9.0 Effectiveness Evaluation

9.1 Introduction

Effectiveness evaluation is an essential, but often neglected, component of any restoration activity. It measures to what extent the work performed has attained the watershed restoration objectives. The benefits of effectiveness evaluation include:

- more efficient allocation of funding and optimization in planning/decision-making (i.e. program benefits);
- a measure of the state of recovery of the watershed (i.e. how much restoration has been achieved, how much more effort is required); and
- technical feedback to refine restoration treatment design and implementation.

This section provides a basis for the development of routine evaluations of hillslope activities for watershed restoration. Program development, including determination of appropriate evaluation objectives, study design, selection of evaluation variables, sampling methodology and suggested sampling protocols are discussed. Analysis of sampling data and presentation of results are also described. Although this section deals only with evaluation of hillslope works, any program of effectiveness evaluation should be an integrated effort whenever practical, combining evaluation of completed hillslope, riparian and stream restoration works.

Effectiveness evaluation addresses three fundamental questions with respect to restoration:

1. Is the restoration work achieving the desired objectives?
2. How can techniques be improved?
3. How can the cost-effectiveness of the work be improved?

Effectiveness evaluations for watershed restoration may range from inspections to validation evaluations (i.e. research) as shown in Figure 9.01.

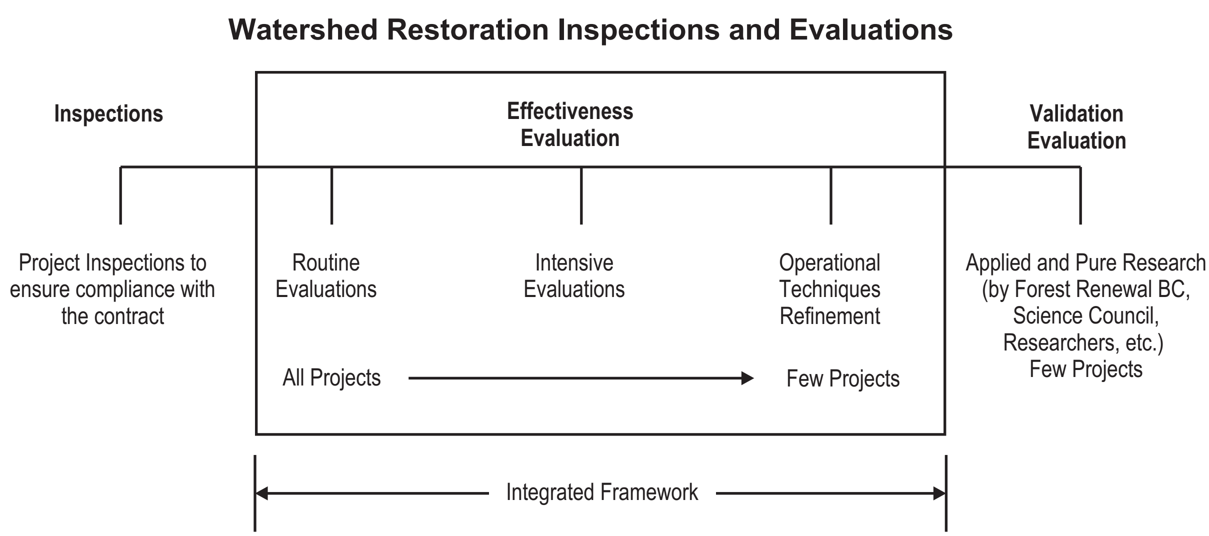

Fig. 9.01 *Inspections and evaluations that may be associated with watershed restoration .*

Any watershed or sub-basin can be segregated into its hillslope, riparian and stream components (Figure 9.02). For the hillslope component, sub-components may include roads, landslides and gullies. Further division to the site level may include features such as cutslopes, fillslopes and cross-ditches for roads, headscarp, transport and deposition zones for landslides, and headwall, sidewall, channel and deposition zones for gullies.

