APPENDIX E

TNSSS SAMPLE DESIGN REPORT

United States
Environmental Protection
Agency

Statistical Design
for the
Targeted National
Sewage Sludge Survey

August 31, 2006

Acknowledgments and Disclaimer

This report has been reviewed and approved for publication by the Office of Science and Technology. The work described in the report was accomplished with the support of Battelle, Westat, and Computer Sciences Corporation, under the direction and review of the Office of Science and Technology.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

1.0 INTRODUCTION

This document describes the statistical design for EPA's Targeted National Sewage Sludge Survey (TNSSS). This survey will collect physical samples of biosolids from publicly owned treatment works (POTWs). The samples will be analyzed for selected chemical, microbial, and other pollutants. In this study design, the term "biosolids" is used interchangeably with "sewage sludge," which is defined in the regulations and used in the statute.

Section 1 provides an overview of POTWs and biosolids, previous surveys, and the decision to conduct the TNSSS. Section 2 describes the statistical design. Section 3 describes data conventions for the data collected by the TNSSS.

1.1 Publicly Owned Treatment Works (POTWs) and Biosolids

Thirty years ago, thousands of American cities dumped their raw sewage directly into our nation's rivers, lakes, and bays. Today, many municipalities have a wastewater treatment plant that incorporates a series of processes to remove pollutants from water used in homes, small businesses, industries, and other facilities. These plants are called "Publicly Owned Treatment Works" or "POTWs." Because of improved wastewater treatment, our waterways have been cleaned up and made safer for recreation and seafood harvest. And, because of the strict Federal and state standards, the treated residuals from wastewater treatment (biosolids) can be recycled. Local governments make the decision whether to recycle the biosolids as a fertilizer, incinerate it or bury it in a landfill.

Biosolids are the nutrient-rich solid, semisolid or liquid organic materials resulting from the treatment of domestic sewage in a treatment facility. Although local municipalities decide how best to manage their biosolids, the U.S. Environmental Protection Agency (EPA) is obligated and continues to provide the public with educational information, based on the best science, about the recycling and disposal of biosolids. Furthermore, Section 405(d) of the Clean Water Act (CWA) requires EPA to identify and regulate toxic pollutants which may be present in sewage sludge in concentrations which may affect public health and the environment. The statute requires EPA to promulgate sewage sludge regulations every two years "for the purpose of identifying additional toxic pollutants and promulgating regulations for such pollutants…" (CWA, Section 405(d)(2)(D)).

1.2 1988 National Sewage Sludge Survey (NSSS) and Round 1 Regulations

In 1989, to obtain additional information about sewage sludge and biosolids, EPA conducted the first National Sewage Sludge Survey (NSSS). In the first phase of the two-phase statistical design, EPA conducted a detailed mail survey to collect information from 462 secondary treatment POTWs about sewage sludge use, disposal, and operational practices. In the second phase, EPA collected samples of final process sewage sludge just prior to use or disposal from 180 POTWs. This document refers to the two-phase survey as the "1988 NSSS."

In 1993, based in part on the information collected by the 1988 NSSS, EPA promulgated regulations (40 CFR Part 503) to establish standards for the use and disposal of sewage sludge. The regulations establish requirements for the final use and disposal of sewage sludge when: 1) applied to the land, including products sold or given away for use in home gardens; 2) disposed on land by placing in surface disposal units; and 3) incinerated. The standards apply to publicly and privately owned treatment works that generate or treat domestic sewage sludge, as well as to any person who uses or disposes of sewage sludge from such treatment works. In the *Federal Register* (*FR*) notice (58 *FR* 9428) promulgating these regulations, the preamble to the regulations acknowledges that the rule was "Round 1" and might not regulate all pollutants that may be present in concentrations that could adversely affect public health and the environment.

1.3 2001 NSSS and Round 2 Activities

In Round 2, pollutants considered but not regulated under Round 1 were again considered. Based on the results of those analyses, three groups of pollutants were placed on the pollutant list for Round 2: polychlorinated dibenzo-p-dioxins (PCDDs, or dioxins), polychlorinated dibenzofurans (PCDFs, or furans), and coplanar polychlorinated biphenyls (PCBs). To obtain new information about the levels of dioxins in biosolids, EPA collected additional biosolids samples from 94 POTWs that participated in the 1988 NSSS analytical survey. This document refers to this study as the "2001 NSSS."

In October 2003, EPA completed its "Round 2" and published the results of a review of land-applied sewage sludge containing dioxin and dioxin-like compounds (68 *FR* 61083) from the 2001 NSSS. EPA gave final notice of its determination that neither numerical limitations nor requirements for management practices are needed to protect human health and the environment from reasonably anticipated adverse effects from dioxin and dioxin-like compounds in land-applied sewage sludge.

1.4 Factors Leading to a Decision for TNSSS

During Round 2 activities, EPA commissioned the National Research Council (NRC) of the National Academy of Sciences to independently review the technical basis of the regulations applicable to sewage sludge that is applied to land. In July 2002, the NRC published a report titled, "Biosolids Applied to Land: Advancing Standards and Practices," that contained assessments and recommendations by NRC's Committee on Toxicants and Pathogens in Biosolids Applied to Land. The NRC report recommended, among other things, that EPA conduct a new national survey of chemicals and pathogens in sewage sludge.

In December 2003, EPA completed a subsequent biennial review and published the results of a review of its sewage sludge regulations (68 *FR* 75531). In this notice, EPA identified additional toxic pollutants in sewage sludge for potential future regulations. Based on a screening assessment of chemical pollutants for which EPA had adequate data (e.g., human health benchmark values, and information on fate and transport in the environment), as well as concentration data in sewage sludge for those pollutants, EPA identified 15 pollutants for possible regulation. EPA then conducted an exposure and hazard assessment of the pollutants using the available information (U.S. EPA 2004). EPA also determined that it needed more recent sewage sludge data for some pollutants, and thus, decided to plan and conduct another study (i.e., the TNSSS). As described in the next section, EPA selected a subset of the total POTW population for the TNSSS using a statistical design. The new concentration data for the selected pollutants will serve as a basis for determining whether to propose amendments to the national 40 CFR Part 503 regulations.

2.0 STATISTICAL DESIGN

This section details the data sources used to develop a sample frame and the statistical basis for the development of the sample design. Section 2.1 describes the data sources used to evaluate the types of POTWs that should be included in the study. Section 2.2 provides the definition of the target population, resulting from the evaluation of the data sources. Section 2.3 describes the sample frame and its coverage relative to the target population. Section 2.4 describes the stratification of POTWs for the sample design. Section 2.5 describes the statistical basis for selecting the 80 POTWs for the study. (Appendix A lists the 80 POTWs.) Section 2.6 estimates the expected bias and precision associated with using the survey data to estimate percentiles of the concentration distribution, for eight pollutants.

2.1 Data Sources

In designing the study, EPA used two data sources to evaluate POTW characteristics. Together, these two databases provide the most complete source of available information about POTWs in the United States.

The primary database originated from the 2004 Clean Watersheds Needs Survey (CWNS), a joint EPA-State survey that collected information regarding water quality programs and projects that may be eligible for funding under the Clean Water State Revolving Fund (CWSRF). The CWNS is conducted once every four years. To supplement information collected in the 2004 CWNS, EPA also used information from the database for the 2000 CWNS. For example, the 2000 CWNS data file occasionally included latitude and longitude information that was missing from the 2004 database for some POTWs.

EPA combined the CWNS database with a database obtained from the Permits Compliance System (PCS), a national computerized information management system for the National Pollutant Discharge Elimination System (NPDES) that contains permit limits, monitoring data, contact information, and other data pertaining to some facilities regulated under NPDES in 2002. EPA used the 2002 version of the PCS database to identify additional POTWs that might not be represented within the CWNS database. For example, federal facilities may be included in the PCS database, but generally are not included in the CWNS database because they are not eligible to participate in the CWSRF. The PCS database also provided additional contact information for some facilities that were included in the CWNS database.

In combining the two databases, EPA applied computer algorithms to eliminate duplicate entries. The resulting "combined" database contains data for a total of 23,090 POTWs. Table 1 displays how these POTWs are distributed based on flow rate, while Table 2 presents the distribution of POTWs according to treatment type.

