PM concentrations. However, due to several factors, these concentration estimates are not necessarily accurate representations of particle mass suspended in the air at the time of sampling (U.S. EPA, 2004a). As noted in the "limitations" of the indicator, ambient PM concentrations may be overestimated or underestimated due to

- neutralization of acid or basic vapors on the filter (e.g., the reaction of SO_2 with basic particles can lead to the formation of sulfate);
- variation in the amount of particle-bound water collected and retained on the PM such that what is measured when the filter is weighed may not represent what was bound to the particle at the moment of sampling;
- variation in the amount of absorbed semivolatile vapors onto collected PM (e.g., organic compounds, nitric acid, and ammonium nitrate) during sampling or postsampling handling, such that what is measured when the filter is weighed may not represent what was bound to the particle at the moment of sampling; or
- variation in pressure, temperature, and relative humidity in the laboratory, especially in cases where the aerosol collected is hygroscopic.

These sources of measurement error are caused by the highly variable nature of ambient PM conditions and the inevitable changes that happen when PM is collected, stored, shipped, handled, and weighed. PM concentrations are sensitive to atmospheric variables such as chemicals that are present in the air at the time of monitoring, temperature, humidity, etc., all of which can change from moment to moment (U.S. EPA, 2004a). A variety of non-Reference/Equivalent sampling methods exist that may capture these atmospheric variables and provide a more faithful quantification of the PM mass that is actually suspended in the air; however, the EPA chose the FRM sampling methods to be the definitional basis of the Ambient Air Quality Standard for $PM_{2.5}$. Thus, for the purpose of this uncertainty analysis, we do not consider these atmospheric variations to contribute to uncertainty in the ROE indicator, nor did we attempt to quantify the potential effect of this source of measurement error on annual mean $PM_{2.5}$ concentrations.

However, there are other sources of measurement variation that rightfully should be considered as sources of uncertainty. PM sampler accuracy depends on the accuracy of the air-flow measurements, the accuracy of the filter weighing process, loss of filter material in handling, and other factors. EPA requires state/local monitoring agencies to conduct testing to quantify the net effect of these sources of variation. One program requires a certain number of sites to have two side-by-side samplers collect simultaneous samplers in an effort to provide an indicator of variation. Because some sources of error could affect both samplers in a pair or affect the handling and weighing of both filters, EPA also requires state/local monitoring agencies to arrange for periodic visits by an independent, highly maintained sampler, the filter from which is handled by separate personnel and weighed at an independent laboratory. The most recent assessment of $PM_{2.5}$ measurement accuracy by EPA, based on these sources of performance data, suggests that monitor measurements can have a percent error of ±6.9, and a bias of about −2.1% (U.S. EPA, 2007a).[3]

Measurement error at a given site will not necessarily cause a large degree of percent error in national or regional annual mean concentration estimates. Because the monitor-specific annual average is an average over many values at that monitor, errors in one direction will tend to cancel out errors in the opposite direction, so the uncertainty about the monitor-specific annual average will be substantially less than the uncertainty about any individual measurement at that monitor. Similarly, errors in opposing directions will offset each other when averaging across sites. However, this does not mean that errors in opposing directions will *completely* cancel one another out, as this would depend on the magnitude and direction of all errors. The small bias at the measurement level, on the other hand, is assumed to carry through to the average over all monitors, and it causes the current PM indicator trendline to shift downwards by about 2% at all points, leaving the slope (expressed as a percentage change over the period) of the trendline unchanged.

2.2.2. Missing Daily Values in the Calculation of Monitor-Specific Annual Averages

An annual average PM concentration at a site can be estimated only if there are at least 11 daily measurements per calendar quarter. This criterion implies, however, that some annual averages may be based on incomplete sets of scheduled sampling values. Consequently, missing data may introduce a bias for annual mean concentration estimates at a given monitoring station if the missing days are not uniformly distributed throughout the year. The PM indicator presented by the ROE mitigates this bias by calculating a "seasonally weighted average," which is simply an average of the four *quarterly* averages.

---

[3]The estimate of bias is a national mean, while the estimate of precision is the upper bound of a 90% confidence interval. The bias estimate is based on site-specific comparisons of regular state/local monitor measurements with independent audit measurements made by a visiting sampler. The precision estimate is based on site-specific comparisons of permanently collocated (paired) monitors run by the state/local agencies. Both bias and precision estimates are based on 3-year averages of SLAMS sites only. For further details, see http://www.epa.gov/ttn/amtic/files/ambient/02_04%20PM2.5%20QA%20Report%20Draft.pdf.

Despite using a seasonally weighted average for the ROE indicator, missing daily values introduce uncertainty. It is possible to quantify this uncertainty by comparing the annual averages for monitor-years in which there are little missing data with the annual averages that would result for these monitor-years if a random sample of days were removed before averaging. In essence, comparing an average based on a complete, or nearly complete, set of daily values with averages based on samples of reduced size allows one to characterize the effect of missing monitor-days.

In such an analysis, based on 2003 data from a single monitoring site in Detroit, MI, samples with missing daily values were created by omitting randomly selected days from the original set of daily values.[4] Samples were created with 25, 40, and 75% of days missing, and corresponding annual average PM concentrations were estimated based on each of these samples. Note that the percent of missing days in this analysis reflects a potential daily sampling schedule, while actual sampling schedules may be on the order of every several days or more. In cases when sampling schedules are less frequent than every single day, it would not be possible to have 75% of days missing, and percent errors from missing monitor-days would likely reflect the results shown for 25% and 40% missing scenarios. Table 2-1 shows the 90% confidence intervals (CIs) for the percent error in the estimate of the monitor-specific annual average based on these reduced samples. The results suggest that missing daily values do not result in any bias in the estimate of a monitor-specific annual average. The results also suggest, not surprisingly, that the amount of uncertainty in an estimated annual average at a monitor will increase as the percentage of missing daily values at that monitor increases. Note that even with as much as 75% of the monitor values missing, the estimated annual average is still within about 9% of the actual value, with 90% confidence.

Table 2-1. Percent error caused by missing monitor-days for a single site (derived by Abt Associates)

Percent of days missing	Lower bound of 90% CI[a]	Upper bound of 90% CI[a]
25%	−3.4%	3.4%
40%	−4.1%	4.1%
75%	−8.6%	8.6%

[a]Based on samples of 46 and assuming no bias.

[4]Monitor data from Detroit, MI were not selected due to any particular criteria; rather, these data were readily available to Abt Associates at the time of uncertainty analysis.

While there will be errors for monitor-years due to missing daily values, it is unlikely that these errors will cause any significant percent error in national point estimates. Again, because some monitor-specific estimates will be overestimates and others will be underestimates, the two types of errors will tend to cancel each other out when the monitor-specific annual averages are averaged across the group of 707 sites for the same year. Moreover, missing monitor days, to the extent that they are random, should not induce any bias to the ROE indicator.

This analysis of uncertainty due to missing monitor days assumes missing daily values occur randomly. However, there may be interdependency among PM concentration values due to the relative locations and times of measurements. For example, it is possible that missed monitoring days are clustered over time because equipment malfunction may lead to some missing daily values occurring in sequence. If PM concentration varies less among monitoring days in sequence than it does among all days in the monitoring year, then, on average, clusters of missing daily values would have a larger impact on the magnitude of the monitor-specific annual average than a set of missing daily values randomly distributed throughout the year. Thus, the values presented in Table 2-1 may underestimate the width of the confidence intervals. The expected value of the annual average based on the sample of days for which values exist should still be unbiased. That is, unless these equipment malfunctions tend to occur when PM levels are high (or low), the expected value of the sample annual average at any given monitor should be the actual annual average at that monitor. To the extent that clustering of missed monitoring days occurs, a similar analysis that randomly selects clusters of days to be omitted could be conducted to quantify the contribution of clustered missed monitoring days to the overall uncertainty in annual PM concentration averages.

2.2.3. Missing Monitor-Specific Annual Averages

A monitor may be included in national averaging as long as it has data for 5 years out of the 6-year time period. If a monitoring site is missing a year of data that falls between year 1 and year 6, a value is calculated using linear interpolation of the two surrounding years (see Figure 2-4). Missing values for years that are at the beginning or end of the 6-year time span are replaced with the value from the nearest recorded year (see Figure 2-5). Both forms of estimation introduce uncertainty into the ROE indicator.

It is possible to characterize the uncertainty caused by linearly interpolated values by identifying all annual averages for a site that also have annual averages for the surrounding years, and comparing those annual averages to the linearly interpolated values between the surrounding years. This provides a distribution of differences between actual and interpolated annual averages, allowing the estimation of uncertainty introduced by linear interpolation. An

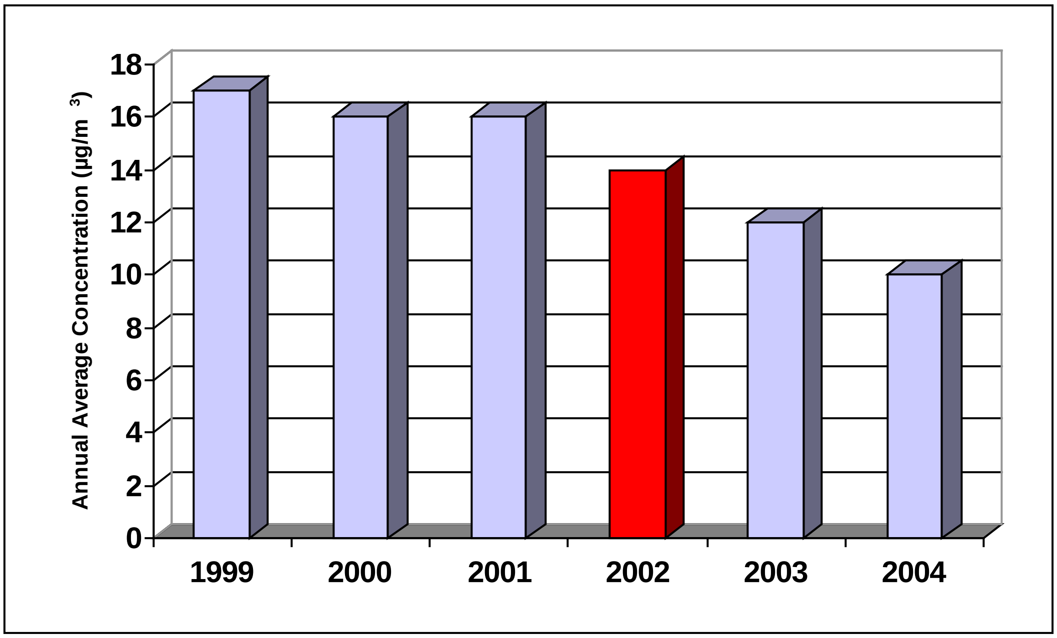


Figure 2-4. Illustration of linear interpolation for missing annual average that is internal to the time-series.

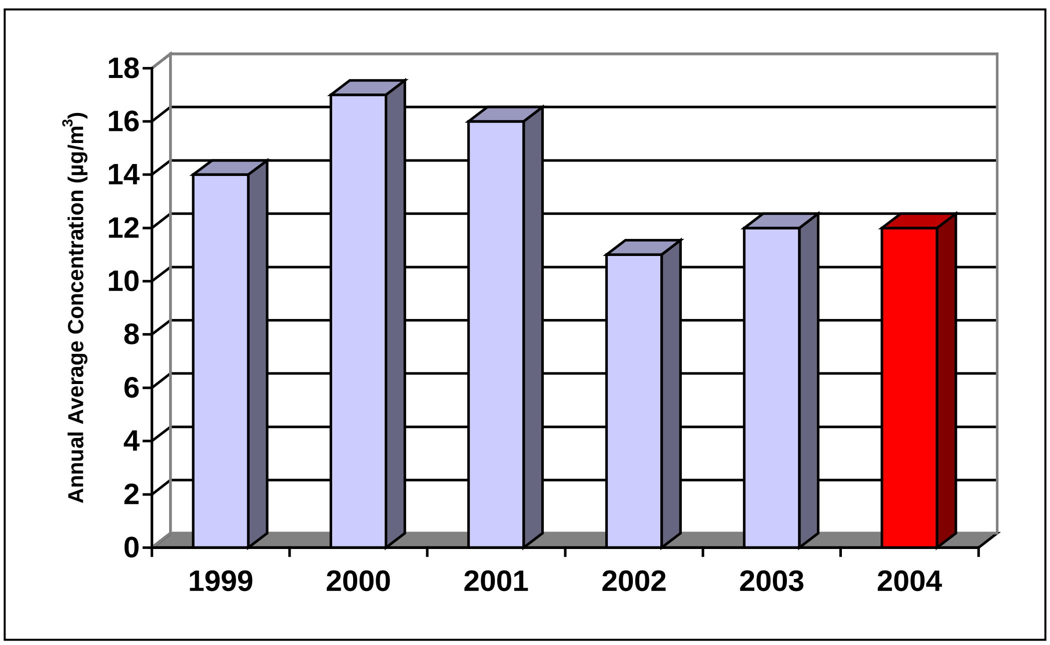


Figure 2-5. Illustration of replacing a missing annual average that is an endpoint of the time-series.

analysis of ROE-specific monitoring data suggests that monitor-specific annual averages estimated by linear interpolation are within 13% of actual average for the monitor in that year.[5, 6]

Similarly, it is possible to characterize the uncertainty caused by replacing missing values at the beginning or end of the 6-year time span by identifying those annual averages that are at the beginning (or end) of the 6-year time span and comparing those averages to averages of the subsequent (or previous) years. This provides a distribution of differences between actual and estimated annual averages, allowing the estimation of uncertainty introduced by replacing missing monitor-year averages at either end of the time series. An analysis of ROE-specific empirical monitoring data suggests that estimated annual averages at the beginning of the time series can have a percent error of ±18%.[7] Likewise, estimated annual averages at the end of the time series can have an error of 16%.[8]

The impact of missing monitor-years on the percent error of the annual mean PM averages in the indicator is likely to have little impact for two reasons. First, in any given year, only a subset of the monitors will have missing values; the smaller this subset, the less influence they have on the overall average across all monitors for the year. Second, as noted above, averaging errors from overestimated and underestimated monitor-years (e.g., as opposed to summing errors) decreases the uncertainty they induce, so that the impact of this source of uncertainty will be dampened once site concentrations are averaged across the total group of 707 sites.

However, it is important to note that using nearest-year replacement as a method of estimating missing years at the beginning and end of the time series does in fact cause a bias in the existing ROE trendline. Specifically, if PM concentrations have been decreasing with time (a negatively sloping trendline), using nearest-year replacement to estimate missing years at either end of the time series will reduce the steepness of the slope, thereby understating the reduction in PM concentrations over time. An analysis of missing years at the ends of the 1999–2004 time series confirms a bias of −1.5% for estimated beginning years and a bias of 4.6% for estimated end years. Hence, while percent error of annual averages for PM concentrations is dampened with national averaging, biases that reduce the steepness of the slope of the trendline are expected to persist in the ROE indicator. It is important to note that the effects of interpolation and end-point substitution are estimates based on a sample of existing data; the actual impact may be greater.

---

[5] The monitoring data used for analyzing uncertainty from missing monitor-specific annual averages do not match exactly those data that were used to create the 2004 indicator. Specifically, data that were received from the Office of Air Quality and Planning Standards for the purpose of this analysis contained monitoring data for only 654 of the potential 781 sites used in the ROE indicator.

[6] Based on a sample of size 2,624, with 90% confidence.

[7] Based on a sample of size 442, with 90% confidence.

[8] Based on a sample of size 752, with 90% confidence.

2.2.4. Data Processing Errors

Data processing errors may occur along the data flow from concentration measurement to presentation of the ROE indicator. While the chances for processing errors are minimized by the use of standardized SAS code for data processing, there is still the possibility for transcription errors in the earlier stages of the data transcription process. The amount of uncertainty that this introduces to the indicator depends on the frequency and magnitude of such data processing errors. However, given the very few manual steps that are taken from measurement of PM at monitors to presenting an indicator, along with the well-defined methods for creating the ROE indicator, there is little reason to believe that data processing errors generate any substantial level of percent error in national point estimates or bias in the national trendline. Moreover, the estimates of precision and accuracy of $PM_{2.5}$ measurements, described above, ought to reflect most, if not all, of the effect of transcription errors—except those that happen systematically on a single day.

2.2.5. Aggregate Effect of Uncertainty on the Report on the Environment (ROE) Indicator

The preliminary analyses discussed above suggest that none of the identified sources of uncertainty is likely to result in substantial error in the point estimate for the national average in each year. This is in part because the percent errors in the monitor-specific annual averages being estimated are relatively small. Specifically, error for a monitor-specific annual average should be of a lesser magnitude than an error in any individual measurement. More importantly, errors in opposite directions will tend to cancel each other out when the monitor-specific annual averages are averaged across many monitors. Hence, uncertainty is dampened rather than amplified as data are processed for the ROE indicator. Table 2-2 summarizes the percent error in monitor-specific annual averages. Due to these reasons, the effect on the ROE indicator point estimates will be insignificant.

It is possible that heterogeneous features in the space-time variation of air pollution, such as variation due to weather effects, would generate larger errors in one direction. To the extent this heterogeneity occurs, stochastic space-time analysis could be used to quantify its contribution to uncertainty in the PM indicator.

Measurement error assessments do suggest that a bias in PM concentrations will persist in the ROE indicator on the order of −2%. Assuming this bias maintains the same direction and magnitude for all 6 years used for the trend, the effect would be to slightly shift the ROE trendline down (lower annual mean concentrations) but leave the slope (in percentage change terms) unchanged. In other words, the degree of PM reductions remains unchanged. Conversely, the bias induced by replacing missing monitor-specific annual averages at either end

Table 2-2. Percent error in monitor-specific annual averages and the effect on the ROE indicator (derived by Abt Associates)

Source of uncertainty	Percent error in monitor-specific annual averages[a]
Measurement error	6.9%
Missing daily values[b]	8.6%
Missing annual averages (linear interpolation)	12.8%
Missing annual averages (beginning-year replacement)	18.2%
Missing annual averages (end-year replacement)	15.9%
Data processing errors	Unknown

[a]With 90% confidence.
[b]Assuming 75% days missing.

of the time series with the nearest valid year does change the slope of the trend, making it less steep. As mentioned, the beginning years of the analyzed PM indicator have a negative bias, and the end years have a positive bias, thus dampening the trend of PM reductions in the ROE indicator. No other sources of uncertainty are expected to induce bias in the indicator, and the degree to which the slope of the trendline actually shifts depends on the method of deriving the trendline (i.e., whether it is linear, etc.). Table 2-3 summarizes the bias in monitor-specific annual averages and the effect on the ROE indicator.

2.3. REGIONALIZATION

In addition to the sources of uncertainty that affect the ROE indicator, several issues arise when national annual averages are compared to annual averages of EPA Regions. In particular, the following issues change the underlying context in which regional and national PM concentrations are to be understood and compared.

The first issue with calculating regional trends in PM concentration concerns the population coverage of FRM monitors. FRM monitors do not cover the entire geographic area of the United States. This suggests that there are populations that are not represented by the national ROE indicator. Similarly, there are populations

Overview: Regionalization

- National and subnational PM concentrations may represent different compositions of population and geographic area (e.g., 70% of the U.S. population lives within 20 km of a FRM monitor, compared to 52% for Region 7).
- Regional calculations may omit information from cross-border monitors that are relevant to regional point estimates and trends.

Table 2-3. Bias induced in monitor-specific annual averages and the effect on the ROE indicator (derived by Abt Associates)

Source of uncertainty	Bias induced in monitor-specific annual averages	Bias effect on existing ROE trendline
Measurement error	−2.1%	Shift trendline −2.1%
Missing daily values[a]	None expected	None expected
Missing annual averages (linear interpolation)	None expected	None expected
Missing annual averages (beginning-year replacement)	−1.5%	Reduce slope of trendline[b]
Missing annual averages (end-year replacement)	4.6%	Reduce slope of trendline[b]
Data processing errors	None expected	None expected

[a]Assuming 75% days missing.
[b]Quantification of slope reduction depends on method of deriving the trendline.

that are not represented by regional point estimates and trends. However, the percent of population covered by FRM monitors within EPA Regions is likely to be different than the national percentage. For example, EPA Region 7 has 52% of its population living within 20 km of a FRM monitor, as opposed to 70% reported nationally. This implies that the degree to which national and regional trends represent the entire population within the Region will be different, and not necessarily comparable. Table 2-4 shows the different percentages of population coverage by distance to a FRM monitor in each EPA Region and nationally.

The first issue with calculating regional trends in PM concentration concerns the population coverage of FRM monitors. FRM monitors do not cover the entire geographic area of the United States. This suggests that there are populations that are not represented by the national ROE indicator. Similarly, there are populations that are not represented by regional point estimates and trends. However, the percent of population covered by FRM monitors within EPA Regions is likely to be different than the national percentage. For example, EPA Region 7 has 52% of its population living within 20 km of a FRM monitor, as opposed to 70% reported nationally. This implies that the degree to which national and regional trends represent the entire population within the Region will be different, and not necessarily comparable. Table 2-4 shows the different percentages of population coverage by distance to a FRM monitor in each EPA Region and nationally.

Table 2-4. Percentages of population coverage by distance to a FRM monitor (derived by Abt Associates)

Distance (km)	**Region**										
	U.S.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
<5	24%	31%	39%	28%	16%	25%	18%	24%	31%	23%	32%
<20	70%	80%	78%	69%	62%	68%	61%	52%	72%	83%	80%
<50	90%	98%	96%	91%	91%	89%	83%	75%	81%	96%	92%
<100	98%	100%	100%	100%	100%	99%	95%	95%	91%	98%	98%
<200	100%	100%	100%	100%	100%	100%	99%	98%	99%	100%	99%
200+	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Another issue regarding calculation of regional trends involves cross-boundary information loss—in particular, when there is a monitor near the border of an EPA region, but not on the other side of that border. The following example portrays a map of a border between EPA Regions 2 and 3 (see Figure 2-6). There are two monitors in New York state, which likely represent the PM conditions for both southern New York state and northern Pennsylvania. However, these two monitors will *not* be used in the calculation of PM concentrations in EPA Region 3 because they are not within the border. This is problematic because regional averages *should* take into account the information collected by those monitors, as it may influence the regional point estimates and trends. This issue illustrates a key limitation to the interpretation of this indicator; the indicator is representative of the average air quality across locations with monitors and may not represent air quality across all locations within a Region or the United States.

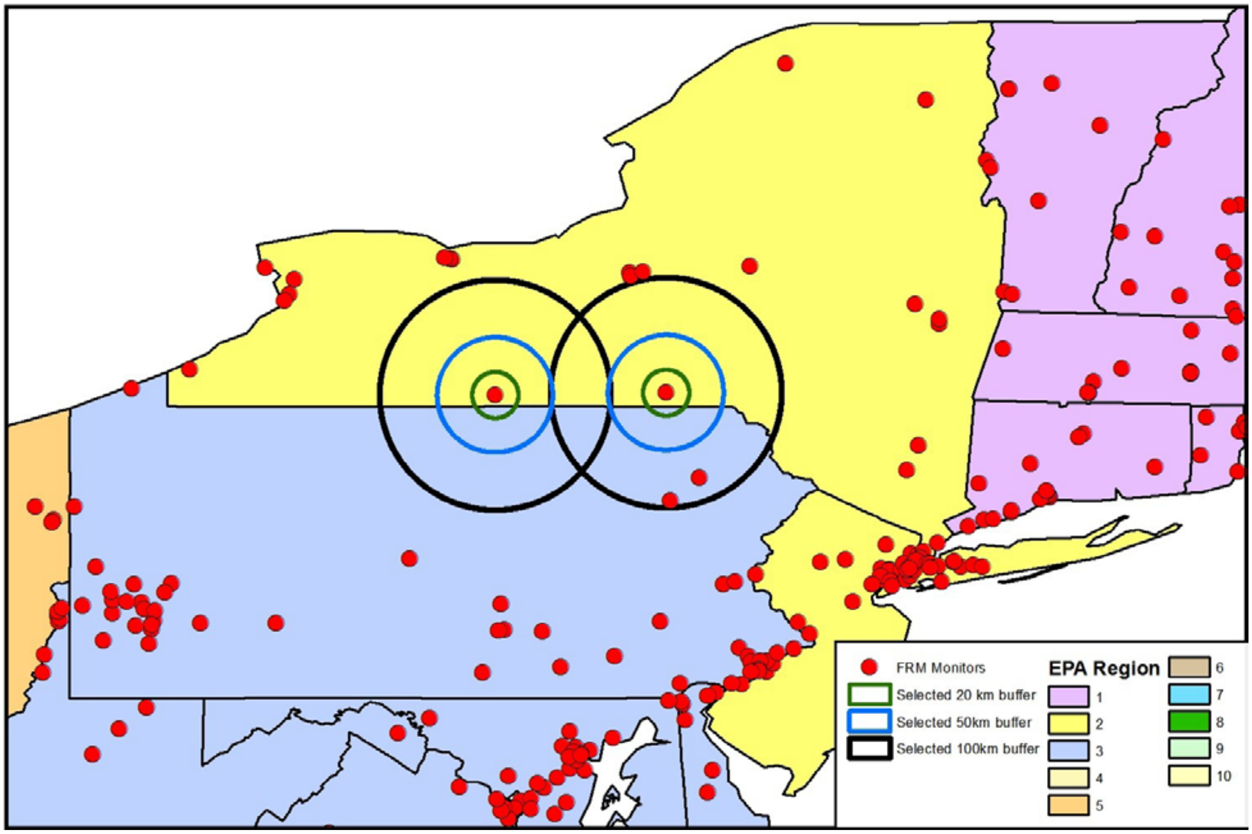


Figure 2-6. Spatial distribution of PM monitor.

3. COASTAL WATER QUALITY INDEX (CWQI) INDICATOR

3.1. DESCRIPTION OF COASTAL WATER QUALITY INDEX (CWQI) INDICATOR

The water quality index in the 2007 draft ROE is called the Trophic State of Coastal Waters. This indicator is identical to the Coastal Water Quality Index (CWQI) from the National Coastal Condition Report (NCCR). The ROE introduces the CWQI indicator with the following description:[9]

> While the presence of many water pollutants can lead to decreases in coastal water quality, four interlinked components related to trophic state are especially critical: nutrients (nitrogen and phosphorus), chlorophyll-*a*, dissolved oxygen, and water clarity. “Trophic state” generally refers to aspects of aquatic systems associated with the growth of algae, decreasing water transparency, and low oxygen levels in the lower water column that can harm fish and other aquatic life. Nitrogen is usually the most important limiting nutrient in estuaries, driving large increases of microscopic phytoplankton called “algal blooms” or increases of large aquatic bottom plants, but phosphorus can become limiting in coastal systems if nitrogen is abundant in a bioavailable form (U.S. EPA, 2003). Nitrogen and phosphorus can come from point sources, such as wastewater treatment plants and industrial effluents, and nonpoint sources, such as runoff from farms, over-fertilized lawns, leaking septic systems, and atmospheric deposition. Chlorophyll-*a* is a surrogate measure of phytoplankton abundance in the water column. Chlorophyll-*a* levels are increased by nutrients and decreased by filtering organisms (e.g., clams, mussels, or oysters). High concentrations of chlorophyll-*a* indicate overproduction of algae, which can lead to surface scums, fish kills, and noxious odors (U.S. EPA, 2004[b]). Low dissolved oxygen levels and decreased clarity caused by algal blooms or the decay of organic matter from the watershed are stressful to estuarine organisms. Reduced water clarity (usually measured as the amount and type of light penetrating water to a depth of 1 meter) can be caused by algal blooms, sediment inputs from the watershed, or storm-related events that cause resuspension of sediments, and can impair the normal growth of algae and other submerged aquatic vegetation.
>
> This indicator, developed as part of EPA’s Coastal Condition Report, is based on an index constructed from probabilistic survey data on five components: dissolved inorganic nitrogen, dissolved inorganic phosphorus, chlorophyll-*a*, daytime dissolved oxygen in bottom or near-bottom waters (where benthic life is most likely to be affected), and water clarity ([U.S.] EPA, 2004[b]). The survey, part of EPA’s National Coastal Assessment (NCA), was designed to provide a national picture of water quality by sampling sites in estuarine waters throughout the contiguous 48 states and Puerto Rico. Each site was sampled once during the

[9]Information about the indicators, their respective metadata, and independent data analyses were extracted and carried out between December 2006 and September 2007.

1997–2000 period, within an index period from July to September. The indicator reflects average condition during this index period.

Key factors like sediment load, mixing processes, and ecosystem sensitivity naturally vary across biogeographic regions and even among estuaries within regions. Thus, reference guidelines for nutrients, water clarity, and chlorophyll-*a* were established based on variable expectations for conditions in different biogeographic regions. For example, due to Pacific upwelling during the summer, higher nutrient and chlorophyll-*a* concentrations are expected in West Coast estuaries than in other estuaries. Water clarity reference guidelines are lower for estuaries that support seagrass than for naturally turbid estuaries. A single national reference range of 2–5 milligrams per liter (mg/L) was used for dissolved oxygen, because concentrations below 2 mg/L are almost always harmful to many forms of aquatic life and concentrations above 5 mg/L seldom are (Diaz and Rosenberg, 1995; U.S. EPA, 2000[a]). The process of classifying individual sites varies by region and is described in detail, along with the regional reference conditions, in [U.S.] EPA (2004[b]).