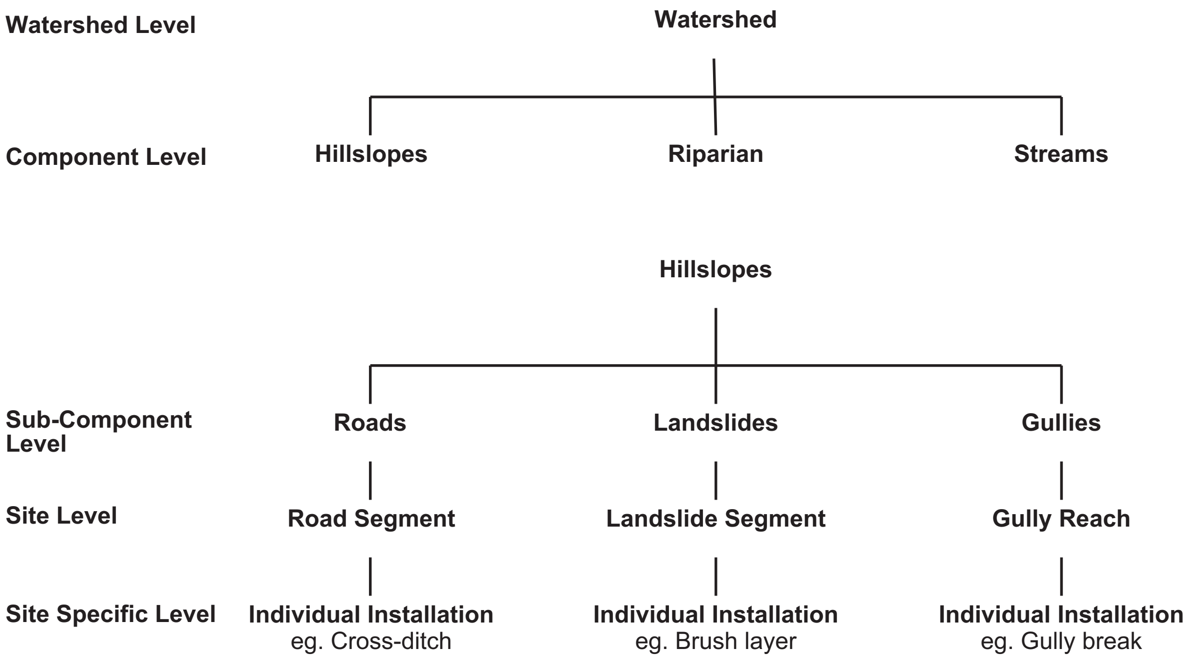

Fig. 9.02 *Watershed, Component and Site Levels*

The evaluation of effectiveness can be performed at specific levels of intensity with benefits and limitations associated with each level. Measured or observed responses may be dependent upon the effect of a specific activity (e.g., an increase in vegetation cover due to seeding) or the combined effects of a number of activities (e.g., the improvement in water quality due to the cumulative effect of all hillslope treatments).

Three levels of effectiveness evaluation are possible: routine evaluation, intensive evaluation and operational techniques refinement. Routine evaluation is the low cost overview evaluation that will be appropriate for the majority of restoration projects. Intensive evaluation and operational techniques refinement are generally specialized studies chosen for selected sites, and are beyond the scope of this document. See WRP Technical Circular No. 12 by Gaboury and Wong (1999) for further information. A comparison between these three levels of evaluation is shown in Figure 9.03

*Fig. 9.03 Comparison of Effectiveness Evaluation Levels**

Evaluation Levels	Scale	Methods	Personnel	Data Type	Timing/Duration	Data Analysis	Outputs
Routine	Watershed to site level; sample locations confined to specific watershed (all treatments and projects to be evaluated)	Aerial and ground-based surveys; non-specialized study design	Technicians + specialist guidance	Qualitative with ratings and Quantitative with or without ratings; (Physical and/or Biological)	Roads: semi-annually to annually for SP and high risk P deactivation for 5 years or until residual risk is acceptable. Landslides and Gullies: every year for first 2 years then every 5 years (until residual risk is acceptable).	Comparative and graphical analysis (means, bar graphs, etc).	Short summary of functional condition of works, failures, remedial works requirements. General treatment performance. Level of residual risk. Evaluate treatments as to meeting objectives; rate or measure variables. Comparisons made with bio-standards, WRP, FPC, policy standards.
Intensive **	Variable; sample locations may be from several watersheds (few treatments or projects to be selected for studies)	Detailed ground sampling using mainly existing protocols; requires specialized study design	Specialists	Quantitative mainly; intensive measurements; Physical and/or Biological	Site specific but generally long term. Variable sampling frequency.	Statistical study design.	Routine Outputs plus: Statistically derived inferences; identification and comparison of relationships between variables; process relationships.
Operational Techniques Refinement **	Variable; sample locations may be from several watersheds (few treatments or projects selected for studies).	Detailed ground; sampling protocols may be existing or require development; requires specialized study design	Specialists, service reps	Quantitative; Physical and/or Biological	May be one time only, dependent on evaluation objectives.	Statistical analysis is driven by study.	Main output is cost-benefit and decision analysis.

