CHAPTER 4. CONDUCTING THE EVALUATION

4.1 INTRODUCTION

This chapter addresses the process of determining whether forestry MMs or BMPs are being implemented and whether they are being implemented according to approved standards or specifications. Guidance is provided on what should be measured to assess MM and BMP implementation, as well as methods for collecting the information, including physical site evaluations, mail- and/or telephone-based surveys, personal interviews, and aerial reconnaissance and photography. Designing survey instruments to avoid error and rating MM and BMP implementation are also discussed.

Evaluation methods are separated into two types: Expert evaluations and self-evaluations. Expert evaluations are those in which actual field investigations are conducted by trained personnel to gather information on MM or BMP implementation. Self-evaluations are those in which answers to a predesigned questionnaire or survey are provided by harvesters and/or landowners associated with the survey site. Self-evaluations might also include examination of materials related to a harvest, such as harvest plans and records of violations of forestry regulations. Extreme caution should be exercised when using data from self-evaluations as the basis for assessing MM or BMP implementation since they are not typically reliable for this purpose. Each of these evaluation methods has advantages and disadvantages that should be considered before to deciding which one to use or in what combination to use them. Aerial reconnaissance and photography can be used to support either evaluation method.

Self-evaluations are useful for collecting information on landowner or harvester awareness of MMs or BMPs, dates of harvest, harvest site conditions, which MMs or BMPs were implemented, and whether the assistance of a professional forester was used. However, the type of or level of detail of information that can be obtained from self-evaluations might be inadequate to satisfy the objectives of a MM/BMP compliance survey. If this is the case, expert evaluations might be called for. Expert evaluations are necessary if information on MM/BMP implementation that is more detailed or more reliable than that that can be obtained with self-evaluations is required, such as an objective assessment of the adequacy of MM/BMP implementation, the degree to which site-specific factors (e.g., slope, soil type, or presence of a water body) influenced MM/BMP implementation, or the need for changes in standards and specifications for MM/BMP implementation. Sections 4.3 and 4.4 discuss expert evaluations and self-evaluations, respectively, in more detail.

Expert evaluations of implementation of forestry MMs or BMPs generally occur after the harvest has occurred (see ***Example***), and direct observation of the adequacy of implementation of many BMPs (e.g., preharvest planning, pesticide applications, or construction of temporary roads) might not be possible. However, evidence of proper BMP implementation is often present at harvest sites. For instance, evidence of excessive skidding in SMAs, vegetation kills due to pesticide use in SMAs, and poorly restored stream banks and stream beds where temporary stream crossings were located is often observable during site evaluations. Supplemental information on aspects of harvest operations that cannot be observed directly during site evaluations might also be obtained from self-evaluations.

Example ... Timing of site evaluations.

> *U.S. Forest Service, Southwest Region*: After logging to within approximately one year after the site is harvested (USDA, 1992).
> *South Carolina*: One year or less than after the site is harvested (Adams, 1994).
> *Florida*: Two years or less after the site is harvested (Vowell and Gilpin, 1994).

Aerial reconnaissance and photography is another means available for collecting information on harvests, though many of the MMs/BMPs employed for forestry are difficult if not impossible to identify on aerial photographs. For this reason, aerial reconnaissance and photography are most useful for identifying potential survey sites and monitoring harvest site regeneration, forest conditions, and some water quality conditions (e.g., sediment runoff, algal blooms). Aerial reconnaissance and photography are discussed in more detail in Section 4.5.

The general types of information obtainable with self-evaluations are listed in Table 4-1. Regardless of the approach(es) used, proper and thorough preparation for the evaluation is the key to success.

4.2 CHOICE OF VARIABLES

Once the objectives of the BMP compliance survey have been clearly defined, the most important factor in the assessment of MM or BMP implementation is the determination of which variable(s) to measure. A good variable provides a direct measure of how well a BMP was implemented. Individual variables should provide measures of different factors related to BMP implementation. The best variables are those which are measures of the adequacy of MM or BMP implementation and are based on quantifiable expressions of conformance with state standards and specifications. As the variables used become less directly related to actual MM or BMP implementation, their accuracy as measures of BMP implementation decreases.

Examples of useful variables include width of streamside management areas, slope of landing areas, and size of culverts, all of which would be expressed in terms of conformance with applicable state standards and specifications. Less useful variables measure factors that are related to BMP implementation but do not necessarily provide an accurate measure of their implementation. Examples of such variables include the number of miles of forest roads constructed and the number of preharvest plans submitted to the state forestry agency, department, or division (hereafter referred to as the state forestry authority). Although these variables relate to MM and BMP

Table 4-1. General types of information obtainable with self-evaluations and expert evaluations.

