ENVIRONMENTAL FINANCE CENTER AT THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL SCHOOL OF GOVERNMENT

REPORT 3

Using a Statistical Sampling Approach to Wastewater Needs Surveys

March 2017

Report to the North Carolina State Water Infrastructure Authority and the North Carolina Department of Environmental Quality Division of Water Infrastructure

Report authors: Jordan Paschal, Shadi Eskaf, and Jeff Hughes.

Executive Summary

Every four years, the EPA conducts the Clean Watersheds Needs Survey (CWNS) in order to estimate the capital infrastructure needs of wastewater utilities. Currently, the EPA requires the State to collect capital infrastructure needs data from as many wastewater systems as possible (a census). This report explores whether taking a sample of systems, rather than a census, would provide equally accurate and precise estimates for the statewide wastewater capital infrastructure needs.

Using capital infrastructure needs data on 230 wastewater systems from the 2012 CWNS, we test nine different sampling methods. Out of the nine, two methods proved more accurate and precise than the rest, and a better choice than attempting a census of needs. These two sampling methods are:

> 1) collecting capital infrastructure needs data from all systems serving a population of greater than 100,000, and stratifying all other systems by population and randomly sampling within each stratum; and,
>
> 2) collecting capital infrastructure needs data from all systems serving a population of greater than 100,000, and stratifying all other systems by population and whether or not the system has a treatment plant and randomly sampling within each stratum.

This report concludes that if the State of North Carolina is given the choice of whether to statistically sample wastewater systems or collect as much data from as many systems as possible (essentially a census), sampling will provide equally accurate and precise estimates while lowering administrative burden. Indeed, for the same or less administrative effort, the State could focus on improving the quality and accuracy of the needs estimates of the sampled systems instead of trying to obtain needs estimates for more and more systems. For example, the State could use cost curves to estimate needs for projects that are left out of Capital Improvement Plans, use a consistent timeframe of analysis for all sampled systems instead of a mix of 5-year and 20-year timeframes as currently documented, etc. The marginal benefits of improving the quality of the data from the sampled systems would likely outweigh the marginal benefits of finding and adding needs estimates from additional wastewater systems in terms of improving the accuracy (and magnitude) of the estimated needs for the entire state.

Contents

Acronyms

CIP	Capital Improvement Plan
CWNS	Clean Watersheds Needs Survey
DWI	Division of Water Infrastructure, North Carolina Department of Environmental Quality
DWINSA	Drinking Water Infrastructure Needs Survey & Assessment
EFC	The Environmental Finance Center at the University of North Carolina, Chapel Hill
EPA	Environmental Protection Agency
SDWIS	Safe Drinking Water Information System
SRF	State Revolving Fund
SWIA	State Water Infrastructure Authority

Introduction

Every four years, the State of North Carolina conducts a survey of government-owned wastewater infrastructure needs for the EPA's national Clean Watersheds Needs Survey. Unlike the Drinking Water Infrastructure Needs Survey & Assessment (DWINSA), which employs a statistical sampling of water systems to estimate the state's drinking water infrastructure needs, the Clean Watersheds Needs Survey (CWNS) is currently designed to employ a census methodology. Specifically, the State attempts to collect documentation of wastewater infrastructure needs from as many government-owned wastewater systems as possible. If a wastewater utility is missed, or documentation of needs is unavailable, its needs are not included in the assessment of needs for the state.

This report presents an analysis of nine statistical sampling options that – if permitted by EPA for the CWNS, or if the State chooses to conduct its own survey – can be made as an alternative to the census method currently employed by the CWNS. The nine sampling methods are based on the sampling methodology used for the DWINSA. While the methodologies and limitations of both national needs surveys are outlined in Report 1, they are briefly discussed in this report in order to provide context to the alternatives presented. Each of the nine sampling options is applied to estimate the state's wastewater infrastructure needs several times using Monte Carlo simulation to test estimates based on different samples of that same population. The range of estimates from each option is compared against the known estimates produced from the census data of 230 active local government wastewater systems that participated in the 2012 CWNS. The results from each sampling methodology option are compared against the known estimates and against one another to test for both precision and accuracy.

The first section of this report provides a brief overview of the different methodologies used in the DWINSA and CWNS, along with a discussion of the CWNS's limitations. The second section outlines the nine alternative methods investigated by the EFC along with individual evaluations of their performance. The report concludes with a discussion of the two most accurate and precise methods along with some factors to consider regarding implementation.

The Methods Currently Used in the EPA Needs Surveys

A detailed description of the DWINSA and CWNS is included in Report 1. This section focuses only on describing and comparing the sampling methodologies currently used by the DWINSA and CWNS.

Statistical Sampling Methodology of the Drinking Water Infrastructure Needs Survey & Assessment (DWINSA)

There are over 73,000 community water systems and non-community water systems across the country for which the DWINSA needs to estimate infrastructure needs. Instead of attempting to establish the needs of each individual water system, DWINSA employs a stratified statistical sampling methodology to estimate the infrastructure needs for each state by only sampling a few of the systems and extrapolating to all non-sampled systems.

Stratified sampling is when the researcher divides the population into groups (i.e. strata), such that the water systems within each group are similar to the other systems in their group. By randomly sampling within each stratum, the average needs of the sampled water systems are assumed to be representative of all other water systems in their stratum, because of their similarity. Not only does this sampling methodology save time and money (compared to surveying all 73,000+ water systems), it improves upon simply choosing a random sample of water systems of all sizes and types. Stratified samples ensure that enough water systems of each size and/or type are included in the survey, otherwise the final total estimated need would be unreliable because a random sample might oversample water systems of one size or type and underrepresent (or completely neglect) systems of another size or type.

The DWINSA uses two criteria to create strata: population served, and water source type (groundwater or surface water/groundwater under direct influence of surface water). In all, there are 12 strata, described in the following three groups:

1) All water systems serving more than 100,000 people are included and sampled (i.e., a census of the large systems is employed). This gives each water system in this stratum a weight of 1.

2) For water systems serving between 3,301 and 100,000 people (medium systems), the survey uses 10 strata, subdividing the group of systems into smaller ranges of service population sizes, and water source type. Within each stratum, the sampled systems are assigned a weight=(total number of water systems in that stratum/sample size of the stratum). For instance, if a stratum has 100 total systems in it, and the DWINSA takes a sample of 30 of that stratum, a system in that sample with needs equal to $10 million would have a weight of 100/30=3.33 (i.e. that one sampled system represents 3.33 systems in that stratum). For this stratum, the weighted estimated need would thus be $33.3 million ($10 million multiplied by 3.33), representing the needs of the 3.33 water systems that the sampled system represents. The weight multiplied by the need would be done for each system in the sample to estimate the total need of all systems in all strata. Then the total needs for each strata would be added together to estimate the total national needs. Each stratum is sampled so that the EPA can be 95 percent confident that the true need for each stratum is plus or minus 10 percent. The total national need is also estimated with a 95 percent confidence level for confidence interval of plus or minus 10 percent.

3) For water systems serving fewer than 3,301 customers, EPA simply uses data collected during the 2007 DWINSA, when a nationally representative sample of small water systems was last conducted, and inflates to present dollars and current numbers of small water systems in that stratum for each state.

Non-community water systems' needs were estimated by simply inflating the needs calculated in the 1999 DWINSA and inflating it to present dollars. Similarly, tribal water systems' needs were estimated using data from the 2011 DWINSA and inflated to present dollars.

It is important to note that each time the DWINSA is administered, the survey consultant alters the sampling frame for water systems between 3,301-100,000 in population. Seventy-five percent of the sample is drawn from a sampling frame (i.e. list of systems from which to sample) that is the same in each survey every four years. Although these systems are randomly selected, inevitably many systems are repeatedly selected each time the survey is conducted. This allows for continuity and comparability between surveys over time. However, it is also important that the survey capture changes. In order to ensure that systems that are new, systems that have moved between strata, or systems that were unselected in the last survey are included in the newest survey, the survey consultant sets aside 25 percent of the sample for these utilities.

More details on the methodology is described in Report 1.

Census Methodology of the Clean Watersheds Needs Survey (CWNS)

This survey is implemented by EPA in order to estimate the capital investment necessary for the nation's publicly-owned treatment works to meet the water quality objectives of the Clean Water Act. Since appropriations for the Clean Water State Revolving Fund are set in federal statute, the CWNS survey findings do not affect state allocations for SRF funds in the same way that the DWINSA state-by-state survey findings affect the allocations for Drinking Water State Revolving Funds. In addition, there are fewer publicly-owned treatment works than there are drinking water systems. At present, the EPA uses a census approach to collecting data on wastewater infrastructure needs, rather than employ a statistical sampling methodology.

In the most recent CWNS for which data are available, the EPA collected data from wastewater systems in 2012. States attempted to seek and collect documentation of infrastructure needs for every publicly-owned treatment works in the state. Next, the States and EPA identified only allowable projects from these documents to include in the state's official needs estimates. Allowable projects are projects that are expected to occur sometime in the next 20 years and have sufficient documentation of cost estimates, description and location of water quantity or water quality public health problem the project addresses, and site-specific information. For wastewater systems serving fewer than 10,000 people, the EPA allows flexibility in terms of documentation required in order to include projects from small wastewater systems. If given enough parameters, the EPA will estimate costs for projects for these systems using cost curves.

In the end, the CWNS estimates of infrastructure needs for a state depends on the availability and comprehensiveness of data and documentation at the system-level, and the ability of the State to collect and include as many of the documents as possible, as well as the ability to distinguish allowable projects. No extrapolation occurs for wastewater systems for which project data and parameter data are absent.

Limitations of the CWNS Census Method

Attempting to take a census of all local government wastewater systems in the state requires significant resources and can provide an unreliable estimate of the true statewide capital investment needs. Firstly, as demonstrated by the EFC in Report 2, there is a large number of wastewater systems that do not participate in the survey, and without extrapolating the calculated needs of participating systems to add on the estimated needs of non-participating systems (using weights), the total needs calculated in the CWNS census is much lower than the true statewide needs. Secondly, if the systems that participate in the survey are different from systems that do not participate, then the final needs estimate will be biased and cannot be used to extrapolate to include the needs of non-participating systems.