Table 1. Combined Database: Distribution of POTWs by Flow Group

Flow Group	Number of Facilities	Cumulative Flow (MGD) for Group	Percent of Total National Flow
>100 MGD	52	14,354	34.8%
>10 MGD and <100 MGD	562	15,298	37.1%
>1 MGD and <10 MGD	2,834	8,954	21.7%
> 0 MGD and <1 MGD	13,402	2,654	6.4%
No Discharge (MGD=0)[1]	23	0	0.0%
TOTAL	**16,873**	**41,260**	100%
Unknown Flow	6,217		
GRAND TOTAL	**23,090**		

[1] The 2000 CWNS Report defines “no discharge” facilities as those that “do not discharge treated wastewater to the Nation’s waterways. These facilities dispose of wastewater via methods such as industrial reuse, irrigation, or evaporation.”

Table 2. Combined Database: Distribution of POTWs by Treatment Type

Treatment Type	Number of Facilities	Cumulative Design Capacity (MGD) for Treatment Type	Percent of Total National Flow
Advanced Treatment II	476	2,966	7.2%
Advanced Treatment I	4,829	17,492	42.4%
Secondary Treatment	11,283	19,328	46.8%
Less than Secondary Treatment	34	195	0.5%
Partial Treatment	60	293	0.7%
Raw Discharge	1	0	0.0%
Unknown Treatment Type	6,407	986	2.4%
GRAND TOTAL	**23,090**	**41,260**	

Note: There are 2 secondary, 3 Advanced I and 18 Unknown Treatment type facilities with flow MGD=0.

2.2 Target Population

A principal task in the development of a sample survey design is establishing a clear, concise description of the target population. Statistical inferences are derived from data collected from a sample and extrapolated to the target population.

After reviewing the information in the combined database, EPA selected a subset of POTWs as the target population. *The target population for the TNSSS is all POTWs that existed in 2002 and/or 2004, and had flow rates greater than 1 MGD, a minimum of secondary treatment except when the final stage is a pond, production of final treated biosolids, and a location in the contiguous United States.*

Because the target population for the 1988 NSSS and 2001 NSSS was different than the TNSSS target population, the results should not be directly compared.[1] The following sections examine each component of TNSSS definition.

2.2.1 Publicly Owned Treatment Works (POTWs)

For purposes of the survey, EPA intends to collect information only from POTWs, defined as local, state, or federal facilities that primarily treat sanitary wastes. As a consequence of the definition, the target population excludes privately-owned, non-publicly owned, and Tribal facilities.

From the information in the combined database, EPA estimates that 2,206 privately-owned and non-publicly-owned facilities provide wastewater treatment. The private sector has served an important role in the effort to control water pollution across the country. When a facility is privatized, the interests of both the local government and the private entity may define normal operating conditions when adjusting to different conditions, such as floods, atypical pollutant levels, or amendments to environmental regulations that increase operating costs. The level of oversight for the private entity may vary to reflect the level of concern that local governments have about the private entity's performance.

The CWNS database has only limited information about Tribal facilities. A more complete data source is likely to be available from the Indian Health Service; however, it was not readily available to EPA during the survey planning process. Of the Tribal facilities reported by some states in the CWNS database, EPA was only able to identify one Tribal facility with a flow greater than 1 MGD. As a consequence, EPA suspects that Tribal facilities generally have very small flows, and thus, would not otherwise be part of the target population (see next section on flows).

EPA also excluded twelve facilities with unknown ownership that had flow rates greater than 1 MGD.

---

[1] As explained in Section 2.5.2, the precision criteria also indicate that the TNSSS results should not be compared to the earlier studies.

2.2.2 Existence in 2002 and/or 2004

The target population consists of POTWs that were in existence in 2002 and/or 2004, because the combined database provides information about POTWs that existed in those two years. In contrast, the 1988 NSSS and 2001 NSSS were based upon information from the 1986 CWNS, and thus, their target population is POTWs that existed in 1986. Thus, the statistical inferences from the TNSSS will provide information about POTWs that existed in 2002 to 2004, while the 1988 NSSS and 2001 NSSS provide information about POTWs that existed in 1986.

2.2.3 Flow Rates Greater than 1 MGD

The target population includes only POTWs with flow rates greater than 1 MGD. Therefore, the target population excludes an estimated 13,402 POTWs with flow rates below 1 MGD, 23 with "no discharge,"[2] and 6,217 with unknown flow rates, based upon the numbers in the combined database. Table 3 indicates that POTWs with flow rates below 1 MGD collectively contribute only about six percent of the total flow among all POTWs in the nation, suggesting that their potential impact to the environment is minor. Thus, EPA determined that it was more appropriate to focus the data collection on larger facilities having higher flow rates. In contrast, the 1988 NSSS and 2001 NSSS retained all facilities with any flow. Thus, statistical inferences for the TNSSS can only be extrapolated to POTWs with flows greater than 1 MGD, while the 1988 NSSS and 2001 NSSS results were extrapolated to POTWs that discharged any flow.

---

[2] The 2000 CWNS Report defines "no discharge" facilities as those that "do not discharge treated wastewater to the Nation's waterways. These facilities dispose of wastewater via methods such as industrial reuse, irrigation, or evaporation."

Table 3. Combined Database: Percentage of POTWs Contributing to Total Flow Rate

Flow Rate Threshold (MGD)	Number of POTWs Exceeding the Threshold	Total Flow Rate (MGD) of These POTWs	% of Total Flow Rate Contributed by These POTWs
0.25	7,101	40,494	98.1%
0.5	5,117	39,800	96.5%
0.75	4,138	39,209	95.0%
1.0	3,552	38,710	93.8%
2.0	2,312	36,977	89.6%
3.0	1,717	35,548	86.2%
4.0	1,372	34,378	83.3%
5.0	1,145	33,375	80.9%
10.0	630	29,812	72.3%
20.0	326	25,662	62.2%
30.0	224	23,271	56.4%
40.0	154	20,874	50.6%
60.0	106	18,539	44.9%
80.0	74	16,325	39.6%
100.0	52	14,354	34.8%

2.2.4 Secondary Treatment

A key provision of the 1972 Clean Water Act established a national policy requiring secondary treatment[3] of municipal wastewater as the minimum acceptable technology supplemented by more stringent water quality-based effluent controls on a site-specific, as-needed basis (U.S. EPA, 2000). As a consequence, almost all POTWs now have secondary treatment. Because POTWs with less than secondary treatment are rare, EPA excluded them from the target population. The 1988 NSSS and 2001 NSSS target population also excluded POTWs with less than secondary treatment (defined as POTWs with “No Discharge”, “Raw Discharge”, and “Advance Primary”).

Of the POTWs with flow rates greater than 1 MGD, the combined database contains six POTWs as having less than secondary treatment. The combined database also included 100 POTWs without a specified treatment type but with flow rates greater than 1 MGD. Because almost all facilities have secondary treatment (or more), EPA assumed that these facilities had secondary treatment.

---

[3] All wastewater first must go through the primary treatment process, which involves screening and settling out large particles. The wastewater then moves on to the secondary treatment process, during which organic matter is removed by allowing bacteria to break down the pollutants (U.S. EPA, 2003b).

2.2.5 Ponds

According to the EPA Wastewater Technology Fact Sheet for Facultative Lagoons (U.S. EPA, 2002), facultative waste stabilization ponds, sometimes referred to as lagoons or ponds, are frequently used to treat municipal wastewater. Although the facultative lagoon concept is land intensive, especially in northern climates, it offers a reliable and easy-to-operate process that is attractive to small, rural communities. The bottom layer of the lagoon includes sludge deposits and supports anaerobic organisms. Typical detention times range from 20 to 180 days depending on the location, with detention times approaching 200 days in northern climates where discharge restrictions prevail. Thus, for POTWs that use ponds as the final treatment step for biosolids, it will be difficult to obtain samples of fully treated biosolids in a timely manner. From the combined database, EPA was able to identify 29 POTWs that use ponds as the final stage of biosolids treatment. Because the combined database generally did not provide detailed information about treatment stages, EPA suspects that the combined database contains more than 29 POTWs with ponds. Thus, as explained in Section 5.6.3, EPA increased the sample size to account for any other POTWs with ponds that might be inadvertently captured into the sample. The 1988 NSSS and 2001 NSSS also encountered the same difficulty, and excluded POTWs with ponds after the original sample was selected. Thus, POTWs with ponds were part of the target population for the 1988 NSSS and 2001 NSSS, but were excluded for logistic concerns.

