Statistical Sampling Approach for Initial and Follow-Up BMP Verification

Purpose

This document provides a statistics-based approach for selecting sites to inspect for verification that BMPs are on the ground (or otherwise continue to be implemented) and performing as expected based on engineering specifications or other applicable criteria. Verification on a BMP-by-BMP basis is emphasized here to both simplify the approach and reflect the need for practical methods to address this large undertaking.

The probabilistic sampling approach described here can be used for both initial and follow-up BMP verification. While the verification guidance (*Chesapeake Bay Program Partnership Agriculture Workgroup's Agricultural BMP Verification Guidance*, as amended) calls for 100% verification of the initial identification of annual or multi-year structural BMPs and plan implementation by trained and certified technical field staff or engineers for most practices, it does allow for statistical sub-sampling to verify single-year BMPs such as tillage practices. The guidance also states that for follow-up BMP verification, states may propose using a sub-sampling approach with documentation as an alternative strategy for review and approval.

Selection of appropriate verification methods at sites selected using this approach is addressed in the agricultural BMP verification guidance (*Chesapeake Bay Program Partnership Agriculture Workgroup's Agricultural BMP Verification Guidance*, as amended). Regardless of the sampling approach used initially for agricultural BMP verification, states should do a post-evaluation of the results and process, updating as necessary.

Background

The need for verification that BMPs are implemented properly and remain functional is documented in the agricultural BMP verification guidance. That guidance also provides information on defining and categorizing agricultural BMPs, defining implementation mechanisms for agricultural BMPs, agricultural BMP verification methods and priorities, and how to develop an agricultural practice verification protocol. In addition, it provides streamlined guidance and an overview of the default verification levels for agricultural BMP verification.

This document supplements the agricultural BMP verification guidance by providing specific information on a statistically-based sampling approach that can be used as part of state efforts to meet verification requirements. The measure of choice for this approach is the proportion (percentage) of implemented BMPs (1) still in place or (2) still performing in accordance with expectations. The approach described here addresses how to compute the sample size necessary to estimate these proportions (i.e., "p" or proportion of "Yes" responses and "q" or proportion of "No" responses) with the desired degree of confidence and a specified acceptable error (±d%) using simple random sampling. No hypothesis testing, comparison of proportions, or trend analysis is considered.

Probabilistic Sampling

Overview

Probabilistic approaches are appropriate for ground verification of agricultural BMPs because they can yield accurate information without having to visit each site. In a probabilistic approach, individuals are randomly selected from the entire group. The selected individuals are evaluated, and the results from the individuals provide an unbiased assessment about the entire group. Applying the results from randomly selected individuals to the entire group is *statistical inference*. Statistical inference enables one to determine, in terms of probability, for example, the percentage of implemented multi-year BMPs that are still in place without visiting every site.

The group about which inferences are made is the population or *target population*, which consists of *population units*. The *sample population* is the set of population units that are directly available for measurement. Statistical inferences can be made only about the target population available for sampling. For example, if only a certain class of BMPs can be ground verified (e.g., cost-shared BMPs), then inferences cannot be made about other classes of BMPs that could not be ground verified (e.g., voluntarily implemented BMPs with no cost-share). States will need to consider carefully how they define their population units for each BMP. See "Defining Population Units" for addition information regarding this very important task.

The most common types of sampling that should be used are either simple random sampling or stratified random sampling. *Simple random sampling* is the most elementary type of sampling. Each unit of the target population has an equal chance of being selected. This type of sampling is appropriate when there are no major trends, cycles, or patterns in the target population. If the pattern of BMP presence or performance is expected to be uniform across the geographic area of interest (e.g., state), simple random sampling is appropriate to estimate the proportion of BMP presence or performance. If, however, implementation is homogeneous only within certain categories (e.g., region of state, cost-shared vs. non-cost-shared), stratified random sampling should be used. See "Sample Size Calculation with Simple Random Sampling" for additional details.

In *stratified random sampling*, the target population is divided into groups called strata for the purpose of obtaining a better estimate of the mean or total for the entire population. Simple random sampling is then used within each stratum. Stratification involves the use of categorical variables to group observations into more units (e.g., cost-shared vs. non-cost-shared), thereby reducing the variability of observations within each unit. In general, a larger number of samples should be taken in a stratum if the stratum is more variable, larger, or less costly to sample than other strata. See "Stratified Sampling" for additional information.