The overall water quality index is a compilation of the five components. For each site, the index is rated high if none of the five components received a score that would be considered environmentally unfavorable (high nitrogen, phosphorus, or chlorophyll-*a* levels or low dissolved oxygen or water clarity), and no more than one component was rated moderate. Overall water quality is low if more than two components received the most unfavorable rating. All other sites receive a moderate index score. If two or more components are missing, and the available components do not suggest a moderate or low index rating, the site is classified as "unsampled." Data from the individual sites were expanded from the probability sample to provide unbiased estimates of the water quality index and each of its components for each EPA Region. Results were also aggregated and weighted by estuarine area for the entire nation (U.S. EPA, 2008a).

The ROE also notes the following limitations to the indicator:

- The coastal areas of Hawaii and a portion of Alaska have been sampled, but the data had not yet been assessed at the time this indicator was compiled. Data are also not available for the U.S. Virgin Islands and the Pacific territories.

- Trend data are not yet available for this indicator. Because of differences in methodology, the data presented here are not comparable with data that appeared in EPA's first National Coastal Condition Report. The data presented here will, however, be able to serve as a baseline for future surveys.

- The National Coastal Assessment (NCA) surveys measure dissolved oxygen conditions only in estuarine waters and do not include observations of dissolved oxygen concentrations in offshore coastal shelf waters, such as the hypoxic zone in Gulf of Mexico shelf waters.

- At each sample location, the components of this indicator may have a high level of temporal variability. This survey is intended to characterize the typical distribution of water quality conditions in coastal waters during an index period from July through September. It does not consistently identify the "worst-case" condition for sites experiencing occasional or infrequent hypoxia, nutrient enrichment, or decreased water clarity at other times of the year. Further, because each site was sampled just once during the index period, the results also may not capture the full range of water quality variability during this period.

- In the original data set, there was an inconsistency in measurements taken for water clarity. Secchi disk depths were measured for all areas and light energy values for some. In order to calculate the water clarity index, light energy values were required for all sites. A model was developed to predict the light energy penetration using the secchi disk depths so that the index could be calculated.

The data used to inform these indicators are collected in a way that attempts to characterize the most degraded state of water quality. Because of this, these indicators can be considered conservative. Data for these five indicators are derived from sampling surveys for EPA's National Coastal Assessment, collected by EPA and supporting government and state agencies. The NCA comprises all the estuarine and coastal sampling done by EPA's Ecological Monitoring and Assessment Program (EMAP) since 1990.[10] All data were collected following protocols outlined in a Quality Assurance Project Plan (QAPP) that was developed specifically for EMAP activities as part of the NCA.[11] A national data set of NCA data surveys was compiled for EPA's NCCR from sampling surveys between 1997 and 2000 (U.S. EPA, 2001a).

In general, areas where distinct ecosystem types meet and transition have relatively greater productivity and diversity. This is especially true for coastal areas where the mixing of fresh and saline water results in some of the most productive ecosystems on the planet. These areas of relatively high productivity include estuaries, coastal wetlands, coral reefs, mangrove forests, and upwelling areas. Some of the critical ecological components of these coastal areas include spawning grounds, nurseries, shelter, and food for finfish, shellfish, birds, and other wildlife (U.S. EPA, 2004b). Coastal areas can also act as sinks for pollutants transported through surface water, ground water (GW), and atmospheric deposition. It is, therefore, critical that these ecosystems be monitored to assess their health. Characterizing coastal areas using the CWQI indicator involves two main steps (U.S. EPA, 2004b) as further described in the next two bullet points. Table 3-1 presents the site-specific decision criteria for each of the five indicators.

---

[10] EPA's NCA (http://www.epa.gov/emap/nca). EPA's Ecological Monitoring and Assessment Program (EMAP: http://www.epa.gov/emap).

[11] In a select few instances, data were used from non-EPA agencies, but only if the sampling stations and data met EMAP criteria. These data are indicator values for some stations in the Chesapeake Bay, collected by the Chesapeake Bay Program, and some chlorophyll-*a* sites originally chosen by New Jersey state agencies.

Table 3-1. Criteria for determining the rating of the five CWQI indicators (U.S. EPA, 2004b, Chapter 1)

Indicator	Region	Rating		
		Good	Fair	Poor
Dissolved inorganic nitrogen [mg/L]	East/Gulf Coast	<0.1	0.1–0.5	>0.5
	West Coast	<0.5	0.5–1.0	>1.0
	Puerto Rico	<0.05	0.05–0.1	>0.1
Dissolved inorganic phosphorus [mg/L]	East/Gulf Coast	<0.01	0.01–0.05	>0.05
	West Coast	<0.01	0.01–0.1	>0.1
	Puerto Rico	<0.005	0.005–0.01	>0.01
Chlorophyll *a* [µg/L]	East/Gulf Coast	<5	5–20	>20
	West Coast	<5	5–20	>20
	Puerto Rico	<0.5	0.5–1.0	1.0
Dissolved oxygen [mg/L]	East/Gulf Coast	<5	2–5	<2
	West Coast	<5	2–5	<2
	Puerto Rico	<5	2–5	<2
Water clarity indicator[a]	East/Gulf Coast	>2	1–2	<1
	West Coast	>2	1–2	<1
	Puerto Rico	>2	1–2	<1

[a]Water Clarity Indicator ratio = (observed clarity at 1-meter depth)/(regional reference clarity at 1-meter depth).

- The first step is to assess conditions at an individual site for each indicator. For each indicator, site condition rating criteria are determined based on existing criteria, guidelines, or the interpretation of scientific literature. For example, dissolved oxygen conditions are considered poor if dissolved oxygen concentrations are less than 2 mg/L (2 milligrams of oxygen per liter of water). This value is widely accepted as representative of hypoxic conditions; therefore, this benchmark for poor condition is strongly supported by scientific evidence (U.S. EPA, 2000[a]; Diaz and Rosenberg, 1995).
- The second step is to assign a regional rating for the indicator based on the condition of individual sites within the region. For example, in order for a region to be rated poor (i.e., low water quality) with regard to the dissolved oxygen

indicator, more than 15% of the coastal area in the region must have dissolved oxygen measured at less than 2 mg/L. The regional criteria boundaries (i.e., percentages used to rate each regional condition indicator) were determined as a median of responses provided through a survey of environmental managers, resource experts, and the knowledgeable public.

Using the site-specific values for each of the five CWQI indicators, the overall value for the CWQI is determined using a set of decision criteria that summarizes the information that collectively exists in the five separate indicators (see Table 3-1). Data from each of the sites are then aggregated to ratings for larger spatial regions (e.g., EPA Regions or Gulf Coast) according to an additional set of decision criteria (see Table 3-2), and regional values are aggregated into a national U.S. rating. Figure 3-1 shows the rating distribution of the CWQI for those EPA Regions where the CWQI is measured.

This case study characterizes the uncertainty related to the CWQI indicator using available information. Figure 3-2 depicts the general sequence of data capture and processing involved in preparing ROE indicators. For the CWQI indicator, we identified the "Data Processing/Aggregation" and "Primary Data Capture Plans and Methods" steps as the major sources of uncertainty pertaining to the ROE end use. For example, we deemed the "Information Processing Rules and Procedures" used in aggregating sampling at specific sites to a national index important sources of uncertainty. The "Data Capture Plans and Methods" were also a major source of uncertainty; although raw sampling data are publically available for the NCA, the sampling results used, the number and timing of replicates, and the number of sampling points within each study area were not available for the CQWI indicator reported in the NCCR. If this information becomes available, a more complete characterization of the uncertainty for this indicator could be conducted. In the absence of the primary data source information, the uncertainty characterization for this case study is, therefore, qualitative.

3.2. OVERVIEW OF UNCERTAINTY

Figure 3-3 shows that the major sources of uncertainty for the two primary ROE end uses of the CWQI indicator were information processing rules and procedures during data aggregation. Since the analysis under this case study is limited to examination of available data, and as underlying data presenting the regional averages for each indicator are not available, no quantitative analysis of uncertainty is presented. There are no recent independent assessments of the quality of the individual indicators for all geographic regions during the index study period.

Table 3-2. Criteria for determining the water quality index rating by site (U.S. EPA, 2004b, Chapter 1)

Rating	Rating criteria
Good/ high	A maximum of one indicator is fair, and no indicators are poor.
Fair/ good	One of the indicators is rated poor, or two or more indicators are rated fair.
Poor/ fair	Two or more of the five indicators are rated poor.

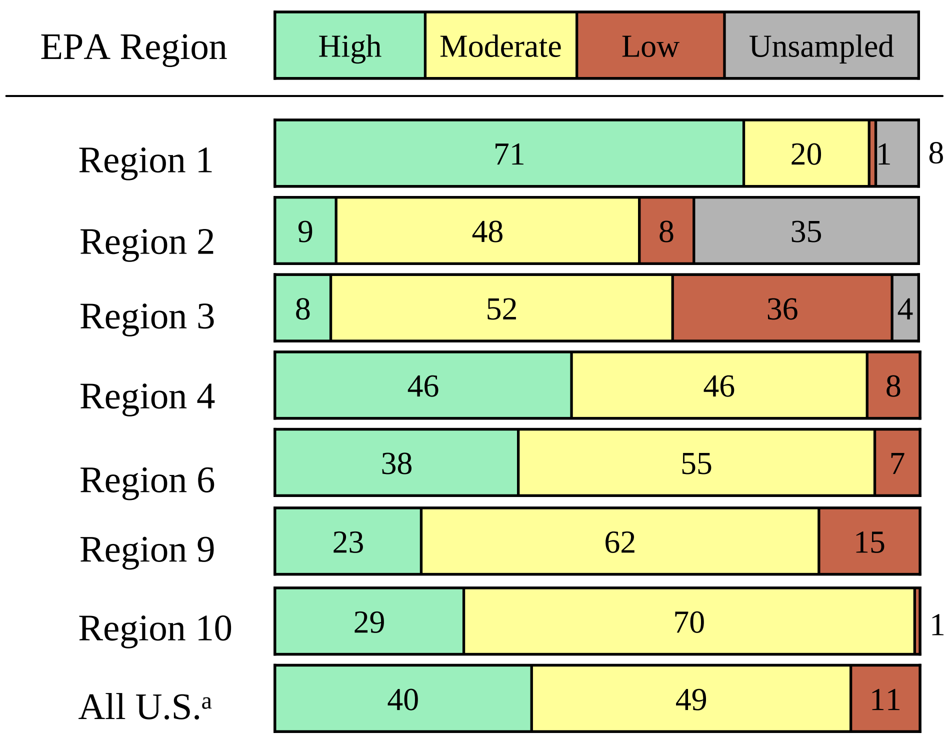


Figure 3-1. National and regional values for CWQI, 1997–2000[a] (U.S. EPA, 2008a, Exhibit 3-21).

[a]The units in the figure are percentages. A rating of **high** corresponds to **good** water quality, **moderate** corresponds to **fair** water quality, and **poor** corresponds to **low** water quality. United States figures reflect the total sampled area. Unsampled areas were not included in the United States calculation.

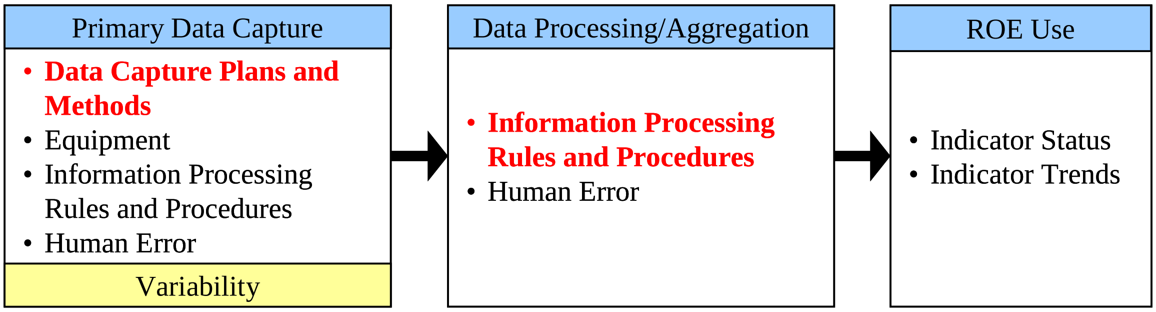


Figure 3-2. General sequence of data capture and processing involved in preparing the coastal water quality index indicator.

In addition to qualitatively addressing the sources of uncertainty, a qualitative assessment of the likelihood that the identified sources of error could affect the value of the aggregated ROE indicator is also presented.

Figure 3-3 presents the data flow that culminates in the ROE indicator presentation, along with sources of uncertainty associated with each step in this data flow. There are three distinct sources of uncertainty in the CWQI-based ROE indicators:

- Limitations to the sampling study designs and methodologies
- Errors in the indicator sample measurements from laboratory equipment
- Errors due to data management, processing, and analysis

Overview: Uncertainty

- The timing and extent of data collection does not provide a robust dataset necessary to quantify temporal uncertainty for each measured indicator, preventing a characterization of the annual indicator uncertainty due to sample design.
- Measurement errors, such as sample collection and laboratory measurements, are not expected to have any significant impact on the CWQI-indicator estimates.

Methods used to aggregate and discretize station indicator data into regional and national CWQI values may over- or underrepresent underlying indicator measurements, increasing the uncertainty in the overall CWQI indicator.

In addition, EPA's presentation of the CWQI data for use by ROE creates the potential for errors. These sources of uncertainty have the potential to play a significant role in the overall uncertainty associated with the CWQI indicator. The extent that each of these contribute to the overall uncertainty of the CWQI is discussed below.

3.2.1. Limitations to Sampling Study Designs and Methodologies

The data used to compute the CWQI were collected according to scientifically based, probabilistic sampling design, meant to distribute sampling locations to best represent the probable water quality of the area (U.S. EPA, 2001b, 1995). However, there are two types of

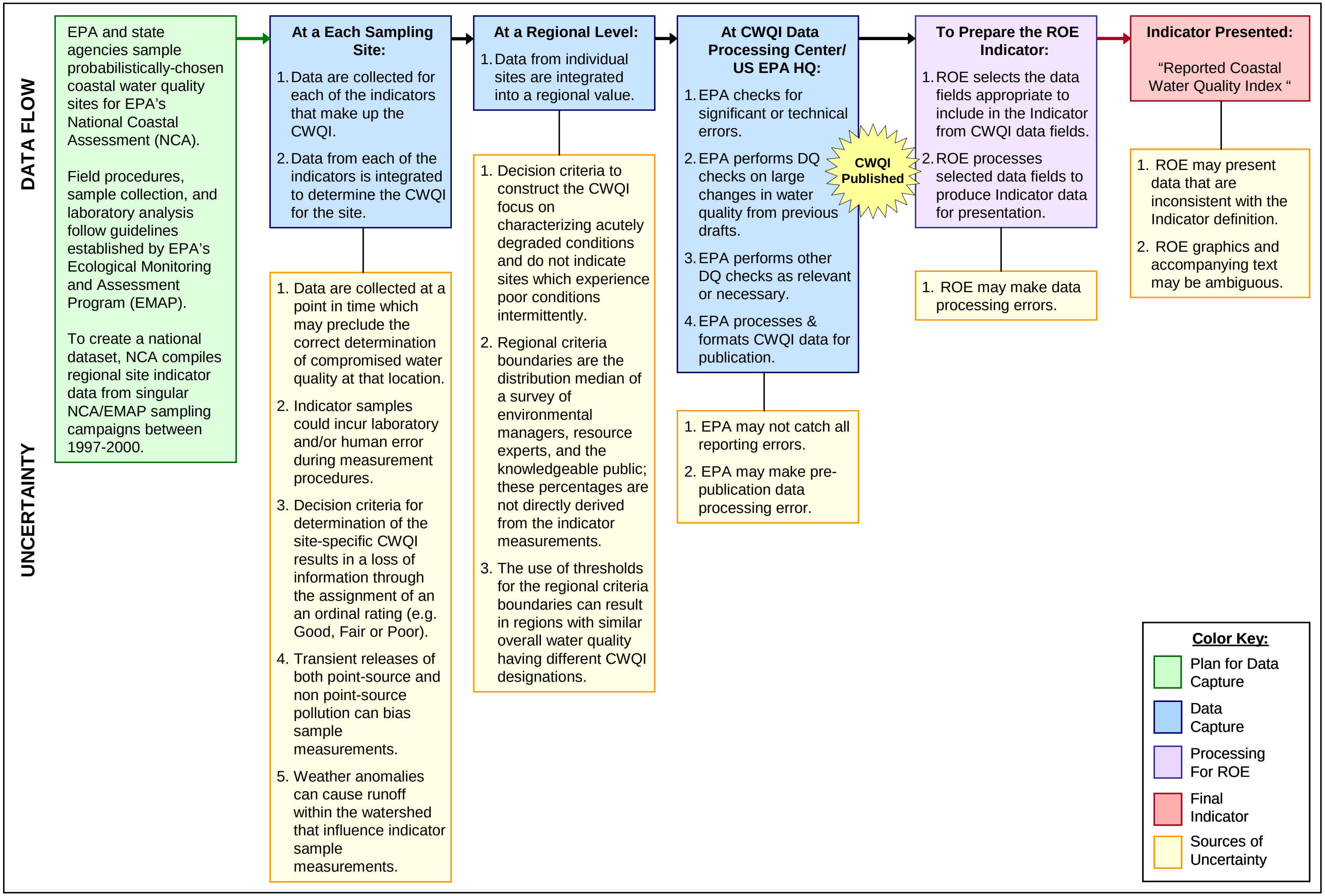


Figure 3-3. Uncertainty and data flow: CWQI composed of water clarity, dissolved oxygen content, organic carbon content of sediments, and chlorophyll concentrations.

design and methodological limitations that can greatly affect the uncertainty associated with EPA's CWQI indicator:

Temporal and Spatial Variability. Aquatic ecosystem processes function across multiple spatial and temporal scales and can be cyclic in nature. These factors create the potential for problems in the sampling activities that are associated with each of the indicators that make up the CWQI. The amplitude of cyclic patterns for rates of ecological processes can be significant and can, therefore, introduce substantial uncertainty to the CWQI indicators. Specifically, the design currently used to obtain the data that comprise the CWQI is constructed to use intra-annual cycles of water quality in an effort to maximize the likelihood that the worst water quality conditions (i.e., maximum trophic state) are detected. If the temporal or spatial variability is such that the worst water quality conditions are missed, which may be highly probable due to only one measurement at most sites, then significant uncertainty can be introduced to the CWQI. Additionally, the fact that these indicators are designed to characterize the worst water conditions needs to be taken into account when interpreting the results at different spatial scales (e.g., site versus national).

- **Transient Weather Events**. Included in this category of internal variability is the influence of weather events across a watershed. For example, anomalous precipitation can interact with nutrient loadings from both point and nonpoint sources to influence water quality downstream. Temperature anomalies can also influence the uptake of applied fertilizers through reductions in primary productivity. This, in turn, can increase the amount of soluble nutrients that are available for runoff into local waterways.

This study design was developed through a process that incorporated input from a number of subject matter experts. Details regarding the principles and implementation of the monitoring design, data collection performed at individual sites, and subsequent analysis can be found through documentation for EPA's EMAP. Nevertheless, additional information regarding these sources of uncertainty is required to accurately characterize and quantify uncertainty in the ROE indicator.

3.2.2. Errors in the Laboratory Analysis

Measurement error that is a consequence of poor data collection techniques can play a substantial role in the analysis of data. This source of uncertainty is easiest to estimate since Quality Assurance (QA) procedures can usually identify and correct errors that are made in both the data-collection and data-processing phases. Although such error can influence the results of data analysis, their potential to substantially influence the CWQI data appears to have been mitigated. To address the potential uncertainty associated with data collection and processing, the protocols outlined in a QAPP guided the data collection. Additionally, we applied the U.S. EPA Data Quality Objectives (DQOs) Framework to ensure that the appropriate type and quantity of data were collected. Within this Framework, the NCA Program developed an a priori, program level DQO for status estimates (see Table 3-3).

Table 3-3. Absolute levels of uncertainty associated with the estimate of proportional area exceeding the indicator criteria (U.S. EPA, 2004b, Appendix A p. 2)[a]

Indicator	NE	SE	Gulf	West	Great Lakes	Puerto Rico	United States
Water quality index	5%	4%	8%	4%	NA	15%	5%
Water clarity	5%	5%	9%	3%	NA	15%	4%
Nitrogen	5%	<1%	5%	3%	NA	14%	3%
Phosphorus	6%	5%	8%	3%	NA	8%	4%
Chlorophyll *a*	5%	4%	9%	4%	NA	14%	5%
Dissolved oxygen	3%	4%	4%	4%	NA	8%	3%

[a]The water quality index is equivalent to the CWQI indicator. Estimates were made a priori by EPA's National Coastal Assessment Program for each region and the U.S. total.

> For each indicator of condition, estimate the portion of the resource in degraded condition within ±10% for the overall system and ±10% for subregions, with 90% confidence based on a completed sampling regime.

As these DQOs do not necessarily define the true statistical uncertainty in the indicator estimates, these values are only estimates and cannot be treated as statistically robust.

3.2.3. Errors due to Data Management, Reporting, and Analysis

An additional source of uncertainty that EPA does not currently address is the loss of information that results from the transformation of continuous data from individual sites (e.g., dissolved oxygen content at a given site) to the categorical indicator that is used to characterize a region (e.g., the dissolved oxygen content in Region 2 is Poor). The potential to introduce error needs to be assessed quantitatively within each of the following steps in the categorical transformation:

- **Data Reporting and Management**. In addition to errors at the time of data entry, unforeseen database issues (e.g., data corruption), and errors associated with other forms of data processing can all contribute to varying degrees of uncertainty in the CWQI. Multilevel steps of QA are performed with each level having its own set of steps that will likely identify errors made in data collection or processing; various forms of these QA procedures are performed at the state, regional, and national level. Examples are completeness checks to ensure data are reported for each station, checking each value recorded against the historical range of parameter values, and consistency checks across data that should possess a certain degree of correlation, as well as comparing data that have been collected by independent methods (e.g., coordinates and water depths). Collectively, these QA procedures should minimize the uncertainty associated with measurement error.
- **Analytical criteria used to transform discrete data to categorical ratings**. At present, the information upon which the CWQI is based is in the form of discretized and aggregated data. Because the CWQI is aggregated hierarchically across spatial units, the sensitivity to changes in a few data points depends on the spatial scale of interest. To some extent, the loss of information that results from transforming the continuous site measurements into categorical regional and national values minimizes the potential impacts of changes in measurements that are not a consequence of actual changes in the indicators. In general, variability that is introduced through errors that are made at the individual site level will exert a relatively greater impact on the regional index than they would on the national index. However, if data processing errors were made while aggregating the regional indices to the national index, then there could be a substantive impact on the national index.

3.2.4. The Impact of Uncertainty on the Report on the Environment (ROE) Indicator

Of the sources of uncertainty that have the potential to affect the CWQI indicator, there are three that can result in changes to the CWQI even if ambient conditions have not changed significantly. Addressed in detail previously, these are

- spatio-temporal variability of ecosystem processes that affect water quality indicators;
- measurement errors associated with collection and processing of data; and
- loss of information due to the transformation of continuous data from individual sites to the categorical indicator that is used to collectively characterize a region.

The potential impact of each of these mechanisms is discussed below.

The indicators that comprise the CWQI have a high level of spatial and temporal variability. The survey design was constructed to identify the water quality in coastal waters during the period from July through September, which is the time period when the most degraded water quality is likely to exist. Because of this, the CWQI indicator is intended, by design, to characterize acutely degraded water quality conditions. The CWQI does not consistently identify the "worst-case" condition for sites experiencing occasional or infrequent hypoxia, nutrient enrichment, or decreased water clarity at other times of the year.

At each sample location, the components of this indicator may have a high level of temporal variability. This survey is intended to characterize the typical distribution of water quality conditions in coastal waters during an index period from July through September. It does not consistently identify the "worst-case" condition for sites experiencing occasional or infrequent hypoxia, nutrient enrichment, or decreased water clarity at other times of the year (U.S. EPA, 2008a).

Further, because of limited sampling, the results cannot capture the full range of water quality variability during this period. It is also questionable that the indicator adequately reflects the average condition during the index period when it was sampled only once in the period 1997–2000.

The impact of measurement errors, which in this context include both the collection and processing of data, has been largely mitigated through the implementation of standardized procedures. These procedures include the DQO process and QAPP implementation.

The CWQI is essentially a function that takes data from each of the sites and maps it to an ordinal categorical variable (i.e., Good, Fair, or Poor) for the region of interest. Table 3-4 presents the rules for aggregating data from site to region. Some information is lost due to the transformation of continuous data from individual sites to the categorical indicator used to collectively characterize a region. Specifically, there are a number of different ways that sites in a region can be assigned a rating of Good, Fair, or Poor.

Table 3-4. Criteria for determining the water quality index rating by region[a] (U.S. EPA, 2004b, Chapter 1)

Rating	**Criteria**
Good/high	Less than 10% of coastal waters are in poor condition, and less than 50% of coastal waters are in combined poor and fair condition.
Fair/good	10% to 20% of coastal waters are in poor condition, or more than 50% of coastal waters are in combined fair and poor condition.
Poor/fair	More than 20% of coastal waters are in poor condition.

[a]Region refers to any geographic area.

If the raw data used for the five indicators were available, this loss of information could be quantified using statistical methods that assess changes in CWQI values and/or a decision endpoint as a consequence of different methods for aggregating the data. One way to assess the uncertainty associated with the CWQI index is to assemble the data used for the ROE determination of the CWQI and construct a log-linear model of the CWQI index as a function of the five indicators (Agresti, 1990). This type of modeling approach would allow for defensible, quantitative estimates of the uncertainty associated with the CWQI. An alternative approach that may be more appropriate, given the hierarchical nature of the variables that influence the CWQI, would be to construct a probabilistic simulation where distributions for each of the variables are estimated using the available data from the ROE. This type of probabilistic approach would allow for the quantification of the importance of the forcing variables in the hierarchical modeling framework that would be built upon the qualitative characterizations of uncertainty presented above. Nevertheless, as data available from the ROE do not include the rated indicators for any regions, such statistical analyses are not possible for this case study.

3.3. REGIONALIZATION

> **Overview: Regionalization**
>
> - Regionalization of the CWQI is obtained by aggregating continuous site data into a nominal, categorical rating (i.e., Poor, Fair, Good) for the region.
> - This aggregation results in some information loss, but in general, identifies regions where sites have experienced decreased water quality.

One of the consequences of aggregation from continuous site data to categorical regional designations is the loss of information. For example, a region where 19% of the sites are rated Poor would be classified as fair, whereas a region with 21% of the sites classified as poor would itself be classified as poor. Despite a relatively insignificant difference in the percentage of sites classified as Poor, the regions would be classified differently on a categorical rating scale, which has only three levels. This problem is somewhat minimized by the used of a five-level classification for some portions of the NCCR. As more levels of classification are used, the approximation of the categorical variable to the continuous variable becomes better, although information is still lost. Nonetheless, Table 3-5 shows this type of a problem, where, for example, in the 2001 report, the NE and the Gulf Coast regions were both rated as Poor although there was a substantial difference in the percentage of area rated as Poor. There are a number of other similar examples in the table, as well as hypothetical scenarios for ratings that are assigned to regions that result in a less-than-clear link back to the values of indicators at the sites within that region. This is essentially a consequence of the nonunique link between possible values of indicators at the sites within a region and the value of the region itself. For example, moving from an indicator rating of 1 to 3 is not the same

Table 3-5. Comparisons of indicators by region for 2001 versus 2004 NCCRs (U.S. EPA, 2004b, Appendix C p. 3)

Indicator	NE Coast		SE Coast		Gulf Coast		West Coast		Great Lakes		Puerto Rico		United States[c]	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Comparison of percent area of poor condition[a] by indicator and region for 2001 vs. 2004 NCCRs														
Water quality index	60	19	13	5	38	9	20	3	–	–	–	9	40	11
Rating scores[b] by indicator and region comparing the 2001 (as published) vs. 2004 NCCRs														
Water quality index	1	2	4	4	1	3	1	5	–	3	–	3	1.7	3.2
Rating scores[b] by indicator and region comparing the 2001 and 2004 NCCRs but calculated with 2004 methods														
Water quality index	1	2	4	4	1	3	1	5	1	3	–	3	1.5	3.2

[a]Percent area of poor condition is the percentage of total estuarine surface area in the region or the nation. Proportional area information is not available for Great Lakes in 2001 or 2004; it is available for selected estuaries in the West Coast in 2001; and in Puerto Rico, it is available only for the 2004 report.
[b]Rating scores are based on a 5-point system, where 1 is Poor, 3 is Average, and 5 is Good (scores for Puerto Rico are available for the 2004 report only).
[c]United States score is based on EPA's calculation of an area-weighted mean of regional scores.

NI = Not included.

as moving from 3 to 5; thus, presenting an average of the regional ratings for the U.S. is misleading. Basically, when information from continuous variables is represented or summarized in a categorical rating, information is lost, and the strength of an analysis, consequently, is decreased.