2.2.6 Production of Final Treated Biosolids

While EPA identified 954 POTWs in the combined database as having partial treatment, only 17 had flow rates greater than 1 MGD. The 2000 Report to Congress defines these POTWs as "provid[ing] some treatment to wastewater and discharg[ing] their effluents to wastewater facilities for further treatment and discharge." Because these facilities are unlikely to produce final treated biosolids which will be measured by this study, EPA did not consider these POTWs to be part of the target population. It is likely that the 1988 NSSS and 2001 NSSS excluded POTWs with partial treatment from the target population.

2.2.7 Contiguous United States

For convenience and budgetary constraints, the target population was restricted to facilities located within the contiguous United States. Of the facilities represented in the combined database, 227 were located outside the contiguous U.S. They are located in Hawaii (24), Alaska (89), American Samoa (2), Guam (9), Puerto Rico (74), Virgin Islands 25), NI (1), N. Mariana Islands (2), and Mexico (1).[4] Of these POTWs, 41 would otherwise have been part of the target population due to having secondary treatment (or better) and flow rates greater than 1 MGD. In the first phase (mail survey) of the 1988 NSSS, the target population included the 50 States, Puerto Rico, and the District of Columbia. For the second phase (physical sampling) of the 1988 NSSS and 2001 NSSS, EPA also restricted the sample to the contiguous United States.

---

[4] One POTW in Mexicali, Mexico is partly funded by the U.S. government, and thus, included in the CWNS database.

2.3 Sample Frame

This section defines the sample frame for the TNSSS and evaluates its coverage of the survey's target population.

2.3.1 Definition

Survey sampling is intended to characterize the entire target population, and hence, all members of the target population must be uniquely identified and have a known chance of being included in the sample. Kish (1965), a classic reference on survey statistics, defines perfect sample frames as listing every element of the target population only once and excluding all non-population elements. The sample frame contains those POTWs within the combined database that meet the definition of the target population. Of the 23,090 POTWs identified in the combined database, 3,337 met the definition for the target population, and thus, the sample frame contains 3,337 POTWs.

2.3.2 Coverage

Differences between the target population and the sample frame exist in virtually all sampling situations, and the extent of differences determines the *coverage*. For example, a sample frame may not include the most recent entries or may contain outdated entries. After evaluating the coverage in the combined database and the 1988 NSSS, EPA concluded that the sample frame was reasonably complete and free from members outside of the target population.

Because of the financial incentives to respond to the CWNS (i.e., to receive funding), EPA considers the CWNS database to be the single best source of information about POTWs. To capture additional state and federal facilities, EPA used the PCS database which includes the discharge information for major dischargers. While the databases identify facilities existing prior to the study year (i.e., 2006), construction of new POTWs is unlikely because of the high cost of building. Also, because POTWs are built to last many years, it is unlikely that the population has decreased significantly in the last few years. Thus, EPA concluded that the combined database of CWNS and PCS data provided good coverage of the target population.

EPA's conclusion is supported by its review of the 1988 NSSS which used the 1986 CWNS database as its sample frame. Of the 208 POTWs in the 1988 NSSS, only five (2%) were classified as "ineligible" or "out of business" a few years after the 1986 CWNS was conducted. Even after further review, EPA was only able to identify one potential issue with coverage. As mentioned previously, because the information is seldom included in the combined database, it is possible that the sample frame includes some POTWs that utilize ponds as the final stage of treatment, although they should not be part of the target population. As a consequence, EPA inflated the sample size as described in Section 2.5.4.

2.4 Stratification

Existing information about the facilities can be used to improve the survey design and the precision estimates. One common technique is to group similar facilities together into mutually exclusive strata. Then, by selecting facilities from each stratum to participate in the survey, it ensures that the sample will include facilities that have the various characteristics that are represented by the different strata. EPA considered stratifying by a series of candidate variables, including the magnitude of the effluent flows, the treatment types, disposal types, and geographic regions. These stratification options were possible because the sample frame contains information about these variables. As explained below, the final sample design stratifies the facilities according to flow rate only.

2.4.1 Flow Rate

Table 4 summarizes the distribution of POTWs across three flow rate groups based on information present in the sample frame. This table shows that the distribution of cumulative flows across the facilities is highly skewed. Of the 3,337 POTWs in the sample frame, 82 percent (i.e., 2,743 POTWs) have flow rates less than 10 MGD and collectively account for only 23.1 percent of the total flow. In contrast, only 51 facilities have flow rates exceeding 100 MGD, but they generate 37.8 percent of the total flow. Because EPA assumes that flow and amount of biosolids are highly correlated, the larger POTWs are considered more likely to generate the disproportionately largest volume of biosolids. Thus, it is particularly important to stratify the population by flow group to ensure that the largest facilities are adequately represented in the sample.

Table 4. TNSSS Sample Frame: Distribution of POTWs by Flow Rate Groups

Flow Group	Number of POTWs	Cumulative Flow (MGD) for Group	Percent of Total Flow
>100 MGD	51	14,188	37.8%
>10 MGD and <100 MGD	543	14,726	39.2%
>1 MGD and <10 MGD	2,743	8,661	23.1%
TOTAL	**3,337**	**37,575**	

2.4.2 Treatment Type

Upon reviewing the distribution of POTWs within the sample frame among four treatment types (Secondary, Advanced Treatment I, Advanced Treatment II, and Unknown), EPA determined that it was unnecessary to stratify by treatment type. Table 5 provides the total number of facilities for each treatment type included in the sampling frame. Because the distinction between Advanced Treatment I and Advanced Treatment II was not clear from the available information,[5] EPA combined the two categories into an “advanced” group in its evaluation of treatment type. The “secondary” and “advanced” groups are both relatively large and about the same size, and thus, the sample can reasonably be expected to include facilities from both

---

[5] The database did not include all available information such as plant schematics.

groups. In addition, the study objectives do not include a comparison of the different treatment groups. Thus, EPA considered it unnecessary to distinguish between the two groups in the sample design.

Table 5. TNSSS Sample Frame: Distribution of POTWs by Treatment Type

TreatmentType	Number of Facilities	Cumulative Design Capacity (MGD) for Treatment Type	Percent of Total Flow
Advanced Treatment II	241	2,869	7.6%
Advanced Treatment I	1,519	16,483	43.9%
Secondary	1,496	17,460	46.5%
Unknown	81	763	2.0%
TOTAL	**3,337**	**37,575**	

2.4.3 Disposal Type

EPA also considered whether the sample should be stratified by disposal type in the TNSSS. Although disposal type was considered in the 1988 NSSS, this new study focuses on the final treated biosolids prior to disposal. EPA has no reason to suspect that the disposal practices would lead to different treatment mechanisms that would affect the characteristics of the final treated biosolids. Thus, EPA determined that it was unnecessary to stratify by disposal type.

2.4.4 Geographic or Spatial Considerations

EPA also considered whether geographic or spatial aspects should be incorporated into the statistical design. For purposes of this survey, there might be benefits to selecting a sample of facilities that covers the entire country and represents different watersheds. Although any random sample could accomplish this in a sense, there might be advantages to placing some spatial constraints on the sample so that the spatial distribution of the sample closely matches the spatial distribution of the population. However, this approach requires accurate location data to place each facility in its spatial context, prior to stratification or any other statistical technique. While latitude and longitude were required parameters in the CWNS database, respondents did not always provide this information. Although latitude and longitude can be approximated, EPA determined that the effort would not be justified, because the study objectives do not include any geographic or watershed comparisons. Instead, to incorporate some spatial context into sample design, a systematic sampling procedure was used as explained in Section 2.5.5 below.

2.5 Sample Design

In the TNSSS, EPA will collect information from a representative sample selected from the sample frame of 3,337 POTWs. In the following sections, references to “sample” and “sample size” mean the number of POTWs from which EPA will collect physical samples of biosolids.

The term “sample design” refers to how EPA selected the sample of facilities from the sample frame. To estimate characteristics of the total population, the sample must be selected using probabilistic methods. It is not necessary that every facility has the same chance for selection or that the sample size be fixed. It is only necessary that all the facilities on the sample frame have a known, non-zero chance of selection.

For the sample design, EPA used a common method for estimating sample sizes that is based upon the binomial distribution as described in Appendix B. Using this sampling approach and precision criteria, EPA determined that a minimum sample size of 80 POTWs (listed in Appendix A) will provide reasonably good estimates of the percentage of the target population achieving a certain outcome. The following sections describe the binomial distribution, precision criteria for the study, the estimated sample size, nonresponse considerations, sample selection, and subsample selection for field duplicates.