This is likely true in North Carolina because the process of including wastewater systems in the CWNS involves locating systems' documentation on their infrastructure needs, such as finding capital improvement plans on local government websites. Smaller local governments are less likely to have a comprehensive capital improvement plan that is easily accessible, and are therefore less likely to participate in the survey, as has happened in the past. Thus, the average needs per system included in the CWNS for North Carolina is likely too high to estimate the needs of the non-participating (smaller) wastewater systems in the state.

The CWNS also collects inconsistent information. First, the depth of the planning documents the survey relies on dictates how high the needs will be estimated. Projects with insufficient documentation are left out, but that does not mean the need does not exist. Second, although the survey seeks to report 20 year needs, the survey underreports the 20 year total need. This happens for a few reasons:

1) Planning documents become less accurate the farther into the future they extend,
2) Forty percent of the CWNS needs were reported by examining Capital Improvement Plans. These plans typically only cover 3-6 years, and very few include projects 10-20 years into the future, and
3) If a wastewater system scales back, pauses, or eliminates projects from its Capital Improvement Plan, the need is not recorded by the CWNS even if it still exists.

Data collection and analysis for the CWNS depends heavily on State participation. The DWINSA has greater participation by the systems themselves and there is a method to account for the needs of water systems that do not participate in the survey. The CWNS simply does not report or include the needs of systems that are not included in the survey.

Specifically with respect to North Carolina, the CWNS has limitations. Needs from more than 180 wastewater systems were not included in the 2012 CWNS. The EFC's analysis indicates that those left out tended to be smaller systems (probably with less documentation) than those included. In Report 2, the EFC estimated that the actual wastewater needs for the State of North Carolina can be up to 50% higher than what the official 2012 CWNS estimated.

New Sampling Method Options for Wastewater Needs Surveys

An alternative method to using a census approach is to take a statistically representative sample of wastewater systems rather than trying to record the needs of the entire population of systems. The State's efforts would then be focused on improving the quality and accuracy of the needs estimates for the sampled systems (e.g. using cost curves to estimate needs for projects that are left out of capital improvement plans, using a consistent timeframe of analysis for all sampled systems instead of a mix of 5-year and 20-year timeframes as currently documented, etc.) instead of on obtaining needs estimates for more and more systems. As a result, for the same amount of effort, the State would have more accurate estimates of needs that can then be reliably extrapolated to estimate the total statewide wastewater needs within a reasonable margin of error (plus or minus 10%), rather than aggregating needs estimates that ignore the needs of non-participating systems.

The DWINSA uses a stratified sample methodology and a similar methodology would be useful for estimating needs in North Carolina for wastewater facilities. In this section, we describe nine different sampling strategies, apply them, and demonstrate the accuracy and precision of the estimates and how they compare to known capital needs estimates. In each of these strategies, only a statistically-representative sample of local government wastewater systems are taken and their needs are used to extrapolate to the total local government wastewater needs for the entire State of North Carolina.

By design, we determine the sample size required in each stratum by intentionally setting a confidence level of 95 percent (a standard in research) with a confidence interval of plus or minus 10 percent. This is the same level of precision that the DWINSA uses when sampling water systems.

Within the stratified sampling methodologies, in any stratum that had 10 or fewer systems, all of the systems' needs were sampled (a census within the stratum). For instance, a stratum may only have six wastewater systems, based on how the strata are designed. Statistically, we need a sample of five out of six to produce the confidence level and confidence intervals desired. In instances like this, we sampled six of six because the marginal effort of including the final wastewater system is worth the increase in precision and accuracy in the estimate.

Demonstrating the accuracy of the different survey methodology strategies

We demonstrate here that statistical sampling of wastewater systems and appropriate use of weights to extrapolate their needs can and does produce a reliable, accurate statewide needs estimate (compared to the current census approach). In North Carolina, the EFC has identified 417 active local government wastewater systems that would be included in the needs survey if all systems are included.

Ideally, we would need to know the total needs of all 417 systems in order to compare those true needs with the needs that we estimate using the sampling methodology. However, as explained previously, only 230 out of the 417 systems participated in the 2012 CWNS survey. For this demonstration purpose, then, we will pretend that those 230 utilities are the only local government utilities in North Carolina, and thus they represent a full census of systems needs in the state. The total reported need for these 230 utilities was $4.966 billion in the 2012 CWNS. In this demonstration, this figure represents the true total statewide wastewater infrastructure needs. The question is: can statistical sampling of the 230 utilities produce a needs estimate that is consistent and close to the $4.966 billion figure?

We employ the nine different sampling methodologies and sample within these 230 utilities and obtain their needs estimates from the 2012 CWNS database. We then use the sampled reported needs to extrapolate (using weights) to a total estimated needs for the state, which we then compare to the "true needs" of $4.966 billion. For each of the methods below, the goal is for the extrapolated needs estimate to be close to the "true needs" of $4.966 billion; this tests the *accuracy* of the method. The difference between the estimated needs from the sample and the true needs are calculated in terms of dollars and percent difference.

Precision of the sampling methodology requires consistent estimates of needs regardless of which wastewater systems are sampled. To test the precision of the nine sampling methodologies below, each methodology is repeated 10,000 times, each time taking a new sample of systems within the 230 systems, and their extrapolated estimated needs compared to the $4.966 billion "true needs" for the state. We provide results of analysis on how different the estimated needs from the sampling methodologies are from the "true needs" estimate, and the range of estimates produced by each method.

The nine different sampling methodologies are summarized in the following table, displaying the different strata used in each option. The strata were created using different combinations of service population, ownership by municipality or non-municipality, and/or presence of a treatment facility. Unless a census is specifically stated in the table, or when the population of wastewater systems is 10 or fewer (in which case a census was employed), random samples were selected in each stratum described below.

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 1	Random Sample			
Method 2	Random Sample			Census
Method 3	Random Sample	Random Sample	Random Sample	
Method 4	Random Sample	Random Sample	Random Sample	Census
Method 5	Stratify by Municipality vs. Non-Municipality	Stratify by Municipality vs. Non-Municipality	Stratify by Municipality vs. Non-Municipality	
Method 6	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality			
Method 7	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	
Method 8	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	
Method 9	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Census

Method 1: Simple Random Sample

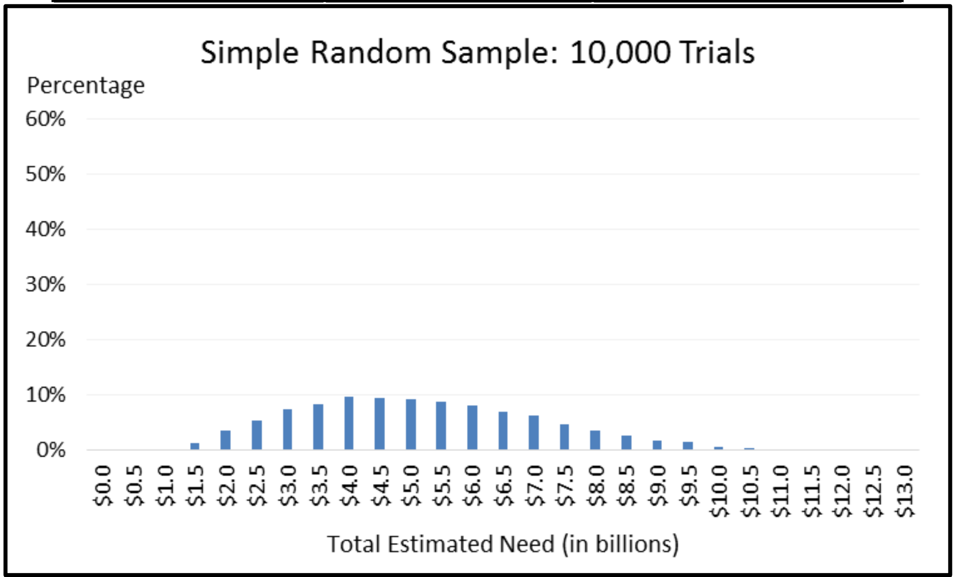
	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 1	Random Sample			

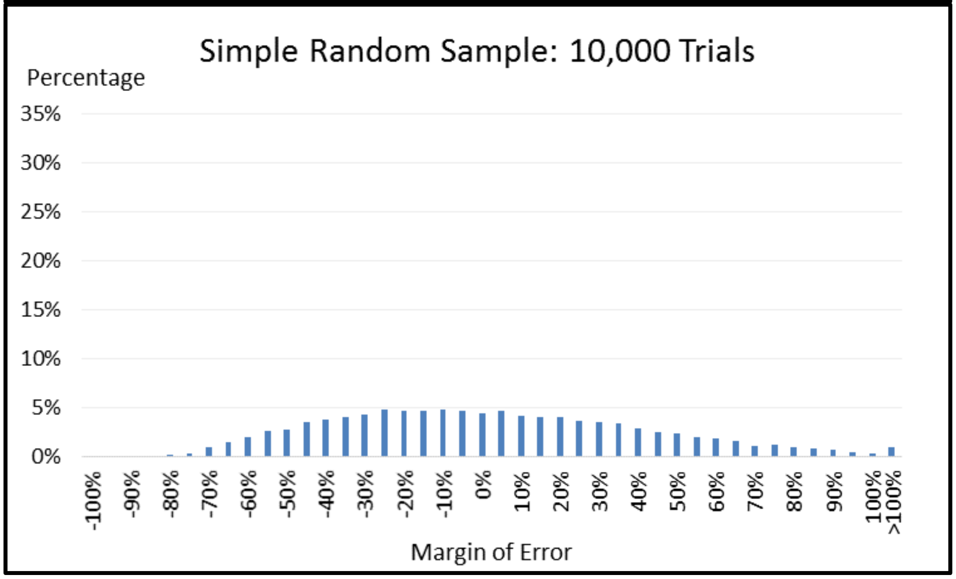
Generally, the gold standard in statistical research is the simple random sample. In this option, a random sample of the 230 wastewater systems is selected, without consideration for system characteristics. With each system in the population having an equal probability of being selected for the sample, those in the sample are statistically assumed to have the same needs, on average, as those not selected. Thus, the average needs from the sample represents the entire population's needs.

When discussing the capital investment needs of wastewater systems, however, it is clear that taking a random sample will not be accurate. A simple random sample that happens to select a disproportionate number of smaller systems will not accurately represent the needs of medium and large systems, and vice versa.

With 230 utilities, a random sample of 68 is required to meet the level of precision defined above. Using Monte Carlo simulations, we simulated taking 10,000 samples of 68 wastewater systems and calculated the statewide needs from the sample each time and the percent difference from the $4.966 billion "true needs" figure. The following table displays the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the "true needs" figure.