Information Obtainable from Self-Evaluations
Harvest applications and management plans associated with the harvest (e.g., preharvest, road, fire, forest chemical) might be available for review prior to site evaluations and can provide much background information, such as: • Type of ownership (private industrial, private nonindustrial, federal, other public) • Total acreage under management • Acres/board feet harvested • Surface water body types on harvest site • Soil type • Ecological characterization of harvested area (e.g., habitat type, significant wildlife) • Presence of critical wildlife habitat • Species harvested • Harvest/management history of the harvested site • Use of cable yarding or ground skidding • Chemicals (e.g., pesticides, fertilizers) applied • Dates of plan preparation and revisions • Locations of roads and road structures, SMAs, loading areas, etc. • MMs and BMPs applied during harvest • Map *Conversations with harvesters and landowners can be used to verify information obtained from applications or can yield supplemental information, such as:* • Dates of harvest • Ambient conditions during applications • Variations from harvest plan • Problems encountered during harvest • Types of equipment used during harvest and in SMAs • Timing, location, and rate of chemical applications
Information Requiring Expert Evaluations
Expert evaluations are necessary to verify information obtained from self-evaluations and records and to assess the actual adequacy of MM and BMP implementation. Expert evaluations are necessary to: • Assess design adequacy • Assess installation adequacy • Assess the appropriateness of operation methods and overall management • Confirm information obtained from self-evaluations

implementation, they provide no real information on whether the MMs and BMPs are actually being implemented or whether they are being implemented properly.

Variables generally will not directly relate to MM implementation since most forestry MMs are combinations of several BMPs. Measures of MM implementation, therefore, usually will be based on separate assessments of two or more BMPs, and the implementation of each BMP will be based on a unique set of variables. Some examples of BMPs related to EPA's Road Construction and Reconstruction Management Measure, variables for assessing compliance with the BMPs, and related standards and specifications that might be required by state forestry authorities are presented in Figure 4.1. Because harvesters choose to implement or not implement MMs/BMPs based on site-specific conditions, it is also appropriate to apply varying weights to the variables chosen to assess MM/BMP implementation to correspond to site-specific conditions. For example, variables related to slope factors might be de-emphasized—and other, more applicable variables emphasized more—on relatively flat harvest sites. Similarly, on a site with a water body, variables related to SMAs, sediment runoff, and chemical deposition (pesticide use, fertilizer use) might be emphasized over other variables to arrive at a site-specific rating of the adequacy of MM/BMP implementation.

The purpose for which the information collected during an MM or BMP implementation survey will be used is another important consideration when selecting variables. An implementation survey can serve many purposes beyond the primary purpose of assessing MM and BMP implementation. For instance, for its 1993 BMP compliance survey, the South Carolina Forestry Commission selected variables that enabled it to assess compliance with each of five categories of BMPs and overall compliance with BMPs. In addition, the Commission analyzed the effect of each of 16 additional variables on BMP compliance (see ***Example***). The purpose of the survey was not only to assess BMP implementation, but also to assess the relationship of various conditions to the level of BMP compliance (Adams, 1994).

Table 4-2 provides examples of useful and less useful variables for the assessment of implementation of the forestry MMs developed by EPA (USEPA, 1993a). The variables listed in the table are only examples, and local or regional conditions ultimately dictate which variables should be used.

4.3 EXPERT EVALUATIONS

4.3.1 Site Evaluations

Expert evaluations are the best way to collect reliable information on MM and BMP implementation. They involve a person or team of people visiting individual harvest sites and speaking with harvest operators and/or landowners to obtain information on MM and BMP implementation. For many MMs, assessing and verifying compliance requires a site visit and evaluation. The

Management Measure for Road Construction and Reconstruction

(1) **Follow preharvest planning (as described in the Preharvest Planning Management Measure) when constructing or reconstructing the roadway.**
(2) **Follow designs planned under the Preharvest Planning Management Measure for road surfacing and shaping.**
(3) **Install road drainage structures according to designs planned under the Preharvest Planning Management Measure and regional storm return period and installation specifications. Match these drainage structures with terrain features and with road surface and prism designs.**
(4) **Guard against the production of sediment when installing stream crossings.**
(5) **Protect surface waters from slash and debris material from roadway clearing.**
(6) **Use straw bales, silt fences, mulching, or other favorable practices on disturbed soils on unstable cuts, fills, etc.**
(7) **Avoid constructing new roads in SMAs to the extent practicable.**

Related BMPs, measurement variables, and standards and specifications:

Management Measure Practice	Potential Measurement Variables	Example Related Standards and Specifications
• Preplan skid trail and landing locations on stable soils and avoid steep gradients and areas that are landslide-prone or erosion-prone, or have poor drainage.	• Soil type or stability along skid trails and at landings. • Gradients along skid trails and at landings.	• Minimum soil stability for skid trails and landings. • Maximum slope for skid trails and landings.
• In moderately sloping terrain, plan for road grades of less than 10%, with an optimal grade between 3% and 5%. Vary road grades frequently to reduce culvert and road drainage ditch flows, road surface erosion, and concentrated culvert discharges.	• Categorization of terrain as flat, moderate, steep. • Road grade estimated over sections of road. • Steep terrain: Average distance between changes in grade.	• Maximum road grade for a given terrain slope. • Maximum distance between changes in grade on steep terrain. • Minimum/maximum distance between drainage features for a given terrain slope.
• Design roads and skid trails to follow the natural topography and contour, minimizing alteration of natural features.	• Natural slope of surrounding terrain. • Slope of skid trails and at landings.	• Maximum slope of skid trails for a given terrain slope.
• Design cut-and-fill slopes to be at stable angles, or less than the normal angle of repose, to minimize erosion and slope failure potential.	• Angle of cut-and-fill slopes. • Stability of soil type where cut-and-fill slopes have been installed.	• Maximum angle for cut-and-fill slopes.

Figure 4-1. Potential variables and examples of implementation standards and specifications that might be useful for evaluating compliance with the Road Construction and Reconstruction Management Measure.

Table 4-2. Examples of variables related to management measure implementation.