2.5.1 Binomial Distribution

The binomial distribution applies to situations in which a particular measurement of interest has only two outcomes (e.g., yes or no), and it is of interest to estimate the percentage of the target population achieving one of these outcomes, known as the *attribute* of interest. The presence or absence of the attribute for a particular POTW is a *dichotomous* or *binary* variable and can represent a response to a yes-no question such as “Is pyrene present in detectable levels at this facility at this point in time?” The binomial distribution models these data, based on the notion of obtaining national estimates of the percentage or fraction of POTWs in the target population (or a subset of the target population) that have a particular attribute. For example, using the study results, EPA might estimate that 40 percent of the POTWs have detectable levels of pyrene. Appendix B provides more information about the application of the binomial distribution in sample size calculations.

2.5.2 Precision Criteria

The procedure for estimating sample sizes that will achieve specified precision criteria are given in Appendix B. As noted in Appendix B, a conservative approach to the sample size calculation is to assume that its true value is 50 percent (e.g., pyrene was detected in the biosolids samples at 50 percent of the facilities). This assumption corresponds to the largest possible variance for the binomial distribution, and therefore, will yield the maximum sample size. If, in fact, the true (unknown) percentage is something other than 50 percent, then the estimate of this percentage calculated from the collected survey data is expected to be even more precise than the precision criteria specify.

Note that the criteria do not contain precision targets associated with comparisons of the TNSSS results to the previous studies (1988 NSSS and 2001 NSSS). To obtain reasonable estimates of the differences between studies, a larger sample size would have been required. Also, as explained in Section 2.2, the TNSSS would have needed to retain the same target population definition as the earlier studies.

For the TNSSS, EPA has adopted the following precision criteria for estimating a percentage of facilities:

(a) <u>Overall Criteria</u>: Across the entire survey, a sample size is necessary to ensure that the 90% confidence interval on the population percentage is no more than +/- 10% if the true value of the percentage was 50 percent (i.e., a confidence interval of (40%, 60%)). In other words, the sample size must ensure that the unknown percentage for the target population is estimated to within 20% of its true value with 90% confidence.

(b) <u>Within Stratum Criteria</u>: Although the survey objectives are to obtain reasonable national, rather than stratum estimates, EPA wanted to ensure that some facilities within each stratum were represented within the sample. (See Section 2.4.1.) Thus, EPA specified that the sample size for each stratum would be large enough to ensure that the 90% confidence interval on the unknown percentage would be no more than +/- 30% if the true value of the percentage was 50 percent (i.e., a confidence interval of (20%, 80%)). In other words, the sample size must ensure that an unknown stratum-specific percentage is estimated to within 60% of its true value with 90% confidence. EPA recognizes that this level of precision will not be sufficient to produce stratum-level estimates.

Initial sample size calculations presented in Appendix B suggest that fewer than 80 POTWs would be needed to sample across all strata under these precision criteria. However, as explained in the next section, EPA incorporated several upward adjustments to the sample size.

2.5.3 <u>Sample Size</u>

Table 6 lists the strata, along with the total number of facilities in the sampling frame (3,337) and within each stratum (N), and the calculated minimum sample sizes (*n*) within each stratum that are necessary to meet Criteria (a) and (b) above. To meet the Overall Criteria (a) of +/- 10% for a 90% confidence interval, a minimum sample size of 68 is needed across all strata. However, if both Criteria (a) and (b) must be achieved, this overall minimum sample size estimate increases to 74. Appendix B shows the sample sizes that would be required for other precision requirements for the national estimates.

Table 6. TNSSS Sample Frame: Sample Sizes by Stratum

Flow Stratum (MGD)	**Number of POTWs in Stratum**	**Sample Size For Overall Estimate** (Criterion b) (Minimum Sample Size to meet overall Criterion a)	**Sample Size for Stratum Estimates** (Criterion a) (Minimum Sample Size to meet within stratum Criterion b)	**Maximum Sample Size for Stratum** (Minimum sample size to meet both Criteria a and b)	**Maximum Adj. to Total 80 Facilities** Maximum Sample Size Per Stratum to Meet Criteria a and b and to account for potential ineligibility)
>100	51	1	7	7	8
>10 and <100	543	11	8	11	12
>1 and <10	2,743	56	8	56	60
TOTAL	**3,337**	**68**	23	**74**	**80**

Once the sample is drawn, EPA expects that a small number of selected facilities will be determined to be ineligible (e.g., some may utilize ponds as explained in Section 2.2.4). To account for this coverage problem in the sample frame, EPA adjusted the sample size upwards by four percent for each stratum. This adjustment resulted in the last column of Table 6, which indicates a total sample size of 80 across all strata. As a result of this adjustment, the estimates can be expected to have slightly higher precision than specified by Criteria (a) and (b).

2.5.4 Nonresponse Considerations

We also considered adjusting the sample sizes for probable nonresponse. However, based upon the information in the 1988 NSSS Statistical Support Document (1992), nonresponse was minimal in the 1988 NSSS. For the analytical survey, of the 208 facilities that were selected, 185 were eligible. The ineligible facilities were primarily those with ponds (i.e., 18 had ponds, and five were out-of-business or otherwise ineligible). Of the 185 eligible facilities, only six were not sampled for logistical reasons. Because the sample design incorporates an adjustment for ineligibility (due to ponds), EPA determined that another minimal adjustment for nonresponse was unnecessary.

2.5.5 Sample Selection

To incorporate some spatial context into the sample, EPA first sorted the facilities within each stratum by EPA Region (e.g., Region 1, Region 2, etc.), and by state name within each Region. EPA then drew a systematic sample from each stratum. Then, for a particular stratum, if N denotes the total stratum size and n denotes the stratum sample size, systematic sampling involves dividing the stratum into n equal-sized subgroups, generating a random number k between 0 and N/n, and selecting the k^{th} facility within each of the n subgroups. By using this systematic approach, EPA ensured that the sample is reasonably diverse from a geographical perspective.

2.5.6 Subsample for Field Duplicates

For quality assurance purposes, EPA determined that field duplicates should be collected for ten percent of the sample. EPA selected a subsample of 8 POTWs using the same statistical sampling approach used to select the sample of 80 POTWs.

As shown in Table 6, the 80 facilities in the sample are distributed among the three strata (determined by flow rate) as follows:

Stratum #1 (>100 MGD): 8 facilities
Stratum #2 (10-100 MGD): 12 facilities
Stratum #3 (1-10 MGD): 60 facilities

As a result, the 8 facilities selected for duplicate sampling is distributed as follows:

Stratum #1: 1 facility (8/80 * 8 = 0.8)
Stratum #2: 1 facility (12/80 * 8 = 1.2)
Stratum #3: 6 facilities (60/80 * 8 = 6.0)

Thus, a ten percent sample was selected from the set of sampled facilities within each stratum, under the restriction that 8 facilities would be selected across all strata. Because strata #1 and #2 each had only one POTW to be selected for duplicate sampling, this POTW was, in essence, selected at random from the set of 8 and 12 POTWs, respectively, that are in the sample from that stratum. Within Stratum #3, the 6 facilities were selected using the procedure specified within Section 2.4.4 (i.e., the 60 facilities were sorted by EPA region and by state, and a systematic sampling approach selected the sample of 6).

2.6 Evaluation of the Sample Design: Bias and Precision of Percentile Estimates

As noted in Section 2.5, the sample design is based upon the binomial distribution and the precision associated with estimating the percentage of the target population having a certain characteristic. However, EPA's survey objectives also include estimating percentiles of the distribution of concentration data in biosolids and the annual quantities of pollutants discharged by POTWs. Thus, for the selected sample design, EPA quantified the expected bias and precision associated with estimating these percentiles, through simulations performed on data from the 1988 NSSS.

2.6.1 Bias

Bias is an important consideration in designing a sample. Bias is the difference between the expected value of an estimate and the true value of a parameter or quantity being estimated. Bias is a characteristic of the data collection process. If the data collection process generates estimates that are consistently (or on average) above or consistently below the true value, the data collection process is biased. The bias may result from a combination of how the facilities are selected, how samples of biosolids are obtained, how the biosolids are analyzed at the laboratory, and the statistical analysis procedures used to summarize the data.

2.6.2 Precision

Another important consideration is the degree of precision that is desired for the estimated values. The precision of the survey estimates depends on both the sample design and the sample size, that is, the number of facilities that are selected. One measure of precision is the half-width of the confidence interval for the estimate. Confidence intervals provide a range of possible values for a particular estimate that would be likely if the study were repeated infinitely many times. Thus, when using 95 percent confidence intervals, 95 percent of such intervals would include the true value, if we could take an infinite number of samples.