	Percentage Error	Total Estimated Need
Mean	-0.70%	$4,932,295,006
Standard Deviation	39.48%	$1,961,039,554
Mean Standard Error	0.39%	$19,610,396
Skewness	0.403	0.403
Minimum	-83.19%	$834,950,559
First Quartile	-30.70%	$3,441,864,753
Median	-4.11%	$4,762,836,388
Third Quartile	26.22%	$6,269,281,090
Maximum	146.13%	$12,224,869,421





The results of the simulation show that the median error of the simple random sample method is to underestimate the total need by 4.11 percent with a standard deviation of 39.48 percent. Although they would be very unlikely results, this method may underestimate the total need by 83 percent or overestimate it by as much as 146 percent.

We conclude that a Simple Random Sample is the least accurate and least precise methodology of those tested in this report. It would be a poor sampling methodology to estimate the statewide wastewater infrastructure needs.

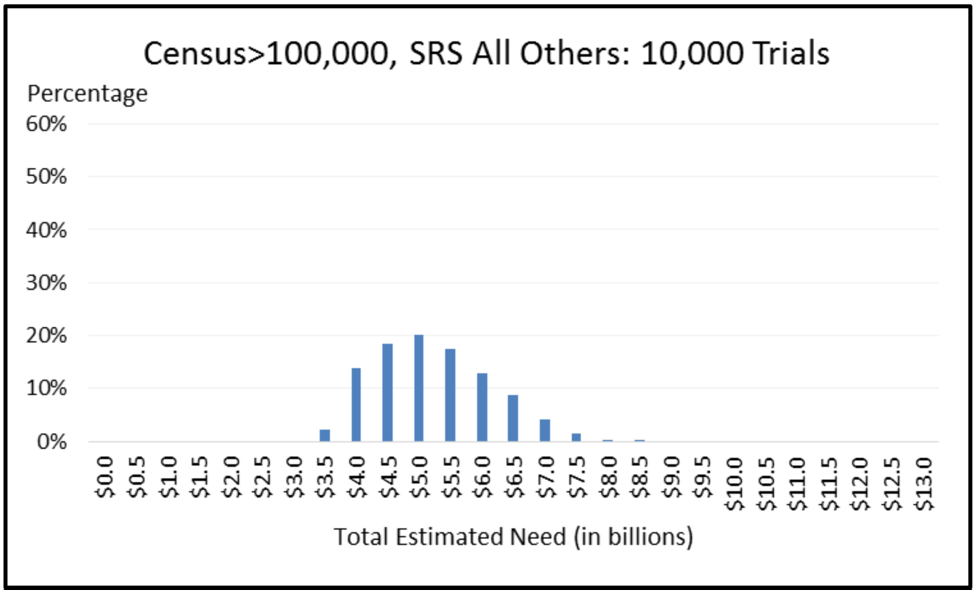
Method 2: Census Systems above 100,000 people; Simple Random Sample all others

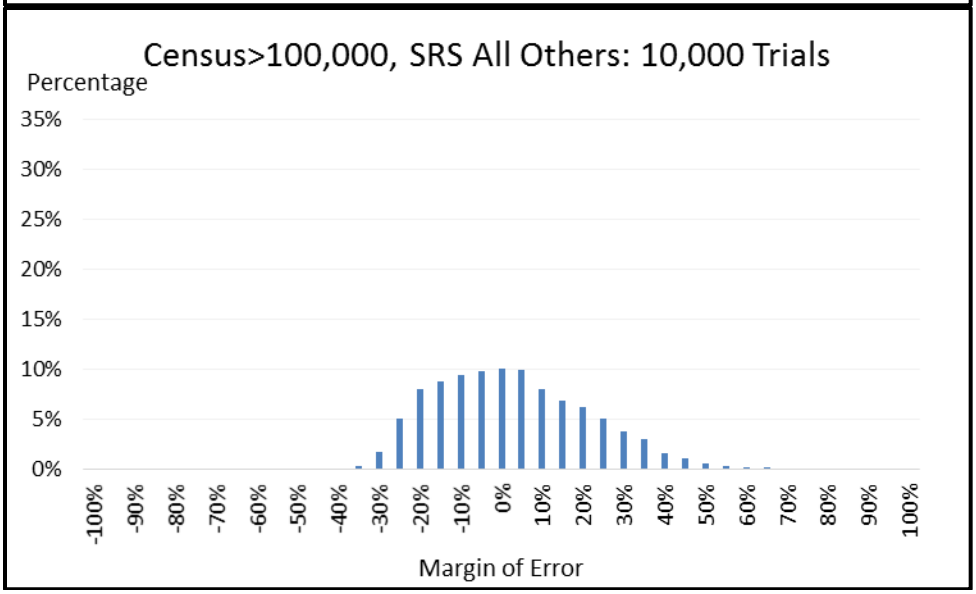
	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 2	Random Sample			Census

As the DWINSA notes in its methods, the capital investment needs of large systems are so large and unique that any survey leaving them out would drastically underestimate the true needs of the state (which is confirmed in the results of Method 1). That is part of the reason the random sample above could be so far off the true need in different sampling iterations. By ensuring that all of the systems serving more than 100,000 people are included and taking a random sample of all other systems, the estimate should become more accurate and more precise. There are 9 systems out of the 230 that serve more than 100,000 people. All nine are included in the sample (a census) under this Method. Out of the remaining 221 systems, 67 were randomly selected into the sample, regardless of system characteristics. This makes a total sample size of 76 systems. The following table and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the “true needs” figure.

	Percentage Error	Total Estimated Need
Mean	0.04%	$4,968,923,392
Standard Deviation	18.38%	$912,979,009
Mean Standard Error	0.18%	$9,129,790
Skewness	0.415	0.415
Minimum	-38.96%	$3,031,903,046
First Quartile	-14.37%	$4,253,460,801
Median	-1.61%	$4,886,826,518
Third Quartile	12.65%	$5,595,273,394
Maximum	62.03%	$8,047,930,430

	Average System Need	Total Statewide Need
Systems >100,000	$268,689,691	$2,418,202,222
All other systems	$11,541,702	$2,550,716,170





The median error dropped to -1.61 percent with a standard deviation of 18.38 percent, also a drop from the first method. The maximum underestimation also dropped from 83 percent to 39 percent and the maximum overestimation decreased from 146 percent to 62 percent.

This method is much more accurate and precise than Method 1. However, overall, we conclude that there are other methods below that are more precise and more accurate.

Method 3: Stratified Sample of All Systems

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	**5,001 - 20,000**	**20,001 - 100,000**	**>100,000**
Method 3	Random Sample	Random Sample	Random Sample	

In this method, the service population strata are further refined into three groups – those serving 1-5,000, 5,001-20,000, and more than 20,000 people. Wastewater systems are randomly sampled within each of the three strata. By stratifying on population served (a key determinant of capital investment needs), the sample estimates of the true total need should become even more precise than the previous two methods. The table below displays the number of systems in each stratum and the sample size needed to meet the desired level of precision defined above. The total sample size is 114.

	Number of Systems	Sample Size	Weight
1-5,000	155	60	155/60
5,001-20,000	38	27	38/27
>20,000	37	27	37/27

Once the estimated need for each stratum is calculated by summing the needs of sampled systems, it is multiplied by the stratum's weight. This gives the final estimated need for that stratum. Adding these weighted estimated needs across the three strata gives the final overall estimated need for all 230 systems.

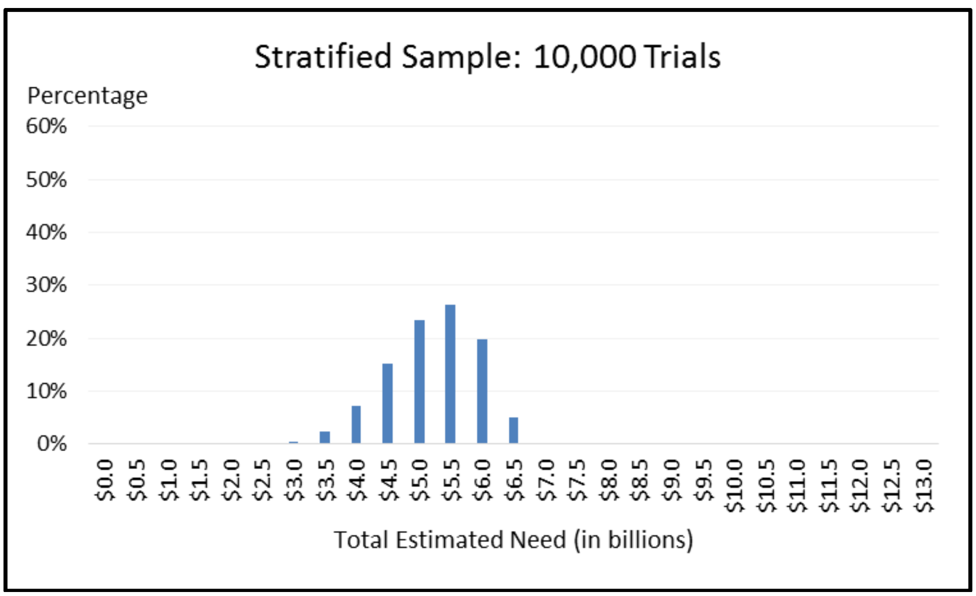
For example, we take a sample of 60 systems out of the 155 that serve fewer than 5,000 people. That sample's needs summed equals $300 million. $300 million is then multiplied by 155/60 (the weight) and equals $775 million, which is the weighted total estimated need for systems below 5,000 in population. These steps are repeated for the other two strata and the three weighted total estimated needs are summed to get the final overall estimated needs for the entire State.