* Details provided are general characteristics, specific projects may vary.
** Recommend that these levels of study be proposal-driven.

Routine effectiveness evaluation can be conducted by qualified technicians with specialists providing minimal guidance in study design and sampling protocol. In general, all WRP projects should be routinely evaluated for a minimum of one year after completion of the work. Routine evaluation may continue on all or a portion of the project area depending upon assessment of residual risk, and on the benefit of evaluating specific sites to provide longer term information. As a general guideline, budgeting for routine monitoring should be approximately 4 % of the cost of the physical restoration works.

The main objectives of routine evaluations are to:

- assess the present configuration and condition of restoration treatments (i.e., are the works still in place);
- qualitatively and/or quantitatively assess whether or not the treatments have been effective in addressing the restoration objectives (i.e. are the works functioning as intended, how well are they functioning);
- determine if remedial work is needed; and
- identify specific hillslope areas which may warrant a more rigorous evaluation or a more specific investigation.

9.2 Routine Evaluation

9.2.1 Evaluation Objectives

Restoration objectives are developed during the initial overview assessment and restoration planning process. Project objectives may initially encompass the broad watershed level, but are further refined at the component and site levels. Watershed, component, and site-level restoration objectives provide the basis for all restoration work conducted within a watershed. Appropriate restoration objectives are specific, obtainable, measurable and represent some desired future condition. They may represent previously established standards or guidelines, or they may represent benchmark conditions from unlogged watersheds. Restoration activities are the physical works conducted to meet restoration objectives. For hillslopes, the main activities typically include road deactivation, landslide rehabilitation and gully rehabilitation. Effectiveness evaluation objectives are intended to determine if the restoration objectives have been met at the appropriate corresponding levels. In Figure 9.04, a hierarchical sequence of restoration objectives and corresponding effectiveness evaluation objectives is shown. Grouping of effectiveness for each site level objective provides a means of determining if sub-component and component level restoration objectives are being met, and ultimately if watershed objectives are being met.

Fig. 9.04 Restoration Objectives, Restoration Activities and Effectiveness Evaluation Objectives

Level	Restoration Objective	Restoration Activity	Effectiveness Evaluation Objective
Watershed	Reduce the generation of sediment related to logging activities.	Take measures to reduce the incidence of logging-related mass-wasting and rehabilitate existing sites of logging-related sediment generation.	Determine if treatments have mitigated sediment generation related to logging activities.
Component	Stabilize and rehabilitate the hillslope areas affected by logging activities.	Stabilize areas of potential failures, provide erosion control for existing sediment sources.	Determine if works have stabilized all areas with potential for failure, determine if erosion control has been successful.
Sub-Component	Reduce mass-wasting and erosion from: Roads Landslides Gullies	Deactivate and revegetate roads Stabilize and revegetate landslides Rehabilitate gullies by reducing incidence or volume of mass-wasting and revegetating.	Determine if mass-wasting hazard has been reduced and erosion control has been effective on deactivated roads, rehabilitated landslides and gullies.

Evaluation objectives may be addressed in even greater detail by focusing on site specific objectives. In Figure 9.05, possible site level variables that may be appropriate for evaluation are shown.

Fig. 9.05 Restoration Objectives, Restoration Activities and Effectiveness Evaluation Variables

Level	Restoration Objective	Restoration Activity	Effectiveness Evaluation Variable
Site	**Road Sites:**	**Road Sites:**	**Road Sites:**
	Reduce volume of fill supported by woody debris, reduce surcharging of slope by fill.	Pullback fillslopes.	Degree of slumping, sliding or other slope instability indicators on fillslope.
	Reduce raveling to allow for establishment of vegetation.	Pullback and reduce angle of fillslope, or install modified brush layers.	% establishment of vegetation.
	Provide toe support for unstable cutslopes.	Buttress unstable cutslopes	Degree of slumping, sliding or other slope instability indicators on cutslope.
	Reduce oversteepening of cutslopes.	Reslope cutslope.	Degree of slumping, sliding or other slope instability indicators on cutslopes.