Management Measure	Useful Variables	Less Useful Variables	Appropriate Sampling Unit
Preharvest Planning	· Agreement between preharvest plan and harvest operation · Inclusion of all required elements in preharvest plan	· Number of preharvest plans developed/approved	· Harvest operation · Preharvest plan
Streamside Management Areas	· Width of SMAs · Leave trees in SMAs meet minimum requirements	· Presence of water body on harvest site · Number of stream crossings in SMA	· 100-ft stretch of SMA
Road Construction/ Reconstruction	· Compaction of fill materials adequate to prevent erosion · Culverts cross streams at right angles	· Miles of road constructed · Number of stream crossings installed	· Fill areas along forest roads · Stream crossings
Road Management	· Culverts free of obstructions · Temporary stream crossings removed	· Completion of road inspections · Number of temporary stream crossings removed	· Culverts · Forest road stream crossings
Timber Harvesting	· Proper slope at landings · Water bodies free of slash materials	· Acres harvested · Number of cable yarding operations	· Landings · 100 yd of stream adjacent to harvest site
Site Preparation and Forest Regeneration	· Adequate distribution of seedlings on prepared sites · Nonmechanical site preparation used in SMAs	· Method of site preparation · Acres revegetated	· 100-yd^2 plots · 100 yd of SMA
Fire Management	· Fire lines constructed to minimize erosion · Intense burning not conducted on steep slopes with high erosion potential	· Acres burned · Size of individual burn areas	· 100 yd of fire line · Burned areas
Regeneration of Disturbed Areas	· Minimum requirements for seedlings per acre met · Erosion-prone areas replanted	· Species planted · Acres revegetated	· Steeply sloped areas · 100-yd^2 plots
Forest Chemical Management	· Mixing and loading areas located away from surface waters · Pesticides applied in accordance with EPA and/or state requirements	· Pounds of chemical applied · Availability of spill contingency plan	· SMAs · Chemical mixing and loading areas · Harvest site
Wetlands Forest	· Any of above with respect to wetlands forest	· As above	· As above

Example ... Variables used by the South Carolina Forestry Commission during the 1993 BMP compliance survey. (Source: Adams, 1994)

Overall BMP compliance and compliance with each of five categories of BMPs were assessed:

- Road systems
- Road stream crossings
- Streamside management zones
- Log decks
- Harvesting operations

and for overall BMP compliance.

Sixteen additional variables were analyzed to determine their effect on BMP compliance:

- Presence of perennial streams
- Terrain type
- Percent slope
- Use of a professional forester
- Required compliance with BMPs
- Physiographic region
- Use of a sales contract
- Percent of site impacted
- Landowner category
- Familiarity of landowner with BMPs
- Logged under wet soil conditions
- Rutting severity
- Road construction applicability
- Soil drainage class
- Presence of jurisdictional wetlands
- Harvest size

following should be considered before expert evaluations are conducted:

- *Obtaining permission from the landowner*. Without proper authorization to visit a site from a landowner, the relationship between landowners and the state forestry authority, and any future regulatory or compliance action could be jeopardized.

- *The type(s) of expertise needed to assess proper implementation.* For some MMs, a team of trained personnel might be required at a site evaluation to determine whether MMs have been implemented properly.

- *The activities that should occur during a site evaluation.* This information is necessary for proper and complete preparation for the site visit, so that the evaluation can be completed in a single visit and at the proper time.

- *The method of rating the MMs/BMPs.* MM and BMP rating systems are discussed below.

- *Consistency among evaluation teams and among site evaluations.* Proper training and preparation of site evaluation team members are crucial to ensure accuracy and consistency.

- *The collection of information while at a site.* Information collection should be facilitated with preparation of data collection forms that include any necessary MM and BMP rating information needed by the evaluation team members.

- *The content and format of postevaluation discussions.* Site evaluation team members should bear in mind the value of postevaluation discussion among team members. Notes can be taken during the evaluation concerning any items that would benefit from group discussion.

Evaluation teams may consist of from a single person suitably trained in silvicultural site evaluation to a group of professionals with various expertise. The composition of an evaluation team will depend on the types of MMs or BMPs being evaluated. Potential team members could include:

- Forester
- State forestry agent
- Hydrologist
- Pesticide specialist
- Soil scientist
- Water quality expert

The composition of evaluation teams will vary from state to state according to the type of forestry practiced, available staff and other resources, and the geographic area covered by the evaluations. All team members should be familiar with the required MMs/BMPs, and each team should have a member who has previously participated in a site evaluation. This will ensure familiarity with the technical aspects of the MMs/BMPs that will be rated during the evaluation and the site evaluation process.

Each evaluation team member will have a different degree of familiarity with the MMs/BMPs and with state requirements. Training may be necessary, therefore, to bring all team members to the level of proficiency needed to conduct the site evaluations. State forestry agents should be familiar with forestry regulations, state BMP standards and specifications, and proper BMP implementation, and therefore are generally well qualified to teach these topics to evaluation team members who are less familiar with them. Forestry agents or other specialists who have participated in BMP implementation surveys should train evaluation team members about the actual conduct of site evaluations. This training might cover identification of BMPs implemented poorly in previous years, analysis of erosion potential, and other aspects of BMP implementation that contain a degree of subjectivity, as well as any standard methods for measurements to judge BMP implementation against state standards and specifications.