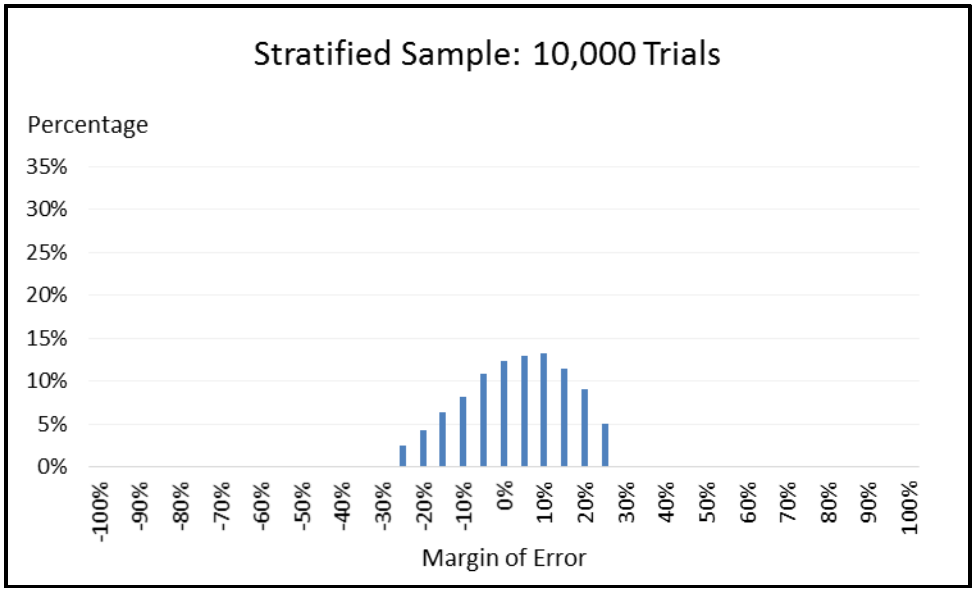
This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in each stratum. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the "true needs" figure.

	Percentage Error	Total Estimated Need
Mean	-0.07%	$4,963,664,000
Standard Deviation	14.18%	$704,307,481
Mean Standard Error	0.14%	$7,043,075
Skewness	-0.428	-0.428
Minimum	-56.59%	$2,156,222,348
First Quartile	-9.51%	$4,494,610,389
Median	1.13%	$5,022,957,154
Third Quartile	10.66%	$5,496,140,808
Maximum	31.10%	$6,511,464,811

	Average System Need	Total Statewide Need
1-5,000	$3,078,679	$477,195,224
5,001-20,000	$11,167,098	$424,349,712
>20,000	$109,787,002	$4,062,119,065

The median error for this method was +1.13 percent with a standard deviation of 14.18 percent, both lower than Method 1 and Method 2. The maximum underestimation got worse by increasing from 39 percent to 57 percent. The maximum overestimation improved by decreasing from 62 percent to 31 percent. The histograms below show the distribution of samples beginning to cluster more around 0 percent error compared to the first two methods, indicating greater accuracy, and the range of estimates is narrower, indicating higher precision than in the previous two methods. Nonetheless, other methods outperform this method in terms of accuracy and precision.





Method 4: Census Systems above 100,000; Stratified Sample for all others

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 4	Random Sample	Random Sample	Random Sample	Census

Repeating the census approach of the largest wastewater systems, those serving more than 100,000 people, should again improve the accuracy of the estimates. Method 3 does include systems serving above 100,000, but only a sample of them. Method 4 samples every wastewater system serving more than 100,000 people, and then stratifies the remainder of systems based on service population, randomly selecting samples within each stratum.

The table below displays the number systems in each stratum and the sample size needed to meet the desired level of precision defined above. The total sample size is 118.

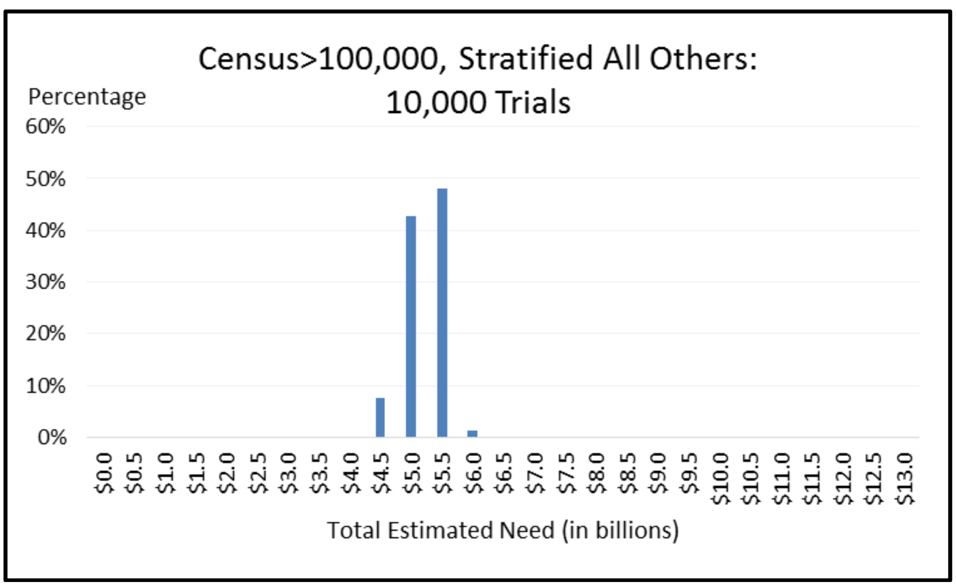
	Number of Systems	Sample Size	Weight
1-5,000	155	60	155/60
5,001-20,000	38	27	38/27
20,001-100,000	28	22	28/22
>100,000	9	9	9/9

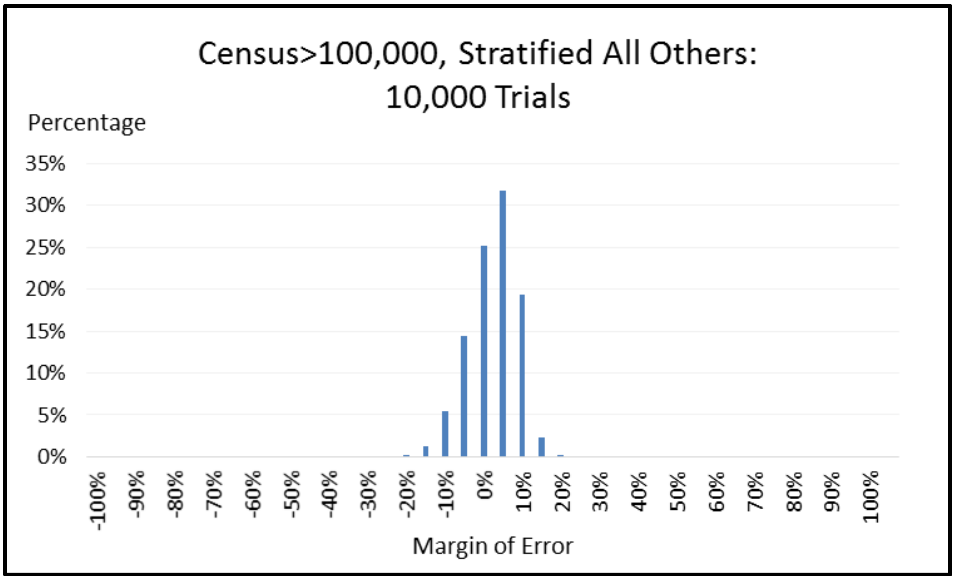
This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in the first three strata, but always including all 9 systems from the largest service population stratum. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the "true needs" figure.

	Percentage Error	Total Estimated Need
Mean	-0.11%	$4,961,595,343
Standard Deviation	6.12%	$304,123,322
Mean Standard Error	0.06%	$3,041,233
Skewness	-0.489	-0.489
Minimum	-24.13%	$3,768,241,303
First Quartile	-4.08%	$4,764,361,581
Median	0.59%	$4,995,739,071
Third Quartile	4.49%	$5,189,780,655
Maximum	15.33%	$5,727,547,912

	Average System Need	Total Statewide Need
1-5,000	$3,083,910	$478,005,972
5,001-20,000	$11,165,807	$424,300,684
20,001-100,000	$58,610,052	$1,641,081,466
>100,000	$268,689,691	$2,418,207,222

The median error for this method was 0.59 percent with a standard deviation of 6.12 percent, both lower than Method 1, Method 2, and Method 3. The maximum underestimation improved from around 56 percent to 24 percent. The maximum overestimation improved from being nearly 31 percent to just over 15 percent. The histogram below shows the distribution of samples continuing to cluster around the true need and an error of 0 percent compared to the first three methods. The range of estimates is particularly narrow, demonstrating a very high precision in the estimates.





Overall, we identify that **Method 4 is the second most accurate and precise sampling methodology in this report** given its median error and standard deviation.

Method 5: Sample stratified by Population and Municipality Status

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	**5,001 - 20,000**	**20,001 - 100,000**	**>100,000**
Method 5	Stratify by Municipality vs. Non-Municipality	Stratify by Municipality vs. Non-Municipality	Stratify by Municipality vs. Non-Municipality	

In an effort to further refine the estimates, another stratum that may be a determinant of capital infrastructure needs can be added. For Method 5, we divide up the systems not only by service population, but also by whether the system is owned by a municipality or not. We use “municipality” as a proxy for population density. Municipal wastewater systems are more likely to serve smaller areas with high population densities than wastewater systems owned by Counties or regional local governments. The more spread out the service population, the greater the per-capita needs of the system, since more miles of collection system pipes are required.

This method functions the same as Method 3, with the only change being there are six strata instead of three. The table below displays the sample size (and total number of systems) of each strata. For instance, there are 145 municipal systems under 5,000 in service population, 58 of which will be sampled. The weight for this group will be 145/58. The total sample size is 128.

	Municipality	Non-Municipality
1-5,000	58 (out of 145)	10 (out of 10)
5,001-20,000	24 (out of 32)	6 (out of 6)
>20,000	20 (out of 25)	10 (out of 12)

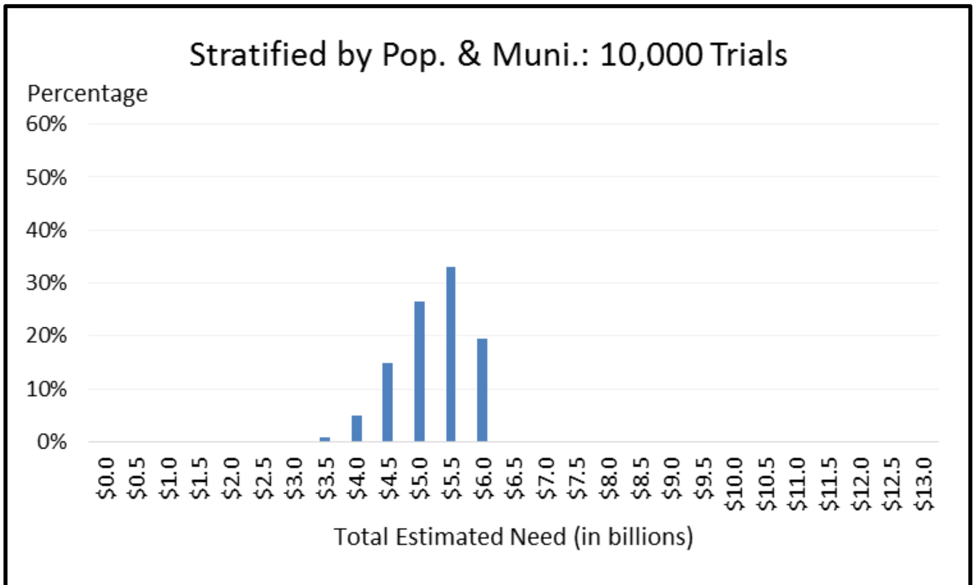
This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in the six strata, except the smallest stratum where all 6 non-municipal systems serving 5,001-20,000 people are always included. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the "true needs" figure.