A careful examination of the rules used to aggregate data from sites to regions clearly shows the emphasis on the identification of the most degraded water quality. Although the intention seems well founded, this type of classification can make it difficult to interpret the ratings. This is especially true when a majority of sampling events for indicators at a given site occurred only once over the 1997–2000 time period. So, a rating of Good could occur in a region where the lowest water quality occurred at a time that was different from when the sampling took place, and a region's water quality could be rated as Poor merely as a consequence of a relatively localized anomaly in water quality that did not persist. Given the focus of the sampling on characterizing the most degraded water quality, the former is certainly more likely than the latter. While it is critical to employ a probabilistic-based sample design in order to generate unbiased estimates of population parameters (e.g., the mean dissolved oxygen concentration across a given spatial region, sampled at locations to represent the probable range of concentrations), the presence of cyclical seasonal patterns in the values of many of the indicators that make up the CWQI could confound attempts to characterize the most degraded water quality conditions.

4. POPULATION SERVED BY COMMUNITY WATER SYSTEMS (CWSs) WITH NO REPORTED VIOLATIONS OF HEALTH-BASED STANDARDS

4.1. DESCRIPTION OF THE CWSs INDICATOR

The ROE introduces the CWSs indicator with the following description:

CWSs, public water systems (PWSs) that supply water to the same population year-round, served over 286 million Americans in fiscal year (FY) 2007 (U.S. EPA, 2007[c])—roughly 95% of the U.S. population (U.S. Census Bureau, 2007). This indicator presents the percentage of Americans served by CWSs for which states reported no violations of EPA health-based standards for over 90 contaminants (U.S. EPA, 2004[c]).

Health-based standards include Maximum Contaminant Levels (MCLs) and Treatment Techniques (TTs). An MCL is the highest level of a contaminant that is allowed in drinking water. A TT is a required treatment process (such as filtration or disinfection) intended to prevent the occurrence of a contaminant in drinking water (U.S. EPA, 2004[d]). TTs are adopted where it is not economically or technologically feasible to ascertain the level of a contaminant, such as microbes, where even single organisms that occur unpredictably or episodically can cause adverse health effects. Compliance with TTs may require finished water sampling, along with quantitative or descriptive measurements of process performance to gauge the efficacy of the treatment process. MCL-regulated contaminants tend to have long-term rather than acute health effects, and concentrations vary seasonally (if at all; e.g., levels of naturally occurring chemical contaminants or radionuclides in ground water [GW] are relatively constant). Thus, compliance is based on averages of seasonal, annual, or less frequent sampling.

This indicator tracks the population served by CWSs for which no violations were reported to EPA for the period from FY1993 to FY2007, the latest year for which data are available. Results are reported as a percentage of the overall population served by CWSs, both nationally and by EPA Region. This indicator also reports the number of persons served by systems with reported violations of standards covering surface water treatment, microbial contaminants (microorganisms that can cause disease), and disinfection byproducts (chemicals that may form when disinfectants, such as chlorine, react with naturally occurring materials in water and may pose health risks) (U.S. EPA, 2004[c]). The indicator is based on violations reported quarterly by states, EPA, and the Navajo Nation Indian Tribe, who each review monitoring results for the CWSs that they oversee [U.S. EPA, 2008f].

The ROE also notes the following limitations of the indicator:

- Noncommunity water systems (typically relatively small systems) that serve only transient populations, such as restaurants or campgrounds, or serving those in a nondomestic setting for only part of their day (e.g., a school, religious facility, or office building), are not included in population served figures.
- Domestic (home) use of drinking water supplied by private wells—which serve approximately 15% of the U.S. population[12]—is not included.
- Bottled water, which is regulated by standards set by the Food and Drug Administration, is not included.
- National statistics based on population served can be volatile, because a single very large system can sway the results by up to 2 to 3%; this effect becomes more pronounced when statistics are broken down at the regional level, and still more so for a single rule.
- Some factors may lead to overstating the extent of population receiving water that violates standards. For example, the entire population served by each system in violation is reported, even though only part of the total population served may actually receive water that is out of compliance. In addition, violations stated on an annual basis may suggest a longer duration of violation than may be the case, as some violations may be as brief as an hour or a day.
- Other factors may lead to understating the population receiving water that violates standards. CWSs that purchase water from other CWSs are not always required to sample for all contaminants themselves, and the CWSs that are wholesale sellers of water generally do not report violations for the population served by the systems that purchase the water.
- Underreporting and late reporting of violations by states to EPA affect the ability to accurately report the national violations total. For example, EPA estimated that between 2005 and 2006, 60% of the data for violations of health-based standards at public water systems was accurate and complete in the Safe Drinking Water Information System-Federal (SDWIS-FED) for all maximum contaminant level and treatment technique rules (excluding the Lead and Copper Rule) (U.S. EPA, 2008j).
- State data verification and other quality assurance analyses indicate that the most widespread data quality problem is underreporting of monitoring and health-based violations and inventory characteristics. Underreporting occurs most frequently in monitoring violations; even though these are separate from the health-based violations covered by the indicator, failures to monitor could mask violations of TTs and MCLs.

The EPA defines PWSs as any drinking water system that serves 25 people or has 15 service connections for at least 60 days a year regardless of whether it is a publicly or

[12] US Environmental Protection Agency. Private drinking water wells. Available at: www.epa.gov/safewater/privatewells/index2.html. Accessed 2009.

privately owned system (U.S. EPA, 2006a). PWSs comprise three types of water systems: CWSs, non-transient non-community water systems (NTNCWSs), and transient noncommunity water systems (TNCWSs). CWSs differ from the other two types of water systems in that they supply water to the same population year round; whereas NTNCWSs and TNCWSs supply water to a population for at most 6 months of the year.

Currently it is estimated there are approximately 55,000 CWSs, and, as of 2004, they served over 272 million Americans, or approximately 92% of the population (U.S. EPA, 2008m, 2004e). The ROE indicator measures the percentage of the United States population served by CWSs with no reported violations of health-based standards. Health-based standards consist of

- **Maximum Contaminant Levels (MCLs):** the highest level at which a contaminant (e.g., heavy metals, microbial contaminants, and disinfection byproducts) can be found in finished (i.e., treated) water (U.S. EPA, 2008m).
- **Treatment Techniques (TTs):** a required treatment process, such as filtration or disinfection, intended to prevent the occurrence of a contaminant in treated tap water, where it is not economically or technologically feasible to measure the level of a contaminant (U.S. EPA, 2008m).

The overall value represented by the indicator reflects the violations (MCL and/or TT) reported to EPA on a quarterly basis by States (including American Commonwealths and Territories) and Tribal drinking water programs for CWSs in which they oversee.

EPA calculates the value represented in the indicator by using data found in the SDWIS/FED, which is the official record of public drinking water systems; the CWSs' violations of state and EPA regulations, and enforcement actions taken by EPA or states that are a result of those violations (U.S. EPA, 2004e).[13] Using these data, EPA identifies the CWSs with reported violations of health-based standards, the population served by CWSs with reported violations of health-based standards, and the total population served by CWSs to obtain the percentage of the population served by CWSs with no health-based standard violations.

We identified Data Capture and Information Processing in SDWIS/FED as the major sources of uncertainty pertaining to the ROE end uses for the "Population Served by CWSs With No Reported Violations of Health-Based Standards" indicator. Because the analysis under this case study is limited to examination of available data, quantitative analysis of information processing errors in SDWIS/FED is used to characterize the uncertainty introduced to the ROE

[13]EPA uses the data underlying the ROE indicator found in SDWIS/FED to calculate several other indicators found in multiple reports to measure the efficacy of EPA's drinking water programs, including: *2006–2011 EPA Strategic Plan*, *EPA's FY 2005 Annual Performance Plan*, *Healthy People 2010*, and *America's Children and the Environment: Measures of Contaminants, Body Burdens and Illnesses*.

indicator. The other major sources of uncertainty, such as data capture inadequacies from unreported water samples, and human error, are characterized qualitatively (see Figure 4-1).

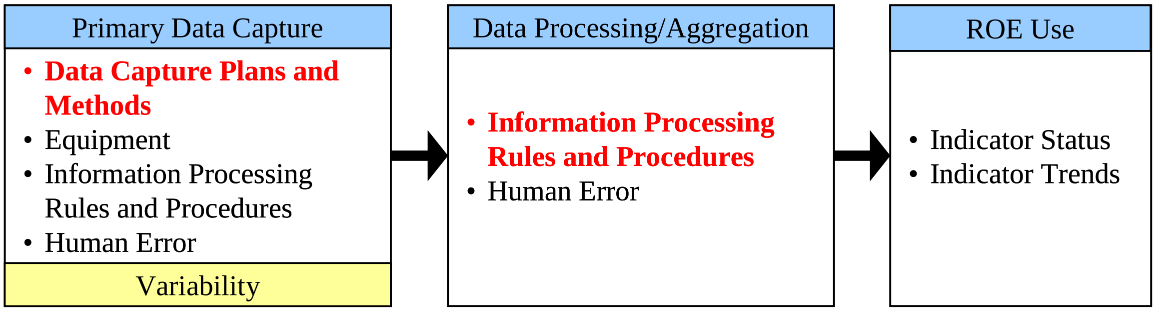


Figure 4-1. General sequence of data capture and processing involved in preparing the population served by community water systems with no reported violations of health-based standards indicator.

4.2. OVERVIEW OF UNCERTAINTY

The data flow that culminates in the ROE indicator presentation, along with elements of uncertainty associated with each step in this data flow, is represented in Figure 4-2. Uncertainty associated with the CWSs indicator, which represents the percentage of the population at the national level served by CWSs with no health-based violations, more than likely occurs because of

- uncertainty and errors in the identification of violations of health-based standards by CWSs and State or Tribal drinking water programs; and
- uncertainty and errors in the reporting of violations of health-based standards by State and Tribal drinking water programs to EPA.

The following discussion describes the areas where errors and uncertainty in the ROE indicator value may arise. The section also includes a calculation of uncertainty using quantitative information, where available. The last part of this section is a discussion of other factors that can contribute to variability in the reported indicator value.

Overview: Uncertainty

- Unreported water samples may result in significant, but unquantified, errors in identifying health-based violations.
- Reporting errors represent the greatest quantified impact on the true value of the ROE indicator. Based on the triennial audits of SDWIS/FED, the percent uncertainty in the indicator is estimated to be 0% to -2% from 1993 to 1998, and 0% to -4% from 1999 to 2004. When the 2006 data quality report is released, uncertainty estimates for 2002–2004 can be revised.

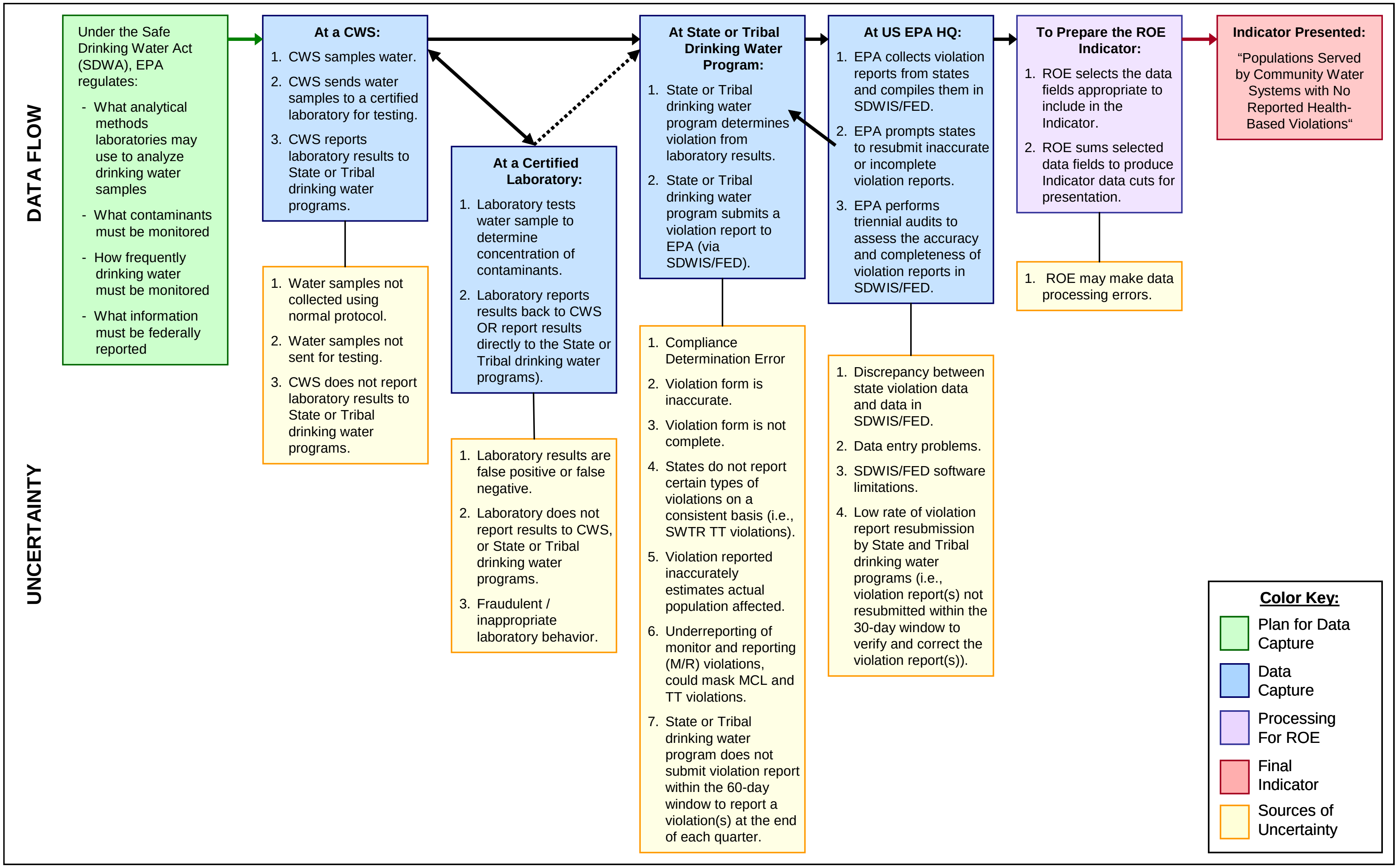


Figure 4-2. Uncertainty and data flow: populations served by CWSs with no reported health-based violations.

4.2.1. Errors in Identifying Health-Based Violations

The process of reporting a health-based violation at a CWS starts with first identifying that a violation has occurred. A CWS monitors water quality at its facility by extracting samples, which are then tested by a certified laboratory. Laboratory results are subsequently reported to an EPA, State, or Tribal drinking water program, where a determination is made whether a violation has occurred.

The first opportunity for uncertainty in the value represented in the ROE indicator arises during the sampling and testing processes. Although the exact percent error associated with incorrectly extracting water samples at the CWS is unknown, and more than likely is extremely low, the possibility of measurement and human error could contribute to an inaccurate representation of the actual population receiving water with no health-based violations. An example of such an error would be incorrectly labeling a sample. Similarly, errors may occur at laboratories that test samples. For example, laboratories may incorrectly identify samples, as false positives or false negatives. However, the likelihood that such an error would significantly alter the true value of the ROE indicator is extremely low.[14] Apart from these forms of human and machine error, uncertainty in the indicator is also associated with the primary data capture—when CWSs do not send water samples to laboratories for testing or when laboratory tests are completed but not sent to the State or Tribal SDWA authority. Since unreported samples do not result in health-based violations, they are not considered a violation in the ROE indicator. However, the fact that such unreported samples could have resulted in health-based violations implies that the ROE indicator may overestimate population served by CWSs with no health-based violations. Recent triennial audits of SDWIS/FED (U.S. EPA, 2006b) indicate that roughly one quarter of those water systems audited did not provide the required samples needed to determine health-based violations.[15] Quantifying the uncertainty associated with missing samples is not conducted in this case study because data on monitoring and reporting violations in SDWIS/FED are not detailed enough to accurately estimate this effect.[16]

---

[14]In 2006, EPA conducted an evaluation of testing procedures in drinking water laboratories (U.S. EPA, 2006d, e). The report did find occurrences of fraudulent and inappropriate laboratory behavior; however, statistics were not available that would allow errors in laboratory testing results to be considered when calculating the overall uncertainty in the ROE CWS indicator.

[15]The *Drinking Water Data Reliability Analysis and Action Plan (2003)* indicates that 49.5% of systems audited had monitoring and reporting violations. Of those violations, almost half were due to situations where a state failed to assess a violation when a system did not sample and the state could not document why it had not assessed a violation. (U.S. EPA, 2004f, p. 16).

[16]GPRA pivot tables of SDWIS/FED data specifies the number of monitoring and reporting violations (M/R violations) for CWS; however, this category includes other types of violations that are not distinguishable from unreported samples. Other violations include failure to report filter malfunctions, failure to perform follow-ups to previous violations, and other procedural violations that may not be related to the required reporting frequency of samples. As a result, assessing uncertainty from all M/R violations would grossly overestimate the degree to which unreported samples affect the ROE indicator values.

4.2.2. Errors in Reporting Health-Based Violations

State and Tribal drinking water programs assess health-based violations and report them to EPA within 60 days of the end of each quarter. Quarterly results are entered into EPA's SDWIS/FED (U.S. EPA, 2006b). State or Tribal drinking water programs are given an additional 30 days to verify and correct information on the submitted violation report(s) (U.S. EPA, 2006b).

Three types of errors may occur when health-based violations are reported to the EPA:

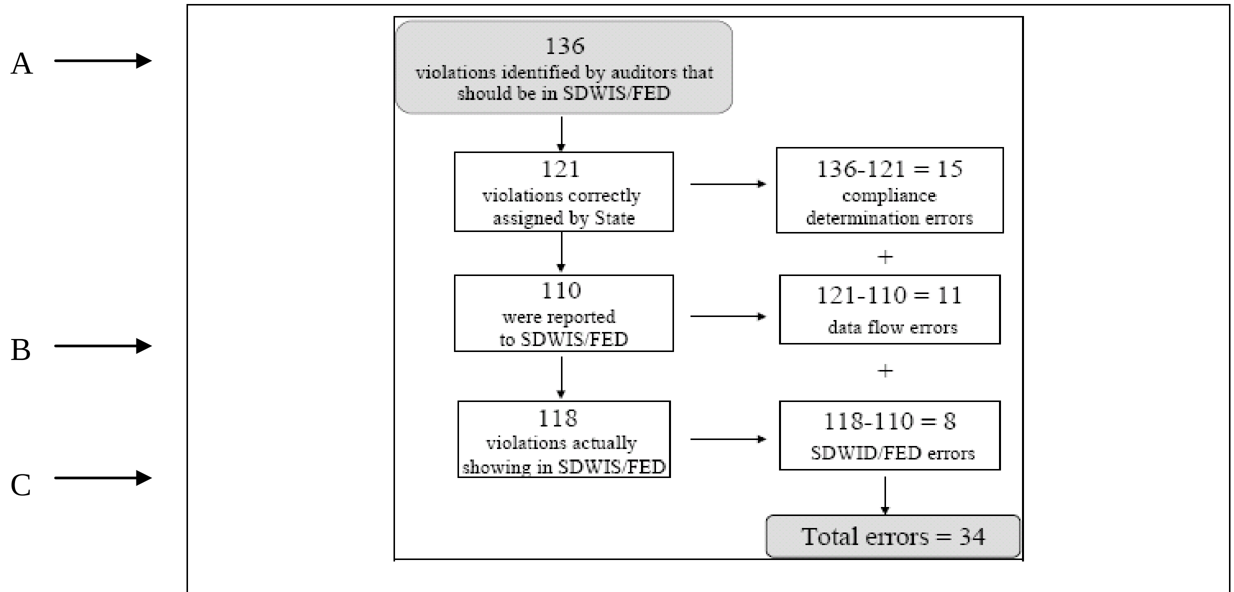
- **Compliance Determination Errors—**When a state fails to cite a violation that should have been assessed.
- **Data Flow Errors—**When the state fails to report a violation that it has correctly assessed to SDWIS/FED.
- **Errors in SDWIS/FED—**Reported violations that should not be in SDWIS/FED, either from assigning a violation where there was none, or failing to remove a rescinded violation. Typographical errors are also included.

Errors that occur during the violation-reporting process are addressed by the EPA in data quality reviews of SDWIS/FED information. EPA undertakes such reviews every 3 years; the first two reviews were conducted in 2000 and 2003, for the periods 1996 to 1998 and 1999 to 2001, respectively.[17] In the reviews, EPA quantifies the number of reporting errors (compliance determination errors, data flow errors, and errors in SDWIS/FED) based on data verification audits conducted by EPA, and presents several related measures of data quality: completeness, accuracy, and SDWIS/FED Data Quality Estimate (DQE). Figure 4-3 explicitly defines and calculates these three measures.

Figure 4-4 presents another graphical representation of these measures (also referencing Figure 4-3).

The measures of completeness and accuracy are important pieces of information for identifying uncertainty in the indicator because they highlight the extent to which violations in SDWIS/FED may be under- or overestimated. Table 4-1 presents the estimated values for these three data quality measures.

[17]The *Data Reliability Analysis of the EPA Safe Drinking Water Information System/Federal Version* (U.S. EPA, 2000b), U.S. EPA (2000b), measured the completeness and accuracy of violations reported by conducting 29 data verification audits in 27 states between 1996 and 1998, which included 1,857 systems. The *Drinking Water Data Reliability Analysis and Action Plan (2003)*, U.S. EPA (2004f), measured the completeness and accuracy of violations reported by data verifications in 31 states between 1999 and 2001, which included 1,890 systems.



	Completeness = % violations that should be in SDWIS/FED that were correct	
B/A =	(correct violations in SDWIS/FED) ÷ (violations that should be in SDWIS/FED)	136 - (15+11) = 110 ÷ 136 = **81% complete**
	Accuracy = % of violations in SDWIS/FED without any errors	
B/C =	(correct violations in SDWIS/FED) ÷ (correct violations in SDWIS/FED + SDWIS/FED errors)	110 ÷ (110 + 8) = **93% accurate**
	Data Quality = % of violations in SDWIS/FED that should be there without any errors	
	1 - (total errors ÷ violations that should be in SDWIS/FED)	1 - (34 ÷ 136) = **75% data quality**

Figure 4-3. Example calculation for violation data quality, synthetic data (U.S. EPA, 2004f).

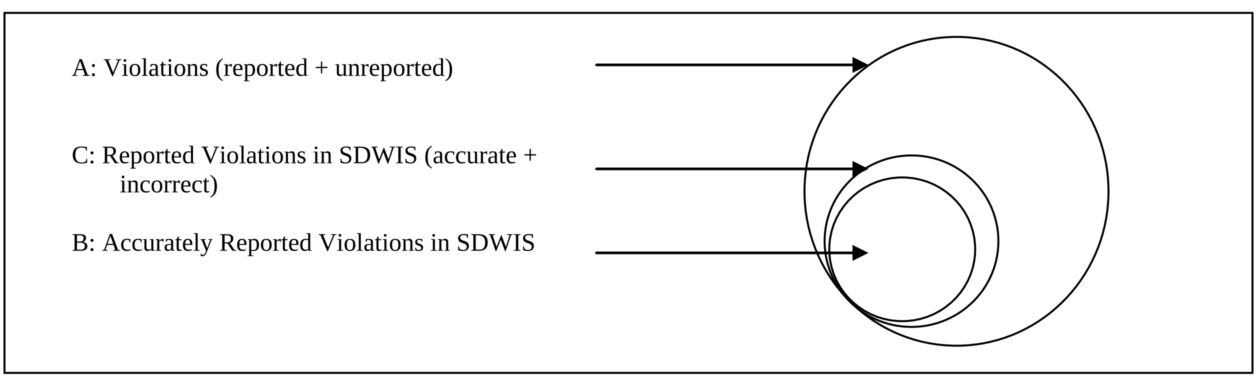


Figure 4-4. Schematic of relationship between violations, reported violations, and accurately reported violations, synthetic data (derived by Abt Associates).

Table 4-1. Calculated 2000 and 2003 SDWIS/FED Data Quality Estimates (DQEs) (derived from U.S. EPA, 2000b, 2004f)[18]

Measure	1996–1999 SDWIS/FED	1999–2001 SDWIS/FED
Completeness	42%	69%
Accuracy	95%	95%
SDWIS/FED DQE	40%	65%

4.2.3. Calculating Uncertainty

In light of the data quality assessments mentioned in the previous section, it is possible to calculate the uncertainty in the ROE indicator by estimating the number of health-based violations present in CWSs (those that are reported and unreported), reconstructing the ROE indicator based on these estimated violations, and comparing the "adjusted" and "unadjusted" indicator values. Specifically, the "adjusted" indicator should indicate the percentage of the population served by CWSs with no health-based standard violations, when accounting for reporting errors (compliance determination errors, data flow errors, and errors in SDWIS/FED).

The reconstruction of the ROE indicator involves the following five steps:

(1) Identifying the number of CWSs, nationally, and the population served by those CWSs.
(2) Identifying the number of CWSs with health-based violations from SDWIS/FED.
(3) Adjusting the number of CWSs with health-based violations in SDWIS/FED to account for reporting errors.
(4) Estimating the population served by CWSs with the "adjusted" number of health-based violations.
(5) Estimating the percent of population served by CWSs with no health-based violations.

One important point in reconstructing the ROE indicator is that the completeness and accuracy measures mentioned in Table 4-1 refer specifically to smaller CWSs. Specifically, the data quality assessments indicated above showed that larger CWSs did not exhibit compliance

[18]Data Quality Estimates are based on audited data from various types of water systems including CWS, NTNCWS, and TNCWS. For the most part, these data are representative of 'small' water systems (10,000 or fewer people served). For example, in the *Drinking Water Data Reliability Analysis and Action Plan (2003),* U.S. EPA (2004f), p.7, only 63 of 1,890 audited systems are considered 'large' (over 10,000 population served).

determination errors (U.S. EPA, 2004f).[19] Therefore, in the reconstruction procedure outlined here, we make a distinction between small and large CWSs, and adjust only the number of health-based violations occurring in small CWSs. The step descriptions below account for this nuance accordingly.

Steps 1 and 2 use data from EPA's published factoids as well as pivot tables containing data collected from the Government Performance and Results Act (GPRA). We segmented the data on the number of CWSs into size categories based on populations served. For this analysis, we used EPA's definition of a small CWSs as one serving fewer than 10,000 people (U.S. EPA, 2006a). Data on population served and the number of CWSs by size category also facilitate the calculation of the average number of people served by small and large CWSs, which is used in later steps.

In Step 3, the number of reported violations (identified in Step 2) is adjusted to reflect findings in EPA's *Drinking Water Data Reliability Analysis and Action Plans*. Specifically, these reports indicate that there are a significant number of small CWSs that do not report violations. In order to adjust the number of small CWSs identified in Step 2, it is necessary to use two measures calculated in EPA's *Drinking Water Data Reliability Analysis and Action Plans*: completeness and accuracy.

In reference to Figures 4-3 and 4-4, Step 2 has provided Item C, Reported Violations in SDWIS. This number should be adjusted to arrive at A, the Number of Violations (reported + unreported).

To estimate Item A, use the definitions:

$$Completeness = B/A\ (\%)$$
$$Accuracy = B/C\ (\%)$$

Transform them to:

$$A = B/Completeness$$
$$B = C * Accuracy$$

[19] In the *Drinking Water Data Reliability Analysis and Action Plan (2003)*, U.S. EPA (2004f), EPA conducted a smaller analysis in which they observed the completeness and accuracy of violations submitted to SDWIS/FED by 30 large systems (>50,000 people) (see p. 19). In this analysis, they found no compliance determination errors. Unfortunately, a statistical analysis could not be performed because regional reporting was not robust. Although the definition of a large system in the *Drinking Water Data Reliability Analysis and Action Plan* (U.S. EPA, 2004f) is slightly different then the one we used throughout the uncertainty discussion in this report, the assumption was expanded to include all large facilities (>10,000 people served).

Now combine the two to make the equation:

*A = (C * Accuracy)/Completeness*

Notice that the result of Step 3 has changed the number of reported small CWSs in violation, identified in Step 2, into the number of estimated small CWSs with violations (reported + unreported). This number, combined with the number of large CWSs with violations, identified in Step 2, gives the total number of CWSs with health-based violations. In the analysis, the number of small CWSs with violations was first adjusted for accuracy, and, then, subsequently, for completeness.

Step 4 uses results from Steps 1 and 3. Using the average population served by small and large CWSs, calculated in Step 1, it is possible to multiply by the corresponding number of CWSs with violations, identified in Step 3, and sum the products to estimate the population served by CWSs with health-based violations.

Step 5 calculates the ROE indicator by subtracting the result of Step 4 from the total population served, identified in Step 1, and then dividing by the total population served, identified in Step 1. This is the percentage of population served by CWSs with *no* health-based violations.

Table 4-2 presents the ROE indicator values for 1993–2004 along with the adjusted ROE indicator values for 1998–2004,[20] while Figure 4-5 presents a graphical representation of both values.[21] The differences between the recreated ROE indicator (at every year) and the reported ROE indicator represent the uncertainty caused by reporting errors.

4.2.4. Additional Sources of Uncertainty

Another source of uncertainty in the ROE indicator arises from potential errors in the reported CWS service populations. While the lack of readily available data regarding the true service populations prevents a quantification of the uncertainty introduced by these estimations, EPA estimates that the addition or removal of a large system from the total number of systems reporting health-based violations can result in a +/−2% change in the indicator value on a yearly

[20] Because DQEs exist for only two assessments (1996–1999 and 1999–2001), the years preceding 1996 use % Completeness and % Accuracy from the former assessment, and the years after 2001 use % Completeness and % Accuracy measures from the latter assessment to adjust the number of health violations.