If the state believes that there will be a difference between two or more subsets of the sites, the sites can first be stratified into these subsets and a random sample taken within each subset. The goal of stratification is to increase the accuracy of the estimated mean values over what could have been obtained using simple random sampling of the entire population. The method makes use of prior information to divide the target population into subgroups that are internally homogeneous. There are a number of ways to "select" sites to be certain that important information will not be lost, or that results will not be misrepresented. One current approach is Generalized Random Tessellation Stratified (GRTS) survey design (Stevens and Olsen 2004).

Sample Size Calculation with Simple Random Sampling

The following are data requirements for the sample size (n) calculations described in this document:

- An initial estimate of both the percent of BMPs still in place and the percent of BMPs still performing as expected. This can be based on previous studies or assumed to be 50% (p=0.5) for a conservative (high) estimate of sample size.
- An allowable error (e.g. ±5% or 0.05). This error (d) can be different for different BMPs based on considerations of BMP importance, risk of BMP abandonment, failure, cost, or other factors.
- A confidence level (e.g., 90% or α=0.10). This is used to determine the 2-sided Z score from the standard normal distribution ($Z_{1-\alpha/2}$), e.g., $Z_{1-\alpha/2}$ is equal to 1.645 for $\alpha = 0.10$. For example, an α=0.10 indicates that the actual proportion of BMPs still in place has a 10 percent chance of being outside the allowable error or calculated confidence interval.
- An estimate of the total population (N) from which the sample is taken (e.g., how many BMPs were installed). This can be based on records of BMP implementation.

In simple random sampling, we presume that the sample population is relatively homogeneous and we would not expect a difference in sampling costs or variability. If the cost or variability of any group within the sample population were different, it might be more appropriate to consider a stratified random sampling approach.

To estimate the proportion of BMPs still in place or still performing as expected (p), such that the allowable error, *d*, meets the study precision requirements (i.e., the true proportion lies between *p-d* and *p+d* with a *1-α* confidence level), a preliminary estimate of sample size (n_0) can be computed with the following equation assuming a large population from which to sample (Snedecor and Cochran, 1980):

$$n_o = \frac{(Z_{1-\alpha/2})^2 pq}{d^2} \qquad (1)$$

In many applications, the number of population units in the sample population (*N*) is large in comparison to the population units sampled (*n*) and the *finite population correction term* (*1-φ*) can be ignored. However, depending on the number of units (e.g., expensive or unique BMPs) in a particular population, *N* can become quite small. *N* is determined by the definition of the sample population and the corresponding population units. If φ is greater than 0.1, the finite population correction factor should not be ignored (Cochran, 1977). Thus, the final sample size (n) can be estimated as (Snedecor and Cochran, 1980)

$$n = \begin{cases} \dfrac{n_0}{1+\varphi} & for\ \varphi > 0.1 \\ n_o & otherwise \end{cases} \qquad (2)$$

where φ is equal to n_o/N.

<u>Terms:</u>

N = total number of population units in sample population

n = number of samples

p = proportion of “yes” responses

q = proportion of "no" responses (i.e., 1-p)

n_0 = preliminary estimate of sample size

φ = n_0/N unless otherwise stated

$Z_{1-\alpha/2}$ = value corresponding to cumulative area of 1-α/2 using the normal distribution

d = allowable error

Practical Sampling Considerations

The best sampling approach will be one that meets statistical objectives and can be performed with maximum ease at minimum cost. Success requires that the information to be used in the equation described above is unambiguous and obtainable within logistical, programmatic, and budgetary constraints.

Defining Population Units

Population units should be defined in a manner that makes enumeration simple. The most promising options for population units are structures (e.g., lagoons), contracts, and plans (e.g., nutrient management plans). States should have access to counts of these population units through federal or state permit programs (e.g., CAFO), federal/state/local cost-share programs, or other sources. In some cases, counts or a portion of counts may need to be obtained from private-sector sources (e.g., nutrient management plans). The use of acreage as a population unit for the purposes of this sampling approach is not considered although acreage might be a useful variable to stratify BMPs (see "Stratified Sampling"). Acreage of practices (e.g., cover crops) inspected through a sampling effort based on contracts can be recorded, however, to provide an additional measure of the extent to which existing practices were inspected. For example, A% of contracts that include cover crops were sampled, covering a total B acres, or C% of existing cover crop acreage in the state.