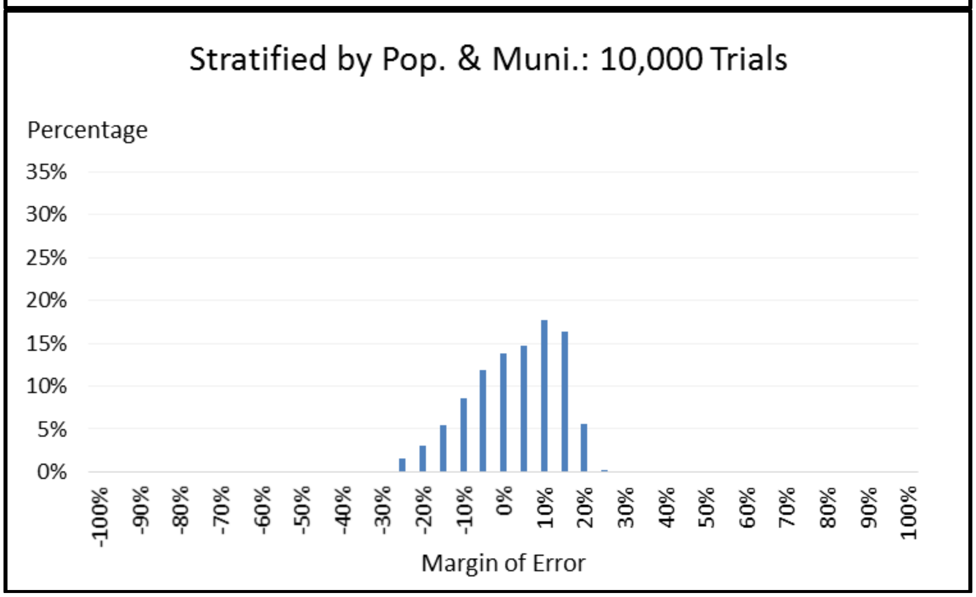
	Percentage Error	Total Estimated Need
Mean	0.08%	$4,970,658,156
Standard Deviation	11.35%	$563,956,867
Mean Standard Error	0.11%	$5,639,569
Skewness	-0.615	-0.615
Minimum	-47.83%	$2,591,108,871
First Quartile	-7.55%	$4,592,231,642
Median	1.65%	$5,048,783,610
Third Quartile	9.24%	$5,425,858,386
Maximum	22.82%	$6,100,300,506

		Average System Need	Total Statewide Need
1-5,000	Municipality	$2,713,765	$393,495,970
	Non-Municipality	$8,608,806	$86,088,061
5,001-20,000	Municipality	$10,511,845	$336,379,039
	Non-Municipality	$14,732,698	$88,396,185
>20,000	Municipality	$119,705,853	$2,992,646,321
	Non-Municipality	$89,471,048	$1,073,652,580

As shown in the table above, there are significant differences between the average needs of systems that are municipally-owned versus other systems. Small, non-municipal systems have larger needs than municipal systems serving the same number of people. Conversely, large municipal systems have higher needs than large non-municipal systems. These differences justify the choice of using municipally-owned vs. non-municipally owned as a defining characteristic of the strata.

The median error for this method was +1.65 percent with a standard deviation of 11.35 percent, slightly higher than Method 4. However this method has similar results to Method 3 and Method 7. The maximum underestimation fell to almost 48 percent while the maximum overestimation increased from around 17 percent in Method 4 to nearly 23 percent in Method 5. Essentially, the estimation has gotten less precise by adding another level of stratification.





Although Method 5 adds another characteristic by which to group the data, since we no longer use a census for the largest systems, the estimates are less accurate.

Method 6: Sample stratified by Municipality and Treatment Status

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 6	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality			

In an effort to further refine the estimates, one more system characteristic is used to define strata. For Method 6, we divide up the systems not only by whether they are owned by a municipality or not, but also by whether they have a wastewater treatment plant or not. Collection-only systems are likely to have much lower needs than wastewater systems that have collection systems as well as treatment plants. In this method, we ignore service population, resulting in only four strata. We randomly select systems within each stratum as shown in the table below, except for the smallest stratum where all 7 non-municipal systems that do not have a wastewater treatment plant are always sampled. For instance, we sample 60 systems out of 160 municipally-owned systems that have treatment plants. The weight for this group will be 160/60. The total sample size is 113.

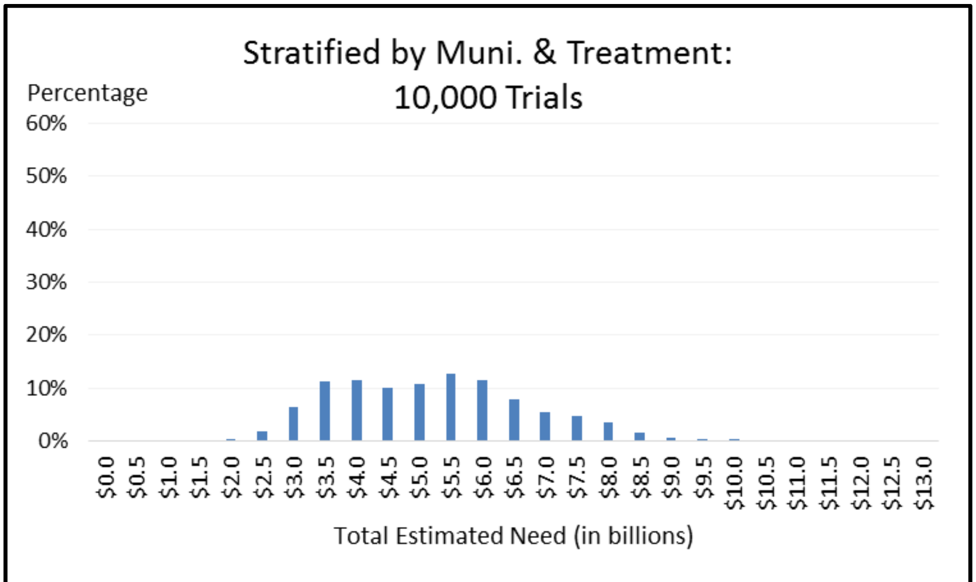
	Municipality	Non-Municipality
Treatment	60 (out of 160)	17 (out of 21)
Non-Treatment	29 (out of 42)	7 (out of 7)

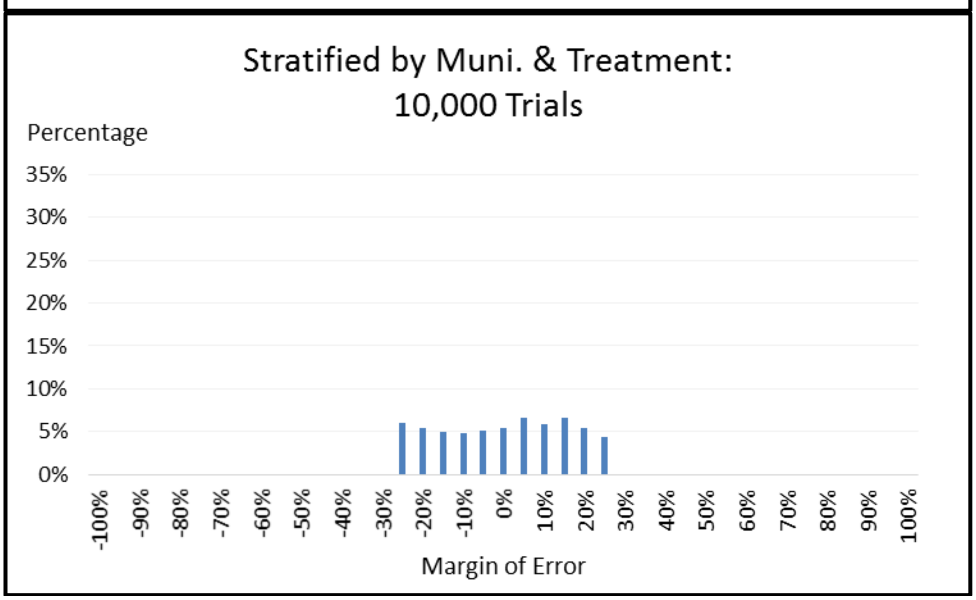
The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the "true needs" figure.

	Percentage Error	Total Estimated Need
Mean	-0.35%	$4,949,414,125
Standard Deviation	29.65%	$1,472,866,494
Mean Standard Error	0.30%	$14,728,665
Skewness	0.306	0.306
Minimum	-65.14%	$1,731,681,894
First Quartile	-25.17%	$3,716,493,347
Median	-0.86%	$4,923,944,234
Third Quartile	19.60%	$5,938,046,244
Maximum	93.60%	$9,616,141,893

		Average System Need	Total Statewide Need
Municipal	Treatment	$22,751,159	$3,640,185,514
	Non-Treatment	$1,509,799	$63,411,546
Non-Municipal	Treatment	$37,814,680	$794,108,273
	Non-Treatment	$64,529,827	$451,708,792

The median error for this method was -0.86 percent with a standard deviation of 29.65 percent. The maximum underestimation is around 65 percent and the maximum overestimation is nearly 94 percent. The histograms demonstrate that although the median error was fairly accurate, the distribution of possible estimates is very large, indicating a lack of precision. This method is less precise than all other methods except for Method 1. The exclusion of system size in the strata, therefore, decreased the precision of the estimates.