Alternatively, if only one or two individuals will be conducting site evaluations, their training in the various specialties, such as those listed above, necessary to evaluate the quality of MM/BMP implementation could be provided by a team of specialists who are familiar with forestry practices and nonpoint source pollution.

In the interest of consistency among the evaluations and among team members, it is advisable that one or more mock evaluations take place prior to visiting selected sample sites as well. These “practice sessions” provide team members with a relaxed atmosphere for familiarizing themselves with MMs and BMPs as they are when they have actually been implemented and under different harvest site conditions, gaining familiarity with the evaluation forms and the meanings of the terms and questions on them, and asking other team members questions in their areas of expertise. Mock evaluations are valuable for ensuring that all evaluators have a similar understanding of the intent of the questions, especially for questions whose responses involve a degree of subjectivity on the part of the evaluator.

Where site evaluation teams are composed of more than two or three people, it might be helpful to divide the various responsibilities for conducting the site evaluations among team members ahead of time to avoid confusion at the harvest site and to be certain that all tasks are completed but not duplicated. Having a spokesperson for the group who is responsible for communicating with the landowner or harvester—prior to the site evaluation, at the site evaluation if they are present, and afterward—might also be helpful. A state forestry agent is generally a good choice as spokesperson because he/she represents the state forestry authority. Newly formed evaluation teams might benefit most from a division of labor and selection of a team leader or team coordinator with experience with site evaluations who will be responsible for the quality of the site evaluations. Smaller teams might find that a division of responsibilities is not necessary, as might larger teams whose members have experience working together. If responsibilities are to be assigned, mock evaluations can be a good time to work out these details.

4.3.2 Rating Implementation of Management Measures and Best Management Practices

Many factors influence the implementation of MMs and BMPs, so it is sometimes necessary to use best professional judgment (BPJ) to rate their implementation and BPJ will almost always be necessary when rating the implementation of MMs or when rating overall BMP compliance at a harvest site. Site-specific factors such as soil type, slope, presence of a water body, and ground cover type affect the implementation of erosion and sediment control BMPs, for instance, and must be taken into account by evaluators when rating MM/BMP implementation. Implementation of MMs will often be based on implementation of more than one BMP, and this makes rating MM implementation similar to rating overall BMP compliance at a harvest site. Determining an overall rating involves grouping the ratings of implementation of individual BMPs into a single rating, which introduces more subjectivity than rating the implementation of individual BMPs based on standards and specifications. Choice of a rating system and rating terms, which are aspects of proper evaluation design, is therefore important in minimizing the level of subjectivity associated with overall BMP compliance and MM implementation ratings.

Individual BMPs, overall BMP compliance, and MMs can be rated using a binary approach (e.g., pass/fail, compliant/ noncompliant, or yes/no) or on a scale with more than two choices, such as 1 to 5 or 1 to 10 (where 1 is the worst) (see ***Example***). The simplest method of rating MM and BMP implementation is the use of a binary approach. Using a binary approach, either an

Example ... of a rating scale (Source: Rossman and Phillips, 1992). More examples are presented in Appendix B.

> Minnesota Division of Forestry uses this 5-choice rating scale for BMP implementation audits:
> 5 = Operation exceeds requirement of BMP
> 4 = Operation meets requirement of BMP
> 3 = Minor departure from BMP
> 2 = Major departure from BMP
> 1 = Gross neglect of BMP
>
> where:
>
> *Minor departure* is defined as "small in magnitude and localized,"*major departure* is defined as "significant magnitude or where the BMPs were consistently neglected" and *gross neglect* is defined as "potential risk to water resources was significant and there was no evidence that any attempt has been made by the operator to apply the BMP."

entire site or individual MMs or BMPs are rated as being in compliance or not in compliance with respect to specified criteria. Scale systems can take the form of ratings from poor to excellent, inadequate to adequate, low to high, 1 to 3, 1 to 5, and so forth.

Whatever form of scale is used, the factors that would individually or collectively qualify a site, MM, or BMP for one of the ratings should be clearly stated. The more choices that are added to the scale, the smaller and smaller the difference between them becomes and each must therefore be defined more specifically and accurately. This is especially important if different teams or individuals rate separate sites. Consistency among the ratings then depends on each team or individual evaluator knowing precisely what the criteria for each rating option mean. Clear and precise explanations of the rating scale can also help avoid or reduce disagreements among team members. This applies equally to a binary approach. The factors, individually or collectively, that would cause a site, MM, or BMP to be rated as not being in compliance with design specifications should be clearly stated on the evaluation form.

Rating sites or MMs/BMPs on a scale requires a greater degree of analysis by the evaluation team than does using a binary approach. Each higher number represents a better level of MM/BMP implementation and/or effectiveness. In effect, a binary rating approach is a scale with two choices; a scale of low, medium, and high (compliance) is a scale with three choices. Use of a scale system with more than two rating choices can provide more information to program managers than a binary rating approach, and this factor must be weighed against the greater complexity involved in using one. For instance, a survey that uses a scale of 1 to 5 might result in one MM with a rating of 1, five with a rating of 2, six with a rating of 3, eight with a rating of 4, and five with a rating of 5. Precise criteria would have to be developed to be able to ensure consistency within and between survey teams in rating the MMs, but the information that only 1 MM was poorly implemented, 11 were below standards, 13 met or were above standards, and 5 were implemented very well might be more valuable than the information that 18 MMs were found to be in compliance with design specifications, which is the only information that would be obtained with a binary rating approach.