Because more values from the sample will lie near the middle of the distribution, the confidence intervals for means and the 50th percentile will be narrower than those for extreme percentiles (e.g., 10th, 99th) where fewer values lie in the distribution. Because the study objectives include the calculation of extreme percentiles, we estimated confidence intervals as part of the evaluation of each design. The width of the confidence intervals depend on the variability (or standard deviation) of the data.

2.6.3 Simulation

In the simulation described in this section, artificial facilities and data are generated that have characteristics similar to the real sample frame and expected measurements. The simulation mimics the population characteristics and the sampling process. In each step of the simulation, a sample of artificial facilities is selected according to the sample design, and percentile estimates are calculated from data generated for this sample. The simulation considers the concentration data measured in the physical samples and the estimated annual mass of biosolids based upon the concentration value (e.g., mg/kg) and annual dry weight of biosolids (i.e., milligrams of pollutants produced on a dry weight basis). By repeating the simulation process many times, the bias and precision of the percentile estimates can be calculated. Details on the simulation approach are given in Appendix C. The results from 400 simulations were used to estimate the expected precision of the percentile estimates.

The simulation was focused on eight pollutants that were initially identified as target pollutants in the health and ecological exposure and hazard screening assessment (see Section 1.4). First, it was necessary to get information on estimated variability in the concentrations of these pollutants. Historical data were available from the 1988 NSSS. Although the TNSSS target population is somewhat different, the 1988 NSSS provides the best source of information for this analysis. In particular, the simulation used summary statistics from an internal document dated January 30, 2003. This printout provides the percent of the values detected in addition to other information published in the statistical support document for eight pollutants. (The ninth pollutant, 4-chloroaniline, was not evaluated in the 1988 NSSS, and therefore, was not included in the simulation.) Table 7 provides a summary of the relevant information from the printout.

Table 7. 1988 NSSS: Summary Statistics for Eight Pollutants

Target Pollutant	Units	Percent Detected	Summary Statistics (with not-detected results replaced by the minimum reported level)				
			Mean	Standard Deviation	Maximum	50th Percentile	CV[1]
Barium	mg/kg	100	873	840	5,570	499	0.96
Beryllium	mg/kg	22	1.84	2.43	21	0.86	1.32
Fluoranthene	ug/kg	5	9,950	13,400	154,000	4,760	1.35
Manganese	mg/kg	100	538	1,040	13,200	276	1.93
Pyrene	ug/kg	5	9,950	13,400	154,000	4,760	1.35
Silver	mg/kg	84	48	112	852	26	2.32
Nitrate	mg/kg	95	1,420	5,040	35,300	97	3.55
Nitrite	mg/kg	83	201	1,210	17,700	13	6.02

[1] Coefficient of variation is equal to the standard deviation divided by the mean.

Second, in evaluating the sample design, it was assumed that the measurements in the new study will have the same median and coefficient of variation (CV) as the 1988 NSSS (Table 7). For a given sample design and percentile, the width of the confidence interval depends primarily on the CV. We chose to use the CV in the simulation because it does not depend on the units being used. For example if units are changed from parts per million to parts per billion, the numerical value of the mean and standard deviation will change by a factor of 1000, but the CV will not change. Thus, we can apply the results from one pollutant to another, even if the units are different or the order of magnitude varies, as long as the values of the CV are similar.

Third, we assumed that the concentration and mass values in the new study would be lognormally distributed. The formula for CV of any particular pollutant based on an assumed lognormal distribution is:

$$CV = \sqrt{\exp(\sigma^2) - 1}$$

Thus, the standard deviation is estimated by solving for σ in the above formula and substituting CV with the estimate given in Table 7. For the mass-based estimates, we used percentiles of concentration weighted by the mass of biosolids represented by each sample. The concentration values were also assumed to be uncorrelated with the mass of biosolids. We simulated the confidence intervals for selected percentiles from the 10th to the 99th to evaluate the precision for the pollutants selected for the study.

Fourth, we simulated 90% confidence intervals for the percentiles. Confidence intervals for percentiles from a lognormal distribution can be expressed as upper confidence interval =

Estimate*R, and lower confidence interval = Estimate/R, where R is a function of the population variability, the sample design and the sample size. As a result, the confidence intervals are not symmetric around the estimate. Instead, the width of the confidence interval can be summarized by the ratio R. For example, the simulated estimate of the 75th percentile of barium concentration measurements is 656 mg/kg, with a 90% confidence interval of (548 mg/kg, 784 mg/kg). The lower confidence bound is below the estimate by 107 mg/kg, while the upper confidence bound is above the estimate by 128 mg/kg. This asymmetry makes the usual summary of the width of a confidence interval (i.e., estimate +/- confidence interval half width) misleading. In this example, R is 1.2. Smaller values of R correspond to more precise estimates and smaller confidence intervals. R will always be greater than 1.0.

2.6.4 Results

The simulations described in the previous section resulted in estimates of the bias and precision associated with estimates of various percentiles of concentrations in biosolids. For selected percentiles ranging from the 10th to the 99th percentile, along with the maximum (denoted by the 100th percentile), Table 8 presents the ratio of the upper 90% confidence interval bound to the estimated percentile.

The upper portion of Table 8 provides precision information for the concentrations, such as the national estimate of the 75th percentile of the concentration values. The lower portion of the table provides the same information for national estimates of mass, such as the 75th percentile of the number of milligrams of pollutants produced on a dry weight basis.

Estimates of percentiles can be biased. If the bias is small compared to the standard deviation, the bias can generally be ignored. However, if the bias is large compared to the standard deviation, the true coverage of the confidence interval will be less than the nominal coverage. Within the shaded cells in Table 8, the bias is greater than one-half of the standard deviation.

As seen from Table 8, the sample design is expected to produce estimates for the concentration-based and mass-based percentiles with relatively good precision. Because most data will be centered around the median (i.e., 50th percentile), the greatest precision occurs when estimating this percentile, and the level of precision decreases (i.e., the confidence interval widths increase) as the percentile of interest moves farther away from the median. When concentration-based data are considered, Table 8 shows that among the seven pollutants, the upper bound of the 90% confidence interval on the median exceeds the estimated median by 21 to 56 percent. For percentiles below the 95th percentile, the upper bound of the 90% confidence interval does not exceed twice the value of the percentile estimate for each of these pollutants. Only when the 99th percentile is reached does the upper bound of the 90% confidence interval exceed twice the value of the estimate for a majority of the pollutants.

The confidence interval widths in Table 8 for mass-based percentile estimates were as much as 85 percent higher than those for concentration-based percentile estimates. For mass-based data, the upper bound of the 90% confidence interval on the median exceeds the estimated median by 44 to 137 percent. For nitrite, which had the highest CV estimate among all of the pollutants in Table 8, the upper bound of the 90% confidence interval was close to four times higher than the estimate for the 95th percentile when mass-based data were considered, while for barium, which had the smallest CV, it was no more than 80 percent higher. While the confidence interval widths are actually lower for mass-based estimates compared to concentration-based estimates for the 99th percentile and the expected maximum, further investigation into potential bias suggests that these estimates are lower than they should be.

When attempting to estimate the maximum expected measurement under both concentration-based and mass-based scenarios, the bias (in absolute value) exceeds one-half of the standard error of the estimate, as noted by the shaded rows in Table 8 when the percentile is specified as 100. This is also true when attempting to estimate the mass-based 99th percentile, as this row is also shaded in Table 8. Under any design approach, bias in estimation is expected to be higher for those parameters that are more associated with the distributional tails, such as large percentiles and the maximum. Typically, only a small amount of information is available to characterize the extreme tails of the distribution, which impacts the ability of the collected data to accurately estimate such parameters. In each case in which it is noted within Table 8 that bias exceeds one-half of the standard error, the bias is negative, indicating that the parameter (i.e., the 99th percentile or the expected maximum) is being under-estimated.

Thus, based upon the precision and bias evaluations, the sample design can be expected to provide more precise estimates of percentiles for concentration rather than mass. However, the sample design appears to provide reasonable estimates under both scenarios.