[21] Although not presented in this analysis, data are available to calculate the ROE indicator for the year 2005.

Table 4-2. ROE indicator values for 1994–2004 and adjusted ROE indicator (U.S. EPA, 2008g; adjusted indicator derived by Abt Associates)

Year	ROE indicator values	Adjusted ROE indicator (accounting for reporting errors)	Percentage difference[a]
1993	79%	79%	0%
1994	83%	81%	−2%
1995	84%	83%	−1%
1996	86%	85%	−1%
1997	87%	87%	0%
1998	89%	89%	0%
1999	91%	91%	0%
2000	91%	90%	−1%
2001	91%	90%	−1%
2002	94%	90%	−4%
2003	90%	89%	−1%
2004	90%	90%	0%

[a]Percentage differences of 0% may be positive values, yet still round to 0% (see Figure 4-5).

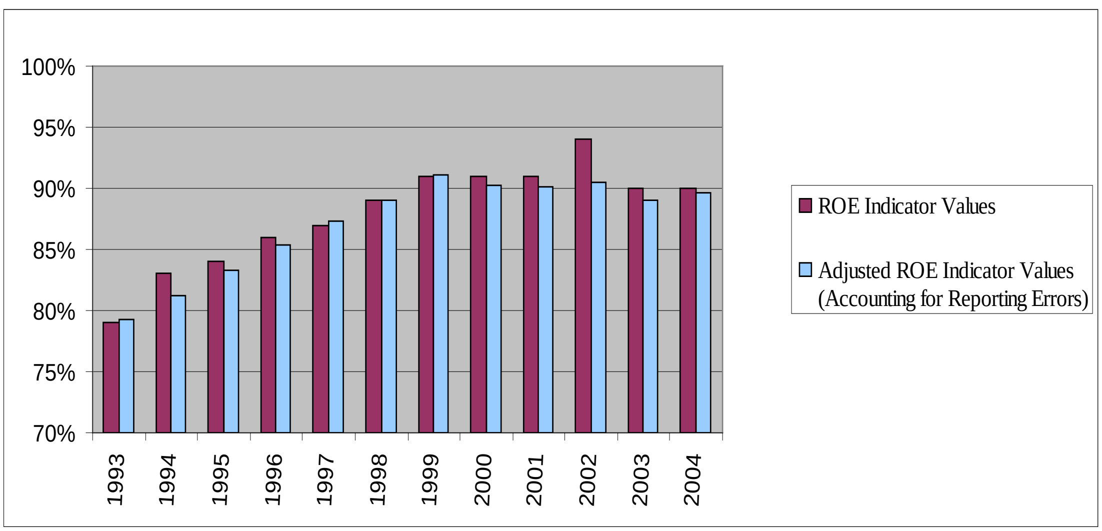


Figure 4-5. Percentage of population served by CWSs with no health-based violations and adjusted estimate (derived from U.S. EPA, 2008g).

basis (U.S. EPA, 2008g).[22] A similar effect on the ROE indicator would not occur in the case of small systems because they serve such a small percentage of the population.

An additional contributor to variability in the ROE indicator value is the introduction of new MCL or TT standards. Up until 2001, CWSs tested for 84 contaminants, but after 2001, the number increased to 91 (U.S. EPA, 2006c, 2004f). The inclusion of new standards over time can potentially shift the indicator value downward as there are more opportunities to fail. As new standards come into effect, the percentage of the population receiving water with health-based violations increases, due to systems' reporting violations of the new standards.[23] In 2003 and 2004, the percentage of the population served by CWSs with no health-based violations for systems complying with only the old standards was consistent with what was observed in earlier years: 91% and 92%, respectively. However, when taking into account systems that reported violations of only the new standards, the indicator values for 2003 and 2004 shifted downward to 90% and 90%, respectively. To account for this variability, associated with changing SDWA program requirements, EPA currently provides two estimates for the ROE indicator: (1) the percentage of the population served with no health-based violations based on the old standards in effect for the entire period presented in ROE (i.e., 1993–2004), and (2) the incremental percentage of the population served with no health-based violations attributed to only the new standards.

4.3. REGIONALIZATION

Overview: Regionalization

- Violation data are available for all CWSs in the United States and can be summarized to the EPA Region and state levels.
- Regions experience varying sensitivity to errors in the ROE indicator due to differences in the distribution of systems with large service populations. Regions 1 and 2 contain several very large systems serving large populations.

EPA requires all state (including American Commonwealths and Territories) and Tribal drinking water programs to monitor CWSs for violations of health-based standards and subsequently report these violations on a quarterly basis. As a result, State and Tribal level data are available to calculate trends in the ROE indicator for all EPA Regions. Currently, EPA's *Factoids: Drinking Water and Ground Water Statistics* reports for FY1998–2004 calculate the percentage of the population served by CWSs with no

[22]Other factors such as distribution of populations served, changes in regulated pollutants, and source water quality all contribute to variability in the ROE indicator as well; however, lack of sufficient data prevents assessments of how these factors affect the ROE indicator.

[23]This effect can be observed when viewing the ROE indicator values for the years 2003 and 2004. Although the compliance date of the new standards was December 31, 2001, the 2002 ROE indicator value did not take into account the impact the new standards had on the indicator value. This could explain why the largest difference between the adjusted indicator value and the actual ROE indicator value was observed for 2002.

health-based violations in each State, in Tribal drinking water programs located in each EPA Region, in American Commonwealths and Territories, and in each EPA Region.

Although the ROE indicator is amenable to regionalization, performing this type of calculation can increase a Region's sensitivity to an error in calculating the population served by CWSs with no health-based violations. In some cases, this may be caused by extremely large systems serving a substantial percentage of the population in a Region.[24] For example, Regions 2, 3, and 9 include multiple systems serving more than one million people located in New York, Maryland, and California, respectively. We expect that these Regions would have similar trends in the indicator over time, given the existence of extremely large systems in each Region. However, this is not the case (see Table 4-3). Unlike Regions 3 and 9, which have fairly high indicator values for 1993–2004, the indicator trend for Region 2 has remained relatively low. More than likely the low indicator value for Region 2 can be attributed to violations being consistently reported for the CWSs located in San Juan, PR, and New York City.[25] Although Region 1 does have one large system serving more than one million people in Eastern Massachusetts, it accounts for only 14% of the total population served in the Region. In Region 1, violations occurring in smaller facilities contribute to the performance of CWSs in the Region. EPA recognizes that these types of observation can be expected when viewing indicators such as the CWS ROE indicator; EPA states in the 2008 ROE Metadata Document that "[n]ational statistics based on population served can be volatile… [and] more pronounced when statistics are broken down at the regional level" (U.S. EPA, 2008[g]). These examples illustrate the importance of considering the distribution of service populations, and large systems in particular, to trends generated at the EPA Regional level and smaller aggregations.

[24]Extremely large CWSs serving greater than one million people serve approximately 29, 20, and 20% of the population in Regions 2, 3, and 9, respectively.

[25]The CWS in San Juan, PR reports a health-based violation almost every quarter (U.S. EPA, 2008k). Over the period 1994–2004, the New York City water system reported numerous TT violations.

Table 4-3. Percentage of population served by CWSs with no health-based violations by EPA Region 1993–2004 (U.S. EPA, 2005a)

Year	Nationality	EPA Region									
		1	2	3	4	5	6	7	8	9	10
2004	90%	93%	80%	85%	93%	96%	92%	92%	92%	86%	93%
2003	90%	89%	54%	95%	93%	95%	93%	93%	92%	98%	93%
2002	94%	88%	81%	98%	96%	94%	93%	95%	97%	99%	91%
2001	91%	65%	77%	98%	95%	92%	96%	90%	94%	97%	83%
2000	91%	62%	76%	97%	95%	95%	96%	95%	94%	94%	83%
1999	91%	75%	61%	98%	95%	95%	95%	95%	94%	97%	94%
1998	89%	64%	60%	97%	95%	95%	95%	94%	93%	95%	89%
1997	87%	62%	55%	97%	93%	92%	93%	95%	91%	95%	74%
1996	86%	60%	53%	92%	93%	92%	94%	95%	92%	91%	74%
1995	84%	57%	52%	91%	92%	92%	88%	95%	90%	88%	75%
1994	83%	57%	55%	87%	90%	88%	87%	94%	91%	90%	87%
1993	79%	60%	56%	85%	90%	77%	92%	93%	92%	69%	85%

5. REPORTED TOXIC CHEMICALS IN WASTES RELEASED, TREATED, RECYCLED, OR RECOVERED FOR ENERGY USE

5.1. DESCRIPTION OF TOXICS RELEASE INVENTORY (TRI) INDICATORS

The release and management of toxic chemicals can occur at every stage in a product's life cycle: extraction and preparation of raw materials, materials processing and refining, product development and use, and disposal and waste management. Approximately 75,000 chemicals are used by industries and businesses in the United States to make the products that our society depends upon. The Toxics Release Inventory (TRI) covers more than 600 specific toxic chemicals and chemical categories listed as TRI reportable due to their acute or chronic health effects and/or significant adverse environmental effects. Facilities that report to TRI are expected to release or manage significant amounts of these toxic chemicals because of their size (as determined by an employment threshold), their industry sector (only certain sectors are required to report to TRI), and chemical use (as determined by activity thresholds).

> Toxic chemicals are contained in waste materials produced by a wide variety of industrial activities, in both public (e.g., sewage treatment plants) and private facilities. These chemical wastes are really a composite matrix of various chemicals, some of which may be hazardous or toxic, and therefore are subject to reporting under the Toxics Release Inventory (TRI) program. Some of these chemicals are released onsite or offsite to air, water, or land (including surface impoundments and underground injection wells). The rest are treated, recycled, or combusted for energy recovery. Reductions in the quantities of TRI chemicals managed at industrial facilities are desirable from both environmental and economic perspectives. TRI chemicals have known toxic properties, rendering them potentially hazardous to workers in both production and waste management facilities, and more generally to ecosystems and human health. As elements of overall business strategies, companies target waste reduction in ways that reduce costs and increase profits.
>
> This indicator tracks trends in the amounts of toxic chemicals in production-related wastes that contain reported TRI chemicals which are either released to the environment or treated, recycled, or combusted for energy recovery. Toxic chemicals in non-production-related waste, such as might be associated with catastrophic events and remedial actions (cleanup), are not included in this indicator because they are not directly related to routine production practices.
>
> TRI contains information on more than 650 chemicals and chemical categories from nine industry sectors, including manufacturing operations, certain service businesses, and federal facilities. Facilities are required to report to TRI if they employ 10 or more employees, are covered by a North American Industry Classification System code corresponding to a TRI-covered Standard Industrial Classification code, and manufacture more than 25,000 pounds, and/or process

more than 25,000 pounds, and/or otherwise use more than 10,000 pounds of a TRI-listed non-persistent, bioaccumulative, toxic (non-PBT) chemical during a calendar year. In addition, EPA has lowered the TRI reporting thresholds for certain PBT chemicals (i.e., to 100 pounds or 10 pounds, except for dioxin and dioxin-like compounds, which have a threshold of 0.1 gram) for certain PBT chemicals and added certain other PBT chemicals to the TRI list of toxic chemicals. These PBT chemicals are of particular concern not only because they are toxic but also because they remain in the environment for long periods of time, are not readily destroyed, and build up or accumulate in body tissue (U.S.EPA, 2002). EPA currently requires reporting of 16 PBT chemicals and four PBT chemical compound categories (U.S. EPA, 2007b). In 2005, 23,500 facilities reported to TRI (U.S. EPA, 2007d).

TRI is national in coverage and includes all U.S. territories. Because the reporting requirements for TRI have varied somewhat between 1998 and 2005 (the most recent year for which annual data reports are available in TRI), chemicals that were reported consistently from year to year over this period are presented separately in this indicator. Facilities that manufacture, process, or otherwise use PBT chemicals have lower reporting thresholds as established in 2000 and 2001; hence these data are depicted separately in the exhibits. Similarly, metal mining sector land releases are analyzed separately because a 2003 court decision altered the scope of TRI reporting of these quantities (U.S.EPA, 2007e, 2008h).

The ROE also notes the following limitations of the indicator:

- TRI data reflect only “reported” chemicals, and not all chemicals with the potential to affect human health and the environment. TRI does not cover all toxic chemicals or all industry sectors. The following are not included in this indicator: (1) toxic chemicals that are not on the list of approximately 650 toxic chemicals and toxic chemical categories, (2) wastes from facilities within industrial categories that are not required to report to TRI, and (3) releases from small facilities with fewer than 10 employees or that manufactured or processed less than the threshold amounts of chemicals.
- TRI chemicals vary widely in toxicity, meaning that some low-volume releases of highly toxic chemicals might actually pose higher risks than high-volume releases of less toxic chemicals. The release or disposal of chemicals also does not necessarily result in the exposure of people or ecosystems.
- Vanadium releases were measured beginning in 2001; because the overall amounts were small relative to the other wastes, they are included in the 2001 to 2005 data for non-PBTs.
- National trends in toxic chemicals in wastes released to the environment are frequently influenced by a dozen or so large facilities in any particular reporting category. These trends may not reflect the broader trends in the more than 23,000 smaller facilities that report to TRI each year.

- Some facilities report off-site transfers for release to other TRI-covered facilities that report these quantities as onsite releases. This double-counting of release quantities is taken into account in the case of release for all sectors in total, but not for releases within individual sectors. This may cause some discrepancy in certain release numbers for specific sectors when compared with release data on all sectors.

This indicator tracks the trends in production-related chemical releases and other waste management of chemicals from the facilities required to report to TRI. This quantity serves as a proxy for the amount of toxic chemicals used in production at TRI reporting facilities, the reduction of which is desirable from both environmental and economic perspectives. The indicator also tracks quantities of TRI chemicals released to air, water, and land and disposed off-site.[26] A reduction in chemical releases, specifically, is particularly desirable from human health and environmental quality perspectives.[27] Note that the scope of this indicator is limited to TRI reporting facilities; it is not meant to represent releases and other waste management at all facilities manufacturing or handling TRI-listed chemicals.

This case study characterizes the uncertainty related to the TRI indicator using available information. Figure 5-1 depicts the general sequence of data capture and processing involved in preparing ROE indicators. We identified the Primary Data Capture step as the major source of uncertainty pertaining to the ROE end use for the TRI indicator. For example, data capture methods and human error were deemed important sources of uncertainty. Historic TRI frozen datasets were readily available, so uncertainty related to human error and evident in late submissions, revisions, and withdrawals of TRI filings was evaluated quantitatively. Quantitative quality assurance studies of data capture, such as audit data representative of the population of TRI filers, were not available. Thus, the frequency and size of errors within TRI submissions could not be estimated quantitatively. If this information becomes available, a more complete characterization of the uncertainty for this indicator could be conducted. Because of the absence of primary data source information, the uncertainty characterization for this aspect of the case study is, therefore, qualitative.

We identified Data capture by individual TRI filers as the major source of uncertainty pertaining to the two primary ROE end uses for the Toxic Chemicals in Wastes Released, Treated, Recycled, or Recovered for Energy Use indicator,. Because the analysis under this case study is limited to examination of available data, we presented a quantitative analysis for filers' revisions, withdrawals, and late filings (which reflect the facility staff's understanding of both

[26]Information about the indicators, their respective metadata, and independent data analyses were extracted and carried out between December 2006 and September 2007.

[27]The text description of the TRI indicators is slightly revised from that presented in the 2007 ROE to reflect the TRI Program Division's terminology and expected interpretation/use of the indicator.

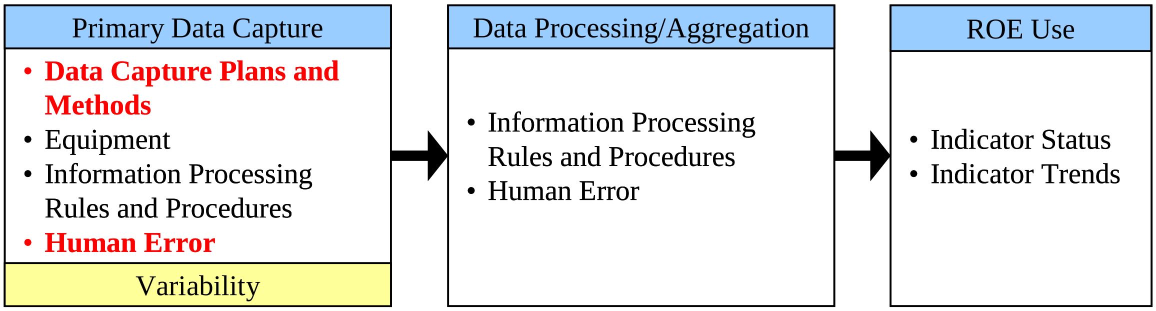


Figure 5-1. General sequence of data capture and processing involved in preparing the report toxic chemicals in wastes released, treated, recycled, or recovered for energy use indicator.

the TRI reporting requirements and the precision of the selected estimation techniques). There are no recent independent assessments of the quality of the estimates facilities submit to the TRI program. The other major sources of uncertainty, such as information processing, are characterized qualitatively. In addition, a qualitative assessment of the likelihood that the identified sources of error could affect the value of the aggregated ROE indicator is also presented.

5.2. OVERVIEW OF UNCERTAINTY

The data flow that culminates in the ROE indicator presentation, along with sources of uncertainty associated with each step in this data flow, are represented in Figure 5-2. There are two types of sources of uncertainty in the TRI-based ROE indicators:

- Uncertainty and errors in the estimates and filing decisions made by staff at facilities required to report to TRI.
- Uncertainty and errors resulting from systems involved in generating or processing the TRI data, which would affect a large proportion of the TRI records and/or the summed values presented by ROE. The sources for these types of potential errors include problems with the filing software, EPA's data processing, and EPA's presentation of the TRI data for use by ROE.

Overview: Uncertainty

- TRI data will change after publication due to facility-submitted <u>late filings, revisions, and withdrawals</u>. National TRI totals are projected to change between -1.1 and +1.8% in the first year after TRI filings are submitted, increase up to +0.4% in the second year, and increase up to +0.5% in the third year based on historical data.
- Because this indicator is a simple sum of values, it is sensitive to <u>errors in large value records</u> (or missing records). While the TRI quality assurance program explicitly focuses effort on large value records, no current information exists to quantify the frequency and size of errors within this subpopulation.

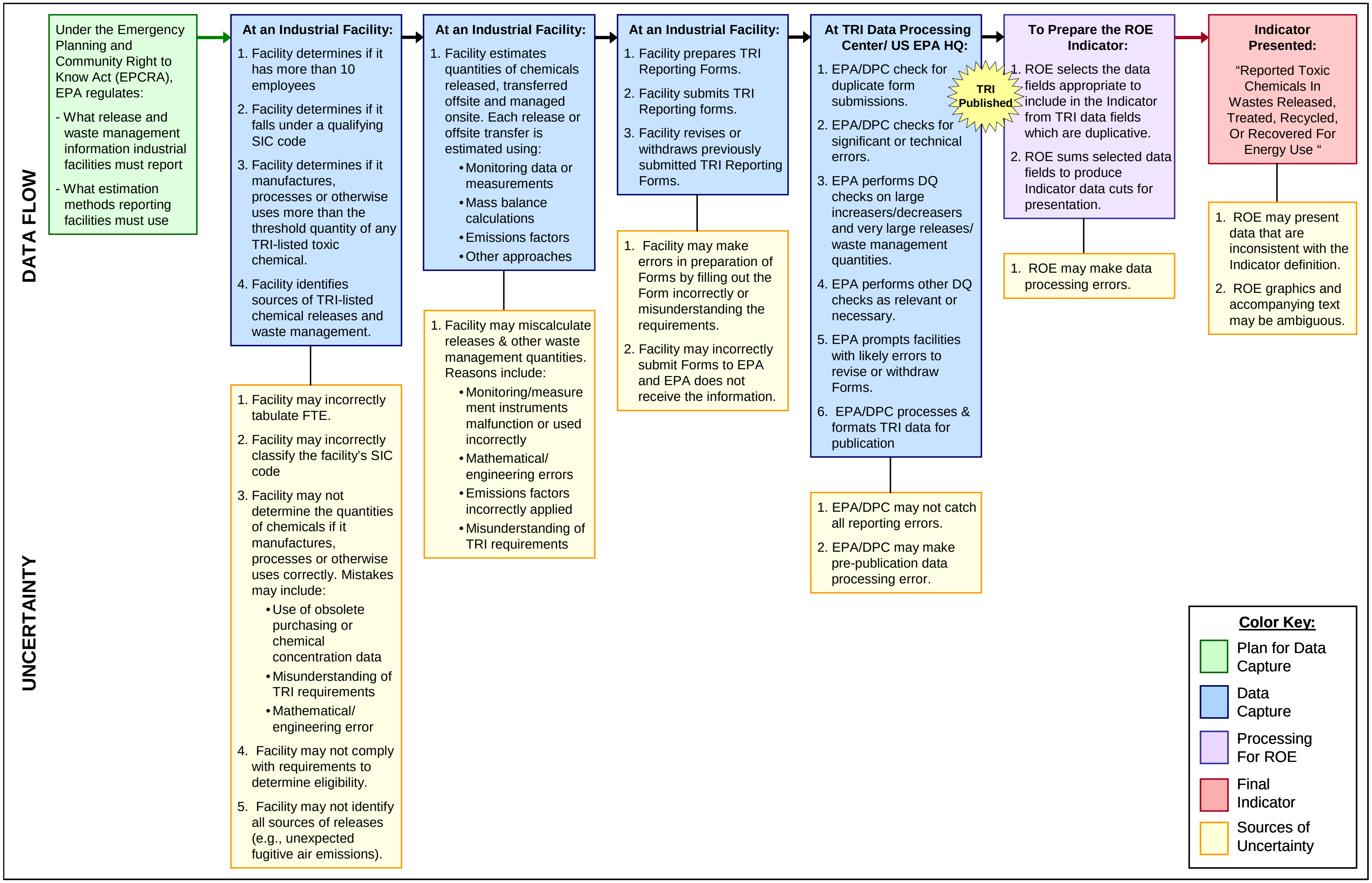


Figure 5-2. Uncertainty and data flow: reported toxic chemicals in wastes released, treated, recycled, or recovered for energy use.

The uncertainty associated with the TRI indicators, which are arithmetic sums at the national level, is composed of two parts: the magnitude of the errors in the reported TRI values and the likelihood of a given error occurring. As discussed below, only one source of uncertainty is likely to have a significant impact on the ROE TRI indicators—that is, large errors in the reported amount of release or total production-related waste fields. The TRI Program Division's data quality efforts flag potential "large" and "very large" errors each reporting year (RY)—filers are notified, and the vast majority of those determined to be in error are revised before the data are published and integrated into the ROE.[28]

5.2.1. Facility-Level Errors

The likely impacts of facility-level errors are difficult to accurately estimate because facilities use a diverse array of methods and information sources when preparing their TRI reports. Errors can occur at any of several points in the reporting process:

- Facility Coverage Determination—determining the number of full-time employees and the primary Standard Industrial Classification code.
- Chemical Threshold Determination—estimating the quantity of each TRI-listed chemical manufactured, processed, or used at the facility in a given year.
- Estimation of Chemical Releases and Other Waste Management Quantities—estimating the quantities of each TRI-listed chemical released to the environment or otherwise managed.
- Preparation and Submission of TRI Reporting Forms—filling out the Forms R and/or submitting them to EPA for incorporation into the database.

Some information is known about the frequency of these types of errors from TRI site visits and data quality efforts. For example, EPA has conducted voluntary, confidential site surveys episodically in the past to identify TRI reporting problems and causes. These investigations have included reviews of the accuracy of threshold determinations, release estimates, and other waste management calculations. It must be noted, however, that some of these reports are now (as of 2007) more than 10 years old, and may no longer be representative of the methods and information sources used by current TRI filers. The findings of these reports,

[28] While not in place for the time period analyzed in this case study, the TRI program has also begun a program with the regions under which a minimum of 500 facilities (about 2.5% of the reporting universe) will be evaluated each year for data quality checks. This is a non-random sample based on information the TRI program has developed to identify potential errors. Regions review facility information and contact facilities as necessary to determine if revisions or other action (e.g., enforcement referral) are necessary.

therefore, may not be representative of the types and magnitude of errors generated by current TRI filers.

The earlier of these site visit investigations, designed to provide a quality assessment of data at the national level, were completed in the early to mid-1990s. They found error rates of $\pm$4% in the United States total values (e.g., releases). The more recent of these investigations were undertaken in the late 1990s and were designed to examine data quality concerns in specific sectors. The sectors surveyed include SICs 25, 281, 285, and 30 (RY94); 26 and 286 (RY95); and 33, 36, and 37 (RY96). Although these studies may not be representative of all TRI filers, they do provide some interesting quantitative assessments of TRI data quality, including

- Depending on the industry, facilities correctly determined their threshold quantity 84–98% of the time.
- Among facilities that did not correctly determine their threshold quantity, facilities were more likely to submit TRI forms despite not exceeding the threshold than they were to not submit TRI forms when they did exceed the threshold.
- Facilities often correctly identified release and other waste management activities that were occurring but reported the wrong type (e.g., fugitive versus stack air release).
- Recycling, both on and off site, was frequently misclassified due to confusion over the definitions of recycling and reuse.
- The percentage difference between facility and site surveyor estimated Total Release and Other Waste Management Quantities was −6.7% for RY94 facilities surveyed (the TRI total was 6.7% less than the auditors' values), −1.2% for RY95 facilities surveyed, and −28% for RY96 facilities surveyed.

These results of the site visits were not intended to be representative of TRI reporting facilities, or even of the sectors studied. For example, in 1996, EPA visited 60 facilities of the more than 3,500 in the sectors of interest. Rather, the site visits were targeted efforts to identify ways to improve EPA's TRI compliance assistance.

Every year, EPA targets at least some TRI filers with potentially large errors in their reporting. In RY00, approximately 14% of the facilities targeted for data quality checks by EPA due to large increase/decrease and/or large waste/release quantities were found to have at least one error in the reported value. In aggregate, approximately 2.5% of all forms submitted to TRI in a given year are later revised and 0.5% are withdrawn, due both to facilities' own reviews, and prompts from EPA's data quality reviews. Revisions and late filings can result in an increase or decrease in the number of pounds of toxic substances released, treated, recycled, or recovered (but not all revisions represent actual changes in pounds—some involve administrative data). The effects of revisions and withdrawals made so far to RY01–RY04 submissions have resulted in a net +1 to −5% change in the national aggregated Total Production-Related Waste quantities for the RYs covered by ROE.

The removal of self-identified errors via the withdrawal process can be expected to introduce uncertainty into the TRI indicator trends. Examining patterns of withdrawals and revisions show that TRI data for more recent RYs likely includes forms that will eventually be revised, withdrawn, or added because of late submissions. Most of these changes occur within the first year of publication, as shown in Table 5-1. Therefore, the initial errors are revised and would not be incorporated into the ROE indicator, which is compiled after the data are "frozen" for the most recent year TRI data are available. For example, in RY02, Red Dog Mine in Alaska withdrew nearly 1 billion pounds in releases due to a reporting error. This withdrawal altered the total production-related waste reported by nearly 3% but was made prior to the RY02 data being published in spring of 2003. In the histogram below (see Figure 5-3), the anticipated change in yearly total production-related waste between the publication of the RY04 and RY05 data sets is represented by the red error bars, as projected using the average change in pounds over time observed between the earlier versions of the published TRI data set.

Table 5-1. Changes to TRI form submissions by time elapsed since Public Data Release (PDR) publication, non-PBTs (U.S. EPA, 2004g)

Forms changed per 1,000 submitted			
	1 Year	**2 Years**	**3 Years**
Withdrawals	0.5	0.2	0.1
Revisions	9.1	4.0	2.4
Late filings	9.0	5.5	4.1
TPRW pounds increased/decreased per million reported			
	1 Year	**2 Years**	**3 Years**
Withdrawals	−13	−106	−265
Revisions	216	176	266
Late filings	3,188	2,580	5,340

TPRW = total production related waste.