If a rating system with more than two ratings is used to collect data, the data can be analyzed either by using the original rating data or by first transforming the data into a binomial (i.e., two-choice rating) system. For instance, ratings of 1 through 5 could be reduced to two ratings by grouping the 1s, 2s, and 3s together into one group (e.g., inadequate) and the 4s and 5s into a separate group (e.g., adequate). If this approach is used, it is best to retain the rating data for the detailed information it contains and to reduce the data to a binomial system only for the purpose of statistical analysis. Chapter 3, Section 3.5, contains information on the analysis of categorical data.

4.3.3 Rating Terms

The choice of rating terms used on the evaluation forms is an important factor in ensuring consistency and reducing bias, and the terms used to describe and define the rating options should be as objective as possible. For a rating system with a large number of options, the meanings of each option should be clearly defined. It is best to avoid using terms such as “major” and “minor” when describing erosion or pollution effects or deviations from prescribed MM/BMP implementation criteria because they might have different connotations for different evaluation team members. It is easier for an evaluation team to agree on meaning if options are described in terms of measurable criteria and examples are provided to clarify the intended meaning. It is also best not to use terms that carry negative connotations. Evaluators are less likely to rate something as having a “major deviation” from an implementation criterion, even if justified, because of the negative connotation carried by the term. Rather than using such a term, observable conditions or effects of the quality of implementation should be listed and specific ratings (e.g., 1-5 or compliant/noncompliant for the criterion) should be associated with the conditions or effects. For example, instead of rating culvert installation as having a “major deficiency,” a specific deficiency should be described and should have an associated rating ascribed to it (e.g., “Culvert as installed does not allow for fish passage = noncompliant”).

Evaluation team members will often have to take specific notes on sites, MMs, or BMPs during the evaluation, either to justify the ratings they have ascribed to variables or for discussion with other team members after the survey. When recording notes about the sites, MMs, or BMPs, evaluation team members should be as specific as the criteria for the ratings. A rating recorded as “MM deviates highly from implementation criteria” is highly subjective and loses specific meaning when read by anyone other than the person who wrote the note. Notes should therefore be as objective and specific as possible.

An overall site rating is useful for summarizing information in reports; for gaining a general idea of the level of compliance with MMs/BMPs, the likelihood that environmental concerns are warranted, and the need for operator training or education; and for conveying information to program managers, who will usually not be as familiar with the MMs or BMPs, the implementation criteria, or the technical specifications as evaluation team members. For the purposes of preserving the information contained in the original ratings of sites, MMs, or BMPs, however, these overall ratings should not replace the original data. Analysis of year-to-year variations in MM or BMP implementation, the factors involved in MM or BMP program implementation, and factors that could improve MM or BMP implementation and MM or BMP program success are possible only if the original, detailed site, MM, or BMP data are used. If each site receives an overall rating and MMs/BMPs are not rated individually, thorough notes on what was seen at the site to justify the overall rating should be kept for future reference and for comparison to future evaluations.

Approaches commonly used for determining final BMP implementation ratings include calculating a percentage based on individual BMP ratings, consensus, compilation of aggregate scores by an objective party, voting, and voting only where consensus on a site or MM/BMP rating cannot be reached. Of course, not all systems for arriving at final ratings are applicable to all circumstances, and the choice of one over others is largely a matter of personal preference.

4.3.4 Consistency Issues

Consistency among evaluators and between evaluations is important. Consistency is likely to be best if only one or two evaluators conduct the site evaluations and the same persons conduct all of the evaluations. If, for statistical purposes, many sites (e.g., 100 or more) need to be evaluated, use of only one or two evaluators might also be the most efficient approach. In this case, a team of evaluators might be useful for revisiting a subsample of the sites evaluated by the one to two persons for quality control purposes. Evaluation teams can also be useful for training the one or two persons who will conduct the site evaluations in their specialties as they relate to MM/BMP implementation and nonpoint source pollution.

If teams of evaluators conduct the evaluations, consistency can be achieved by keeping the membership of the teams constant. Differences of opinion, which are likely to arise among team members, can be settled through discussions held during evaluations, and the experience of team members who have done past evaluations can help guide decisions. Pre-evaluation training sessions, such as the mock evaluations discussed above, will help ensure that the first few site evaluations are not “learning” experiences to such an extent that sites must be revisited to ensure that they receive the same level of scrutiny as sites evaluated later.

If different sites are visited by different teams of evaluators or if individual evaluators are assigned to different sites, it is especially important that consistency be established before the evaluations are conducted. For best results, discussions among evaluators should be held periodically during the evaluations to discuss any potential problems. For instance, evaluators could visit some sites together at the beginning of the evaluations to promote consistency in ratings, followed by site evaluations conducted by individual evaluators. Then, after a few site or MM evaluations, evaluators could gather again to discuss results and to share any knowledge gained to ensure continued consistency.

As mentioned above, consistency can be established during mock evaluations held before the actual evaluations begin. These mock evaluations are excellent opportunities for evaluators to discuss the meaning of terms on rating forms, differences between rating criteria, and differences of opinion about proper MM/BMP implementation. A member of the evaluation team should be able to represent the state's position on the definition of terms and clarify areas of confusion.