States will need to choose population units that make the most sense for those BMPs they verify. Structural BMPs, for example, could be enumerated on the basis of actual structural units or contracts with the structure. If contracts are used as the population unit it is recommended that the total number of structural BMP units inspected on the sampled farms is recorded as well (e.g., if contracts can include more than one structure).

Stratified Sampling

Because some BMPs provide a greater pollutant load reduction than others, states may want to place priority on verification of those BMPs. If, for example, nutrient management plans (NMPs) have yielded the greatest nitrogen and phosphorus load reductions, it might be appropriate to emphasize these practices in the BMP verification program to provide results with better precision. For example, a smaller confidence interval (e.g., ±5%) and greater confidence level (e.g., 95%) might be appropriate for these BMPs. Less important BMPs, with respect to nutrient reduction, could be verified with a larger confidence interval (e.g., ±15%) and/or lower confidence level (e.g., 80%).

Alternatively, if state reports have indicated that livestock operations, for example, yield a greater load reduction than cropland farms (or vice versa) for a particular BMP, the state may want to use a stratified random sampling approach. A separate population for livestock operations and cropland farms would be developed for the BMP, with perhaps even a different confidence interval or confidence level applied to the two strata. The intent of this approach would be to provide the best verification data on a targeted basis within the resource constraints of the state. The same logic would apply to stratification by geographic region, BMP delivery program (e.g., permits, cost-share, voluntary), farm size (e.g., large vs. small), or risk (e.g., BMPs most likely to be abandoned or implemented poorly vs. BMPs that are more reliably implemented and maintained).

Grouping

If the count for a specific BMP is so low that it would be difficult to achieve a reasonably precise estimate of verification via sampling, a state may consider combining similar BMPs to increase the number of population units and increase the precision of the verification estimate. Similarity of BMPs could be judged on the basis of nutrient reduction credits provided by the Bay model. For example, if BMP A is credited with a 10% reduction in nitrogen load and BMP B is credited with a 12% reduction in nitrogen load (per unit applied), it may be reasonable to combine the two BMPs for the purpose of verification. This approach would be most appropriate for BMPs that account for a smaller share of the state's load reductions attributed to agricultural BMPs.

Field Verification Methods

States will need to establish field protocols that address the type of information to be collected and consistency between different field technicians or groups collecting the data. Specific verification methods and the need for quality assurance procedures are discussed in the agricultural BMP verification guidance. Essential to the statistical approach described in this document is determination and documentation of how "yes" and "no" responses will be assigned for the two basic questions:

- Is the BMP there?
- Is the BMP functioning properly?

States may have existing verification programs that go beyond simple yes/no determinations. For example, a state may have a third, gray area response between yes and no indicating that the BMP is partially functional or could be functional after tweaking by the landowner. This may be very important information for purposes other than verification using this statistical approach, but the data will need to be reduced down to yes/no to apply the method described here. A simple approach to reducing data down to yes/no responses is that anything not "yes" is "no." Using this approach, BMPs checked off as "gray area" BMPs would be added to the "no" tally.

States should consider performing initial field testing as part of their overall plan for agricultural BMP verification. This will help identify issues that can be resolved before the program is launched.

Timeframe for Sampling

Field inspections should be scheduled to provide the best opportunity to observe the features of a BMP that best indicate its presence and whether it is functioning properly. Cover crops, for example, may need to be observed both at planting and later to determine if seeds have germinated and cover has been established. The number of sites to be examined would remain the same, but the number of site visits would double in this case. States will need to consider when each BMP should be examined to establish a cost-efficient inspection schedule that can be achieved with existing resources.

Level of Effort

Resources committed to verification will most likely come from resources that could be used for other purposes such as technical and financial assistance for BMP implementation. Scheduling of staff activities will be an essential element to ensure that verification and other program functions are carried out successfully. The efficiency with which staff are deployed may be increased if states can find opportunities to piggyback verification work with other tasks while visiting individual farms. The establishment of standard operating procedures for verification site visits, creative use of modern technology, and other innovative approaches may help reduce the time required for inspections and the recording and management of verification data.

Application to Chesapeake Bay Program

There are currently 47 agricultural BMPs and interim BMPs subject to verification under the Chesapeake Bay Program, and this number will increase over time. States may track even more BMPs before having them translated into BMPs recognized by the Bay model. While there may be interest in designing a single, comprehensive sampling approach that addresses all BMPs that must be verified at specified levels of precision and confidence, such an approach is not recommended because it might become logistically impractical. Keep it simple.