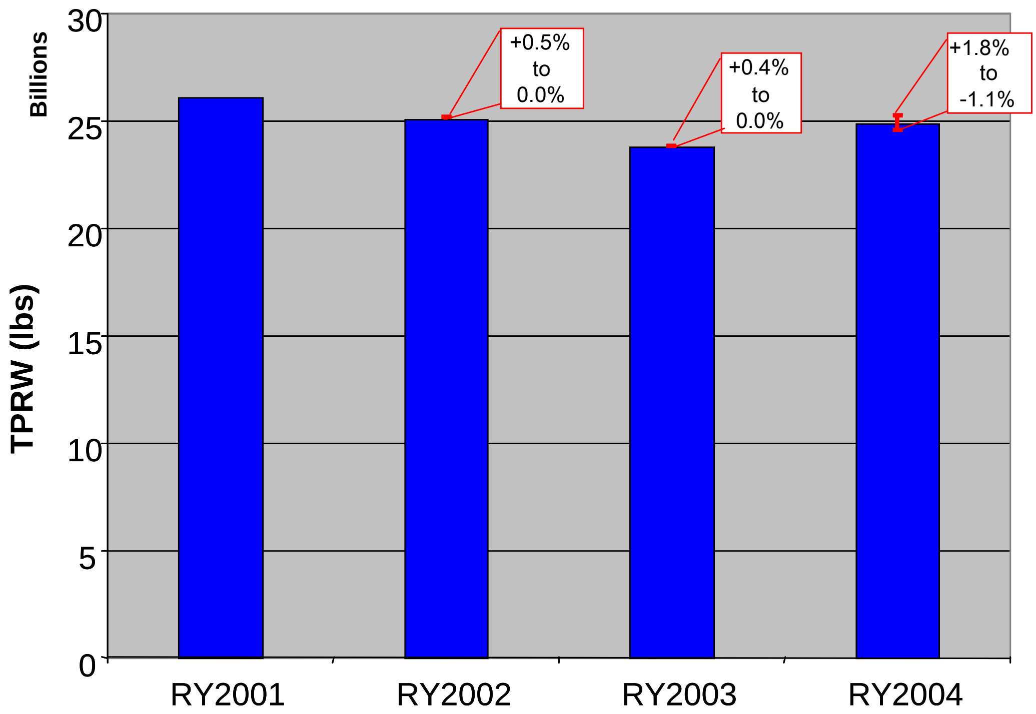


Figure 5-3. Projected change in TPRW based on historical revisions (U.S. EPA, 2004g).

5.2.2. Systemic Errors

The capture, processing, and presentation of TRI data by EPA have the potential to contribute significant uncertainty to the aggregated ROE indicator. If EPA makes a data processing error while reading in or preparing the TRI data for publication, the impacts could affect many records in the TRI. If present, this type of error could significantly change the total pounds reported in the TRI indicators. Likewise, if data fields are misinterpreted, omitted, or records are double-counted in the aggregation of the TRI indicator, the ROE indicator could contain significant errors. While the potential[29] impacts of these sources of uncertainty are very large, such errors are unlikely to occur. At the EPA-level, duplicative data quality checks make the identification and remedy of such errors inevitable.

The following tables show the distributions of total production-related wastes (see Table 5-2) and releases (see Tables 5-3 and 5-4) among non-PBT forms submitted to TRI for RY03 and RY04. The three tables mirror the data presentations used in ROE for this indicator.

[29]The TRI Program's various quality assurance efforts, such as electronic reporting including use of Central Data Exchange and Web submissions, built-in data quality logic in reporting software, and more recently, electronic facility data profiles, automation of revisions, and withdrawals reduce the likelihood and impacts of systemic errors.

Table 5-2. Distribution of TPRW among non-PBT Forms (RY03–04 average) (derived by Abt Associates from RY04 TRI PDR [U.S. EPA, 2004g])

TPRW range of pounds	Avg. forms/yr	Percent of forms	Percent of TPRW
0–10	3,544	5.6%	0.00002%
10–100	2,649	4.2%	0.0005%
100–1,000	7,108	11.3%	0.01%
1,000–10,000	13,189	20.9%	0.2%
10,000–100,000	23,340	37.1%	3.5%
100,000–1,000,000	10,157	16.1%	12.9%
1,000,000–10,000,000	2,597	4.1%	32.1%
10,000,000–100,000,000	360	0.6%	31.5%
100,000,000–1,000,000,000	21	0.03%	19.8%

Table 5-3. Distribution of releases among non-PBT Forms—no metal mining (RY03–04 average) (derived by Abt Associates from RY04 TRI PDR [U.S. EPA, 2004g])

Releases range of pounds	Avg. forms/yr	Percent of forms	Percent of releases
0–10	10,734	17.0%	0.0004%
10–100	6,534	10.4%	0.006%
100–1,000	13,822	22.0%	0.15%
1,000–10,000	14,447	22.9%	1.4%
10,000–100,000	13,126	20.8%	11.3%
100,000–1,000,000	3,684	5.9%	25.8%
1,000,000–10,000,000	581	0.9%	37.0%
10,000,000–100,000,000	29	0.0%	12.6%
100,000,000–1,000,000,000	2	0.003%	11.6%

Table 5-4. Distribution of releases among non-PBT forms—metal mining only (RY03–04 average) (derived by Abt Associates from RY04 TRI PDR [U.S. EPA, 2004g])

Releases range of pounds	Avg. forms/yr	Percent of forms	Percent of releases
0–10	35	8.3%	0.00001%
10–100	22	5.2%	0.0001%
100–1,000	36	8.6%	0.00%
1,000–10,000	72	17.3%	0.0%
10,000–100,000	104	24.9%	0.5%
100,000–1,000,000	100	24.0%	4.5%
1,000,000–10,000,000	40	9.5%	15.7%
10,000,000–100,000,000	7	1.7%	21.7%
100,000,000–1,000,000,000	2	0.5%	57.6%

As shown in Table 5-2, over 50% of the pounds of total production-related waste reported are represented by less than 1% of the forms submitted.

From a probabilistic perspective, only errors in forms or facilities with large TPRW amounts have the potential to alter the indicator value or trend at the national level. For this reason, EPA focuses its data quality efforts on facilities reporting large year-to-year increases and decreases. Given the rate of form revision after the TRI data are first published is 0.9% in the first year after TRI public data release (PDR) and is less for earlier years, the chance that the ROE indicator or the trend is significantly different than its value after all facility-based changes occur is unlikely. As the Red Dog Mine example illustrates, and based on program records for the past 15 years, such a large error is rare, but facility error is the most likely source of uncertainty that could significantly affect the TRI indicator.

5.3. REGIONALIZATION

The primary purpose of TRI is to inform communities and citizens of chemical hazards in their areas. As such, TRI reporting is required for all facilities that meet the reporting thresholds throughout all United States and territories. This results in data that are very amenable to regionalization and scaling.

Generating the EPA Regional cuts of the TRI indicators, however, brings up what will be an important issue for this indicator—how off-site releases are calculated. For the national ROE indicator, EPA removes those transfers that are sent to hazardous waste treatment storage and disposal facilities (TSDFs) that are themselves reporting to TRI. While there is a time lag, which may affect the year in which chemical releases are reported by the TRI shipper and received by the TSDF, this approach avoids double counting some off-site releases. For regional trends, this method may need to be modified to account for inter-regional transfers of hazardous wastes, which may be a substantial percentage of the total regional values.

For example, in Region 3, off-site releases totaled 71.3 million pounds in 2004. Nationally, between 10–20% of off-site releases are removed to avoid double counting amounts also reported by TRI-reporting TSDFs. In Region 3, however, there does not appear to be significant double-counting. Onsite releases at TSDFs (SIC 4953 and 7389) totaled 1.5 million pounds, but more than 99% was associated with non-Resource Conservation and Recovery Act (RCRA) surface impoundment disposal at one facility (Max Environmental in Yukon, PA; 15698MLLSRCEMET), whose releases may include shipments from non-TRI facilities in the Region and facilities in other Regions. One solution to this issue that mimics the logic and meaning of the national ROE indicator is to exclude chemicals shipped for treatment, storage, or disposal (some of which is calculated to be released) within the Region but include off-site shipments outside of the Region.

Overview: Regionalization

- TRI data are reported by all covered facilities throughout the United States.
- Some processing steps applied to the national-level TRI indicator may require re-evaluation at the Regional level. For example, for the national ROE indicator, EPA removes those transfers that are sent to hazardous waste TSDFs that are themselves reporting to TRI. For regional trends, this method may need to be modified to account for inter-regional transfers of hazardous wastes, which may be significant.
- Regional trends are driven by the largest reporters (e.g., metal mines in Region 8).

Regionalization of the TRI indicators also brings to light a somewhat confusing aspect of the presentation of the ROE indicator that might be remedied. Indicator presentations include those for both total production-related waste (TPRW; includes waste management activities) and total releases. In the presentation of total production-related waste, releases and waste management quantities from the metal mining sector are included. The first release table is titled: Releases of Chemicals (without metal mining and PBT chemicals), but only Land Releases are labeled as having metal mining values removed. The subsequent metal mining table for releases presents all releases: air, water, land, and off-site.

When the indicator data are scaled to a Regional level, the inconsistencies in these presentations become more apparent. The TPRW data in Regions with significant metal mining activities are dominated by the metal mining land releases. If all metal mining releases are removed from the presentation of TPRW and the first releases table, the ROE indicator data might be more straightforward.

Trends at the Regional level tend to be heavily influenced by the trend among the top reporters, as shown in Table 5-5. The Regional profile of waste management (how quantities are split between releases and other waste management) can be distorted by these largest reporters, as well. For example, in Region 10, the top reporter—Red Dog Mine in Kotzebue, Alaska—released over 300 million pounds of non-PBT chemicals in 2005 and did not report any quantities of waste recycled, recovered, or treated. As a result, Region 10 is the only EPA Region with more than 50% of TPRW being released.

Table 5-5. Regional impact of top facilities (derived by Abt Associates)

EPA Region	Percent of TPRW from top five facilities
1	24%
2	34%
3	27%
4	19%
5	29%
6	27%
7	29%
8	56%
9	40%
10	57%
All United States	10%

6. MIGRATION OF CONTAMINATED GROUND WATER UNDER CONTROL AT HIGH-PRIORITY CLEANUP SITES

6.1. DESCRIPTION OF MIGRATION OF CONTAMINATED GROUND WATER (GW) UNDER CONTROL AT HIGH-PRIORITY CLEANUP SITES INDICATOR

The ROE introduces the Migration of Contaminated Ground Water Under Control at High Priority Clean Up Sites indicators with the following description:

> The EPA Superfund and Resource Conservation and Recovery Act (RCRA) Programs conduct a number of activities to address the nation's most severely-contaminated lands. The Programs investigate and collect data on potentially-contaminated sites to determine whether they are contaminated and require cleanup. When a potentially-hazardous waste site is reported to EPA, trained inspectors determine whether the site presents a hazard to human health and the environment. Sites that pose the greatest threat are placed on the National Priority List (NPL) or RCRA Corrective Action high-priority list.
>
> One of the priorities for both the NPL and the high-priority Corrective Action sites is preventing the continued spread of contaminated ground water [GW], often referred to as plumes of contaminated ground water. Protecting the ground water is especially important in those areas where ground water is the primary source for drinking water and irrigation, and where the population could be exposed to contaminants (e.g., vapors).
>
> EPA and state officials determine that the migration of contaminated ground water is not continuing above levels of concern when ongoing monitoring shows that the contaminant plume is not expanding or negatively impacting surface waters (U.S. EPA, 1999[a]). Preventing further migration of contaminated ground water may result from an action taken, such as installation of a pump and treat or subsurface barrier system, or because of natural attenuation of the contaminants. A determination of whether migration has been prevented is based on monitoring data (usually hundreds of analytical samples) collected from ground water wells located within and surrounding the spatial extent of the ground water plume (U.S. EPA, 1999a, 2008j).
>
> This indicator describes the percentage of NPL and RCRA corrective action sites where government officials have determined that ground water is not continuing to spread above levels of concern (i.e., that exceed the appropriate drinking water standards). This indicator covers both final and deleted NPL sites, and all 1,714 RCRA corrective action sites on the Government Performance and Results Act (GPRA) Cleanup Baseline [CB]. The percentage of sites where ground water contamination continues to spread is also noted, as well as the number of sites where there are insufficient data to make a finding …. The intention of the indicator is not to capture an action or administrative determination on the part of EPA, but to convey the underlying pressure on the environment and potential for

human health effects resulting from contaminated ground water [U.S. EPA, 2008i].

The ROE also notes the following limitations of the indicator:

- The NPL does not represent all of the contaminated or potentially contaminated sites listed in the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) database, which contains information on thousands of hazardous waste sites, potential hazardous waste sites, and remedial activities across the nation. A small percentage (less than 1%) of the total number of final and deleted NPL sites are excluded from the NPL Indicator Baseline for reasons of consistency.
- The indicator covers the 1,714 RCRA Cleanup Baseline sites tracked from 2000 to 2005 and the 1,968 sites tracked in 2006 and 2008, and not the entire group of 3,746 hazardous waste management sites currently believed to need cleanup or investigation under the RCRA Corrective Action Program (i.e., initial assessments, and if needed more thorough investigations and cleanups).
- The extent to which people have been affected, or could be affected, by the contaminated ground water at NPL or RCRA Cleanup Baseline sites is not considered in this indicator, but is addressed in the Current Human Exposures Under Control at High-Priority Cleanup Sites indicator.
- The indicator does not address ground water contaminated at other types of sites, such as sites with leaking underground storage tanks and other sites being addressed solely by state cleanup programs.
- Concentrations of toxic and hazardous contaminants in ground water that must not be exceeded to designate a site as under control vary somewhat from state to state, though they fall within a range determined to be acceptable to EPA (U.S. EPA, 2008l).
- This indicator is based on the certification by a responsible official that the criteria necessary to designate whether contaminated ground water is continuing to spread above levels of concern have been met (U.S. EPA, 1999a, 2008l). Trends in the number of sites where the spread of contaminated ground water has been shown to occur above levels of concern may be underestimated to the extent that certification lags behind the migration of contaminated ground water or certification is delayed due to insufficient or outdated information.

This case study characterizes the uncertainty related to the Migration of Contaminated Ground Water Under Control at High-Priority Cleanup Sites indicator using available information. Figure 6-1 depicts the general sequence of data capture and processing involved in preparing ROE indicators. We identified the Information Processing Rules and Procedures step as a key source of uncertainty pertaining to the ROE end use for the Migration of Contaminated Ground Water Under Control at High-Priority Cleanup Sites indicator. For example, differing site determination criteria used by RCRA and the Superfund Programs means

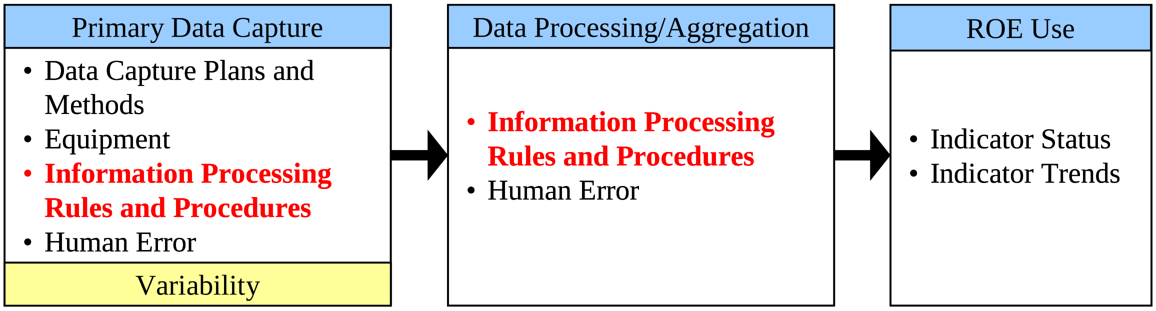


Figure 6-1. General sequence of data capture and processing involved in preparing the migration of contaminated ground water under control at high priority cleanup sites.

that "migrating groundwater" is not defined in the same way by the two programs (U.S. EPA, 1996). Because of these dissimilar reporting requirements, cross-program analysis or aggregation is not possible without introducing additional uncertainty or controlling for this definitional difference.[30] Within each program, however, well defined site characterization and QA protocols result in indicator values with minimal uncertainty. Because the analysis under this case study is limited to examination of available data, uncertainty from varying programmatic definitions is characterized qualitatively, while a brief quantitative analysis is presented to demonstrate the inclusion uncertainty contributed by RCRA's cumulative programmatic design.

6.2. OVERVIEW OF UNCERTAINTY

Overview: Uncertainty

- Differences in programmatic definitions preclude cross-program analyses and aggregation, as additional uncertainty would be introduced.
- Within each program, well defined site characterization and QA protocols result in indicator values with minimal uncertainty.

Figure 6-2 represents the data flow that culminates in the ROE indicator presentation, along with elements of uncertainty associated with each information-processing step. The GW Contamination ROE indicator has several potential sources of uncertainty:

- Monitoring and analysis used to characterize GW at RCRA corrective action and the Superfund sites.
- Determination of GW status.
- Program definitions affecting GW status determination, or the population of sites included in the ROE indicator.

[30] Information about the indicators, their respective metadata, and independent data analyses was extracted and analyses carried out between December 2006 and September 2007.

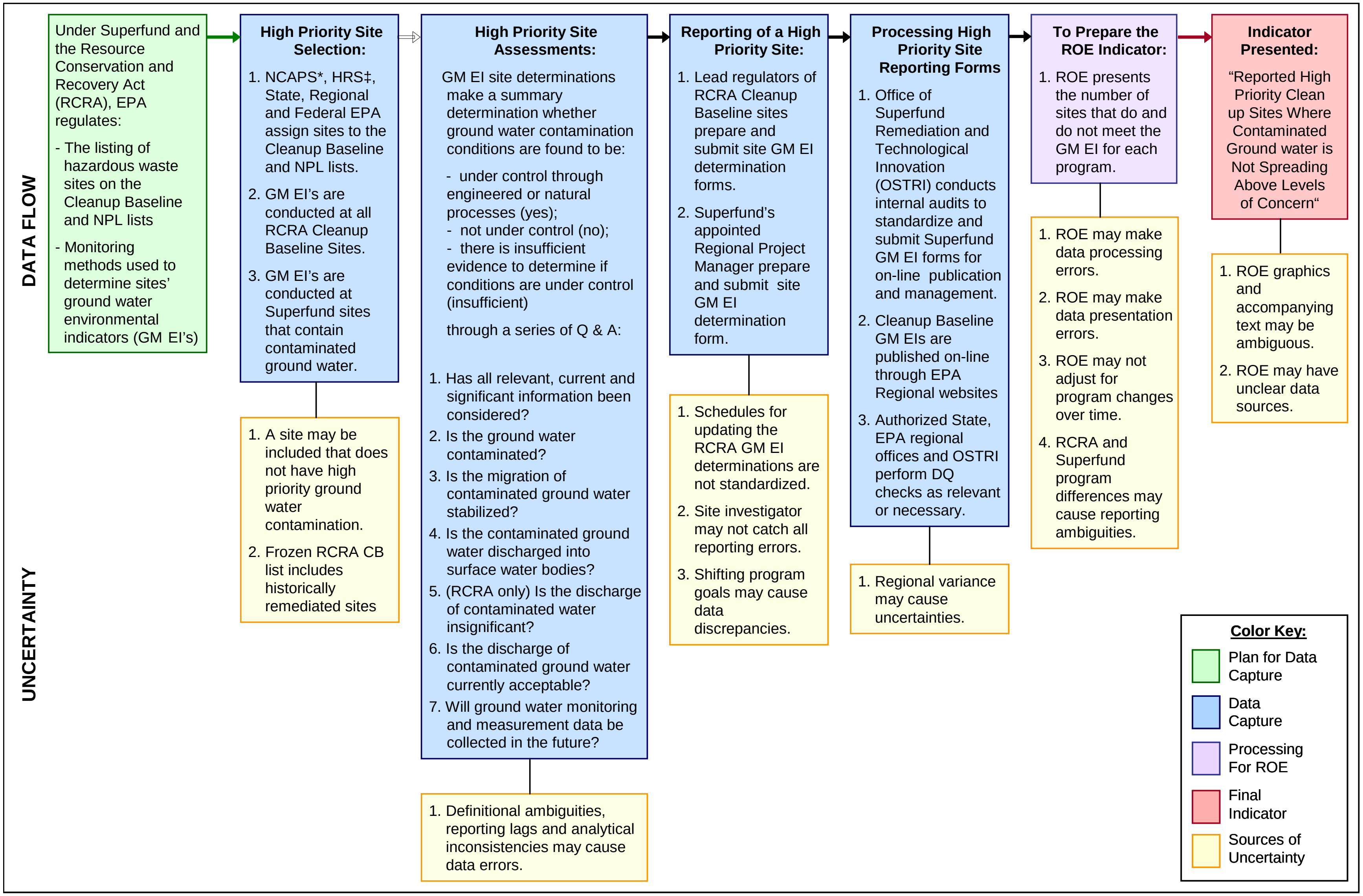


Figure 6-2. Uncertainty and data flow: reported high priority clean-up sites where contaminated ground water is not spreading above levels of concern.

This section examines the types and magnitude of uncertainty associated with the calculation and presentation of the ROE indicator values for both the RCRA and the Superfund Programs.

6.2.1. Determination of Ground Water (GW) Status

The RCRA and the Superfund Programs use a standard GW monitoring environmental indicator (GM EI) (U.S. EPA, 2007f, g). The GM EI indicates whether contamination is below protective, risk-based levels, or, if not, whether the migration of contaminated GW is stabilized and no unacceptable discharge to surface water and monitoring is occurring (see Figures 6-3 and 6-4).

GM EI protocols include a series of seven questions for the RCRA program and six questions for the Superfund Program. The specific questions are included in the *Superfund_GM_EI* and *RCRA_GM_EI* determination flow charts. For each evaluation question, the site manager, appointed or approved by the respective programs, must answer the question and document the reasons for that answer. In addition to the questions, the site manager must make a summary determination whether GW contamination conditions are one of the following:

- not migrating through engineered or natural processes (yes);
- migrating through engineered or natural processes (no); or
- lacking sufficient evidence to determine if GW is migrating (insufficient).

A technically rigorous process determines eligibility for Corrective Action and NPL lists. For the RCRA program, qualification is possible with the National Corrective Action Prioritization System (NCAPS) numerical prioritization system. For the Superfund Program, with the Hazardous Ranking System[31] numerical prioritization system, and then further selection is conducted through state and national nomination processes subject to EPA review and input. Because of these nomination processes, the population of sites assessed with the GM EI is stable and well defined, and, therefore, resilient to unusual circumstances that could cause uncertainty in the determination process.

One GAO Report "tested the accuracy of the data in EPA's Superfund database on the progress of sites through the cleanup process for a statistically random sample of 98 National Priorities List sites" (U.S. GAO, 1998). On the basis of these sample results, the report determined "that the cleanup status of National Priorities List sites reported by the Superfund database as of September 30, 1997, was accurate for 95% (plus or minus 4.4%) of the sites."

---

[31]Calculations based upon the data from U.S. Environmental Protection Agency (EPA). Superfund Information Systems. Available at: http://cfpub.epa.gov/supercpad/cursites/srchsites.cfm. Based on most recent data in database.

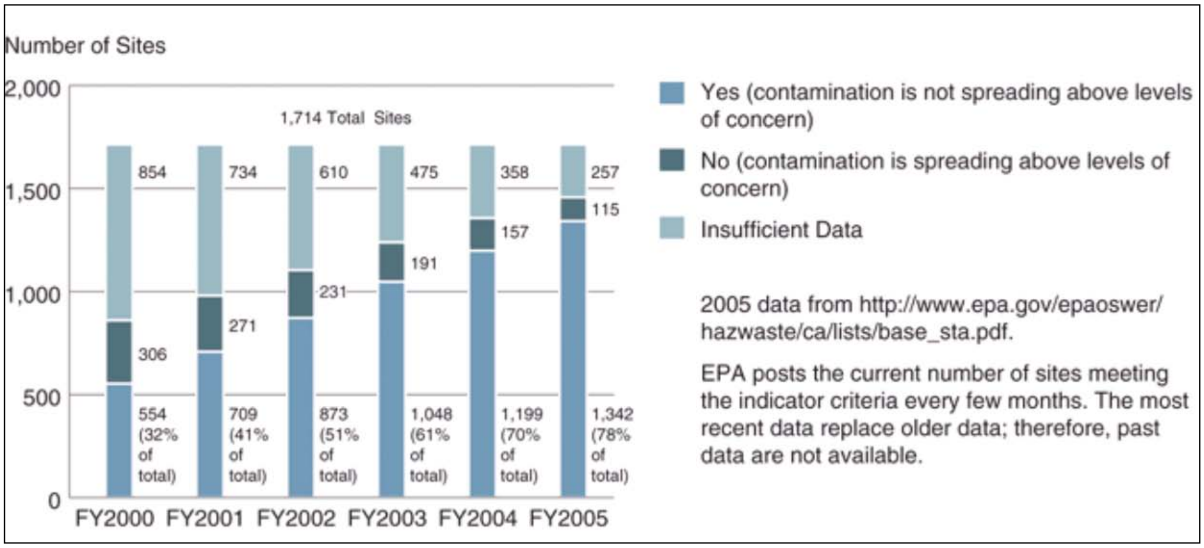


Figure 6-3. RCRA CB sites where GW is not continuing to spread above levels of concern (U.S. EPA, 2008j).

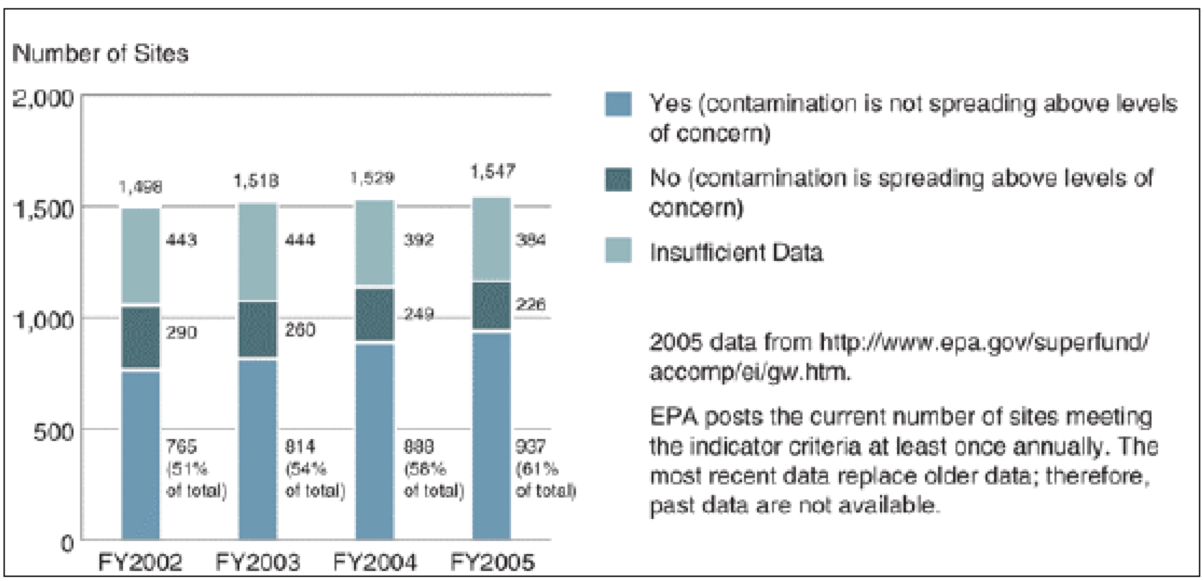


Figure 6-4. Superfund NPL sites where GW is not continuing to spread above levels of concern (U.S. EPA, 2008j).

Similar data quality reviews for the RCRA Corrective Action GM EI site determinations were not identified; however, reviews of the GM EI indicated that while the GM EI is “a good indicator of progress," it has timeliness limitations that lead to uncertainty about the RCRA GM EI's ability to represent a site's current conditions (U.S. EPA, 2000c).

A 2000 EPA study reviewing records before the ROE indicator period attempted to reproduce 62 RCRA EI determinations[32] receiving a status of "yes" for Regions 4, 5, and 6 (U.S. EPA, 2000c). Of the 62 determinations reviewed, 37 had EI determinations that were different than, did not support, or were not included in RCRIS. Of these 37, 8 were found to be incomplete because they were not dated or because an accurate conclusion was not selected on the EI form, 24 were found to have dates that differed from the date in RCRIS, and 5 were not included in RCRIS at the time of the analysis, which was 8 months after the EI had been completed at the site. Only the 2000 data from this EPA study are included in the 2000–2005 ROE analysis. Since this report, many of the recommendations from the GAO and EPA studies have been incorporated into the RCRA program, thus, the current reporting errors are likely significantly lower.

Overall, despite the RCRA GM EI's timeliness limitations, the Superfund and RCRA GM EIs are robust indicators conveying accurate programmatic site characterizations. Documented processes and controls are in place to ensure that the GM EIs monitor stable and well-defined population sizes with frequent reporting requirements, standardized methods, opportunity for national, state, and Regional oversight, and multiple QA protocols when making their determinations and documenting these determinations into the federal data systems. However, additional sources of error in data capture, and in particular machine error, may exist; thus, it may be important to undertake both data quality review and data error estimation analyses to further define uncertainty within this indicator.

6.2.2. Program Definitions Affecting Ground Water (GW) Status Determination or Population of Sites

6.2.2.1. *Site Selection for the Resource Conservation and Recovery Act (RCRA) Corrective Action Program*

While the ROE indicator correctly tallies contaminated GW sites that have been selected for the RCRA Corrective Action program, using the NCAPS methodology, not all sites in the RCRA Corrective Action program have received or need contaminated GW management (U.S. DOE, 1996). The NCAPS system assigns a priority status to sites receiving a high score for any one individual migration pathway or when the site's overall score is greater than or equal to 52. It is likely that there are some sites in the RCRA Corrective Action universe that received high rankings in other migration pathways, or had overall scores of at least 52, without having high-priority GW contamination. Since all RCRA CB sites are monitored with the GM EI, this

[32]These determinations include the other RCRA EI, CA725, for *Human Exposures Under Control*.

means that some sites are monitored that do not have significant or high-priority GW contamination, and these sites receive a “yes” GM EI determination in the ROE indicator (U.S. EPA, 2007h). Including such sites results in a higher percentage of sites categorized as “yes” but a smaller percentage change due to RCRA program actions to remediate contaminated GW. In contrast, the Superfund sites are subject to GW monitoring only if the GM EI establishes there is confirmed GW contamination at the site.