Table 8. Simulation: Estimated Width of 90% Confidence Intervals on Selected Percentiles for Eight Pollutants (Shading indicates significant bias)

Percent-iles	Barium (CV= 0.96, σ = 0.81)	Beryllium (CV=1.32, σ = 1.0)	Fluoran-thene and Pyrene (CV=1.35, σ = 1.02)	Manganese (CV=1.93, σ = 1.25)	Nitrate (CV= 3.55, σ = 1.62)	Nitrite (CV = 6.0, σ = 1.90)	Silver (CV= 2.32, σ = 1.36)
Concentration							
10	1.28[1]	1.36	1.37	1.47	1.65	1.80	1.52
25	1.22	1.28	1.29	1.36	1.49	1.60	1.40
50	1.21	1.27	1.27	1.34	1.46	1.56	1.38
75	1.23	1.29	1.30	1.37	1.51	1.62	1.41
90	1.30	1.39	1.39	1.50	1.69	1.86	1.56
95	1.40	1.51	1.52	1.67	1.95	2.19	1.75
99	1.81	2.10	2.12	2.50	3.29	4.05	2.73
100	2.00[2]	2.36	2.39	2.91	3.99	5.09	3.21
Mass							
10	1.67	1.89	1.91	2.20	2.78	3.33	2.37
25	1.51	1.67	1.68	1.89	2.28	2.64	2.00
50	1.44	1.58	1.59	1.76	2.09	2.37	1.86
75	1.51	1.67	1.68	1.89	2.28	2.63	2.00

Percentiles	Barium (CV= 0.96, σ = 0.81)	Beryllium (CV=1.32, σ = 1.0)	Fluoranthene and Pyrene (CV=1.35, σ = 1.02)	Manganese (CV=1.93, σ = 1.25)	Nitrate (CV= 3.55, σ = 1.62)	Nitrite (CV = 6.0, σ = 1.90)	Silver (CV= 2.32, σ = 1.36)
90	1.69	1.93	1.94	2.25	2.87	3.45	2.43
95	1.79	2.06	2.08	2.45	3.19	3.92	2.66
99	1.75	2.01	2.03	2.38	3.08	3.75	2.58
100	2.00	2.36	2.39	2.91	3.99	5.09	3.21

[1] Values are always greater than 1.0. Smaller values correspond to more precise estimates and smaller confidence intervals.
[2] Within each shaded cell, the bias is greater than one-half of the standard deviation.

3.0 DATA CONVENTIONS

The laboratories will express results of the analyses either numerically or as not quantitated (sometimes referred to as ‘not detected’) for a pollutant in a sample. If the result is expressed numerically, then the pollutant was quantitated in the sample. For the non-quantitated results, for each sample, the laboratories reported a “sample-specific quantitation limit.” When results are below the minimum level of quantitation but are detected, laboratories will be required to report the actual calculated result, regardless of its value. Section 3.1 describes the adjustments to concentrations for water content in the physical samples. Section 3.2 describes EPA’s approach to mathematically combining the data when more than one sample is collected from a POTW.

3.1 Adjustments for Water Content in the Physical Samples

The different wastewater treatment processes used in POTWs across the U.S. can result in biosolids that vary widely in physical characteristics. One of the most obvious and important variations is the amount of water that is removed from the biosolids prior to disposal. Some POTWs produce residuals that are pourable liquids, containing 1% solids or less, while other POTWs produce residuals that are 90% or greater solids by volume. For example, data from 94 facilities participating in the 2001 NSSS included samples with solids contents ranging from 0.2% to 96%.

Because of the differences in water content of the biosolids, it is difficult to compare the amounts of a contaminant in samples from different facilities using concentration values expressed in either mass per unit volume (e.g., μg/L) or mass per unit mass (e.g., μg/kg). In addition, the varying water content of the samples can affect the overall analytical sensitivity (e.g., detection limits) of the analytical methods used to determine the contaminants.

One approach for addressing the differences in water content is to express all of the results in mass per unit mass, but on a dry-weight basis. In other words, the unit mass of sewage sludge is a unit of dry solids, with all of the moisture removed. This is a simple calculation in which the concentration determined from a “wet” or “as received” sample is divided by the percentage of the sample represented by the solid material. For example, if the concentration of a contaminant measured using a mass of wet sample is determined to be 25 μg/kg, and the sample is 60%

solids, then the “dry weight concentration” is 25/0.6 = 41.67 µg/kg. Even results for samples that are pourable liquids with <1% solids can be expressed in terms of dry weight concentrations.

The simple mathematical conversion to dry weight concentrations does not address the limitations placed on analytical sensitivity when samples of widely different water content are analyzed. Many analytical methods applicable to biosolids instruct the laboratory to take a specific known weight of material for analysis (e.g., 30 grams). Samples that are pourable liquids are generally treated as if they are wholly aqueous samples, and a known volume (e.g., one liter) is used for the analysis. These differences mean that any measure of method sensitivity will depend on the initial mass or volume chosen for analysis and its water content. In national studies such as the 1988 NSSS, EPA has been able to minimize the sensitivity differences by instructing the laboratories performing the analyses to determine the percent solids of the samples first, then using that information to select a portion of the sample for the analysis that contains some set dry weight of solids.

3.2 Reporting Values for Multiple Biosolids Samples per POTW

In general, EPA intends to collect only one biosolids sample per POTW. In two cases, EPA will collect more than one sample at a POTW. First, EPA will collect field duplicate samples at eight POTWs. Second, EPA will collect one sample of the biosolids (i.e., treated sludge) from each system at POTWs with multiple treatment systems.

Regardless of the number of biosolids samples collected at a POTW, the laboratories will provide individual results for each biosolids sample. EPA then will mathematically combine the results of the chemical analyses to obtain one value for each pollutant at that facility. EPA is considering an approach which will aggregate the results on a case-by-case basis. For example, consider a facility that treats approximately the same amount of untreated sludge through each system, with the relative amounts of (treated) biosolids differing due to the way the different treatment systems process the sludge. Because the untreated amounts are the same, EPA might arithmetically average the biosolids results (i.e., weight the results equally in obtaining one single result for the facility). In another example, consider a facility with two treatment systems, with one system handling twice the volume of untreated sewage sludge. In this case, in calculating a value for that facility, EPA might weight the result from the larger system twice as much as the results from the other system.

4.0 REFERENCES

Cochran, William G. 1963. *Sampling Techniques*, John Wiley and Sons, Inc., New York.

Kalton, Graham. 1983. *Introduction to Survey Sampling*, SAGE Publications, Newbury Park, California.

Kish, L. 1965. *Survey Sampling*. New York: Wiley.

U.S. Environmental Protection Agency. 1988. "Supporting Statement for The National Sewage Sludge Questionnaire Survey." Submitted to Office of Management and Budget as part of the Information Collection Request.

U.S. Environmental Protection Agency. 1992. *Statistical Support Documentation for 40 CFR, Part 503: Final Standards for the Use or Disposal of Sewage Sludge*. Volume I, Volume II.

U.S. Environmental Protection Agency. 1998. *South Florida Ecosystem Assessment: Vol I. Final Technical Report: Phase I.* EPA-904-R-98-002.

U.S. Environmental Protection Agency. 2000. *Progress in Water Quality: An Evaluation of the National Investment in Municipal Wastewater Treatment*. EPA-832-R-00-008. Located at http://www.epa.gov/OW-OWM.html/wquality/benefits.htm.

U.S. Environmental Protection Agency. 2002a. *Statistical Support Document for the Development of Round 2 Biosolids Use or Disposal Regulations*. EPA -822-R-02-034. Located at http://www.epa.gov/waterscience/biosolids/adssummary.pdf.

U.S. Environmental Protection Agency. 2002b. *Wastewater Technology Fact Sheet: Facultative Lagoons*. Located at http://www.epa.gov/OW-OWM.html/mtb/faclagon.pdf.

U.S. Environmental Protection Agency. 2003a. *Clean Watersheds Needs Survey 2000: Report to Congress*. EPA-832-R-03-001. Located at http://www.epa.gov/owm/mtb/cwns/2000rtc/toc.htm.

U.S. Environmental Protection Agency. 2003b. *Where Does All the Dirty Water Go?* EPA-832-F-03-008. Located at http://www.epa.gov/npdes/pubs/centralized_brochure.pdf.

U.S. Environmental Protection Agency. 2004. *Technical Background Document Refined Biosolids Exposure and Hazard Assessment Interim Final*. Contract No. 68-C-04-006. Office of Water, Washington, DC. October 8, 2004.

U.S. Environmental Protection Agency. 2005. *Draft for External Peer Review: Study Plan for the2005 National Sewage Sludge Survey*. Office of Water, Washington, DC.

APPENDIX A

LIST OF 80 POTWs SELECTED FOR THE TNSSS

Sorted by State and City

Field Duplicates will be collected at shaded POTWs

In this public version of the document, the listing has been removed to protect confidentiality of the survey participants.