A simple approach to sampling is to:

1. Estimate sample sizes for the priority BMPs,
2. Choose the largest "n" value from the set of priority BMPs,
3. Randomly select the farms to inspect for the priority BMPs,
4. Check records for the non-priority BMPs at the selected farms to determine the respective "n" values for non-priority BMPs,
5. Estimate confidence intervals for the non-priority BMPs based on the "n" values
6. Do either:
    - Increase random sample size for priority BMPs as needed to reach suitable confidence intervals for the non-priority BMPs and repeat steps 3-5 until a suitable confidence interval is reached for all BMPs of interest, or
    - Develop a separate sampling approach for non-priority BMPs by carrying out steps 1-3 for the non-priority BMPs. This creates two sampling approaches, but there may be overlap on sites visited.

This approach is illustrated with an example featuring five priority BMPs (Table 1) and five non-priority BMPs (Table 2) that must be verified by the state. Equations 1 and 2 are applied to the data in Table 1 to estimate sample sizes required for each priority BMP.

Table 1. Example: Priority agricultural BMPs for verification.

BMP	Population Unit	N	d	α	P (a priori)	n	% Sampled
Nutrient Management Plans	plan	350	.05	.10	.70	139	40
Cover Crops	contract	750	.05	.10	.65	186	25
Conservation Tillage	contract	2,000	.05	.10	.90	98	5
Prescribed Grazing	contract	155	.05	.10	.85	74	48
Forest Buffers	contract	900	.05	.10	.90	89	10

In this case, the state would need to inspect 186 farms to satisfy the precision and confidence level requirements for cover crops (Table 1). The state would then randomly select 186 farms from the set of farms with contracts including cover crops. Next, the state would check the contracts for those 186 farms to see if they also included nutrient management plans, conservation tillage, prescribed grazing, or forest buffers. For illustrative purposes, assume that the state found that the 186 farms selected based on cover crop contracts had the following counts for the other four priority BMPs:

- Nutrient Management Plans: 145 plans
- Conservation Tillage: 132 contracts
- Prescribed Grazing: 55 contracts
- Forest Buffers: 93 contracts

With the exception of prescribed grazing, sample sizes are also adequate for the other four priority BMPs. A sample size of 55 for prescribed grazing would yield a confidence interval of ±7% at α=.10.

The state can now choose to:

- Accept the slightly larger confidence interval for prescribed grazing, or
- Increase the sample size for cover crops and see if the prescribed grazing “n” value reaches the target of 74 (this would likely require an increase of at least another 60 farms based on the ratio of prescribed grazing to cover crop contracts), or
- Randomly select an additional 19 sites with prescribed grazing contracts from the 100 (155-55) prescribed grazing contract sites not captured in the cover crops sample. The total sample size would now be 205, a slight over-sampling for cover crops.

Assuming the state decides to add 19 sites for prescribed grazing contracts, the state now estimates the required sample sizes for non-priority BMPs, assuming a larger confidence interval (d=.10) and same confidence level (α=.10).

Equations 1 and 2 are also applied to the data in Table 2 to estimate sample sizes needed for each non-priority BMP. Note that the value of d is greater than used for Table 1 while the value for α is kept at 0.10. These choices and those made for Table 1 are judgment calls that the state must make.

Table 2. Example: Non-priority agricultural BMPs for verification.

BMP	Population Unit	N	d	α	P (a priori)	n	% Sampled
Land Retirement	contract	65	.10	.10	.90	19	29
Barnyard Runoff Control	contract	125	.10	.10	.95	12	10
Poultry Phytase	contract	475	.10	.10	.95	13	3
Crop Irrigation Management	contract	33	.10	.10	.85	17	52

Reviewing the contracts for the 205 farms selected based on cover crop and prescribed grazing contracts yielded the following counts for the non-priority BMPs:

- Land Retirement: 47 plans
- Barnyard Runoff Control: 15 contracts
- Poultry Phytase: 2 contracts
- Crop Irrigation Management: 47 contracts

Comparing these numbers with the results in Table 2 it can be seen that in this case all but poultry phytase would be adequately sampled. The simplest approach at this point would be to randomly select 11 additional contracts (13-2) from the 473 (475-2) poultry phytase contracts not captured in the cover crops/prescribed grazing sample, yielding 216 farms to inspect to meet statistical requirements for all tracked BMPs included in this example.