6.2.2.2. *Freezing of the Resource Conservation and Recovery Act (RCRA) Cleanup Baseline (CB) Population*

An additional source of uncertainty arises from freezing the RCRA CB population. For programmatic tracking purposes, RCRA CB tracked 1,714 sites for the period 2000−2005. Included in the RCRA CB is a group of 255 sites with a status of “yes” that had already completed the RCRA cleanup process prior to being nominated in 1999 (U.S. EPA, 1999b). Technically, this group of 255 sites was included as a part of the RCRA program because the Corrective Action program is cumulative, and once a site has been added to the high-priority list, it remains on the list even if it is has been remediated (U.S. DOE, 1996). Table 6-1 presents the shifts in success rates resulting from excluding these historical sites from the 2000−2005 data set.

In contrast, once a NPL site has been remediated in the Superfund Program, it is deleted from the NPL. Therefore, the denominator in the Superfund ROE presentation shifts with real time: increasing as sites are added to the program, and decreasing as sites are deleted from the program. NPL sites are nominated and deleted from the list on a continual basis, whereas RCRA sites are only periodically nominated to the RCRA CB. Because the NPL program actively updates its site status each year, the NPL indicator does not contain the inclusion uncertainty that the RCRA program has.

6.2.3. Frequency of Ground Water (GW) Status Assessment and Reporting

RCRA requires that a site record in RCRAInfo must be updated if and when the site conditions change, but has no required regular assessment or reporting (U.S. EPA, 2007i). Any changes that occur at a Superfund site using the GM EI are required to be entered into the WasteLAN database 30 days within knowing of the EI status change. The requirement to regularly update the GM EIs reduces the uncertainty surrounding the sites’ determination, forces managers to become familiar with the Superfund reporting procedures, and minimizes several types of errors that could otherwise occur in the underlying data used to make a site determination.

Table 6-1. Effect of including decontaminated sites in the RCRA CB percentage of sites receiving a "yes" for GW contamination control (derived by Abt Associates)

	2000	2001	2002	2003	2004	2005
ROE indicator	32%	41%	50%	60%	69%	77%
Year to year change	---	+9%	+18%	+28%	+37%	+45%
ROE indicator excluding the 255 sites remediated prior to 2000	20%	30%	42%	53%	63%	73%
Year to year change	---	+10%	+22%	+33%	+43%	+53%

Because the RCRA CB does not require periodic updating of GM EI determination like the Superfund Program, it is possible that the most recently recorded GM EI for an RCRA site is not current. Conditions that could cause this change in site determination include

- GW contamination leaching from historic sediment contamination and local underground storage tank leachate (U.S. EPA, 2007h, k).
- Natural attenuation at a contaminated site.
- Changes in Maximum Contaminant Level (MCL) standards (U.S. EPA, 2007j).
- Relapse of contaminant migration to a site that was already deemed to have contained its GW contamination migration (U.S. EPA, 2007h).
- Failure of remediation technology.

Although no record exists of error resulting from these changing conditions, to the extent that RCRA GM EI is not regularly updated, a given year's site determination may not reflect the site's actual conditions at that time.

6.3. REGIONALIZATION

While the ROE indicators are compiled at a national level, similar trend summaries can be generated at the EPA Regional level because the indicator is a simple arithmetic sum of all sites within a geopolitical unit (e.g., United States or EPA Region). For both the RCRA and Superfund Programs, however, the regional calculations of High-Priority Cleanup Sites Where Contamination is not Continuing to Spread Above Levels of Concern are more sensitive to the uncertainty associated with mischaracterizing the status of a single site, because the total population of sites in a given Region (or state) is smaller at the Regional level than at the national level (see Table 6-2). For example, in Region 9, an error or change in the GM EI at only one site would affect the percentage of sites receiving a GM EI determination of “yes” from between 0.3% (Region 5) to 1.6% (Region 8). For the same reason, the effect of including remediated sites in a trend analysis may be more significant at the Regional level.

Table 6-2. Distribution of 2005 Superfund sites with GM EI assessments[a] (U.S. EPA, 2007k)

EPA region	Number of sites	Number of sites receiving a GM EI of “yes”	Percent of sites receiving a GM EI of “yes”
1	123	66	54
2	292	172	59
3	194	110	57
4	182	120	66
5	313	197	63
6	128	76	59
7	66	41	62
8	61	38	62
9	109	64	59
10	81	54	67
Total	**1,549**	**938**	**61**

[a]The regional proportions for these data are from the 2007 Comprehensive Environmental Response, Compensation and Liability Information System (CERCLIS) database, extrapolated to the Superfund sites listed as of 2005.

Another source of variability that affects both RCRA and Superfund ROE indicators is the state-to-state variations in the contamination levels used for GM EI determinations. RCRA and the Superfund Programs allow the lead agency/department and Regional Project Manager (RPM), respectively, to select appropriate “levels” for the GM EI determinations (U.S. EPA, 2008b, 2008p). Although the Superfund Program provides guidance for determining what contamination levels to use, the RPM has ultimate authority in making the determination of yes, no, or insufficient for the site, based on the recommended guidance documents (U.S. EPA, 20058p). In both RCRA and the Superfund Program, states can use a different contamination standard for determining their GM EIs. Where state standards differ, the percent “yes” for a given Region will be affected by the distribution of sites across the states.

7. BLOOD LEAD-LEVEL INDICATOR

7.1. DESCRIPTION OF THE BLOOD LEAD-LEVEL INDICATOR

The ROE introduces the Blood Lead-Level indicator with the following description (U.S. EPA, 2008n):

> Lead is a naturally occurring metal found in small amounts in rock and soil. Lead has been used industrially in the production of gasoline, ceramic products, paints, metal alloys, batteries, and solder. While lead arising from the combustion of leaded gasoline was a major source of exposure in past decades, today lead-based paint and lead-contaminated dust from paint are the primary sources of lead exposure in the home. Lead levels can be measured in blood or urine.
>
> Lead is a neurotoxic metal that affects areas of the brain that regulate behavior and nerve cell development (NRC, 1993). Its adverse effects range from subtle responses to overt toxicity, depending on how much lead is taken into the body and the age and health status of the person (CDC, 1991). Lead is one of the few pollutants for which biomonitoring and health effect data are sufficient to clearly evaluate environmental management efforts to reduce lead in the environment.
>
> Infants, children, and fetuses are more vulnerable to the effects of lead because the blood-brain barrier is not fully developed in them (Nadakavukaren, 2000). Thus, a smaller amount of lead will have a greater effect on children than on adults. In addition, ingested lead is more readily absorbed into a child's bloodstream, while adults absorb only 10%. Because of lead's adverse effects on cognitive development, the Centers for Disease Control and Prevention (CDC) have defined an elevated blood lead level as equal to or greater than 10 micrograms per deciliter (µg/dL) for children under 6 years of age (CDC, 2005[a]).
>
> This indicator is based on data collected by the National Health and Nutrition Examination Survey (NHANES). NHANES is a series of surveys conducted by CDC's National Center for Health Statistics [NCHS] that is designed to collect data on the health and nutritional status of the civilian, non-institutionalized U.S. population using a complex, stratified, multistage, probability-cluster design. CDC began monitoring blood lead in 1976 as part of NHANES II, which covered the period from 1976 through 1980. Blood lead was also monitored in NHANES III, which covered the period between 1988 and 1994. CDC's National Center for 1Environmental Health [NCEH] conducted the laboratory analyses for the biomonitoring samples. Beginning in 1999, NHANES became a continuous and annual national survey, visiting 15 U.S. locations per year and surveying and reporting for approximately 5,000 people annually.

The ROE notes the following limitations to the indicator:

- Because the data from NHANES 1999−2000 and 2001−2002 represent only two survey periods, changes in estimates between the two time periods do not necessarily reflect a trend. Earlier data sets are available (e.g., NHANES III), but the data are not directly comparable to NHANES 1999−2002. As CDC releases additional survey results (e.g., 2003−2004), it will become possible to more fully evaluate trends (CDC, 2002).

Lead is an environmental toxicant that adversely affects the nervous, hematopoietic, endocrine, renal, and reproductive systems (CDC, 2007). Over time, the United States population has been exposed to lead through numerous avenues, including but not limited to lead-based paint and leaded gasoline. For adults, lead exposure can primarily be attributed to occupational and recreational sources; whereas, the major source of exposure for children is currently from deteriorated lead-based paint and the resulting dust and soil contamination (CDC, 2005a). Regardless of the source of exposure, lead can result in adverse health effects in children and adults; however, children are more susceptible to the neurotoxic effects of lead due to an undeveloped blood-brain barrier (U.S. EPA, 2008n). To estimate the population's exposure to lead, EPA uses data obtained from the NHANES.

The CDC's NCHS conducts NHANES as an assessment of the health and nutritional status of the nation. Using a complex, stratified, multistage, probability cluster design CDC selects a representative sample of the civilian, non-institutionalized population for health interviews and detailed physical examinations (U.S. EPA, 2008o). Initially, NHANES consisted of a 4- to 6-year survey; in 1999, NHANES was altered and became a continuous and annual national survey of adults and children 1 year and older.[33] As part of this change, the sample size and number of locations used to generate national estimates of health and nutrition were reduced, resulting in NHANES visiting only 15 locations per year and interviewing and reporting for approximately 5,000 people annually (U.S. EPA, 2008n).[34]

The locations used in NHANES were chosen using two of the four panels of primary sampling units (PSUs), which were defined as single counties not used for the National Health Interview Survey (NHIS) (CDC, 2002). CDC selected 15 locations for each year of the survey, 1999−2004, out of the remaining pool of 200 PSUs using the following approach:

[33]This change in NHANES made it possible to obtain biennial blood lead level estimates, and explains why the ROE blood lead level indicator does not include estimates prior to 1999.

[34]Blood lead level estimates are now more subject to the limits of increased sampling error due to the smaller number of individuals and geographic locations sampled. It has been approximated that the standard errors of estimates generated using data from the new annual NHANES (e.g., NHANES 1999−2000), as was done for the ROE blood lead level indicator, are roughly 70% greater than those calculated in NHANES III, which included 6 years worth of data (CDC, 2002).

> In order to create six annual national samples, 120 of the 200 NHIS PSUs were selected using a measure of size related to 1990 Census county-specific information on the percent Mexican American, percent Black, and the NHIS PSU-selection probability. 20 PSUs were randomly assigned to each year in 1999−2004. For each year, a subset of 15 PSUs was selected with the remaining 5 PSUs held in reserve (CDC, 2002).

This random sampling process results in a new subset of locations being used for each year of the survey. It is important to keep in mind that NHANES was designed to measure the national health status, not environmental exposure. The locations selected represent nationally representative samples in terms of age, gender, race, ethnicity, and income based on U.S. Census distributions. As a result, sample locations were not chosen to represent ranges of environmental exposure nor any other factor relating to a specific health outcome.

At each location, surveys and physical examinations are conducted for each person selected to participate in NHANES. Blood samples are taken from every person over 1 year of age during the physical examination, and each blood sample is sent to CDC's NCEH for analysis.[35] Upon analysis from all locations, the resulting blood lead data are used to calculate the national geometric mean blood level, distribution percentiles, and 95% CIs by age, by sex, and by race/ethnicity. CDC reports these estimates in their *National Report on Human Exposure to Environmental Chemicals*. EPA directly uses these geometric mean blood lead level estimates and distribution percentiles presented in the *National Report on Human Exposure to Environmental Chemicals* as the ROE Blood Lead-Level indicator.[36] Table 7-1 presents the geometric mean blood lead levels and the distribution percentiles presented by EPA for the ROE indicator.

This case study characterizes the uncertainty related to the Blood Lead-Level indicator using available information. Figure 7-1 depicts the general sequence of data capture and processing involved in preparing ROE indicators. We identified "Data Capture Plan and Methods" as the major source of uncertainty pertaining to the ROE end use for the Blood Lead-Level indicator. For example, we deemed the sampling design to be an important source of uncertainty, but the distribution of geographic locations and the number of sampling points across representative environments from the annual surveys underlying the indicator were not available for further analysis. If this information becomes available, a more complete

[35]Physical examinations are performed in mobile examination centers (MECs). The only analyses that the MEC laboratories can perform are the complete blood count and pregnancy analysis of the physical examination; as a result, all remaining analyses are conducted elsewhere by approximately 28 CDC NCEH laboratories across the United States (CDC, 2007).

[36]Information about the indicators, their respective metadata and independent data analyses were extracted and carried out between December 2006 and September 2007.

Table 7-1. ROE Blood Lead-Level indicator and distribution percentiles (U.S. EPA, 2008n)

Stratification	Survey years	Geometric mean	Distribution percentiles	
			50th	75th
Total, age 1 and older	1999–2000	1.7	1.6	2.4
	2001–2002	1.5	1.4	2.2
Age group				
1–5 years	1999–2000	2.2	2.2	3.3
	2001–2002	1.7	1.5	2.5
6–11 years	1999–2000	1.5	1.3	2.0
	2001–2002	1.3	1.1	1.6
12–19 years	1999–2000	1.1	1.0	1.4
	2001–2002	0.9	0.8	1.2
20 years and older	1999–2000	1.8	1.7	2.5
	2001–2002	1.6	1.6	2.2
Gender				
Male	1999–2000	2.0	1.8	2.9
	2001–2002	1.8	1.7	2.7
Female	1999–2000	1.4	1.3	1.9
	2001–2002	1.2	1.1	1.8
Race/ethnicity				
Mexican Americans	1999–2000	1.9	1.7	2.8
	2001–2002	1.7	1.6	2.5
Non-Hispanic blacks	1999–2000	1.8	1.8	2.7
	2001–2002	1.5	1.5	2.2
Non-Hispanic whites	1999–2000	1.6	1.6	2.4
	2001–2002	1.4	1.4	2.1

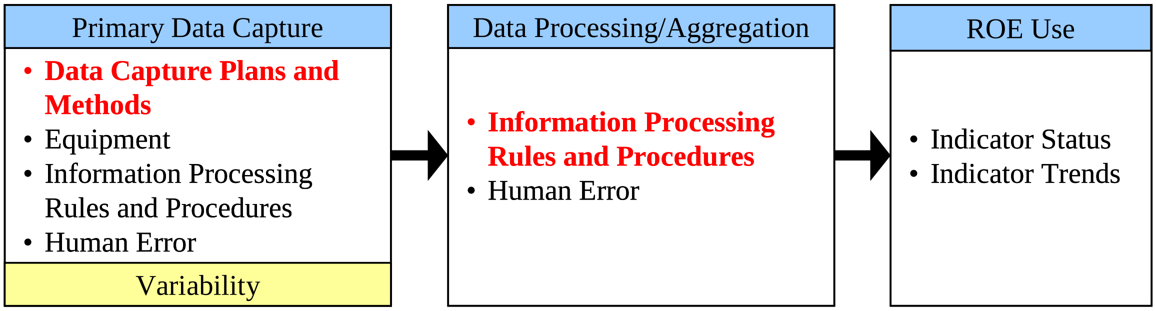


Figure 7-1. General sequence of data capture and processing involved in preparing the blood lead level indicator.

characterization of the conditional uncertainty for this indicator could be conducted. While statistical measurements of confidence intervals are provided from NHANES in association with its survey and laboratory measurements, the absence of additional primary data source information at this time requires some uncertainty characterization for this case study to also be qualitative.

7.2. OVERVIEW OF UNCERTAINTY

A more detailed data flow that culminates in the ROE indicator presentation, along with sources of uncertainty associated with each step in this data flow, are represented in Figure 7-2. Potential uncertainty associated with the ROE Blood Lead-Level indicator, which depicts the national blood lead level by age, by sex, and by race/ethnicity, exists because of

- limitations in the NHANES study design;
- errors in the collection of blood samples during NHANES examinations;
- errors in the laboratory analysis of blood samples; and
- errors in the calculation of national blood lead levels.

The following discussion describes the areas where uncertainty in the ROE indicator may arise. The last part of this section is a discussion of other factors that may contribute to variability in the reported indicator value.

Overview: Uncertainty

- Limitations in the study design of NHANES, particularly the number and distribution of sampling locations, are expected to represent the greatest proportion of the uncertainty in the NHANES calculated confidence interval; these limitations also contribute to additional uncertainty as to whether the study is representative of environmental exposure to lead.
- Human errors in the collection of blood samples during NHANES examinations, equipment errors in analyzing the lead levels of blood samples, and information processing errors in calculating the national blood lead level by age, by sex, and by race/ethnicity are not expected to significantly contribute to uncertainty in the BLL indicator estimates.

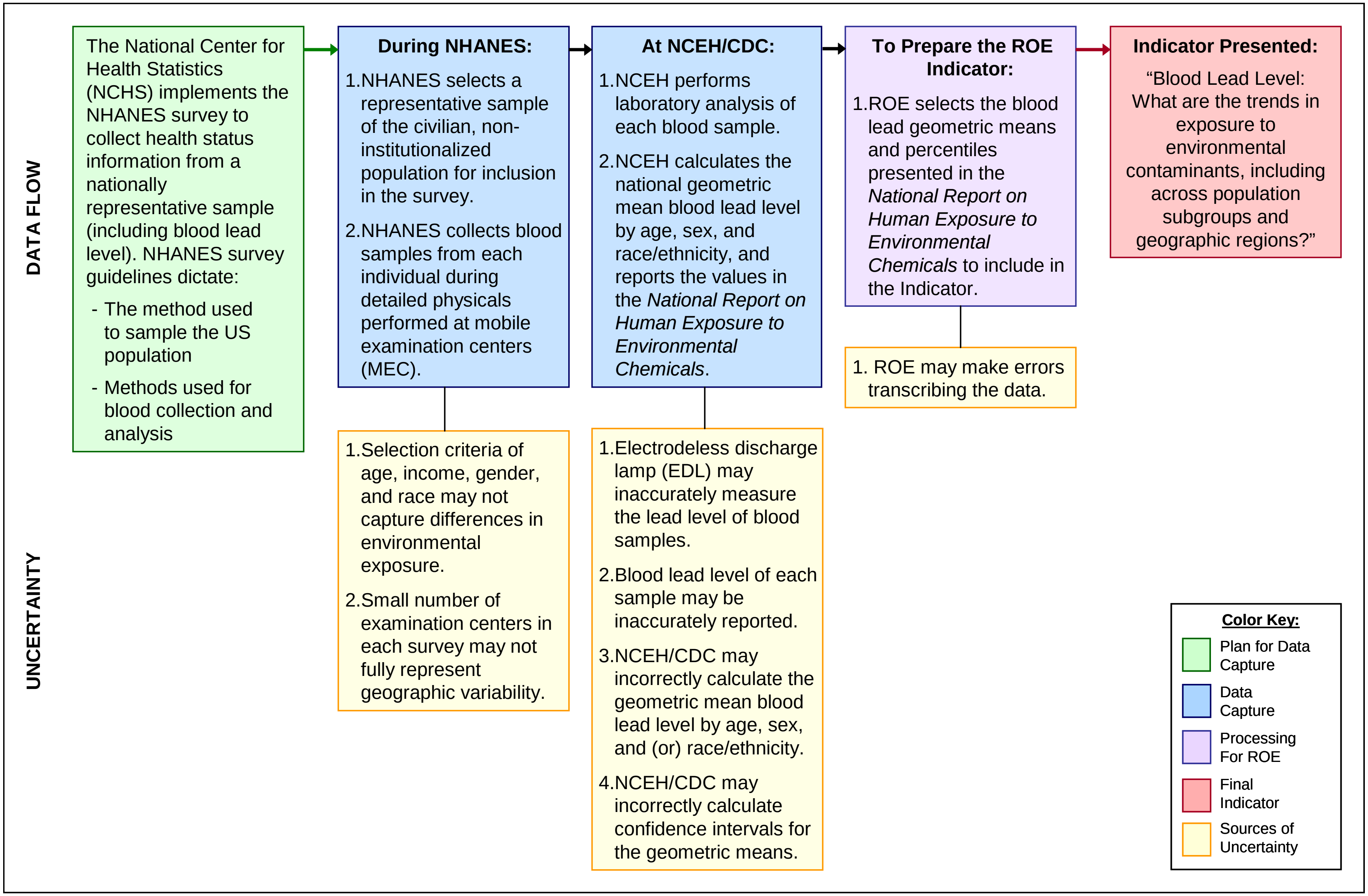


Figure 7-2. Uncertainty and data flow: blood lead level—trends in exposure to environmental contaminants, including across population subgroups and geographic regions.

7.2.1. Limitations in the National Health and Nutrition Examination Survey (NHANES) Study Design

The study design of NHANES may contribute to the uncertainty surrounding the ROE Blood Lead-Level indicator, primarily due to the number and distribution of sampling units used to generate national estimates. Although NHANES conducts physical examinations of approximately 8,000 people annually, the current annual survey is based in only 15 locations (U.S. EPA, 2008n).[37] The small number of sample locations means that the data collected may not represent the true geographic variability in lead exposure, resulting in uncertainty surrounding the true value of the ROE indicator (U.S. EPA, 2008n). The resulting uncertainty in the indicator would cause the greatest concern when viewing estimates for narrowly defined demographic groups or other specific subgroups (U.S. EPA, 2008o). For example, lead exposure is known to significantly correlate with both older house age and lower income communities (Pirkle, et al, 1998). Thus, NHANES blood lead level averages may be conditionally biased by not explicitly accounting for environmental exposure variables, such as house age or geographic region, as a sample selection criterion or post-stratification weight.

7.2.2. Errors in the Collection of Blood Samples

NHANES monitors the health and nutritional status of the United States population through a series of surveys and physical examinations. During the physical examinations, mobile examination centers (MECs) collect blood samples from persons aged 1 year or older, providing the quantitative data necessary to calculate the national geometric mean blood lead level by age, by sex, and by race/ethnicity. Errors that occur during the collection process, such as the misidentification of participants and/or the mislabeling of samples, could result in erroneous records being retained for a particular subpopulation. However, the CDC has extensive procedures in place to minimize these errors. Recording procedures are automated, and data are sent directly to computer databases; in addition, transcribed data are required to be proofread by laboratory employees. Because of these protocols, it is unlikely that these types of errors occur frequently, and therefore, the uncertainty in the ROE indicator generated by these errors is expected to be small.

7.2.3. Errors in the Analysis of Blood Samples

Upon completion of the physical examinations, each MEC sends the blood samples collected to laboratories at CDC's NCEH for analysis. NCEH determines the lead concentration

[37] According to the 2000 Census, the United States population was 281,421,906 (http://www.census.gov);.7,970 samples were collected in the 1999-2000 survey, while 8,945 samples were collected in the 2001-2002 survey (CDC, 2005a).

of each blood sample by measuring the light absorbed at 228.8 nm by ground state atoms of lead from an electrodeless discharge lamp (EDL) source (U.S. EPA, 2008o). Errors in the analysis of blood samples may potentially arise as a result of an incorrectly calibrated or malfunctioning EDL source. However, with the measurement of blood lead concentrations occurring at approximately 28 CDC NCEH laboratories across the United States, an incorrectly calibrated or malfunctioning EDL source at one laboratory would affect, at most, 1/28th of the samples (or roughly 360 samples per NHANES reporting period) in the ROE indicator (CDC, 2007). Additionally, the CDC has extensive procedures and protocols in place for calibrating instruments between laboratories and managing incorrectly measured samples. Even if one laboratory experienced a systematic measurement error, it is very likely that this inconsistency would be detected and resolved quickly, long before an entire NHANES (2-year) cycle passes. Therefore, a corruption or measurement error at one lab is likely to affect much less than 1/28th of the samples. Although we do not have data to quantify the frequency or magnitude of machine errors, we expect their effect on the uncertainty in the ROE indicator to be quite small, due to the protocols that are in place.

7.2.4. Errors in the Calculation of National Blood Lead Levels

After the blood lead concentration of each sample has been identified, NCEH calculates the national geometric mean blood lead level by age, by sex, and by race/ethnicity for the defined years of the survey.[38] Although the majority of patient and sampling information is recorded automatically by the NHANES computerized databases, there are a few points of manual data input into the system. While this may create the potential for transcription errors, the CDC requires routine checks of data, outliers, and inconsistencies in their inventories. As a result, recording errors are expected to be very limited and unlikely the result of bias. Therefore, uncertainty in the ROE indicator arising from transcription errors in reported blood lead concentrations or the calculation of the ROE indicator can be assumed to be near zero.

7.2.5. Overall Uncertainty in Report on the Environment (ROE) Blood Lead-Level Indicator

CDC quantifies the uncertainty in the national geometric mean blood lead levels by calculating 95% confidence intervals (CIs). CDC presents the resulting geometric mean blood lead levels stratified by age, by sex, and by race/ethnicity; their corresponding 95% CI; and their distribution percentiles in the *National Report on Human Exposure to Environmental Chemicals.*

[38]Although CDC began monitoring blood lead levels in 1976, ROE includes blood lead levels for only those years, 1999–2000 and 2001–2002, when NHANES became a continuous and annual survey (U.S. EPA, 2007o).

EPA uses the geometric mean blood lead levels and their associated distribution percentiles as the ROE Blood Lead-Level indicator. Figures 7-3 through 7-6 present the geometric mean blood lead levels by age, by sex, and by race/ethnicity and their corresponding 95% CI to depict the uncertainty in the ROE Blood Lead-Level indicator. As the blood lead level measurements do not include environmental exposure either as a selection criteria or a post-stratification weight, the reported confidence intervals do not incorporate this additional (and unknown) degree of uncertainty.

Figure 7-3. ROE Blood Lead-Level indicator and 95% CIs: total, age 1 and older (1999–2002) in µg/dL (CDC, 2005a).

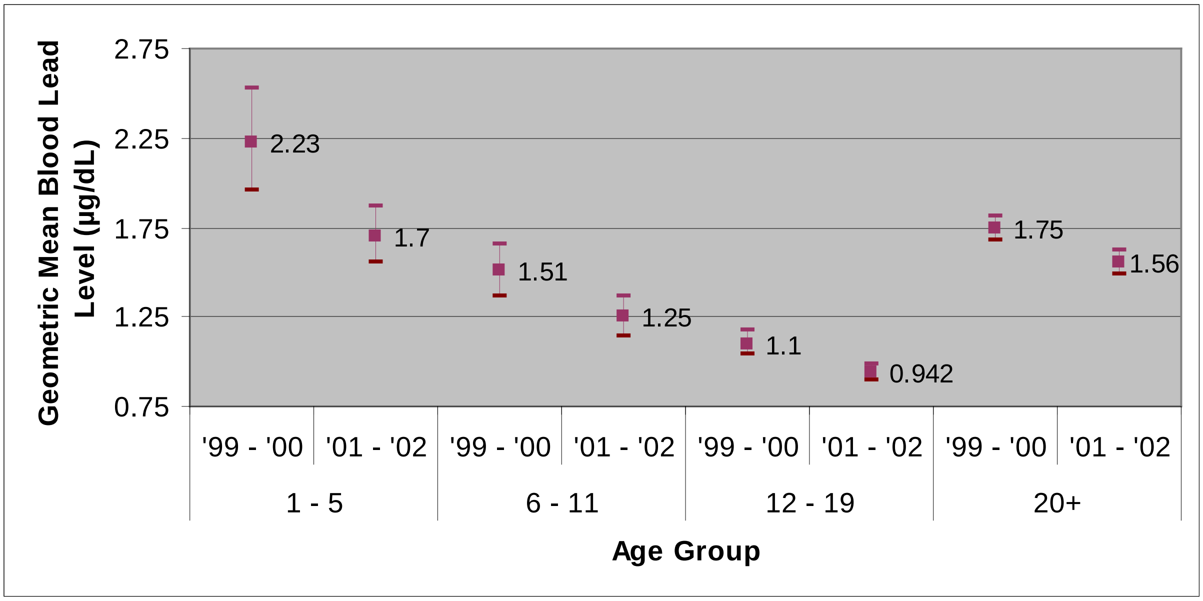


Figure 7-4. ROE Blood Lead-Level indicator and 95% CIs: age group (1999–2002) (CDC, 2005a).