Level	Restoration Objective	Restoration Activity	Effectiveness Evaluation Variable
Site	**Road Sites:**	**Road Sites:**	**Road Sites:**
	Reduce pore water pressure in road fills and native soils by dispersing flow or discharging into established channels.	Install appropriate water management structures at required locations.	Proportion of water management structures functioning at various efficiencies (nominal ranking of efficiency from 1 to 5).
	Reduce erosive capacity of water.	Install water management structures to direct water into non-erodible channels.	Proportion of water management structures functioning at various efficiencies (nominal ranking of efficiency from 1 to 5).
	Increase protection from rain splash, increase interception, reduce surface runoff velocity.	Seeding + plant; apply erosion control blankets, rip rap, mulch; increase slope roughness.	% vegetation cover established, degree of invasion by native species (i.e. have native species established in treated area), degree of rilling (ordinal scale).
	Provide site conditions conducive to establishment of commercial tree species (i.e. reduce soil compaction, increase soil nutrients, provide favourable microsites).	Scarify soil, replace organic soil, create mounds.	% survival and growth rate of commercial tree species.
Site	**Landslide Sites:**	**Landslide Sites:**	**Landslide Sites:**
	Reduce oversteepening of headscarp/sidescarps.	Reslope headscarps/sidescarps to mechanically stable angle.	Extent of headscarp/sidescarp retrogression.
	Reduce pore water pressure in headscarp area.	Divert upslope water away from headscarp area.	Reduction in seepage from headscarp.
	Reduce erosive capacity of surface water in channels.	Install armour, check dams, or soil bioengineering structures along surface channels.	Surface flow confined to intended channels; durability of installed structures (ordinal scale).
	Increase protection from rain splash, increase interception, reduce surface runoff velocity by establishing adequate vegetative cover.	Seeding + plant + soil bioengineering.	% vegetation cover established, degree of invasion by native species (i.e. have native species established in treated area), degree of rilling (ordinal scale).
	Increase soil strength provided by woody species (buttressing, arching, anchoring).	Planting with woody species, installing soil bioengineering structures (wattle fences, brush layers, live smiles, etc).	% survival and growth rate of woody species, survival of soil bioengineering structures and establishment of cuttings.
	Reduce volume of potential debris flow.	Clean logging introduced woody debris.	Volume of subsequent debris flows reduced (as compared to natural events nearby).

Level	Restoration Objective	Restoration Activity	Effectiveness Evaluation Variable
Site	**Gully Sites:**	**Gully Sites:**	**Gully Sites**
	Reduce erosion or likelihood of debris flow/flood initiation from channel.	Remove channel obstructions (i.e. road fill from gully crossings, logging-related woody debris jams, failsafe road crossings).	Amount of sediment transport, frequency of debris flow/flood events.
	Increase soil strength provided by woody species on sidewalls and headwalls.	Plant woody species.	% survival and growth rate of woody species.
	Reduce transport of debris flow/flood material to downstream values.	Construct mitigation structures (berms or basins).	Structure has functioned as intended following event (i.e. debris contained or deflected).
	Increase protection from rain splash, increase interception, reduce surface runoff velocity on denuded sidewalls/headwalls by establishing adequate vegetative cover.	Seeding + plant + soil bioengineering.	% vegetation cover etablished on headwalls and sidewalls, degree of invasion by native species (i.e. have native species established in treated area).
	Reduce potential for avulsion and subsequent erosion on fans.	Excavate channels on fan to increase channel incision, planting, soil bioengineering or riprap to stabilize streambanks.	Frequency of channel avulsion.

Planning for effectiveness evaluation should begin as early as possible in a restoration project since the collection of pre-restoration data may provide useful benchmarks for comparison. Ideally, the collection of baseline data should be incorporated into both the overview/inventory and assessment/prescription phases of work where necessary. A list of possible evaluation variables or parameters should be compiled, the distribution of sample sites considered and limitations (e.g. technical and logistical) identified. Site level evaluation objectives have the most direct linkage to restoration activities and generally require detailed information that is typically obtained during the prescription stage. This may include field verification of the suitability of prospective evaluation sites and evaluation variables. Selection and establishment of photo stations, for instance, might be appropriate at the overview stage and the prescription stage, respectively. To meet budgetary and logistical constraints, the number of evaluated sites may have to be reduced through a screening or stratification process. For completed watershed restoration projects where no evaluation sites were selected, some field verification should be included as a component of the study design. It may be possible to gather some pre-restoration data through review of historical information, such as a compilation of landslide history in a watershed based on review of historical aerial photographs.