Currently, we do not have any information to suggest that selecting BMPs in this way (i.e., based on largest n value for priority BMPs) would result in a biased sampling of other BMPs. However, it should be an issue that is discussed within states based on knowledge of BMP implementation patterns.

Generalized Example

By executing Equations 1 and 2 over a wide range of scenarios we are able to construct generalized tables that indicate appropriate sample sizes within the established constraints. This begins with forming a precision statement that includes an allowable error term, ±d, and a confidence level. For example, a state may want to estimate the percentage of manure sheds passing the verification process to within ±10% at the 95% confidence level. Here is where the state might think about identifying different goals for different types of programs or BMPs. For example, some practices might be of a higher or lower importance to the Bay model in terms of loading while other practices might be of higher or lower risk of meeting the implementation requirements.

The state would also want to use a priori knowledge about the likely proportion of "yes" responses. One way to factor in this knowledge might be to establish a few categories or levels of expected implementation. For example, states may choose to set an "excellent" level of expected maintenance at 85%. Similarly, a 70% level could be set for "good," and 50% could be used if no information is available. These would essentially be the starting point assumptions of p to be used in equation 1. We can then combine these levels of BMP maintenance with a few choices of allowable error and confidence levels. In this example, we chose allowable error values of ±5, 10, 15, 20 and 25 percent and confidence levels of 90 and 95 percent.

Table 3 shows the results of those calculations. The top panel is for a 95% confidence level and the bottom panel is for 90% confidence level. The left-most columns show the expected level of BMP maintenance and allowable error, respectively. The Large N column represents the sample size without correction for finite populations; and the remaining six columns represent the adjusted sample sizes for a variety of population sizes. For example, to estimate the proportion of 200 BMPs successfully passing through the validation process assuming a 90% confidence level, assuming a likely percentage of BMPs equal to 85%, and an allowable error of ±10%, results in a sampling requirement of 30 as shown by the orange star. The blue bars represent a histogram of sample size.

Table 3. Generalized example: calculation of n.

95% Confidence Level

p		±d	Large N	100	200	600	1000	1,500	2,000
No Information	50%	5%	385	80	132	235	278	307	323
	50%	10%	97	50	66	84	89	92	93
	50%	15%	43	31	36	41	42	42	43
	50%	20%	25	20	23	24	25	25	25
	50%	25%	16	14	15	16	16	16	16
Good Maintenance	70%	5%	323	77	124	210	245	266	279
	70%	10%	81	45	58	72	75	77	78
	70%	15%	36	27	31	34	35	36	36
	70%	20%	21	18	20	21	21	21	21
	70%	25%	13	12	13	13	13	13	13
Excellent	85%	5%	196	67	99	148	164	174	179
	85%	10%	49	33	40	46	47	48	48
	85%	15%	22	19	20	22	22	22	22
	85%	20%	13	12	13	13	13	13	13
	85%	25%	8	8	8	8	8	8	8

90% Confidence Level

p		±d	Large N	100	200	600	1000	1,500	2,000
No Information	50%	5%	271	74	116	187	214	230	239
	50%	10%	68	41	51	62	64	66	66
	50%	15%	31	24	27	30	31	31	31
	50%	20%	17	15	16	17	17	17	17
	50%	25%	11	10	11	11	11	11	11
Good Maintenance	70%	5%	228	70	107	166	186	198	205
	70%	10%	57	37	45	53	54	55	56
	70%	15%	26	21	24	25	26	26	26
	70%	20%	15	14	14	15	15	15	15
	70%	25%	10	10	10	10	10	10	10
Excellent	85%	5%	138	58	82	113	122	127	130
	85%	10%	35	26	30	34	34	35	35
	85%	15%	16	14	15	16	16	16	16
	85%	20%	9	9	9	9	9	9	9
	85%	25%	6	6	6	6	6	6	6

Recognizing that sampling percentage can be the focal point for verification efforts, we can take Table 3 and divide through by the population size. Table 4 contains the same results as Table 3 but we display the results based on sampling percentage and use a 4-color stop light coding scheme. Sampling levels greater than 20% are coded black, 10 to 20% are coded red, 5-10% are coded yellow, and less than 5% are coded green. Table 4 therefore provides a quick visual assessment of sampling percentages needed to meet verification expectations. For example, where N is small (e.g., 100), nearly all sampling levels need to be greater than 20% for an allowable error of ±15% or smaller at the 90 and 95% confidence levels.