Descriptions of MMs and BMPs should be detailed enough to support any ratings given to individual features and to the MM or BMP overall. Sketching a diagram of the MM or BMP helps identify design problems, promotes careful evaluation of all features, and provides a record of the MM or BMP for future reference. A diagram is also valuable when discussing the MM or BMP with the landowner or identifying features in need of improvement or alteration. Landowners can also use a copy of the diagram and evaluation when discussing their operations with state forestry agents. Photographs of MM or BMP features are valuable reference material and should be used whenever an evaluator feels that a written description or a diagram could be inadequate. Photographs of what constitutes both good and poor MM or BMP implementation are valuable for explanatory and educational purposes; for example, for presentations to managers and the public.

4.3.5 Postevaluation Onsite Activities

It is important to complete all pertinent tasks as soon as possible after the completion of a site evaluation to avoid extra work later and to reduce the chances of introducing error attributable to memory error or confusion. All evaluation forms for each site should be filled out completely before leaving the site. Information not filled in at the beginning of the evaluation can be obtained from the landowner if necessary. Any questions that evaluators had about the MMs/BMPs during the evaluation can be discussed, and notes written during the evaluation can be shared and used to help clarify details of the evaluation process and ratings. The opportunity to revisit the site will still exist if there are points that cannot be agreed upon among evaluation team members.

Also, while the evaluation team is still on site, the landowner should be informed about what will follow; for instance, whether he/she will receive a copy of the report, when to expect it, what the results means, and his/her responsibility in light of the evaluation, if any. Immediately following the evaluation is also an excellent time to discuss the findings with the landowner if he/she was not present during the evaluation.

4.4 SELF-EVALUATIONS

4.4.1 Methods

Self-evaluations, while often not a reliable source of MM or BMP implementation data, can be used to augment data collected through expert evaluations or in place of expert evaluations where the latter cannot be conducted. In some cases, state forestry authority staff might have been involved directly with a harvest and will be a source of useful information even if an expert evaluation is not conducted. Self-evaluations are an appropriate survey method for obtaining background information from harvesters and landowners.

Mail, telephone, and mail with telephone follow-up are common self-evaluation methods. Mail and telephone surveys are useful for collecting general information, such as the location of harvest operations, species harvested, methods used, and dates of harvest. Also, harvest application or notification records can provide useful background information, including any special conditions applied to the harvest by the state forestry authority. Recent advances in and increasing access to electronic means of communication (i.e., e-mail and the Internet) might make these viable survey instruments in the future.

Mail surveys with a telephone follow-up and/or site visit are an efficient method of collecting information. To ensure comparability of results, information collected as part of a self-evaluation—whether collected through the mail, over the phone, or during site visits—should be collected in a manner that does not favor one method over the others. Ideally, telephone follow-up and site visit interviews should consist of no more than reading the questions on the questionnaire, without providing any additional explanation or information that would not have been available to those who responded through the mail. This approach eliminates as much as possible any bias associated with the different means of collecting the information. Questionnaire design is discussed in Section 4.4.3.

It is important that the accuracy of information received through mail and phone surveys be checked. Inaccurate or incomplete responses to questions on mail and/or telephone surveys commonly result from survey respondents misinterpreting questions and thus providing misleading information, not including all relevant information in their responses, not wanting to provide some types of information, or deliberately providing some inaccurate responses. Therefore, the accuracy of information received through mail and phone surveys should be checked by selecting a subsample of the harvesters and/or landowners surveyed and conducting follow-up site visits.

4.4.2 Cost

Cost can be an important consideration when selecting an evaluation method. Site visits can cost several hundred dollars per harvest operation depending on the complexity of the operation, the information to be collected, and the number of evaluators used. Mail and/or telephone surveys can be an inexpensive means of collecting information, but their cost must be balanced with the type and accuracy of information that can be collected through them. Other costs need to be figured into the overall cost of mail and/or telephone surveys as well, including follow-up phone calls and site visits to make up for a poor response to mailings and for accuracy checks. Additionally, the cost of questionnaire design must be considered, as a well-designed questionnaire is extremely important to the success of self-evaluations. Questionnaire design is discussed in the next section.

The number of evaluators used for site visits has an obvious impact on the cost of a MM/BMP implementation survey. Survey costs can be minimized by having one or two evaluators visit harvest sites instead of having multiple-person teams visit each survey site. If the expertise of many specialists is desired, it might be cost-effective to have multiple-person teams check the quality of evaluations conducted by one or two evaluators. This can usually be done at a subsample of harvest sites after the sites have been surveyed.

An important factor to consider when determining the number of evaluators to include on site visitation teams, and how to balance the use of one to two evaluators versus multiple-person teams, is the objectives of the survey. Cost notwithstanding, the teams conducting the site evaluations must be sufficient to meet the objectives of the survey, and if the required teams would be too costly, then the objectives of the survey would need to be modified.

Another factor that contributes to the cost of a MM/BMP implementation survey is the number of sites to be surveyed. Once again, a balance must be reached between cost, the objectives of the survey, and the number of sites to be evaluated. Generally, once the objectives of the study have been specified, the number of sites to be evaluated is determined statistically to meet required data quality objectives. If the number of sites that is determined in this way would be too costly, then it would be necessary to modify the study objectives or the data quality objectives. Statistical determination of the number of sites to evaluate is discussed in Section 2.3.