Method 7: Sample stratified by Population and Treatment Status

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	**5,001 - 20,000**	**20,001 - 100,000**	**>100,000**
Method 7	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	

This method is similar to Method 5, using service population to define strata, but instead of using municipal ownership, the presence of a treatment plant was also used to define the six strata. The table below displays the sampling size of each stratum. For instance, 52 out of 114 systems with treatment plants serving fewer than 5,000 people were sampled. The weight for this stratum will be 114/52. The total sample size is 138.

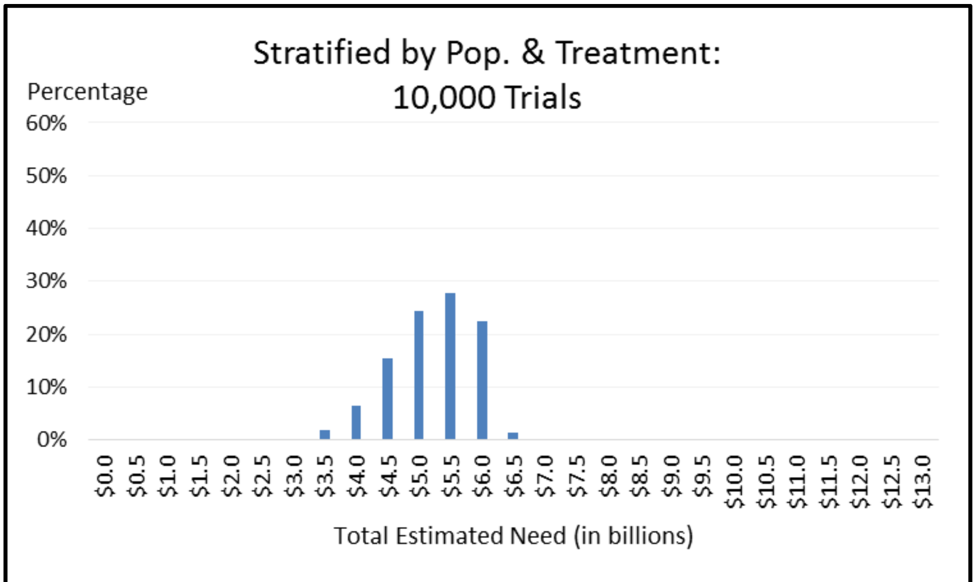
	Treatment	Non-Treatment
1-5,000	52 (out of 114)	29 (out of 41)
5,001-20,000	26 (out of 36)	2 (out of 2)
>20,000	23 (out of 31)	6 (out of 6)

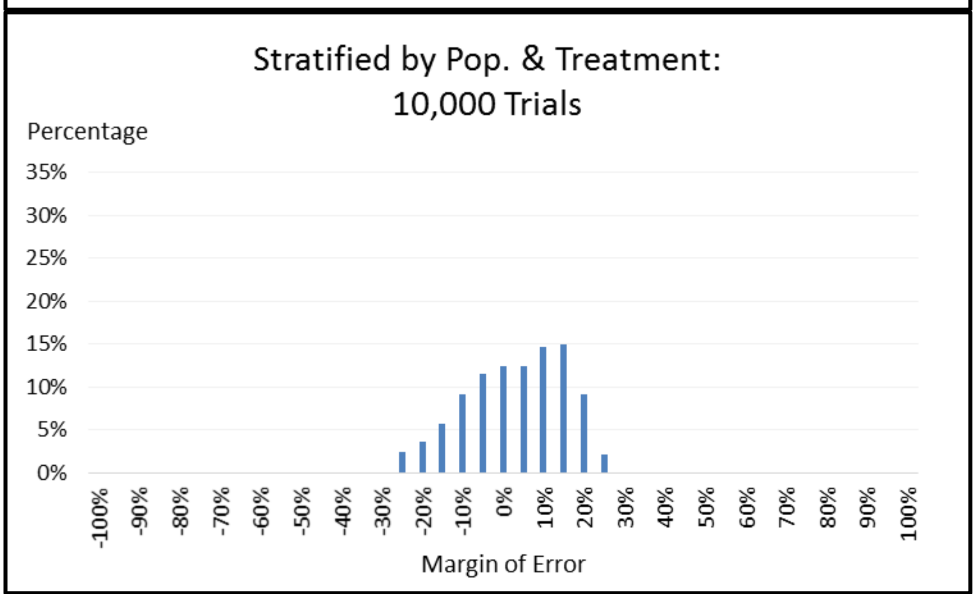
This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in four strata, but always including the 8 systems without treatment plants that serve more than 5,000 people. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the “true needs” figure.

	Percentage Error	Total Estimated Need
Mean	-0.10%	$4,962,123,999
Standard Deviation	12.98%	$644,912,310
Mean Standard Error	0.13%	$6,449,123
Skewness	-0.523	-0.523
Minimum	-49.22%	$2,522,056,465
First Quartile	-8.90%	$4,524,836,410
Median	1.38%	$5,035,228,866
Third Quartile	10.35%	$5,480,985,570
Maximum	27.82%	$6,348,646,169

		Average System Need	Total Statewide Need
1-5,000	Treatment	$3,695,690	$421,308,627
	Non-Treatment	$1,423,420	$58,360,237
5,001-20,000	Treatment	$11,602,995	$417,707,836
	Non-Treatment	$3,272,699	$6,545,398
>20,000	Treatment	$116,376,172	$3,607,661,331
	Non-Treatment	$75,090,095	$450,540,570

As shown in the table above, there are significant differences between the average needs of systems that have treatment plants versus those that do not. Unlike in Method 6, when comparing the needs of systems to those of similar service populations, systems with treatment plants were found to have, on average, higher needs than systems that do not have treatment plants, which is as expected. These differences justify the choice of using treatment plant presence as a defining characteristic of the strata, at least when differentiating between system sizes.





The median error for this method was 1.38 percent with a standard deviation of 12.98 percent. The maximum underestimation is around 49 percent and the maximum overestimation is 28 percent. This method is better than Methods 1, 2, 6, and 8, but Methods 4 and 9 are still superior as their distribution is tighter (more precise) and closer to an error of 0 percent (more accurate) than Method 7.

Method 8: Sample Stratified by Population, Municipality, and Treatment

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 8	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	Stratify by Treatment vs. Non-Treatment & Municipality vs. Non-Municipality	

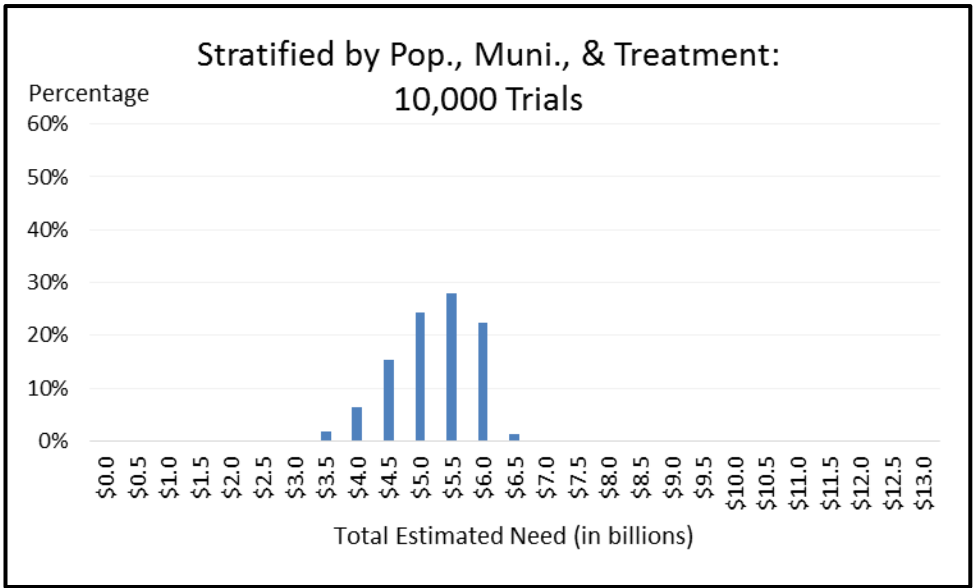
Method 8 uses service population, ownership by municipality, and presence of a treatment plant all together to produce 12 different strata – the most of all of the methods investigated in this report. Theoretically, the addition of more strata should make the needs estimates more precise than in Methods 3, 5, and 7. The table below displays the sample size of the 12 strata. The total sample size is 151. Two of the strata did not contain any wastewater system among the 230 being studied, and 5 strata included very few wastewater systems for which we always included all of their wastewater systems in the sample.

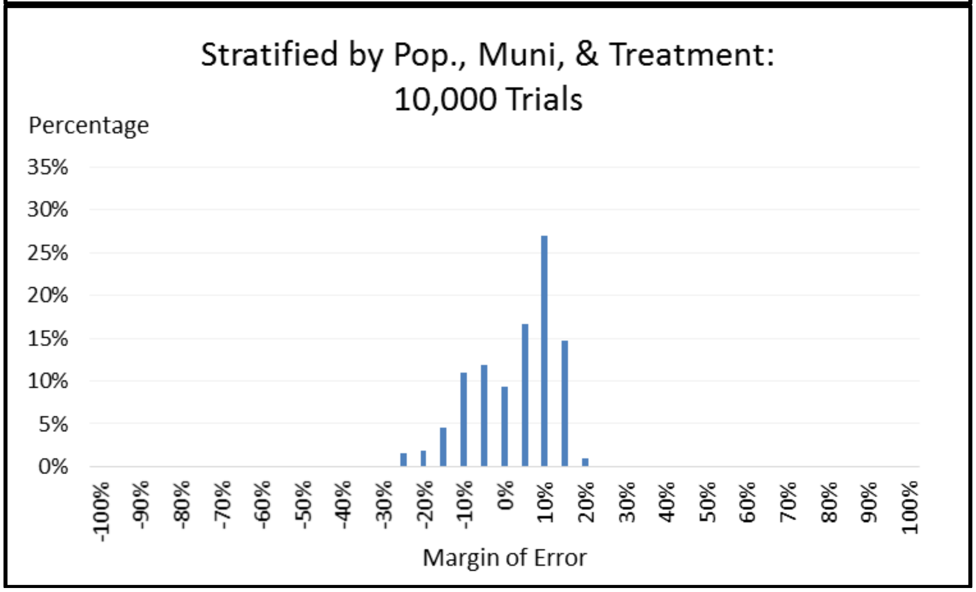
		Treatment	Non-Treatment
1-5,000	Municipality	50 (out of 105)	28 (out of 40)
	Non-Municipality	9 (out of 9)	1 (out of 1)
5,001-20,000	Municipality	23 (out of 30)	2 (out of 2)
	Non-Municipality	6 (out of 6)	0 (out of 0)
>20,000	Municipality	20 (out of 25)	0 (out of 0)
	Non-Municipality	6 (out of 6)	6 (out of 6)

This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in four strata, but always including the 30 systems in the six small-sample strata. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the “true needs” figure.