9.2.2 Study Design

Evaluation study design will reflect the individual restoration project's set of unique characteristics and objectives. Field verified site level evaluation objectives, as shown in Figure 9.05, will be the major guide for study design.

For hillslope activities, routine evaluation at the lowest intensity may consist of a visual qualitative assessment conducted on all restoration structures and treatments carried out in the project area. This assessment could be carried out either from an aircraft or from review of low level aerial photographs or video. Qualitative assessments use indices of response rather than more expensive and time-consuming direct measures. For example, an indication of road surface erosion might be represented by an estimate of the degree of surface rutting rather than by detailed measures of sediment generation. Establishment of fixed photo points or a detailed video record can be a cost-effective means of demonstrating the change in site conditions over time. At the highest intensity, routine evaluation may be predominantly a quantitative assessment conducted at selected sites within the project area. Measurement of key variables might be carried out on the ground following set protocols. For example, vegetation cover on a landslide or deactivated road could be quantitatively assessed by establishing random vegetation plots. For most studies, some combination of low intensity qualitative and higher intensity quantitative evaluation will be appropriate. Higher intensity evaluations may be warranted for sites where there was a high initial risk, where innovative treatments were applied, or where prescribed work could not be completed (e.g. due to safety concerns, logistical limitations, etc.) or was not attempted (e.g. site was risk managed).

9.2.3 Selection of Evaluation Variables

The evaluation variables selected for a routine level of evaluation should be practical (i.e. obtainable and measurable), yet adequately address the site level evaluation objectives. Some guiding principles for variable selection as described by MacDonald et al. (1991) are:

- The selection of evaluation variables will differ from watershed to watershed as objectives and site characteristics vary;
- Select variables that are the best indicators of change and measure them in the appropriate areas that are the most responsive to change;
- Select variables that provide the best evidence/inference that objectives are being met and are least influenced by extraneous events (i.e. natural variability);
- Select variables that are easily observed, measured and qualified and/or quantified to help ensure consistent data collection; and
- Select variables that are the least influenced by location of sampling or time of sampling.

For watershed restoration, four types of evaluation variable can be defined based on initial conditions and the correlation between variable value and the desired direction of change:

Type I Variable - has an initial value prior to the restoration activity and is negatively correlated with the desired direction of change expressed by the restoration objective.

> *Examples:* number of road fill failures on deactivated road or number of debris flows following gully rehabilitation; variable values should decrease following treatment (high initial value, expected lower value after treatment).

Type II Variable - has no initial value prior to the restoration activity and is negatively correlated with the desired direction of change expressed by the restoration objective.

> *Examples:* degree of erosion along constructed cross-ditch or the number of soil bio- engineering structures that have failed; variable values expected to decrease over time.

Type III Variable - has an initial value prior to the restoration activity and is positively correlated with the desired direction of change expressed by the restoration objective.

> *Examples:* vegetation cover on landslide following hydroseeding or annual growth of conifers in riparian zone after alder brushing; variable values should increase after treatment (low initial value, expected higher value after treatment).

Type IV Variable - has no initial value prior to the restoration activity and is positively correlated with the desired direction of change expressed by the restoration objective.

> *Examples:* vegetation cover on pulled back road fillslope after handseeding, or flow through outlet of installed French drain; variable values expected to increase over time.

9.2.4 Sampling Methodology

Sampling methodology is guided by the overall study design and encompasses sampling aspects such as the nature and number of sample sites (i.e. whether they are continuous or discrete sites), the sampling regime (i.e. duration and frequency of sampling) and the sampling protocols (standardized sampling procedures). Some guidance by specialists may be appropriate in developing the sampling methodology.

As stated, routine evaluation may range from review of all treated sites at a relatively low intensity, to review of a selected subset of sites at variable intensity. Some possible methods of subset selection may be:

- based on residual risk following treatment (i.e. subset may be biased toward higher residual risk sites to determine as soon as possible if remedial action is required);
- via a sampling of specific site characteristics such as terrain type, treatment technique, or initial site conditions (this has the potential to provide the greatest amount of beneficial information); or
- through a random sampling of stratified sites to allow for higher intensity sampling at representative sites.