APPENDIX B

DETERMINING SAMPLE SIZE NECESSARY TO OBTAIN AN ESTIMATE OF THE PROPORTION WHERE THE BOUND ON ERROR OF ESTIMATION IS NO GREATER THAN ±0.05 WITH 90% CONFIDENCE

To determine sample size, EPA used formulas associated with a normal approximation to the binomial distribution, assuming an infinite population size. EPA also assumed that the sample design was a simple random sample, rather than a stratified sample. EPA made this assumption because its objective was to calculate national (overall) estimates, rather than stratum-specific estimates. EPA allocated the sample to the different strata to ensure that each stratum had some POTWs in the sample. As described in the peer review report (2006), EPA selected this approach based upon a comparison of several different approaches.

The approach used to select the sample first considers the confidence interval on a proportion for a simple random sample taken from an infinite population. In particular, if p denotes the unknown proportion, $\hat{p}$ is an estimate of the proportion made from the collected data, and n is the sample size, then a 90% confidence interval on the proportion is given by

$$\hat{p} \ \pm \ 1.645 * \sqrt{\hat{p}(1-\hat{p})/n}$$

The value $B = 1.645 * \sqrt{p(1-p)/n}$ is labeled as the “bound on error of estimation.” Solving for n, the sample size that achieves a bound on error of estimation equal to B_0 is equal to:

$$n = (1.645 / B_0)^2 * p(1-p)$$

Note that n depends on the unknown proportion p and is maximized when $p = 0.5$. Thus, sample sizes represent the “worst case” scenario when they are calculated assuming the unknown proportion equals 0.5. Such sample sizes are conservative, meaning they will be larger than necessary when the value of the proportion does not equal 0.5.

Table B-1 provides the sample sizes that would be required for different precision requirements for the national estimates. This table shows, for example, that a sample size of 11 is necessary to ensure that a proportion is estimated to within ±0.25 with 90% confidence (i.e., a 90% confidence interval of (0.25, 0.75) when the unknown proportion equals 0.5). EPA selected a precision requirement of ± 0.10, and thus selected an overall sample size of 68. EPA then applied a precision requirement of ± 0.30 to each stratum, and adjusted the sample sizes accordingly.

Table B-1. Sample Sizes Necessary to Achieve Precision Criteria with 90% Confidence, for a Range of Criteria As Noted by the Bound on Error of Estimation

Bound on Error of Estimation of the Unknown Proportion	Sample Size
± 0.05	271
± 0.06	188
± 0.07	139
± 0.08	106
± 0.09	84
± 0.10	68
± 0.15	31
± 0.20	17
± 0.25	11
± 0.30	**8**

The assumption of an infinite population that was made in developing this design is generally adequate even if the population size is finite but large. While the large size assumption appeared to be adequate for the target population (3,337) and for two of the three strata (543 and 2,743), the size of the stratum for facilities with flow rates exceeding 100 MGD was smaller (51). Thus, an investigation was done to determine how the sample sizes may change if finite population and stratum sizes were assumed. This involved adjusting the above sample size formulas to account for finite population and stratum sizes. The outcome resulted in no change to the stratum-specific sample sizes and a decline of one sampling unit in the minimum sample size across the target population. Thus, because there was no adverse effect to the design when accounting for finite population and stratum sizes, the sample sizes associated with the infinite population assumption were retained.

EPA recognizes that the level of precision will not be sufficient for comparisons between the TNSSS and the earlier studies (i.e., 1988 NSSS and 2001 NSSS). In addition, because the target populations differ, the results from the TNSSS and the earlier studies should not be compared.

APPENDIX C
METHODS TO USING COMPUTER SIMULATIONS TO EVALUATE PRECISION OF PERCENTILE ESTIMATES

When a sample is selected from a specified target population and data are collected from the sample in order to estimate some distributional parameter for the population, a confidence interval is often calculated with the estimate in order to quantify the level of precision associated with the estimate. If the sample selection process was to be repeated many times and a 90% confidence interval was calculated for each such sample, it is assumed that 90 percent of these confidence intervals would contain the true value of the distributional parameter. Thus, confidence intervals provide a measure of consistency (or variability) associated with the estimate generated from the collected sample data.

For the TNSSS, 90% confidence intervals will be calculated on estimated percentiles of the distribution of concentrations in biosolids for specific target pollutants. The process of calculating these confidence intervals can involve either:

1. Using standard formulas that rely on certain assumptions, or
2. Simulating concentration data across POTWs that are consistent with certain assumptions, estimating distributional percentiles from the simulated data, and repeating the simulation many times in order to estimate the precision of the percentile estimates.

If the assumptions are consistent between these two approaches, both should yield essentially the same results. However, the simulation approach has the advantage of being able to estimate bias associated with the percentile estimates (i.e., the difference between the estimate and the actual value of the percentile within the target population of POTWs). This appendix details the approach taken to performing the simulation approach, whose results are presented in Section 5.7 of this study plan.

The process of estimating percentiles of the distribution of target pollutants across POTWs within the TNSSS involves the following steps:

1. A stratified sample of 80 POTWs is selected from the sampling frame of 3,337 POTWs. (The process of selecting this sample is discussed in Section 5.6.4, and the list of POTWs selected for this sample is given in Appendix A.) As noted in Section 5.6, each POTW had an equal probability of selection within each stratum.

2. At each selected POTW, a representative sample of biosolids is obtained, along with an estimate of the annual mass of biosolids produced by the POTW.

3. The concentration of various chemical constituents is measured within each biosolids sample.

4. Summary statistics are calculated on these measured concentrations, including percentiles of the distribution of concentrations across POTWs.

5. Two sets of percentile estimates are calculated from the measured concentrations, with the two sets distinguished by how each measurement is weighted in the calculations:

    a. One set in which each measurement is weighted by the number of POTWs represented by the sampled POTW, and
    b. One set in which each measurement is weighted by the estimated annual mass of biosolids produced by the sampled POTW.

    Using the 75th percentile as an example, the first set of percentile estimates will represent the concentration for which 75 percent of POTWs have a lower biosolids concentration and 25 percent have a higher concentration. The second set of estimates will represent the concentration for which 75 percent of biosolids mass has a lower concentration and 25 percent has a higher concentration.

For each of eight pollutants identified in the health and ecological exposure and hazard screening assessment (Section 1.4), a simulation was performed to characterize the precision associated with estimating specified percentiles of the chemical concentration. The simulation procedure assumed the following:

- The annual mass of biosolids generated by a POTW is proportional to its flow rate.
- The chemical concentration measurements have a lognormal distribution.
- The chemical concentrations and the mass of biosolids are statistically independent.

Because the chemical concentrations are assumed to have a lognormal distribution, the log-transformed concentrations for a given pollutant are assumed to have a normal distribution with mean μ and standard deviation σ. For a given random variable X and for some number p between 0 and 100, let $Q_{PW}(X)$ denote the weighted p^{th} percentile of the distribution of X, where different values of X are weighted according to a specified set of weights W. Then, if Z denotes a random variable having a standard normal distribution, the weighted p^{th} percentile of the distribution of concentrations C for a given target pollutant is given by:

$$Q_{PW}(C) = \exp(Q_{PW}(\ln(C))) = \exp(Q_{PW}(Z\sigma + \mu)) = \exp(Q_{PW}(Z)\sigma + \mu) = \exp(\mu)\exp(Q_{PW}(Z))^{\sigma}$$

If Q_{PWU} denotes the upper bound of a confidence interval on Q_{PW}, then.

$$Q_{PWU}(C) = \exp(Q_{PWU}(\ln(C))) = \exp(Q_{PWU}(Z\sigma + \mu)) = \exp(\mu)\exp(Q_{PWU}(Z))^{\sigma}$$

The value $Q_{PWU}(Z)$ is determined from standard normal distribution tables that are found in many statistics textbooks. (For example, if a 90% confidence interval is of interest, then $Q_{PWU}(Z) = 1.645$).