4.4.3 Questionnaire Design

Many books have been written on the design of data collection forms and questionnaires (e.g., Churchill, 1983; Ferber et al., 1964; Tull and Hawkins, 1990), and these can provide good advice for the design of simple questionnaires to be used for a single survey. However, for complex questionnaires or ones that will be used for initial surveys as part of a series of surveys (i.e., trend analysis), it is strongly advised that a professional in questionnaire design be consulted. Although it might seem that designing a questionnaire is a simple task, small details such as the order of questions, the selection of one word or phrase over a similar one, and the tone of the questions can significantly affect survey results. A professionally designed questionnaire can yield information beyond that contained in the responses to the questions themselves, while a poorly designed questionnaire can invalidate the results.

The objective of a questionniare, which should be closely related to the objectives of the survey, should be extremely well thought out prior to its being designed. Questionnaires should also be designed at the same time as the information to be collected is selected to ensure that the questions address the objectives as precisely as possible. Conducting these activities simultaneously also provides immediate feedback on the attainability of the objectives and the level of detail of information that can be collected. For example, a researcher might want information on protection of habitat near surface waters, but might discover while designing the questionnaire that the desired information cannot be obtained through the use of a questionnaire, or that the information that could be collected would be insufficient to fully address the chosen objectives. In such a situation the researcher could revise the objectives and questions before going further with questionnaire design.

Tull and Hawkins (1990) identified seven major steps in questionnaire construction:

1. Preliminary decisions
2. Question content
3. Question wording
4. Response format
5. Question sequence
6. Physical characteristics of the questionnaire
7. Pretest and revision

Preliminary decisions include determining exactly what type of information is required, determining the target audience, and selecting the method of communication (e.g, mail, telephone, site visit). These subjects are addressed in other sections of this guidance.

The second step is to determine the content of the questions. Each question should generate one or more of the information requirements identified in the preliminary decisions. The ability of the question to produce the necessary data needs to be assessed. "Double-barreled" questions, in which two or more questions are asked as one, should be avoided. Questions that require the respondent to aggregate several sources of information should be subdivided into several specific questions or parts. The ability of the respondent to answer accurately should also be considered when preparing questions. Some respondents might be unfamiliar with the type of information requested or the terminology used. A respondent might have forgotten some of the information of interest, or might be unable to verbalize an answer. Consideration should be given to the willingness of respondents to answer the questions accurately. If a respondent feels that a particular answer might be embarrassing or personally harmful (e.g., might lead to fines or increased regulation), he or she might refuse to answer the question or might deliberately provide inaccurate information. For this reason, answers to questions that might lead to such responses should be checked for accuracy whenever possible.

The next step is to decide on the specific phrasing of the questions. Simple, easily understood language is preferred. The wording should not bias the answer or be too subjective. For instance, a question should not ask whether groundskidding led to erosion during the harvest. Instead, a series of questions could ask whether groundskidding was used, the slope of land on which it was used, which BMPs were used initially to control erosion from groundskidding, and what additional measures were used to control erosion from groundskidding (if erosion occurred). These questions all request factual information of which a forest operator should be knowledgeable, and the questions progress from simple to more complex. All alternatives and assumptions should be clearly stated on the questionnaire, and the respondent's frame of reference should be considered.

Fourth, the type of response format should be selected. Various types of information can best be obtained using open-ended, multiple-choice, or dichotomous questions. An open-ended question allows respondents to answer in any way they feel is appropriate. Multiple-choice questions tend to reduce some types of bias and are easier to tabulate and analyze; however, good multiple-choice questions can be more difficult to formulate. Dichotomous questions allow only two responses, such as "yes-no" or "agree-disagree." Dichotomous questions are suitable for determining points of fact, but they must be very precisely stated and unequivocally solicit only a single piece of information.

The fifth step in questionnaire design is the ordering of the questions. The first questions should be simple to answer, objective, and interesting in order to relax the respondent. The questionnaire should move from topic to topic in a logical manner without confusing the respondent. Early questions that could bias the respondent should be avoided. There is evidence that response quality declines near the end of a long questionnaire (Tull and Hawkins, 1990). Therefore, more important information should be solicited early. Before presenting the questions, the questionnaire should explain how long (on average) it will take to complete and the types of information that will be solicited. The questionnaire should not present the respondent with any surprises.

The layout of the questionnaire should make it easy to use and should minimize recording mistakes. The layout should clearly show the respondent all possible answers. For mail surveys, an attractive appearance is important for securing cooperation.

The final step in the design of a questionnaire is the pretest and possible revision. A questionnaire should always be pretested with members of the target audience. This will preclude expending large amounts of effort and then discovering that the questionnaire produces biased or incomplete information.

4.5 AERIAL RECONNAISSANCE AND PHOTOGRAPHY

Aerial reconnaissance and photography can be useful tools for gathering harvest site information quickly and comparatively inexpensively. For the purposes of forestry BMP compliance surveying, aerial reconnaissance can be useful for selecting survey sites and evaluating harvest sites. In Florida, survey sites for each county are selected from fixed-wing aircraft flown in a predetermined pattern. This approach reduces bias in selecting survey sites. The selected sites are then visited by foresters for BMP compliance evaluations (Vowell and Gilpin, 1994). Survey sites are selected from fixed-wing aircraft in South Carolina as well (Adams, 1994). In addition, aerial photography has been proven to be helpful for forest regeneration assessment (Hall and Aldred, 1992; Hudson, 1988); forest inventory and analysis (Hackett, 1988); terrain stratification, riparian area delineation, vegetation mapping, stream morphology characterization, inventory site identification, planning, and monitoring in mountainous regions (Born and Van Hooser, 1988; Hetzel, 1988); rangeland monitoring (BLM, 1991); and agricultural conservation practice identification (Pelletier and Griffin, 1988). Factors such as the characteristics of what is being monitored, scale, and camera format determine how useful aerial photography can be for a particular purpose.