	Percentage Error	Total Estimated Need
Mean	0.09%	$4,971,406,705
Standard Deviation	10.35%	$514,117,493
Mean Standard Error	0.10%	$5,141,175
Skewness	-0.779	-0.779
Minimum	-37.41%	$3,109,012,882
First Quartile	-7.92%	$4,573,734,556
Median	3.27%	$5,129,316,137
Third Quartile	8.32%	$5,379,901,779
Maximum	17.50%	$5,836,107,921

		Average System Need	Statewide Need
1-5,000	Municipality & Treatment	$3,203,993	$336,419,242
	Municipality & Non-Treatment	$1,427,150	$57,086,016
	Non-Municipality & Treatment	$9,435,538	$84,919,839
	Non-Municipality & Non-Treatment	$1,168,222	$1,168,222
5,001-20,000	Municipality & Treatment	$10,965,320	$328,959,597
	Municipality & Non-Treatment	$3,272,699	$6,545,398
	Non-Municipality & Treatment	$14,732,698	$88,396,185
	Non-Municipality & Non-Treatment	N/A	N/A
>20,000	Municipality & Treatment	$119,861,990	$2,996,549,745
	Municipality & Non-Treatment	N/A	N/A
	Non-Municipality & Treatment	$103,470,315	$620,821,891
	Non-Municipality & Non-Treatment	$75,090,095	$450,540,570





The median error for this method was +3.27 percent with a standard deviation of 10.35 percent. The maximum underestimation fell to almost 37 percent while the maximum overestimation was nearly 18 percent. Although Method 8 adds all three characteristics by which to group the data (and thus, in theory, decreases statistical variance) and is more precise than most of the other methods described above, the needs estimates are less accurate than others.

Method 9: Census above 100,000; All Others Stratified by Population and Treatment

	Service Population of Publicly-Owned Treatment Works			
	1 - 5,000	5,001 - 20,000	20,001 - 100,000	>100,000
Method 9	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Stratify by Treatment vs. Non-Treatment	Census

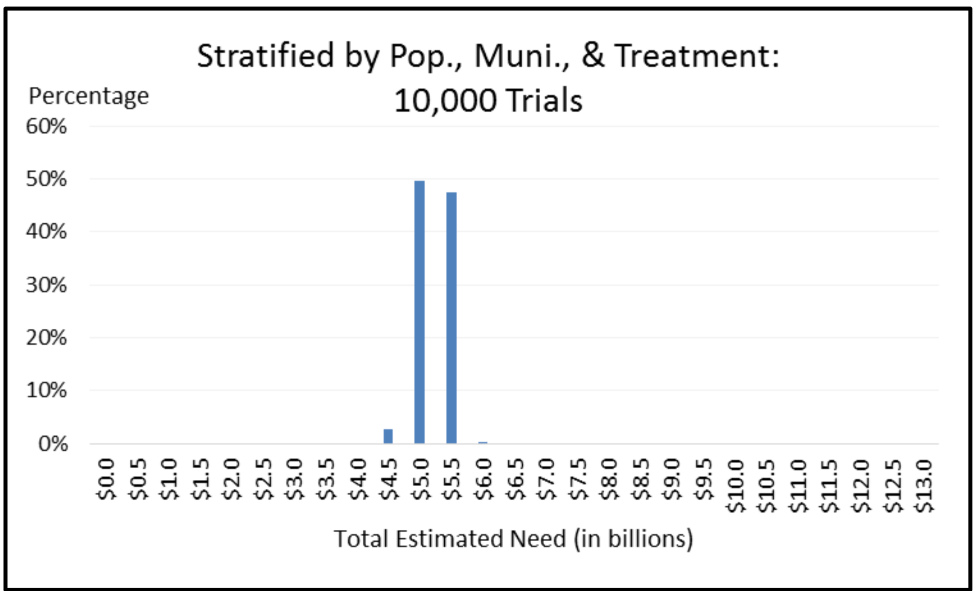
Method 7 demonstrated that stratifying by service population and by presence of a treatment plant could make the estimated needs more precise and accurate, but it still underperformed compared to Method 4, which had stratification of system size but also sampled all wastewater systems serving more than 100,000 people. In Method 9, we combine these two sampling strategies by stratifying the systems based on service population and presence of treatment plants, but by ensuring fully sampling all wastewater systems serving more than 100,000 people. The table below displays the sampling size of each of the 7 strata. The total sample size is 141.

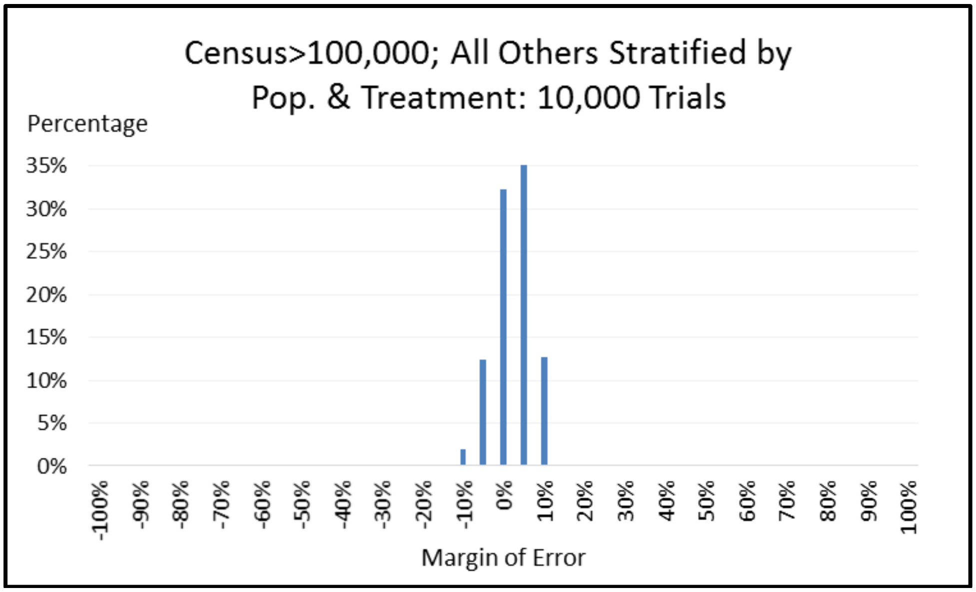
	Treatment	Non-Treatment
1-5,000	52 (out of 114)	29 (out of 41)
5,001-20,000	26 (out of 36)	2 (out of 2)
20,001-100,000	18 (out of 23)	5 (out of 5)
>100,000	All 9 systems are sampled	

This method is repeated 10,000 times, each time by drawing new samples of wastewater systems in four strata, but always including the 7 systems without treatment plants that serve 5,001 – 100,000 people and all 9 systems serving more than 100,000 people. The following tables and graphs display the descriptive statistics of the 10,000 estimates of the statewide wastewater infrastructure needs and their percentage errors (or difference) from the “true needs” figure.

	Percentage Error	Total Estimated Need
Mean	-0.01%	$4,966,204,871
Standard Deviation	4.48%	$222,605,086
Mean Standard Error	0.05%	$2,226,051
Skewness	-0.366	-0.366
Minimum	-16.15%	$4,164,634,537
First Quartile	-3.08%	$4,814,078,007
Median	0.41%	$4,987,074,160
Third Quartile	3.27%	$5,129,249,837
Maximum	11.45%	$5,535,668,984

		Average System Need	Total Statewide Need
1-5,000	Treatment	$3,698,896	$421,674,159
	Non-Treatment	$1,420,826	$58,253,865
5,001-20,000	Treatment	$11,607,702	$417,877,272
	Non-Treatment	$3,272,699	$6,545,398
20,001-100,000	Treatment	$52,070,738	$1,197,626,970
	Non-Treatment	$89,203,997	$446,019,985
>100,000	Census	$268,689,691	$2,418,207,222





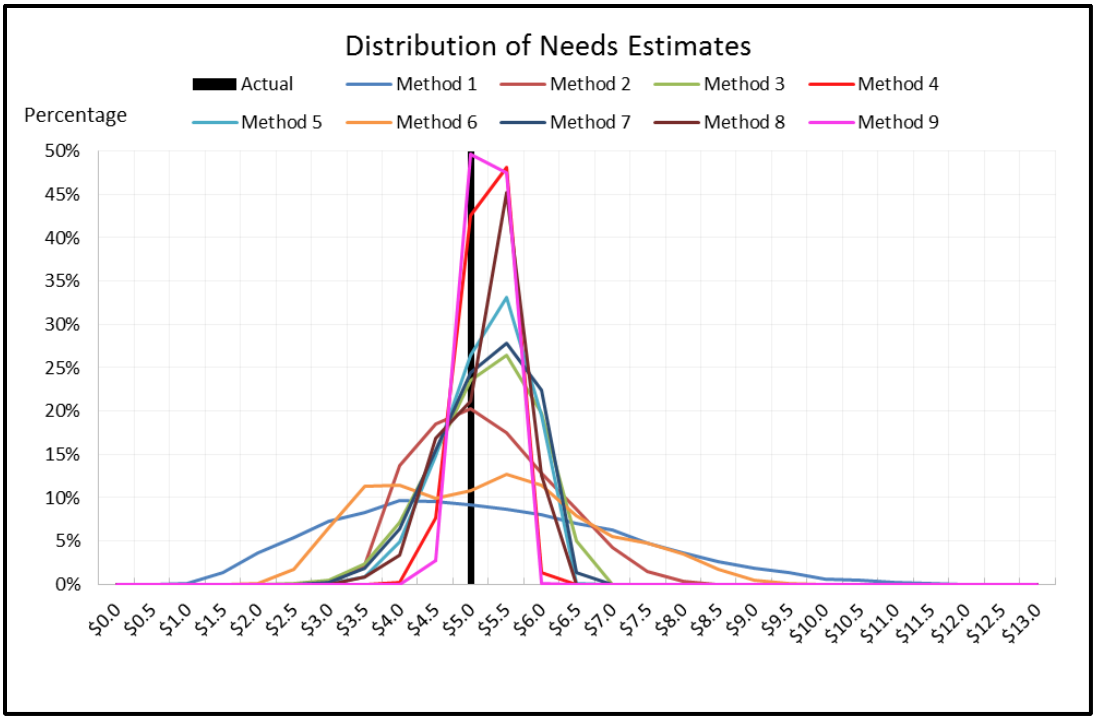
The median error for this method was +0.41% percent with a standard deviation of 4.48 percent. The maximum underestimation is around 16 percent and the maximum overestimation is slightly above 11 percent. By ensuring that all systems above population of 100,000 are included and by stratifying the sample by population and treatment, the needs estimate has significantly improved, both in terms of accuracy (smallest median error) and in precision (smallest standard deviation of the errors). The range between the minimum and maximum error is the smallest, between -16 percent and 11 percent, and centers very closely around 0. The histograms show how closely clustered all 10,000 samples are around an error of 0 percent.