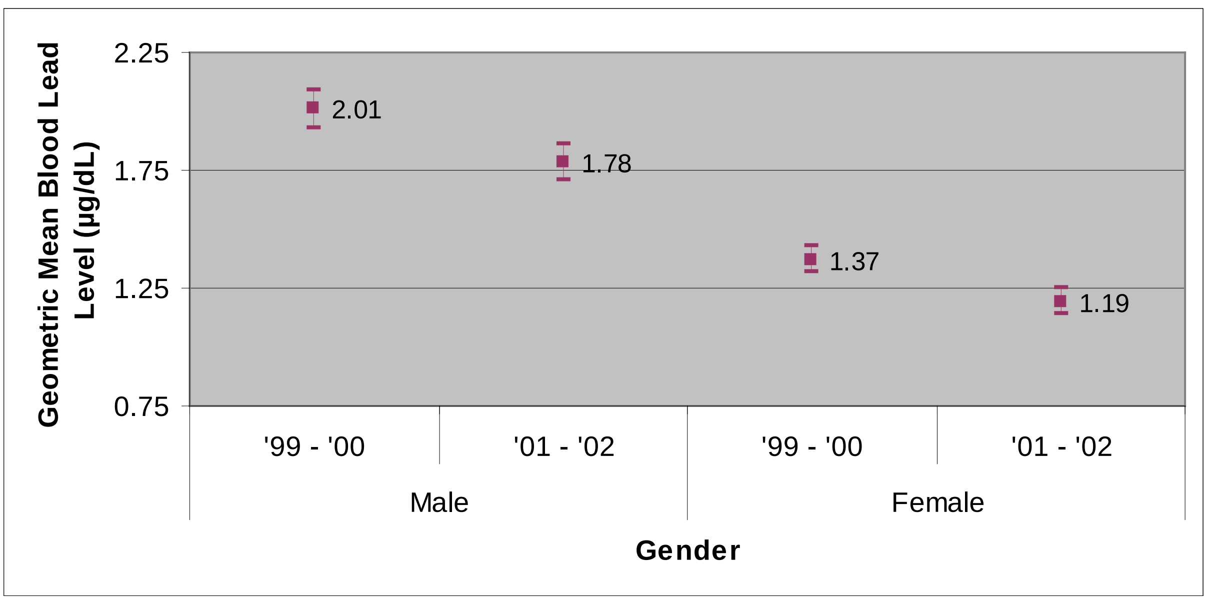


Figure 7-5. ROE Blood Lead-Level indicator and 95% CIs: gender (1999–2002) (CDC, 2005a).

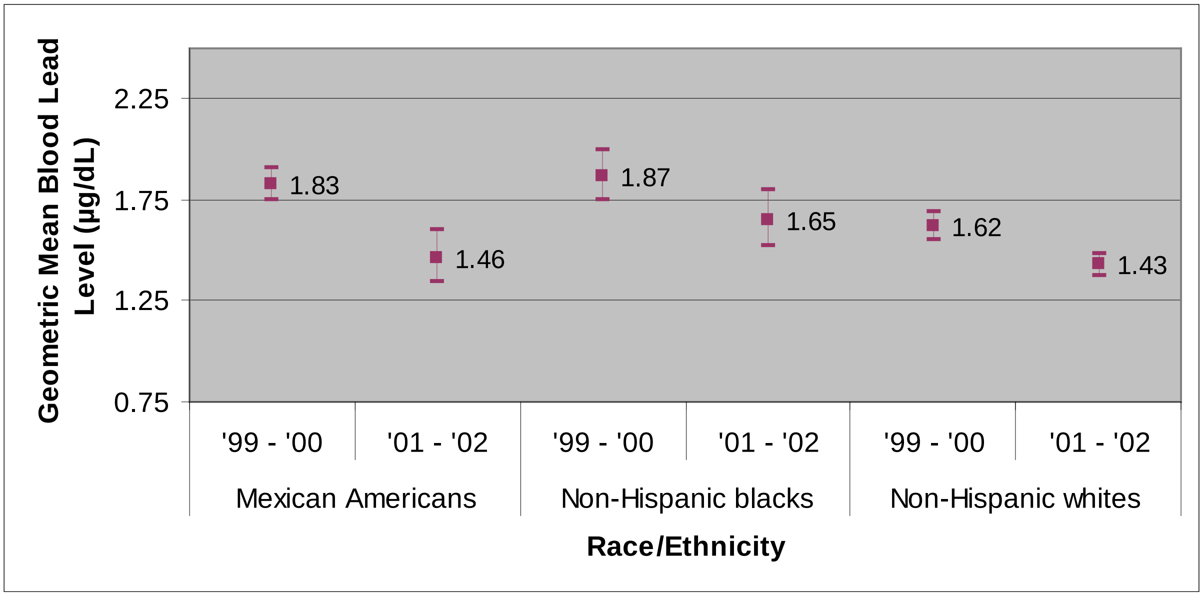


Figure 7-6. ROE Blood Lead-Level indicator and 95% CIs: race/ethnicity (1999–2002) (CDC, 2005a).

7.2.6. Other Sources of Uncertainty

The Blood Lead-Level indicator provides an estimation of national blood lead levels for *only* two time periods, 1999–2000 and 2001–2002. Until more surveys are completed, conclusions regarding the difference between the two periods could be made, but not regarding the trend in blood lead levels.

By observing the ROE indicator data presented in Table 7-1 and Figures 7-3 through 7-6, it is obvious that blood lead levels differ between 1999–2000 and 2001–2002. However, it is unclear whether the differences reflect actions taken on a national level to reduce lead exposure, or whether the differences are due to variability between the persons and locations sampled in the two time periods. The locations were chosen to be nationally representative of the United States population in terms of age, gender, race, ethnicity, and income. In analyzing national estimates for the two time periods, differences between the two estimates may easily arise when using a small, rotating sample size of 15 locations, because of inter-individual and variability in environmental exposures.

For the ROE Blood Lead-Level indicator, the difference in blood lead level estimates observed for the two time periods may be due to variability between the samples as a result of one or more of the following:

- Individuals being exposed to different levels of lead.[39]
- Differences in behavior between persons that result in the inhalation or ingestion of lead.
- Differences in persons to the degree which a given ingestion or inhalation increases their blood lead level.
- Individuals having a past history of lead exposure.
- Biological differences between persons.

One or any combination of the above-referenced factors can ultimately result in the differences observed in blood lead levels between the two time periods represented in the ROE indicator. However, corrections for these factors must be made at the level of the sampling unit or household. As NHANES does not select sampling locations to account for environmental exposure, and since the magnitude of a conditional bias in the national estimates cannot be estimated without accounting for possible conditional bias in the NHANES study, it is not possible to determine to what extent the survey-to-survey difference on average blood lead levels is attributable to differences between the sampled locations/populations used in each NHANES survey.

7.3. REGIONALIZATION

EPA directly uses the data presented in the *National Report on Human Exposure to Environmental Chemicals* for the ROE Blood Lead-Level indicator. Although the ROE Blood Lead-Level indicator is calculated using data obtained from NHANES, which is a nationally representative survey, the current design does not make it possible to obtain estimates for smaller geographic areas (CDC, 2005b).

Overview: Regionalization

- National blood lead level estimates derived from NHANES data cannot be extrapolated to the regional level because of the small sample size and small number of geographic locations used in the survey.
- Regional differences in blood lead levels do exist, but the data collected by NHANES do not allow the calculation of regional estimates.

The NHANES data do not make it possible to obtain regional blood lead level estimates, but regional differences in lead exposure, and subsequently, blood lead levels have been observed. Regional differences in blood lead levels can occur through a variety of avenues including—but not limited to—the following:[40]

[39]This can be attributed to the fact that environmental exposure measures, such as exposure to lead, tend to vary geographically (U.S. EPA, 2008n).

[40] Potential regional differences in blood lead levels were obtained during a discussion with Penny Schafer (personal communication, July 10, 2007).

(1) Exposure to lead-based paint: Exposure to lead-based paint is prevalent in areas of the country where a large percentage of the housing stock consists of pre-1960 or pre-1978 homes. This is due to homes built prior to 1960 and 1978 containing lead-based paint. Although all homes built during this time period may have contained lead-based paint, numerous renovation and remodeling activities have removed the lead hazard from some homes. Unfortunately this is not the case for all housing facilities, and as a result, lead exposure has been found to be the greatest in housing facilities in low-income areas due to their inadequate upkeep (i.e., the presence of lead-based paint and in some cases deteriorating lead-based paint). Detailed analyses of the national housing stock has found regional differences in blood lead levels due to exposure to lead-based paint. Persons residing in the northeast and north central parts of the country are exposed to lead-based paint more than those in the south and southwest part of the country, where the housing stock tends to be much newer.

(2) Exposure to the residuals of leaded gasoline: Small regional differences in blood lead levels have been observed in individuals that reside in urban areas near highways, due to exposure to lead-contaminated soil. Although lead was phased out of gasoline, lead residuals have been found in the soil next to highways because lead is non-mobile in soil.

(3) Exposure to lead-containing ethnic products: Elevated blood lead levels have been observed in areas of the country where individuals from Mexico or India reside, such as Southern California. In both cultures lead-containing medicine, candy, and pottery are sometimes used, inadvertently exposing the population to lead on a continuous basis.

The above examples highlight that regional differences in lead exposure do exist; unfortunately the current design of NHANES does not make it possible to calculate regional estimates. In order for the ROE Blood Lead-Level indicator to be extrapolated to the regional level, the data used to generate the national estimate would need to come from a larger sample—and, more importantly, from respondents representative of different levels of environmental exposure to lead.

8. FISH FAUNAL INTACTNESS INDICATOR

8.1. DESCRIPTION OF FISH FAUNAL INTACTNESS INDICATOR

The ROE introduces the Fish Faunal Intactness indicator with the following description:

Intactness, the extent to which ecological communities have retained their historical composition, is a critical aspect of the biological balance of the nation's ecological systems (NRC, 2000). It is of particular importance in freshwater systems that are impacted by pollution, habitat alteration, fisheries management, and invasive species.

This indicator tracks the intactness of the native freshwater fish fauna in each of the nation's major watersheds by comparing the current faunal composition of those watersheds with their historical composition. In this case, historical data are based on surveys conducted prior to 1970. The indicator specifically measures the reduction in native species diversity in each 6-digit U.S. Geological Survey hydrologic unit code (HUC) cataloguing unit in the 48 contiguous states. Intactness is expressed as a percentage based on the formula:

reduction in diversity = 1– (# of current native species ÷ # of historical native species).

The native species diversity indicator proposed by the National Research Council (NRC, 2000) compared expected native species diversity (projected from species-area-curve models) with observed diversity. This "Fish Faunal Intactness" indicator makes use of empirical, rather than modeled, data sets, and it focuses on a well-known group of organisms with a fairly strong historical record.

Reductions in watershed diversity may be due either to the overall extinction of a species (at least 12 U.S. freshwater fish species are known to be extinct, and another three species are known only from historical records and may be extinct) or, more commonly, to the extirpation of a species from selected watersheds. In the case of regional extirpations, opportunities may exist for restoring a species to watersheds in its historical range.

The fish distributional data underlying this indicator were gathered by NatureServe, a nonprofit research organization, and are derived from a number of sources, including species occurrence data from state Natural Heritage Programs, a broad array of relevant scientific literature (e.g., fish faunas), and expert review in nearly every state. These data were assembled during the 1997–2003 period. The underlying data include distributions for 782 native freshwater fish species across small watersheds (8-digit HUC). For this indicator, data were pooled and reported by larger 6-digit HUCs to reduce potential errors of omission in the smaller watersheds (U.S. EPA, 2008c).

The ROE also notes the following limitations of the indicator:

- The incomplete historical record for freshwater fish distributions and inconsistent inventory records for contemporary fish distributions are sources of uncertainty.
- Although NatureServe has attempted to compile the most complete distributional information possible for these species at the 8-digit HUC level, these data are dynamic; new records frequently are added and existing records are revised as new information is received and as taxonomic changes occur.

This indicator tracks the intactness of the native freshwater fish fauna in each region of the nation by comparing the current faunal composition of those watersheds with their historical composition. The indicator specifically measures the reduction in native species diversity in each 6-digit United States Geological Survey (USGS) HUC in the 48 conterminous states. Information on the current distributions of fish species is primarily based on data that were aggregated at the state level in 1970. The historic distribution information dates back to the nineteenth century (see Figures 8-1 and 8-2).

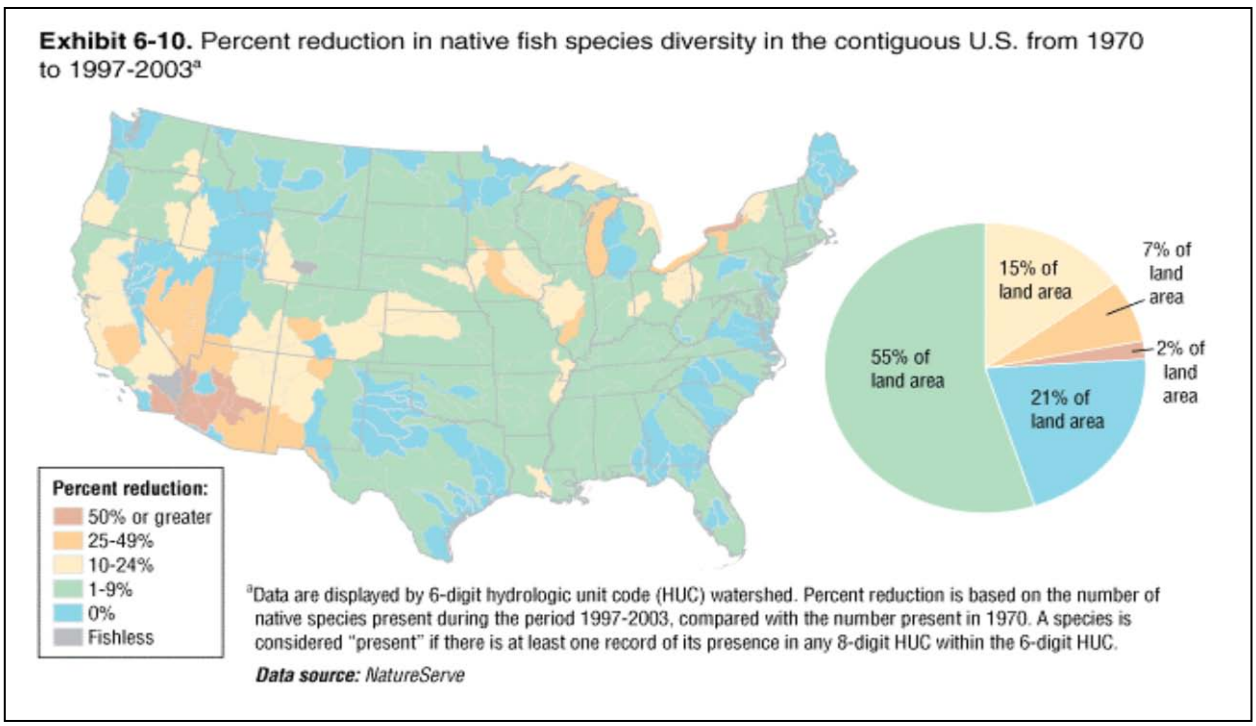


Figure 8-1. Percentage reduction in native diversity in the contiguous U.S. from 1997–2003[a] (U.S. EPA, 2008c).

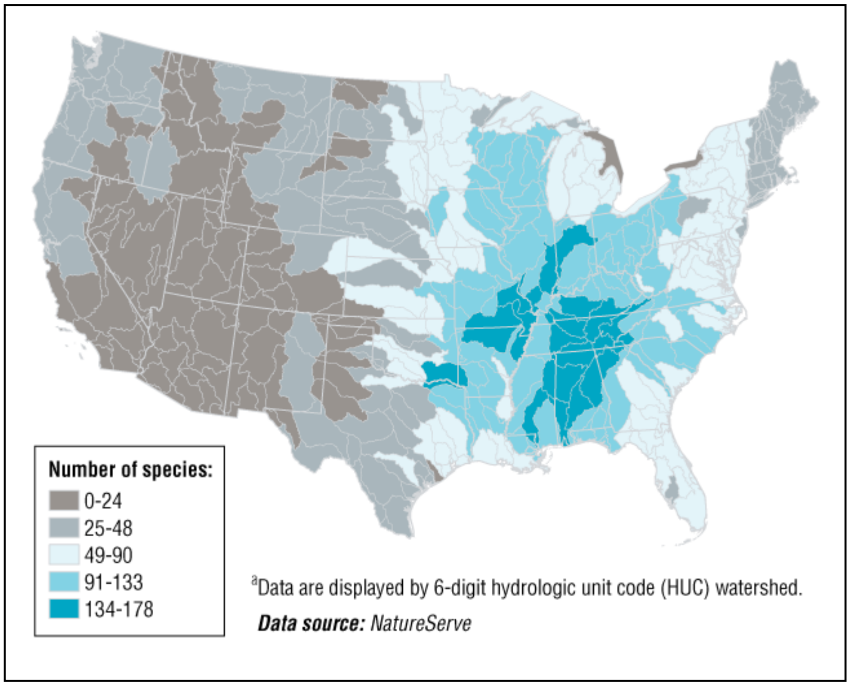


Figure 8-2. Historical diversity of native fish species in the contiguous U.S., 1970[a] (U.S. EPA, 2008c).

Intactness is expressed as a percentage based on the following formula:

reduction in diversity = 1 - (# of current native species/# of historic native species)

This Fish Faunal Intactness indicator makes use of empirical, rather than modeled, data sets, and it focuses on a well-known group of organisms with a fairly strong historical record. Fish have historically been surveyed more extensively than any other aquatic fauna: systematic sampling of fish in the United States began in the nineteenth century. This indicator relies heavily on species spatial distribution data that have been compiled in state "Fishes of" books. Primary data were quality-controlled by experts in every state, primarily academics and scientists from the USGS. Additional information describing the status and location of rare and endangered species was provided by the Network of State Heritage Programs through NatureServe.[41,42]

[41]Information about the indicators, their respective metadata, and independent data analyses were extracted and carried out between December 2006 and September 2007.

[42]Species found are considered present whether they are stocked yearly or an established species (P. Fuller, personal communication, May 2007).

This case study characterizes the uncertainty related to the Fish Faunal Intactness indicator using available information. Key information necessary for more complete and quantitative characterization of uncertainty was not available to the authors of this case study. Figure 8-3 depicts the general sequence of data capture and processing involved in preparing ROE indicators. We identified the Primary Data Capture step as the major source of uncertainty pertaining to the ROE end use for the Fish Faunal Intactness indicator. For example, the Data Capture Plans and Methods used in primary studies were deemed important sources of uncertainty, but the sampling methods, the number and timing of replicates, and the number of sampling points across representative environments within a reporting area from the individual studies underlying the indicator were not available. If this information becomes available, a more complete characterization of the uncertainty for this indicator could be conducted. Because of the absence of primary data source information, the uncertainty characterization for this case study is, therefore, qualitative.

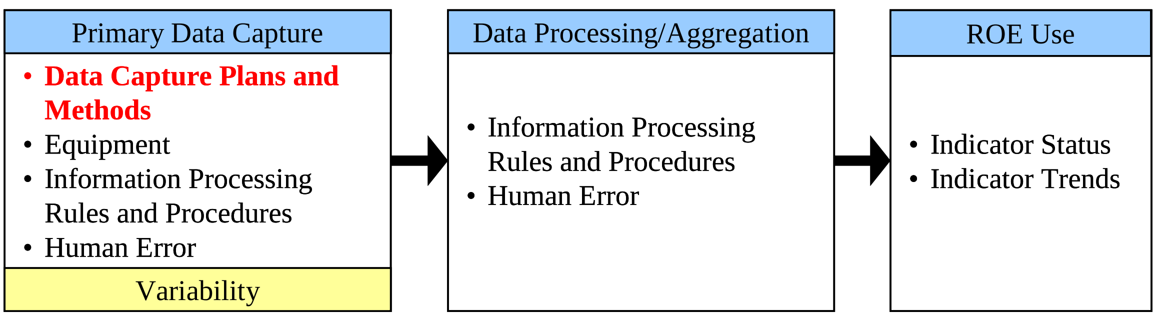


Figure 8-3. General sequence of data capture and processing involved in preparing the fish faunal intactness indicator.

8.2. OVERVIEW OF UNCERTAINTY

Overview: Uncertainty

- Incompleteness of sample data for any given study contributes to uncertainty in individual studies' characterization of native fish species' intactness and at the HUC level.
- Errors in species characterization and determination of presence for individual studies lead to mischaracterization of species' presence at the HUC level.
- The methodology used to determine that a species is present in a HUC results in the potential for over-reporting of the actual range for any given species.

Historical data provide a conservative inventory of the richness of fish populations. These historical data generally represent incomplete samples of any particular site, and are subject to gaps in geographic coverage. The mechanism by which the indicator is calculated partially addresses this issue by assuming that all contemporary species were also present historically (and included in the denominator of the indicator formula). This assumption, however, does not account for species purposefully or accidentally introduced after the historical data were obtained, as discussed in Section 8.2.2. In general, the accuracy of historical fish distributions varies by state according to the comprehensiveness and quality of their catalogs and recordkeeping. However, this case study does not identify an assessment of the accuracy of historical records.

Conversely, there are undoubtedly instances in which a species continues to exist in a watershed, but has not been documented since 1970, or those reports were not captured in the creation of this data set. This would have the effect of overstating apparent reductions in diversity.

Three specific categories of uncertainty relevant to the Fish Faunal Intactness indicator are

- Uncertainty in completeness of sample data.
- Uncertainty and errors in characterization of species taxonomy and status.
- Uncertainty in representation of species spatial distribution.

Each category includes a number of specific sources of uncertainty. Figure 8-4 shows the data flow relationship to the types of uncertainty.

Table 8-1 lists categories and sources of uncertainty, along with their potential impact on the ROE indicator.

8.2.1. Completeness of Sample Data

8.2.1.1. *Completeness of a Species Inventory at a Site (Within a Sample and Across Time)*

Two types of omission errors occur with species inventories at the site level. The first type occurs within a site sample, as there is a high likelihood that not all species will be represented. Fish are difficult to capture, and rarer species will be less likely to be caught in any given sample. Different sampling techniques are more effective than others, and increasing the numbers of replicate samples will increase the proportion of species captured. The second type of omission error is temporal. A single sampling event will not capture the full species composition at any site, as there are seasonal and annual cycles of species and population fluxes. A thorough site inventory must include replicated sampling across the relevant temporal cycles for each aquatic system. Information detailing the sampling methods, the number of replicates, and the timing of these replicates was not available for this analysis, but would help to provide uncertainty attributes for site-level sampling events.

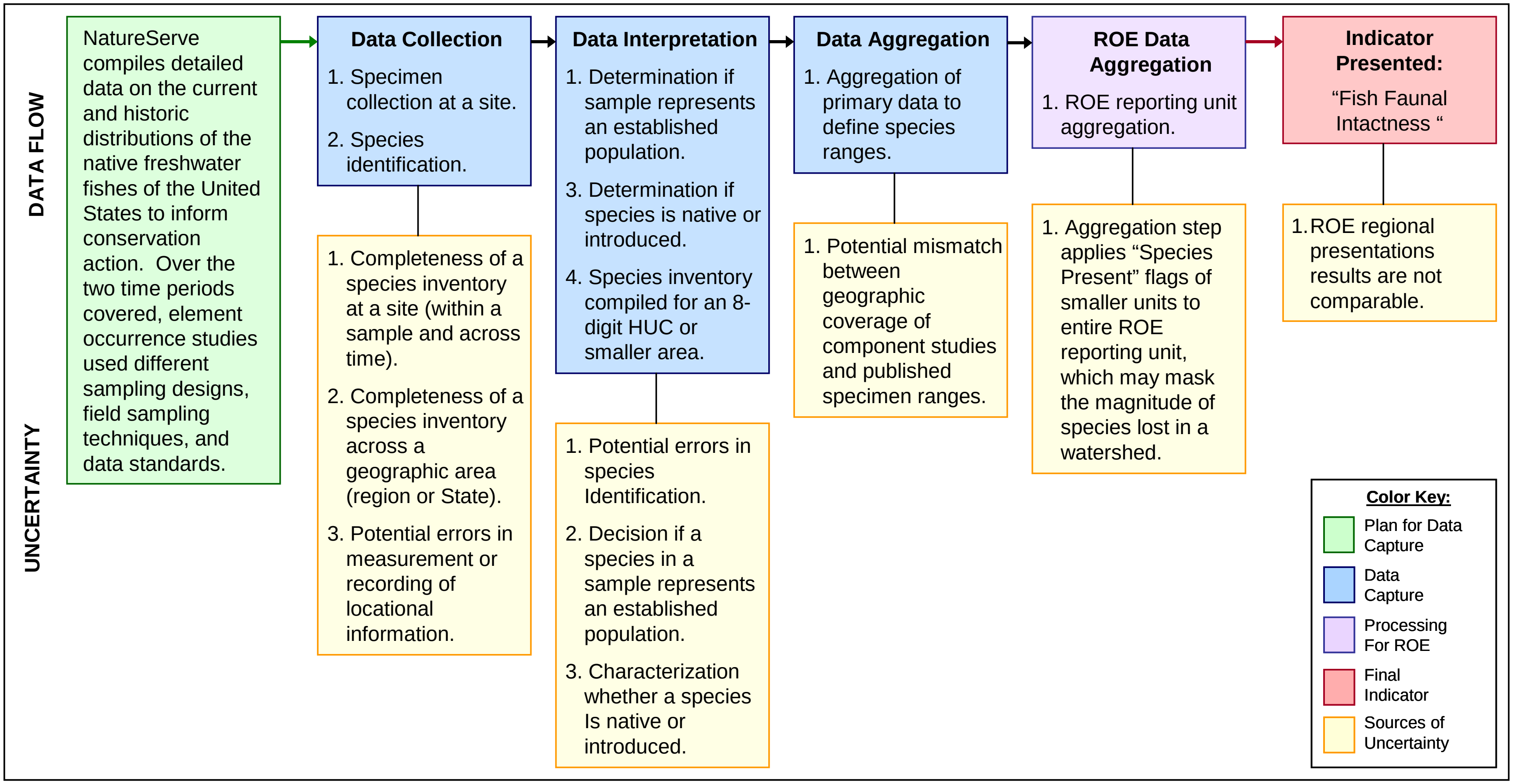


Figure 8-4. Uncertainty and data flow: fish faunal intactness.

Table 8-1. Sensitivity of the indicator to potential sources of uncertainty (derived by Abt Associates)

Potential sources of uncertainty		Sensitivity of the indicator
1. Completeness of sample data		
1a	Completeness of a species inventory at a site (within a sample and across time)	High
1b	Completeness of a species inventory across a geographic area (e.g., state)	High
2. Characterization of species taxonomy and status		
2a	Species identification	Low
2b	Determining whether a species in a sample represents an established population	High
2c	Determining whether a species is native or introduced	Medium
3. Representation of species spatial distribution		
3a	Locational information	Medium
3b	Aggregation of primary data to represent species ranges	Medium
3c	Aggregation of species range maps into ROE reporting units	Medium

8.2.1.2. *Completeness of a Species Inventory Across a Geographic Area*

Each reporting region (e.g., state, ecoregion, administrative region, watershed) contains a variety of different aquatic environments. These environments represent different physical, chemical, and biological conditions that, in turn, determine which species could be present. All representative environment types must be sampled to obtain a complete picture of the species composition within each reporting region. Information on the number of sampling points across representative environments within the reporting area was not available for this analysis, but having such information would provide valuable uncertainty metadata for this indicator.

Summary: Experts rate the completeness of the sample data as a source of uncertainty that has a direct impact on the measure of native fish species intactness for a given HUC.[43] The historical information was collected over a long period of time and assumed to be a reasonably

[43] Telephone communication between L.L. Master (NatureServe) and Andrew Stoeckle (Abt Associates) on April 30, 2007.

complete representation of the fish fauna. The current site-level measurements will have a tendency for under-representation based on the errors discussed above. As a result, this indicator will have a tendency to report a lower level of fish intactness than may truly exist.

8.2.2. Characterization of Species Taxonomy and Status

8.2.2.1. *Species Identification*

There are two sources of error related to the identification of species. The first type of error stems from the incorrect classification of a captured species. Museum collections provide some information that can be used to determine the extent of misclassification errors. The second type of identification error results from the use of different approaches for the classification of fish fauna. There are generally a small percentage of species that ichthyologists will classify differently based on different taxonomic approaches. The advent of genetic classification techniques has clarified some of these taxonomic controversies. Species identification errors are not expected to create a large level of uncertainty in this indicator.

8.2.2.2. *Determining Whether a Species in a Sample Represents an Established Population*

Ideally, this indicator would incorporate information on only those species whose populations were established in the reporting region. Many species that are represented in any sample have been purposefully or accidentally introduced and do not represent stable established populations that could sustain a presence in the watershed. It is necessary to document the repeated presence of introduced species over time at a site without successive re-introduction to determine that it is an established population. Species that are stocked annually should be considered within this analysis because their effect is the same as if they were an established breeding population at this site. Experts believe that this source of uncertainty could have a high impact on the development of this indicator at the HUC level.

8.2.2.3. *Determining Whether a Species is Native or Introduced*

This indicator measures changes in the ratio of native to introduced species for any region over time. In some cases, scientists are not unanimous as to whether a particular species has been introduced or is native. The importance of this source of uncertainty is ranked as medium relative to its ability to skew the indicator results.

Summary: The greatest source of error in this category occurs when native species are tallied that are truly not established in the reporting unit; the indicator will then demonstrate a higher level of faunal intactness than is appropriate. A similar bias will occur when introduced species are listed as native.

8.2.3. Representation of Species Spatial Distribution

8.2.3.1. *Locational Information*

The locational information associated with historic samples rarely complies with current standards for locational precision, definition, and documentation. We applied expert judgment to determine the appropriate use of historical sampling data with imprecise locational accuracy in many cases. Such expert determinations are difficult to document and incorporate into summary analyses when aggregating multiple samples from multiple sources.

8.2.3.2. *Aggregation of Primary Data to Represent Species Ranges*

The precision of the primary (site-specific) data regarding species presence has a strong effect on whether a fish is included in a regional species list. This indicator relies on watershed, regional, or state-scale species range maps that were compiled by experts using locational information from multiple primary data sources. These lists end up including (a) species that are well established across the entire area; (b) geographically peripheral species that occupy a slight percentage of the entire reporting region; and (c) species that are incorrectly assigned to a region as a result of ambiguous locational information in the primary sample data. Uncertainty in the primary data's spatial coverage generally results in the exaggerated representations of species with range maps.