Table 4. Generalized example: calculation of sampling percentage.

95% Confidence Level

	p	±d	Large N	100	200	600	1000	1,500	2,000
No Information	50%	5%	385	80%	66%	39%	28%	20%	16%
	50%	10%	97	50%	33%	14%	9%	6%	5%
	50%	15%	43	31%	18%	7%	4%	3%	2%
	50%	20%	25	20%	12%	4%	3%	2%	1%
	50%	25%	16	14%	8%	3%	2%	1%	1%
Good Maintenance	70%	5%	323	77%	62%	35%	25%	18%	14%
	70%	10%	81	45%	29%	12%	8%	5%	4%
	70%	15%	36	27%	16%	6%	4%	2%	2%
	70%	20%	21	18%	10%	4%	2%	1%	1%
	70%	25%	13	12%	7%	2%	1%	1%	1%
Excellent	85%	5%	196	67%	50%	25%	16%	12%	9%
	85%	10%	49	33%	20%	8%	5%	3%	2%
	85%	15%	22	19%	10%	4%	2%	1%	1%
	85%	20%	13	12%	7%	2%	1%	1%	1%
	85%	25%	8	8%	4%	1%	1%	1%	0.4%

90% Confidence Level

	p	±d	Large N	100	200	600	1000	1,500	2,000
No Information	50%	5%	271	74%	58%	31%	21%	15%	12%
	50%	10%	68	41%	26%	10%	6%	4%	3%
	50%	15%	31	24%	14%	5%	3%	2%	2%
	50%	20%	17	15%	8%	3%	2%	1%	1%
	50%	25%	11	10%	6%	2%	1%	1%	1%
Good Maintenance	70%	5%	228	70%	54%	28%	19%	13%	10%
	70%	10%	57	37%	23%	9%	5%	4%	3%
	70%	15%	26	21%	12%	4%	3%	2%	1%
	70%	20%	15	14%	7%	3%	2%	1%	1%
	70%	25%	10	10%	5%	2%	1%	1%	1%
Excellent	85%	5%	138	58%	41%	19%	12%	8%	7%
	85%	10%	35	26%	15%	6%	3%	2%	2%
	85%	15%	16	14%	8%	3%	2%	1%	1%
	85%	20%	9	9%	5%	2%	1%	1%	0.5%
	85%	25%	6	6%	3%	1%	1%	0.4%	0.3%

Summary

A robust sampling effort begins with clear identification of the target population and enumeration of the population units (i.e., N). States will need to define the appropriate population unit for a large number of agricultural BMPs. Use of structural units, contracts, or plans is recommended.

Appropriate sample size for verification is driven by N, the desired margin of error (e.g., ±10%), the desired level of confidence (e.g., 95%), and the proportion of the sampled population that will have a positive result (p). States will need to apply their judgment in making decisions on the values for d and α. Improved precision (smaller d) or greater confidence (smaller α) will require increased sampling, while reduced sampling levels will result in lower confidence levels or increased allowable errors.

A priori knowledge is important in setting sample sizes; 50% is a conservative value with respect to sample size calculations. That is, absent knowledge of the likely proportion of positive responses, a p value of 0.5 is used in the calculation, resulting in a larger sample size than would result from using values of p greater or smaller than 0.5. It will benefit states to check for records on BMP compliance to use in the calculation of sample sizes.

The error associated with setting sample sizes for small populations can be large. In these cases it might be appropriate to group BMPs into classes rather than accept margin of errors that are too large to be helpful.

Field assessments of BMPs will require “yes” or “no” determinations for this statistical approach to be applicable. This may involve performing an additional step for states with existing verification approaches, but should not interfere with achievement of other objectives the state may have. States will need to strive for consistency among field staff making these assessments.

Finally, with limited resources states will need to seek optimal scheduling for field visits by considering appropriate timing to inspect different types of BMPs, multiple site visits for some BMPs, other staff commitments, and the potential for achieving multiple objectives during each site visit. Development and application of standard protocols for field assessments may also save time.

References

Cochran, W.G. 1977. Sampling Techniques. 3rd ed. John Wiley and Sons, New York, New York.

Snedecor, G.W. and W.G. Cochran. 1980. *Statistical methods*. 7th ed. The Iowa State University Press, Ames. Iowa.

Stevens, D. L., Jr. and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262-278.