One measure of precision associated with an estimate of an unknown distributional parameter, such as a percentile, is the ratio of an upper confidence bound on the estimate to the estimate's value. For a percentile, this ratio is determined as follows:

$$\frac{Q_{PWU}}{Q_{PW}} = \left(\frac{\exp(Q_{PWU}(Z))}{\exp(Q_{PW}(Z))} \right)^{\sigma} = R_{PW}^{\sigma}$$

The standard deviation of the log-transformed concentration measurements, σ, will vary from one pollutant to another. However, note that the factor R_{PW} depends only on the underlying sample design, which is defined by the sample size taken from each stratum and the selection probability of each POTW within each stratum, which is based upon some "measure of size" associated with the POTW. This factor is calculated as:

$$R_{PW} = \exp\left(\frac{\ln\left(\frac{Q_{PWU}}{Q_{PW}} \right)}{\sigma} \right)$$

The calculation of the confidence interval associated with the *p*th percentile Q_{PW} assumes that $\ln(Q_{PW})$ has a normal distribution. This calculation depends on the underlying standard error in the estimate of $\ln(Q_{PW})$. This standard error is determined by simulation that involves the following steps:

1. For each POTW in the sample frame:

   - Estimate its mass of biosolids (*MBS*) by its reported flow rate (*MGD*) times a proportionality factor (*G*):

     *MBS = MGD * G* .

     However, because *MBS* is used only as a weighting factor in this simulation, and because the proportionality factor *G* is constant across all POTWs, the value of *G* does not impact the outcome of the simulation. Thus, for simplicity, *G* was set equal to 1 in these calculations (i.e., *MBS = MGD*).

   - Determine its measure of size (*MOS*):

     $MOS = MGD^{c}$ (for some value of *c* determined by the sample design)

     Because this formula allows a POTW's measure of size to be related to its flow rate, it accounts for sample designs in which POTWs are selected with probability proportional to flow rate raised to some power *c* (e.g., *c*=1 or *c*=0.5). Here, the sample design selected for the TNSSS features selecting POTWs randomly from each stratum with equal probability. Therefore, $MOS = 1$ for each POTW (i.e., $c=0$).

2. For each POTW in the sample frame and for each pollutant of interest, simulate the pollutant concentration in a sample of biosolids from the POTW, assuming lognormality in the distribution of concentrations:

$$C = \exp(Z\sigma + \mu)$$

where Z is a random observation from a standard normal distribution. However, based upon the intended use of these simulated concentrations in estimating precision within the procedure given below, the values of σ and μ do not impact the precision calculations. For example, the term μ cancels out in the equations below. Thus, for simplicity, the values of μ and σ were set to 0 and 1, respectively, in these calculations.

3. Within each stratum, select a systematic sample of n_i POTWs (where n_i is the target sample size for the i^{th} stratum, given in the last column of Table 9), taking into account the value of *MOS* for each POTW. Within the i^{th} stratum, the sampling process occurs as follows:

- POTWs are sorted in increasing order of its value of *MOS*.

- A total of n_i POTWs are selected using a systematic sampling approach with a random start. For the j^{th} POTW within the i^{th} stratum, the probability of selection (*Selprob*) was calculated as

$$Selprob_j = \frac{MOS_{ij} \bullet n_i}{\sum_j MOS_{ij}}$$

If this calculation exceeded one for a given POTW, then it was reset to equal 1.0 (i.e., the POTW was selected with certainty), and the denominator was adjusted to equal the sum across all POTWs in the stratum that were not selected with certainty.

Note that because $MOS=1$ under the sampling design used in the TNSSS, the POTWs do not need to be ordered within each stratum based on some measure of size. Instead, they are ordered based on criteria specified within Section 5.6.4 (i.e., by EPA region, then by state name within region). Furthermore, the value of *Selprob* simplifies to the stratum's sampling fraction (i.e., the ratio of the targeted sample size to the stratum size).

4. Calculate the following two sets of percentiles for the population of POTWs in the frame, using the simulated log-transformed concentrations generated in Step 2:

   - $Pop_{PF} = Q_{PF}(\ln(C))$ = the p^{th} percentile of the distribution of log-transformed concentrations for the population, under the assumption that each POTW represents only itself within the sampling frame.

   - $Pop_{PM} = Q_{PM}(\ln(C))$ = the p^{th} percentile of the distribution of log-transformed concentrations for the population, under the assumption that each POTW is weighted by its value of *MBS* (i.e., mass of biosolids).

   Because the log-transformed chemical concentrations were simulated for all POTWs in the frame within Step 2, these population percentiles were calculated by applying the UNIVARIATE procedure in SAS to the simulated data. Possible values of *p* included all integers from 1 to 99, as well as the distribution's minimum and maximum values.

5. Calculate the following two sets of percentiles for the sample of POTWs selected within Step 3, using the simulated log-transformed concentrations for these POTWs that were generated in Step 2:

   - $Samp_{PF} = Q_{PF}(\ln(C))$ = the p^{th} percentile of the distribution of log-transformed concentrations within the sample, with each POTW weighted by the number of POTWs which it represents within the population (i.e., $W = 1/\,Selprob$).

   - $Samp_{PM} = Q_{PM}(\ln(C))$= the p^{th} percentile of the distribution of log-transformed concentrations within the sample, with each POTW weighted by the mass of biosolids which it represents within the population (i.e., $W = MBS\,/\,Selprob$).

   These percentiles represent sample-based estimates of their respective percentiles within the population distribution. As in Step 4, these population percentiles were calculated by applying the UNIVARIATE procedure in SAS to the simulated data for the sampled POTWs. Possible values of *p* included all integers from 1 to 99, as well as the distribution's minimum and maximum values.

6. For each value of *p* and under both sets of percentiles, calculate the difference between the sample-based estimate of the percentile (from Step 5) and the population estimate (from Step 4):

$$\Delta_{PF} = Samp_{PF} - Pop_{PF}$$
$$\Delta_{PM} = Samp_{PM} - Pop_{PM}$$

   Note that the sample estimate is assumed to be unbiased if $\Delta_{PF} = 0$.

7. Repeat Steps 2 through 6 a total of 400 times, thereby obtaining 400 different estimates of Δ_{PF} and Δ_{PM}.
8. Calculate the means ($\overline{\Delta}_{PF}$ *and* $\overline{\Delta}_{PM}$) and standard deviations ($\hat{S}_{PF}$ *and* $\hat{S}_{PM}$) of the 400 sample estimates of Δ_{PF} *and* Δ_{PM}, respectively.

9. Assuming that the sample-based percentile estimates are unbiased (i.e., $\overline{\Delta}_{PF} = 0$), a 90% confidence interval for the *p*th percentile, in log-transformed units, is calculated as:

$$Samp_{PF} \pm 1.645 \bullet \hat{S}_{PF}.$$

Using this formula, the ratio R_{PF} of the upper bound of this confidence interval to the estimate of the percentile, weighted by the number of POTWs in the target population that each sampled POTW represents, is:

$$R_{PF} = \exp\left(\frac{\ln\left(\frac{Q_{PFU}}{Q_{PF}}\right)}{\sigma}\right) = \exp\left(\frac{\ln\left(\frac{\exp\left(Samp_{PF} + 1.645 \bullet \hat{S}_{PF}\right)}{\exp\left(Samp_{PF}\right)}\right)}{\sigma}\right) = \exp\left(\frac{1.645 \bullet \hat{S}_{PF}}{\sigma}\right)$$

Note that this formula is based on the formula for R_{PW} presented earlier in this appendix, with subscript *PW* replaced by *PF* to emphasize that the weights are based on the number of POTWs. By replacing subscript *PF* with *PM* within this equation, it is assumed that the POTWs are weighted by the mass of biosolids.

Although different assumptions could be made on the value of the standard deviation of log-transformed concentrations σ, R_{PF} and R_{PM} were initially calculated with $\sigma = 1.0$. Then, the result was raised to the σ power (i.e., R_{PF}^{σ} and R_{PM}^{σ}) once σ was estimated from information obtained from the 1988 NSSS Statistical Support Document (1992). In doing so, it was assumed that the measurements to be made in the TNSSS will originate from a distribution having the same median and coefficient of variation (CV) as was observed in the 1988 NSSS. Because the concentration and mass values in the TNSSS are assumed to be lognormally distributed, the CV of a particular pollutant concentration is calculated as:

$$CV = \sqrt{\exp(\sigma^2) - 1}$$

Thus, σ was estimated by solving this formula for σ and substituting CV with the estimate given in the last column of Table 10 of this study plan for the given pollutant. R_{PF}^{σ} and R_{PM}^{σ} were computed for the 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentiles, in order to evaluate the precision associated with the estimates of these percentiles. Those percentiles whose estimated bias exceeded one-half of its estimated standard error (i.e., $\overline{\Delta}_{PF} > \hat{S}_{PF} / 2$) were noted.

For the mass-based estimates, concentration percentiles were weighted by the mass of biosolids represented by each sample. The concentration values were assumed to be uncorrelated with the mass of biosolids.