8.2.3.3. *Aggregation of Species Range Maps into Report on the Environment (ROE) Reporting Units*

The development of this indicator showing the present and historical composition of native fishes for each ROE reporting unit was completed through the aggregation of range map information to the 6-digit HUC code level. The errors associated with the original data within specific range maps are carried into these aggregated ROE reporting units. Aggregation to the ROE reporting units will tend to expand the assumed range for any given species, which is demonstrated in Figure 8-5. When aggregated to the 6-digit level, the species is portrayed as intact throughout the unit. However, given species may not be studied or could be disappearing from the smaller HUCs that make up the larger ROE reporting unit. This situation is not seen at the aggregated level until the species disappears from the unit entirely. Thus, while the indicator may depict an intact population exists within a ROE 6-digit HUC, it does not imply that the species is present in each of the smaller HUCs. The influence of aggregating to ROE reporting units is a moderate source of uncertainty in the overall ROE indicator value.

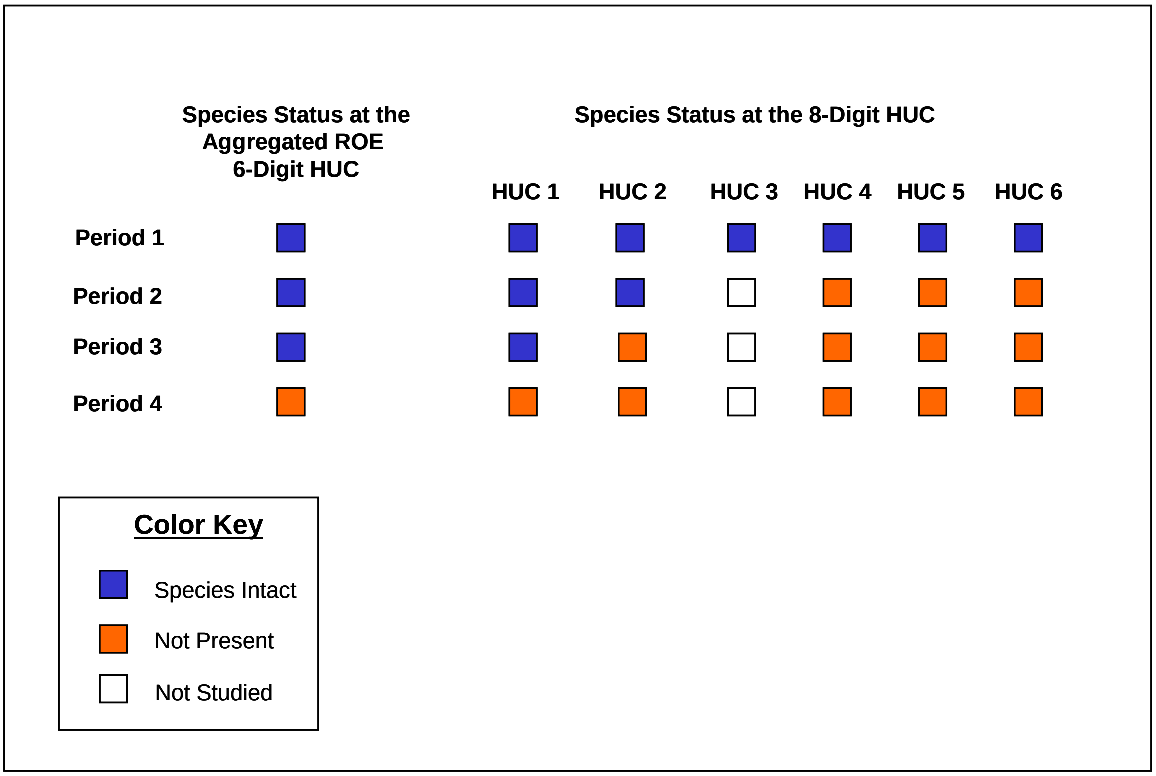


Figure 8-5. Illustration of species status at the 6- and 8-digit HUCs.

Summary: In most cases, imprecise original locations and the aggregation of smaller reporting units into larger units results in the potential for over-reporting of the actual range for any given species. The use of very large watersheds as reporting units is effective in reducing errors associated with over counting numbers of species within a reporting unit, but this also provides an indicator that is sufficiently geographically generalized that it provides limited practical use for analysis and management.

8.3. REGIONALIZATION

This indicator is based on distributional data for 782 native freshwater fish species across 8-digit HUC watersheds. Because of the uncertainties listed above, the 8-digit HUC data were pooled, and the Fish Faunal Intactness Indicator is currently reported by larger 6-digit HUC units to reduce potential error of over-reporting the actual numbers of species within a HUC. It is generally agreed that reporting at the 8-digit scale would provide a more suitable scale for the interpretation and use of this indicator, but there is concern that the current state of information and aggregation processes will not result in a robust indicator at this finer scale. The 6-digit HUC units represent large heterogeneous watershed units that will generally over-represent actual species ranges. One would need to go back to the original records for specific site locations, as well as include more-recent data in order to report at the 8-digit HUC level with a similar level of presence/absence confidence. Experts suggest that this is both feasible and desirable and would allow robust aggregation of these units into broader regional representations.

At the 6-digit HUC, there are approximately 25–30 watersheds per EPA Region, whereas with an 8-digit HUC, there are closer to 2,000 watersheds. It is not advised that EPA further aggregate and disaggregate the current 6-digit HUC indicator information to create additional data sets and reports at the EPA Regional level. Not only would this have little practical utility for the Agency relative to protection and management of the resource, this would also introduce an additional level of uncertainty in the robustness of the indicator. Judgment calls would need to be made at the edges of EPA Regions and the HUC units regarding the appropriate attribution of range, which cannot be completed without reference to the original data sets. To maintain data integrity, EPA should strive to report this indicator at the 8-digit HUC level, and then can evaluate specific situations where any individual HUC is split between multiple EPA Regions.

9. CONCLUSIONS

Examination of the seven case studies reveals a number of implications for more systematically quantifying uncertainty in the ROE indicators. This overview assessment of the seven case studies addresses four questions concerning whether sufficient information was available to quantify uncertainty and if so, if the uncertainty was large enough to affect interpretation of the time series trend or change as presented in ROE.

For two of the seven indicator case studies (ambient PM concentrations and blood lead), uncertainty is both known and small enough to detect changes in the indicators over time (see Figure 9-1). For two other indicators (the CWQI and Fish Faunal Intactness), uncertainty is largely unknown. In the first case, the unknown uncertainty stems from the fact that the design of the surveys may under-sample concentrations of the chemicals that contribute to the index during transient events, rather than uncertainty in the indicator value itself. In the second case, the unknown uncertainty has to do with the unknown accuracy of the data that support the indicator. For the remaining three indicators, we quantified some—but not all—of the sources of uncertainty using readily available information. Unquantified sources of uncertainty for these three indicators are thought to be significant, which might alter the interpretation of trends and current status.

The indicators chosen for the pilot study are typical of many, but not all, of the other 85 indicators in the ROE. For example, the Ambient PM indicator is typical of three other ambient air indicators measured in a similar way (although there are differences in the time intervals, instrumental methods, and the number and location of sampling sites). The Blood Lead-Levels indicator is typical of five other tissue concentration indicators measured by the NHANES program. The use of the index period, single annual visit sampling design is typical of at least nine other indicators in the ROE, but many of these measure aspects of the systems that are less transient than water quality. The other case study indicators have aspects in common with other indicators in the ROE, but not as consistently as those cited above.

In general, the project suggests that systematically quantifying uncertainty in the ROE indicators will be challenging, but that in most cases, at least some of the uncertainty can be quantified, and, thus, reduced. The study has identified many of these challenges and opportunities, and will greatly facilitate the addition of quantitative uncertainty estimates to the indicators in the future. The implications of the quantifiable and unquantifiable uncertainty on the utility of the indicators for informing EPA and the public will have to be determined on an indicator-by-indicator basis. This was not an objective of the pilot study.

Report on the Environment (ROE) Indicator Case Studies Assessment

	Ambient Particulate Matter Concentrations	Coastal Water Quality Index	Population Served by Community Water Systems with no Reported Violations of Heath-Based Standards	Reported Toxic Chemicals in Wastes Released, Treated, Recycled, or Recovered for Energy Use	High Priority Clean-up Sites where Ground Water is not Continuing to Spread Above Levels of Concern	Fish Faunal Intactness	Blood Lead Level
Is quantified uncertainty [small enough or too large] to detect change in the indicator values over time?	Small Enough	Uncertainty is largely unknown	POSSIBLY Small Enough	POSSIBLY Small Enough	Small Enough	Uncertainty is largely unknown	Small Enough
Are we confident that we have identified all major sources of uncertainty and variability?	YES	YES	YES	YES	YES	YES	YES
Can the major sources of uncertainty in current status be quantified with the data currently available?	YES	YES	YES	NO	YES	NO	YES
Can the major sources of uncertainty and variability in trends be quantified with the data currently available?	YES	YES	YES	NO	YES	NO	YES

Figure 9-1. Overview of uncertainty findings.

These case studies also found that the information needed to characterize and quantify uncertainty, variability, and limitations is neither consistently nor readily available for ROE indicators.

9.1. FURTHER ANALYSES

U.S. Environmental Protection Agency staff, peer reviewers, and content experts generously contributed their time and thoughts during the review of this document. As part of this review process, additional documentation related to uncertainty, limitations, and scaling were identified. In addition, the reviewers identified the analyses presented below as useful in further quantifying uncertainty associated with the ROE indicators.

9.1.1. Ambient Particulate Matter (PM) Concentration Indicators

- It is possible that heterogeneous features in the space-time variation of air pollution, such as variation due to weather effects, would generate larger errors in one direction of the PM indicator. To the extent this heterogeneity occurs, stochastic space-time analysis could be used to quantify its contribution to uncertainty in the PM indicator.
- For each monitor, if missed monitoring days are clustered over time, then there would be a larger impact on the magnitude of the monitor-specific annual average than if missing daily values are randomly distributed throughout the year. To the extent that clustering of missed monitoring days occurs, a similar analysis to that conducted in Section 2.2.2, which randomly selects clusters of days to be omitted, could be conducted to quantify the contribution of clustered missed monitoring days to the overall uncertainty in annual PM concentration averages.

9.1.2. Coastal Water Quality Index (CWQI) Indicator

- For each indicator in the CWQI, heterogeneity due to temporal (i.e., seasonal or annual) variability would likely generate different uncertainty estimates for each Region. It is noted in the limitations to the ROE indicator that temporal variability cannot be accounted for with the current dataset of only one sample per station, but may be possible if additional data become available for the stations sampled in the CWQI. To the extent this heterogeneity occurs in a long-term dataset, stochastic space-time analysis could be used to quantify the contribution of heterogeneity in the uncertainty in the CWQI indicator.
- The presentation of the CWQI indicator as a categorical variable may create a loss of information across the different approaches to aggregating data at the indicator, regional, and national levels. To the extent that imprecision is created by the index aggregation, three approaches could be used to estimate uncertainty in the overall index:
    - a log-linear model of the CWQI index as a function of the five measured indicators;

- a probabilistic simulation of the CWQI index using the distribution of the component variables; and/or
- construction of confidence intervals for the national aggregate database of indicator measurements.

9.1.3. Community Water Systems (CWSs) Indicator

- According to recent triennial audits of SDWIS/FED (U.S. EPA 2006b), approximately one quarter of all CWSs do not provide the required samples needed to determine health-based violations. Unreported samples result in an overestimation of the ROE indicator for CWSs since potential health-based violations, which would lower the national indicator values, are not identified. If information regarding the rate of health-based violations in unreported samples becomes available, further analysis should be conducted to assess the overestimation of population served by CWSs with no health-based violations.
- Populations serviced by CWSs are based on estimations rather than true service populations. Such estimations may overestimate or underestimate ROE indicator values. If information regarding the true service populations becomes available, further analysis should be conducted to assess the level of uncertainty caused in the ROE indicator by estimating service populations.

9.1.4. Toxics Release Inventory (TRI) Indicators

- The use of heuristic methods for estimates and filing decisions may lead to errors in the quantities or releases and waste management reported to TRI on individual forms. If heuristics cause TRI reported quantities to be consistent overestimates (or underestimates), the national and Regional totals presented in the TRI-based ROE indicators would be influenced in the same direction. To the extent the use of heuristics affects the directionality of errors in TRI reporting, algorithms could be used to quantify their contribution to uncertainty in the TRI-based ROE indicator values.

9.1.5. Ground Water (GW) Indicator

- Sources of error in data capture, in particular from machine error, may result in additional uncertainty within the indicator. Additional analyses of data capture and a further review of raw data may significantly contribute to the reduction of this uncertainty.

9.1.6. Blood Lead Level (BLL) Indicator

- It is possible that the differences between the 2 years of data presented for this indicator are based on differences in the sample populations, particularly for reasons such as environmental lead exposure. To the extent that the sample populations are significantly different, and to the extent that the underlying data and their weights could be obtained by CDC, statistical methods such as models or principal component analysis could be used to deconvolute blood lead level differences due to population characteristics (such as age, gender, race, house age, population density, etc.). As more data are presented in future years, the degree to which these characteristics differ could be used to quantify the uncertainty in BLL indicator trends.

9.1.7. Fish Faunal Intactness Indicator

- Include additional information on the completeness of data with respect to (1) sampling frame and spatial and temporal coverage, (2) sample selection of sites within a geographic areas and, (3) field sampling methods within sites.
- Highlight the potential differences in the methods used for historical monitoring versus recent monitoring.

REFERENCES

Agresti, A. (1990) Categorical data analysis. New York: John Wiley and Sons; p. 558.

CDC (Centers for Disease Control and Prevention). (1991) Preventing lead poisoning in young children. Accessed November 21, 2004. Available online at http://aepo-xdv-www.epo.cdc.gov/wonder/prevguid/p0000029/p0000029.asp.

CDC (Centers for Disease Control and Prevention). (2002) NHANES 1999-2000 addendum to the NHANES III analytic guidelines. Last update 8/30/02. Available online at http://www.cdc.gov/nchs/data/nhanes/guidelines1.pdf.

CDC (Centers for Disease Control and Prevention). (2005a) Third national report on human exposure to environmental chemicals. NCEH publication no. 05-0570. Available online at http://www.cdc.gov/exposurereport/report.htm.

CDC (Centers for Disease Control and Prevention). (2005b) Morbidity and mortality weekly report (MWWR). Blood lead levels—United States, 1999-2002. Available online at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5420a5.htm.

CDC (Centers for Disease Control and Prevention). (2007) NHANES 2001-2002 data release. Available online at http://www.cdc.gov/nchs/data/nhanes/nhanes_01_02/cvx_b_doc.pdf.

Diaz, RJ; Rosenberg, R. (1995) Marine benthic hypoxia: a review of its ecological effects and the behavioral responses of benthic macrofauna. Oceanogr Mar Biol Ann Rev 33:245–303.

Nadakavukaren, A. (2000) Our global environment: a health perspective. 5th Ed. Prospect Heights, IL: Waveland Press, Inc.

NRC (National Research Council). (1993) Measuring lead exposure in infants, children, and other sensitive populations. Washington, DC: National Academies Press.

NRC (National Research Council). (2000) Ecological indicators for the nation. Washington, DC: National Academies Press. Available online at http://www.nap.edu/openbook/0309068452/html/.

Pirkle, J. L., Kaufmann, R.B., Brody, D. J., et al. (1998) "Exposure of the U.S. Population to Lead, 1991-1994." *Environmental Health Perspectives* Volume 106, Number 11.

U.S. Census Bureau. (2007) Monthly population estimates for the United States: April 1, 2000 to November 1, 2007. NA-EST2006-01. Available online at http://www.census.gov/popest/national/tables/NA-EST2006-01.xls [accessed October 2007]. Available online at http://search.census.gov/search?q=Monthly+population+estimates+for+the+United+States%3A+April+1%2C+2000+to+September+1%2C+2007&spell=1&client=subsite&access=p&ie=UTF-8&subtitle=hhes-xyz&output=xml_no_dtd&proxystylesheet=subsite&filter=0.

U.S. DOE (Department of Energy). (1996) National corrective action prioritization system (NCAPS). RCRA information brief. Office of Environmental Policy and Assistance, Washington, DC. EH-413-070/1296 . Available online at http://homer.ornl.gov/nuclearsafety/env/guidance/rcra/ncaps.pdf..

U.S. EPA (Environmental Protection Agency). (1995) Environmental monitoring and assessment program (EMAP): laboratory methods manual—estuaries. Volume 1: biological and physical analyses. Office of Research and Development, Narragansett, RI; EPA/620/R-95/008. Available online at http://www.epa.gov/emap/html/pubs/docs/groupdocs/estuary/field/lab_man.pdf.

U.S. EPA (Environmental Protection Agency). (1996) Coordination between RCRA corrective action and closure and CERCLA site activities. Memo from Steven A. Herman, Office of Enforcement and Compliance Assurance and Elliot P. Laws, Office of Solid Waste and Emergency Response, to RCRA/CERCLA National Policy Managers. Available online at http://www.epa.gov/fedfac/documents/924memo.htm.

U.S. EPA (Environmental Protection Agency). (1999a) Interim—final guidance for RCRA corrective action environmental indicators. Available online at http://www.epa.gov/epaoswer/hazwaste/ca/eis/ei_guida.pdf.

U.S. EPA (Environmental Protection Agency). (1999b) Press release: EPA and states to accelerate cleanups at more than 1,700 hazardous waste management sites. Available online at http://www.epa.gov/history/topics/cercla/09.htm.

U.S. EPA (Environmental Protection Agency). (2000a) Ambient aquatic life water quality criteria for dissolved oxygen (saltwater): Cape Cod to Cape Hatteras. Office of Water, Washington, DC; EPA/822/R-00/012. November 2000. Available online at http://nepis.epa.gov/EPA/html/Pubs/pubtitleOW.htm.

U.S. EPA (Environmental Protection Agency). (2000b) Data reliability analysis of the EPA safe drinking water information system/federal version (SDWIS/FED). Office of Water, Washington, DC; EPA 816-R-00-020. Available online at http://www.epa.gov/ogwdw/data/pdfs/reports_draap_final_2000.pdf.

U.S. EPA (Environmental Protection Agency). (2000c) RCRA corrective action. RCRA corrective action focuses on interim priorities—better integration with final goals needed. Available online at http://www.epa.gov/oig/reports/2000/rcraaction.pdf.

U.S. EPA (Environmental Protection Agency). (2001a) National coastal assessment quality assurance project plan, 2001-2004. Environmental Monitoring and Assessment Program (EMAP). Office of Research and Development, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL; EPA/620/R-01/002. Available online at http://www.ecy.wa.gov/programs/eap/psamp/EMAP/EMAPQAPP2001.pdf.

U.S. EPA (Environmental Protection Agency). (2001b) National coastal assessment: field operations manual. Office of Research and Development, National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, FL; EPA 620/R-01/003. Available online at http://nepis.epa.gov/EPA/html/Pubs/pubtitleORD.htm.

U.S. EPA (Environmental Protection Agency). (2002) 2000 toxics release inventory (TRI) public data release report, executive summary. Office of Environmental Information, Washington, DC; EPA/260/S-02/001. Available online at http://www.epa.gov/tri/tridata/tri00/pdr/execsummary_final.pdf.

U.S. EPA (Environmental Protection Agency). (2003) Mid-Atlantic integrated assessment, MAIA—estuaries 1997-98, summary report. Office of Research and Development, National Health and Environmental Effects Research Laboratory, Atlantic Ecology Division; EPA/620/R-02/003. Available online at http://nepis.epa.gov/EPA/html/Pubs/pubtitleORD.htm.

U.S. EPA (Environmental Protection Agency). (2004a) Air quality criteria document for particulate matter. Volume 1, Oct. 2004. Office of Research and Development, Washington, DC; EPA/600/P-99/002aF, pp. 2–55. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=2832.

U.S. EPA (Environmental Protection Agency). (2004b) National coastal condition report II. Office of Research and Development/Office of Water, Washington, DC; EPA/620/R-03/002. Available online at http://www.epa.gov/owow/oceans/nccr/2005/.

U.S. EPA (Environmental Protection Agency). (2004c) Safe drinking water act 30th anniversary fact sheet: drinking water standards and health effects. Office of Water, Washington, DC; EPA/816/F-04/037. Available online at http://www.epa.gov/safewater/sdwa/pdfs/fs_30ann_standards_web.pdf.

U.S. EPA (Environmental Protection Agency). (2004d) Safe drinking water act 30th anniversary fact sheet: glossary. Office of Water, Washington, DC; EPA/816/F-04/035. Available online at http://www.epa.gov/safewater/sdwa/pdfs/fs_30ann_glossary_web.pdf.

U.S. EPA (Environmental Protection Agency). (2004e) Factoids: drinking water and ground water statistics for 2003. January 2004, EPA 816-K-03-001. Available online at http://www.epa.gov/safewater/data/pdfs/data_factoids_2003.pdf.

U.S. EPA (Environmental Protection Agency). (2004f) Drinking water data reliability analysis and action plan (2003)—for state reported public water system data in the EPA safe drinking water information system/federal version (SDWIS/FED). March 2004, EPA 816-R-03-021. Available online at http://www.epa.gov/ogwdw/data/pdfs/reports_draap_final_2003.pdf.

U.S. EPA (Environmental Protection Agency). (2004g) Toxics release inventory (TRI) public data release report. Toxic Release Inventory Program, Washington, DC; EPA/260/R-06/001. Available online at http://www.epa.gov/tri/tridata/tri04/brochure/2004brochure.pdf

U.S. EPA (Environmental Protection Agency). (2005a) Factoids: drinking water and ground water statistics for 2004. Office of Water, Washington, DC; EPA 816-K-05-001. Available online at http://www.epa.gov/safewater/data/pdfs/data_factoids_2004.pdf.

U.S. EPA (Environmental Protection Agency). (2006a) Regulating public water systems and contaminants under the Safe Drinking Water Act. .Available online at http://www.epa.gov/safewater/standard/setting.html [accessed November 2006].

U.S. EPA (Environmental Protection Agency). (2006b) Safe drinking water information system/federal version. Available online at http://www.epa.gov/safewater/databases/sdwis/index.html. [accessed February 2006].

U.S. EPA (Environmental Protection Agency). (2006c) 2006-2011 EPA strategic plan: charting our course. Pre-publication copy. September 29, 2006. EPA-190-R-006-001. Available online at http://www.epa.gov/ocfo/plan/2006/entire_report.pdf.

U.S. EPA (Environmental Protection Agency). (2006d) Promising techniques identified to improve drinking water laboratory integrity and reduce public health risks. Office of the Inspector General, Washington DC; 2006-P-00036. Available online at http://www.epa.gov/oig/reports/2006/20060921-2006-P-00036.pdf.

U.S. EPA (Environmental Protection Agency). (2006e) Supplemental report: survey results: EPA Regional Drinking Water Laboratory certification authorities, Office of the Inspector General, Washington DC, October 5, 2006. Available online at http://www.epa.gov/oig/reports/2006/20061005-2006-P-00036A.pdf.

U.S. EPA (Environmental Protection Agency). (2007a) 3-Year quality assurance report for calendar years 2002, 2003 and 2004. Table 1. Office of Air Quality Planning and Standards, Ambient Air Monitoring Group, Research Triangle Park, NC. Available online at http://www.epa.gov/ttn/amtic/files/ambient/PM2.5%202002_2004%20final%20report.pdf.

U.S. EPA (Environmental Protection Agency). (2007b) EPA's 2007 report on the environment: science report. Chapter 2, exhibit 2-24, p. 2–47. National Center for Environmental Assessment, Washington, DC. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=140917.

U.S. EPA (Environmental Protection Agency). (2007c) How to access local drinking water information. Safe drinking water information system (SDWIS). Office of Water, Washington, DC. Accessed October 2007. Available online at http://www.epa.gov/safewater/data/getdata.html.

U.S. EPA (Environmental Protection Agency). (2007d) 2005 Toxics release inventory (TRI) public data release. Available online at http://www.epa.gov/tri/tridata/tri05/.

U.S. EPA (Environmental Protection Agency). (2007e) EPA analysis of decision in Barrick Goldstrike Mines, Inc. v. Whitman. Available online at http://www.epa.gov/triinter/lawsandregs/mining/barrick_lawsuit_epa_analysis.htm [Accessed November 28, 2007].

U.S. EPA (Environmental Protection Agency). (2007f) Superfund environmental indicators. Available online at http://www.epa.gov/superfund/accomp/ei/ei.htm.

U.S. EPA (Environmental Protection Agency). (2007g) Corrective action. Available online at http://www.epa.gov/epawaste/hazard/correctiveaction/index.htm.

U.S. EPA (Environmental Protection Agency). (2007h) Environmental indicators - frequently asked questions. Available online at http://www.epa.gov/epawaste/hazard/correctiveaction/eis/faqs.htm.

U.S. EPA (Environmental Protection Agency). (2007i) RCRA corrective action EI guidelines. Available online at http://www.epa.gov/epawaste/hazard/correctiveaction/eis/ei_guida.pdf.

U.S. EPA (Environmental Protection Agency). (2007j) Science report (SAB review draft). Report on the Environment (ROE). Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=140917.

U.S. EPA (Environmental Protection Agency). (2007k) Comprehensive Environmental Response, Compensation and Liability Information System (CERCLIS) database. Available online at http://cfpub.epa.gov/supercpad/cursites/srchsites.cfm.

U.S. EPA (Environmental Protection Agency). (2008a) Trophic state of coastal waters. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=47&subtop=274&lv=list.listByChapter&r=188229 [accessed on August 29, 2008].

U.S. EPA (Environmental Protection Agency). (2008b) RCRA corrective action environmental indicators. Accessed 2008. Available online at http://www.epa.gov/epawaste/hazard/correctiveaction/eis/slides/overview.pdf.

U.S. EPA (Environmental Protection Agency). (2008c) Fish faunal intactness. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=50&subtop=211&lv=list.listByChapter&r=188264.

U.S. EPA (Environmental Protection Agency). (2008d). Ambient concentration of particulate matter. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=46&subtop=341&lv=list.listByChapter&r=188192 [accessed on August 29, 2008].

U.S. EPA (Environmental Protection Agency). (2008e). Ambient concentrations of particulate matter: metadata. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewMeta&ch=46&lShowInd=0&subtop=341&lv=list.listByChapter&r=188192. [accessed 2008].

U.S. EPA (Environmental Protection Agency). (2008f) Population served by community water systems with no reported violations of health-based standards: metadata. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=47&subtop=203&lv=list.listByChapter&r=188226 [accessed on August 29, 2008].

U.S. EPA (Environmental Protection Agency). (2008g) Population served by community water systems with no reported violations of health-based standards. Report on the environment (ROE): Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=47&subtop=203&lv=list.listByChapter&r=201564.

U.S. EPA (Environmental Protection Agency). (2008h) Toxic chemicals in production-related wastes combusted for energy recovery, released, treated, or recycled: indicators. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=48&subtop=228&lv=list.listByChapter&r=188241 [accessed on August 29, 2008].

U.S. EPA (Environmental Protection Agency). (2008i) Migration of contaminated ground water under control at high-priority cleanup sites: indicators. Report on the Environment. Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&ch=48&subtop=313&lv=list.listByChapter&r=188232 [accessed on August, 29, 2008].

U.S. EPA (Environmental Protection Agency). (2008j) Migration of contaminated ground water under control at high-priority cleanup sites: charts and graphs. Report on the Environment. Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewMidImg&ch=48&lShowInd=0&subtop=313&lv=list.listByChapter&r=188232 [accessed 2008].

U.S. EPA (Environmental Protection Agency). (2008k) Safe drinking water information system (SDWIS): overview. Envirofacts. Washington, DC. Available online at http://www.epa.gov/enviro/html/sdwis/index.html.

U.S. EPA (Environmental Protection Agency). (2008l) Superfund environmental indicators guidance: human exposure revisions. Office of Superfund Remediation and Technology Innovation, Washington, DC. Available online at http://www.epa.gov/superfund/accomp/ei/pdfs/final_ei_guidance_march_2008.pdf [accessed 2009].

U.S. EPA (Environmental Protection Agency). (2008m) Drinking water. Report on the Environment (ROE). May 2008. Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=list.listBySubTopic&lv=list.listByChapter&ch=47&s=203.

U.S. EPA (Environmental Protection Agency). (2008n) Blood Lead Level. Report on the Environment. Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewInd&lv=list.listByAlpha&r=188246&subtop=208 [accessed 2008].

U.S. EPA (Environmental Protection Agency). (2008o) Blood lead level: metadata. Report on the Environment (ROE). Available online at http://cfpub.epa.gov/eroe/index.cfm?fuseaction=detail.viewMeta&ch=49&lShowInd=0&subtop=208&lv=list.listByChapter&r=188246.

U.S. EPA (Environmental Protection Agency). (2008p) Superfund environmental indicators guidance human exposure revisions. March 2008. Available online at http://www.epa.gov/superfund/accomp/ei/pdfs/final_ei_guidance_march_2008.pdf.

U.S. GAO (General Accounting Office). (1998) Information on the status of Superfund sites. Report to Congress Requesters, United States General Accounting Office. Available online at http://www.gao.gov/archive/1998/rc98241.pdf.