Photographic scale and resolution must be taken into consideration when deciding whether to use aerial photography, and a photographic scale that produces good resolution of the items of importance to the monitoring effort must be chosen. Born and

Van Hooser (1988), investigating the usefulness of aerial photography for the classification of inventory and monitoring sample points and locating the sample points on the ground, found that a scale of 1:58,000 (i.e., 1 unit on a photograph represents 58,000 units on the ground) was marginal for use in forestry resource inventorying and monitoring. Hetzel (1988) and Mereszczak (1988), using a large-format camera (see below), found that at a scale of 1:30,000 riparian areas were easily distinguishable and could be delineated with 100 percent accuracy, and cover types could be delineated with 83 percent accuracy. Mereszczak (1988) found that aerial photography was especially useful for monitoring riparian areas because changes in their ecological condition in response to management practices are evident over time frames of 10 years or less. Reutebuch and Shea (1988) reported that photographs taken at a scale of 1:12,000 or larger have a typical resolution of less than 1 foot on the ground. Hall and Aldred (1992) were able to clearly delineate and map cutovers and nonforest areas (water bodies, roads, landings, clearings, brush areas) at a photographic scale of 1:10,000, and could detect conifer seedlings 30 cm or taller, if not hidden beneath other trees, at scales of 1:800 to 1:500. The Bureau of Land Management (BLM) uses low-level, large-scale (1:1,000 to 1:1,500) aerial photography to monitor rangeland vegetation (BLM, 1991). BLM reports that scales smaller than 1:1,500 (e.g., 1:10,000, 1:30,000) are too small to monitor the classes of land cover (shrubs, grasses and forbs, bare soil, rock) on rangeland.

Pelletier and Griffin (1988) investigated the use of aerial photography for the identification of agriculture conservation practices. They found that practices that occupy a large area and have an identifiable pattern, such as contour cropping, strip cropping, terraces, and windbreaks, were readily identified even at a small scale (1:80,000) but that smaller, single-unit practices, such as sediment basins and sediment diversions, were difficult to identify at a small scale. They estimated that 29 percent of practices could be identified at a scale of 1:80,000, 45 percent could be identified at 1:30,000, 70 percent could be identified at 1:15,000, and over 90 percent could be identified at a scale of 1:10,000.

Camera format is another factor that must be considered. Large-format cameras are generally preferred over small-format cameras (e.g., 35mm), but are more costly to purchase and operate. The large negative size (9 cm x 9 cm) produced using a large-format camera provides the resolution and detail necessary for accurate photo interpretation. Large-format cameras can be used from higher altitudes than small-format cameras, and the image area covered by a large-format image at a given scale (e.g., 1:1,500) is much larger than the image area captured by a small-format camera at the same scale. Small-format cameras can be used for identifications that involve large-scale features, such as mining areas, the extent of burning, and large animals in censuses, and they have definite applications in forestry as well. Owens (1988) recommends 35mm photography for monitoring small areas ($\leq$ 3 mi^2) at low altitude. A particularly useful application of 35mm photography is mapping private landowner parcels and change monitoring (Owens, 1988). Owens (1988) used large-format photographs as baseline data and in subsequent years used 35mm photographs to monitor timber harvests with much success.

Small-format cameras are limited in the resolution that they provide when photographs are enlarged (Owens, 1988).

BLM recommends the use of a large-format camera because it provides flexibility to increase sample plot size, it permits modest navigational errors during overflight, and the images provide the photo interpreter with more geographical reference points (BLM, 1991). Large-scale photographs have advantages over topographic maps. Specifically, they have much higher resolution, contain many more features and ground characteristics, and—when viewed in stereo—provide an accurate, 3-dimensional model of an area, complete with vegetative cover information and land-use characteristics (Reutebuch and Shea, 1988). Also, large-format photography equipment is standard equipment for most photo contractors, so one could be hired to take the photographs in lieu of purchasing the equipment.

A drawback to the use of aerial photography is that forestry BMPs that do not meet implementation or operational standards but are similar to practices that do are indistinguishable from ones that do in an aerial photograph (Pelletier and Griffin, 1988). Also, practices that are defined by managerial concepts rather than physical criteria, such as preharvest planning or forest chemical management, cannot be detected with aerial photographs.

Regardless of scale, format, or item being monitored, it is useful for photo interpreters to receive 2 or 3 days of training on the basic fundamentals of photo interpretation and important that they be thoroughly familiar with the vegetation and landforms in the areas where the photographs that they will be interpreting were taken (BLM, 1991). A site visit to the field sites in the photographs is recommended to improve correlation between the interpretation and actual site characteristics. Usually, after a few site visits and interpretations of photographs of those sites, photo interpreters will be familiar with the photographic characteristics of the vegetation in the area and site visits can be reserved for verification of items in doubt. A change in type of vegetation or physiography in photographs normally requires new site visits until photo interpreters are familiar with the characteristics of the new vegetation in the photographs.