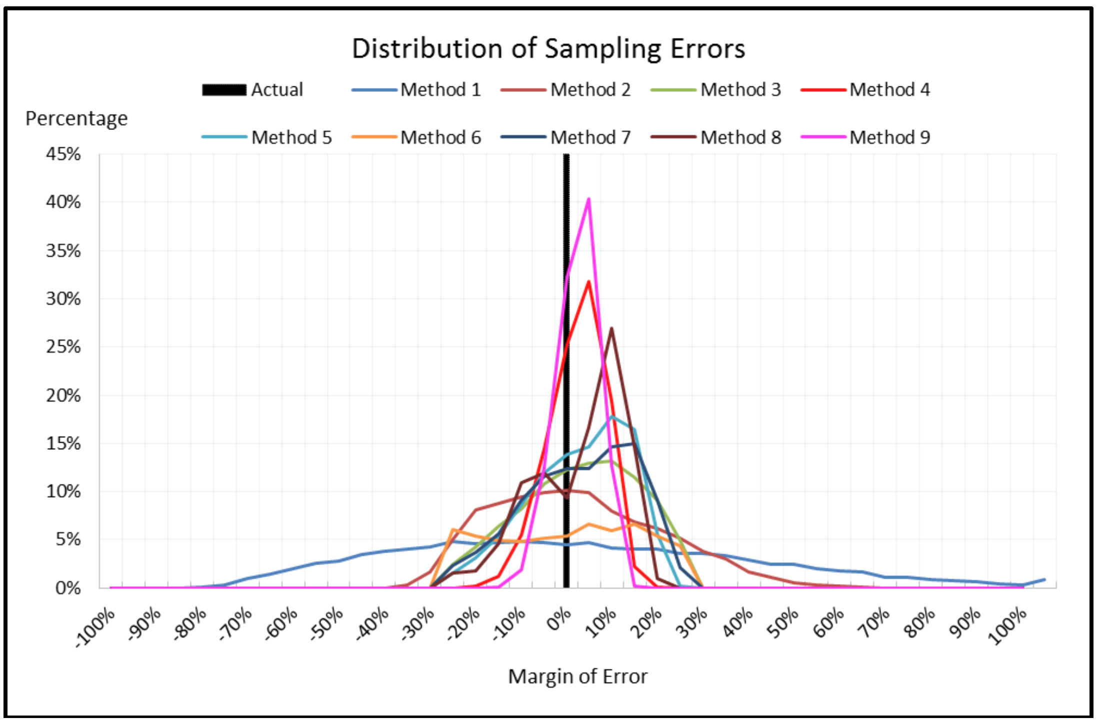
We determine that Method 9 is the most accurate and the most precise method assessed in this report.

Comparing all Nine Methods

The table below displays the summary statistics of the errors (or difference) between the nine different methods' estimates of the statewide wastewater infrastructure needs from the "true needs" figure. The table is arranged from a combination of higher accuracy and/or higher precision to lower accuracy and/or precision. The graphs below display the distribution of needs estimates, and their errors from the "true needs", as determined from the 10,000 runs of each methodology. The narrower the distribution, the more precise the estimates, and the closer the center is to the black line, the more accurate the estimates.

	Method								
	# 9	# 4	# 7	# 5	# 3	# 2	# 8	# 6	# 1
Sample Size	141	118	138	128	114	76	151	113	68
Median Error	0.41%	0.59%	1.38%	1.65%	1.13%	-1.61%	3.27%	-0.86%	-4.11%
Standard Deviation	4.48%	6.12%	12.98%	11.35%	14.18%	18.38%	10.35%	29.65%	39.48%
Average Estimated Statewide Need	$4.966 B	$4.962 B	$4.962 B	$4.971 B	$4.964 B	$4.969 B	$4.971 B	$4.949 B	$4.932 B





Method 9 produces the most accurate and the most precise estimates of needs, compared to all other methods, as demonstrated by the lowest median error and standard deviation of the errors, respectively. Method 4 produces the second most accurate and second most precise estimates of needs. All other methods produce results that are significantly less accurate or significantly less precise than Methods 9 and 4.

Conclusion

Every four years the EPA conducts the Clean Water Needs Survey in order to estimate the capital infrastructure needs of wastewater utilities. Currently, the EPA requires the State to collect capital infrastructure needs data from as many systems as possible (a census). This report explores whether nine different sampling methods, which would require fewer administrative resources, would provide an estimate of the statewide wastewater capital infrastructure needs that is as accurate and as precise as a census of needs.

We established a universe of 230 wastewater systems in the state since their capital infrastructure needs are individually identified and reported in the 2012 CWNS database. We applied the nine different sampling strategies, calculating the needs projected from the samples (using weights within each stratum), and used Monte Carlo simulations to select 10,000 different samples for each method. We compared the needs estimates to the “true needs” of the 230 CWNS-surveyed wastewater systems’ needs to determine which was the most accurate (a median statewide needs estimate close to the “true needs”) and the most precise (a narrow distribution around the median estimate of statewide needs).

Out of the nine methods, Method 9 (collecting capital infrastructure needs data from all systems serving a population of greater than 100,000 and stratifying all other systems by service population and whether or not the system has a treatment plant) produced the most accurate and the most precise estimates of needs. Method 4 (collecting capital infrastructure needs data from all systems serving a population of greater than 100,000 and stratifying all other systems by service population) produced the second most accurate and the second most precise estimates of needs. Both options provide a better choice for estimating the state’s needs than attempting a census of needs of all wastewater systems, as currently conducted following EPA’s CWNS methodology.

In weighing whether Method 4 or Method 9 is more feasible to implement, there are a few factors to consider. First, Method 9 is more accurate and more precise. However, Method 9 requires knowledge of whether or not wastewater systems include treatment plants, before sampling and data collection begins. If DWI already has access to a centralized, and complete, database of all wastewater systems and their characteristics (i.e. service population and treatment plant status), meaning that the administrative burden would not be much higher than implementing Method 4, then Method 9 is the superior choice. If treatment plant status is not readily available for all wastewater systems, then Method 4 may be preferable because the loss in accuracy and precision is small.

Second, although Method 9 is more accurate and precise than Method 4, Method 9 may simply be more accurate because it samples 23 more systems than Method 4. Because Method 9 divides the systems into more groups than Method 4, each group is smaller. When sampling smaller groups, the sample size must be larger to achieve the level of precision necessary (95% confidence level with margin of error of ±10 %). Thus, Method 9, which has more, smaller groups, has a larger total sample size than Method 4. It is unclear whether Method 9 is better than Method 4 because it is simply a more accurate method or because it is sampling more systems, or both.

This additional sampling in Method 9 would add administrative burden to DWI. DWI would have to decide if slightly more accuracy and precision would be worth the extra burden. In this report, out of 230 wastewater systems, the extra burden is sampling 23 more systems in Method 9 than in Method 4. Out of the full 417 wastewater systems, the extra burden might rise to sampling an additional 40 systems statewide, requiring the needs data collection of 40 more systems.

In the future, if DWI is given the choice to sample wastewater systems rather than conduct a census, this report shows that a census of the largest wastewater systems and stratified sampling of small and medium sized systems will provide equally accurate and precise estimates compared to a census of all wastewater systems, while lowering administrative burden. Sampling would provide administrative capacity to focus on improving the quality and accuracy of the needs estimates for the sampled systems (e.g. using cost curves to estimate needs for projects that are left out of capital improvement plans, using a consistent timeframe of analysis for all sampled systems instead of a mix of 5-year and 20-year timeframes as currently documented, etc.) instead of on obtaining needs estimates for more and more systems.

Appendix

To demonstrate what sample sizes would be needed for the preferred sampling methods, the table below is an approximation of the sample sizes needed for each stratum for Method 4.

The EFC identified 417 active wastewater systems in the state. The service population data for 409 of these wastewater systems is known primarily from the 2012 CWNS's "present service population receiving collection service". Where that data was missing, the municipal population estimate of a wastewater service area was determined from the Census Bureau and used as a surrogate. If neither of those data were available, the equivalent water system's service population estimate from SDWIS (downloaded in January 2016) was used. The 8 wastewater systems for which we currently lack service population data are:

1) Avery Creek Sanitary District
2) Davidson County
3) Efland
4) Henderson County
5) Lower Cape Fear Water & Sewer Authority
6) Maggie Valley
7) Proctorville
8) Southeast Brunswick Sanitary District

In order to achieve a 95% confidence level and ±10 % confidence interval in the statewide needs estimates of the 417 (or, actually, 409) wastewater systems in North Carolina, the following sample sizes would be required for Method 4's strata. A total sample size of 152 would be required. Thus, DWI would need to determine the capital needs data for 152 wastewater systems statewide, instead of attempting to collect data on all 417 systems (or even the 230 systems that North Carolina achieved in collecting data from in 2012). This demonstrates the value of statistical sampling. For less effort in data collection, the State could nearly as accurately estimate the needs of the state's wastewater systems. In fact, the State could produce even more accurate results, by spending more time checking the quality of the needs data as documented in the systems' CIPs and other documentation.

	Number of Systems	**Sample Size**	**Weight**
1-5,000	287	72	287/72
5,001-20,000	69	40	69/40
20,001-100,000	43	30	43/30
>100,000	10	10	10/10