Sampling should be conducted to provide an initial pre-work site characterization if possible, a site characterization immediately post-work, and post-work evaluation at specified intervals. As the frequency and duration of evaluations for any project may be modified based on the results of initial evaluations, there should be a certain amount of flexibility in planning. For routine evaluation, it may be appropriate to conduct post-work evaluation of high risk sites initially twice in the first year following work, or after extreme weather events. Depending on results, it may be appropriate to decrease evaluation frequency to once a year for subsequent years. Evaluation of sites may continue for approximately five years or until the residual risk is deemed acceptable (e.g. until evaluation variables have reached apparent equilibrium with the environment). For lower risk sites, initial results may indicate a reduced evaluation frequency to once every second year or longer, or to cease evaluations entirely.

For water management structures and slope stability treatments, evaluations during wet periods are generally most appropriate. For vegetation, sampling during peak growth periods should be considered.

A key consideration of sampling methodology is the collection of easily measured, objective, field-based data using repeatable methods that provides consistent data. Development of standardized sampling protocols, rating systems and field data collection forms are important to ensure:

- a reduction in the subjectivity and bias inherent in assessments conducted by different field personnel;
- a standardized valuation of observations and measurements into well-defined categories; and
- that data can be compiled, summarized and compared objectively with results from other projects.

For routine evaluations, qualitative assessments of structures or treatment can be rated using a binary classification (yes/no), an ordinal scale (good, fair, poor), or a numerical scale (1,2,3,4). Rating levels should be established prior to the field evaluation to ensure consistency in the grading of observations and measurements by field personnel. A certain level of training will be required to ensure consistent data collection.

Sampling protocols for effectiveness evaluation of hillslope restoration have not yet been developed for the watershed restoration in B.C. An example of a possible nominal rating system for evaluating the effectiveness of cross-ditches and other water management structures is as follows:

5 - All water confined to intended channel, no significant erosion (functioning as intended)

4 - All water confined to intended channel, minor erosion evident.

3 - All water confined to intended channel but significant downcutting and sidewall erosion.

2 - Some water diverted from intended channel, partial failure of ditchblock or sidewall; possible remedial action required.

1 - Most of water diverted from intended channel or channel obliterated; remedial action required (total failure).

If such structures are also intended to provide vehicle access, each structure could also be rated for accessibility in a similar manner, or by simply using a binary classification (Yes - accessible/No - not accessible).

A possible rating system for evaluation of the effectiveness of road fill pullback may consist of the following effectiveness rating criteria (ordinal scale):

Good - one or less prominent field indicators of instability or erosion; field indicators selected for use as evaluation variables included evidence of displacement, tension cracks, slumping, non-retrieved road fill, slope failure, incision, water flow obstruction, surface erosion.

Fair - two to three prominent field indicators of instability or erosion;

Poor - greater than three prominent field indicators of instability or erosion.

For routine evaluation of a treated landslide, a reassessment of the site following the LRAP procedure may provide adequate qualitative information. For higher intensity sampling vegetation plots or transects could be established to provide quantitative data.

Rating categories for measured values should be based on comparisons with published guidelines or field procedures.

Establishment of photos stations at selected sites following an established protocol, such as provided in the USDA Watershed Restoration Effectiveness Monitoring Protocol (June 1996), may be a cost-effective means of recording changes in conditions over time. Video records can play an equally important role if procedures are carefully documented and linked to GPS stations.

9.2.5 Analysis and Presentation of Data

Some typical outputs from routine evaluations include summaries of:

- the number, type and location of restoration treatments and the intensity of their evaluation;
- the structural condition, stability, establishment/growth of the restoration treatments;
- treatments that need remedial measures and what measures are recommended;
- treatments that require more detailed evaluations (routine or intensive);
- recommendations for improvement in treatment design, technique or implementation;
- residual risk; and
- recommendations for future frequency and duration of evaluation.

Data analyses for routine evaluations may include:

- simple graphical displays of effectiveness rating frequencies by structure or treatment type from visual or ground based measurements;
- comparative analyses of how data in treatment areas differ over time;
- tabular inventories of effectiveness ratings for treatments or structures; and
- comparative analysis of key variable data, before and after treatment.

Interpretation of results aids in determining if the stated restoration objectives are being met. Results that are inconclusive, however, suggest that either:

- the evaluation protocol needs to be modified to provide more precise results;
- the evaluation objectives need to be modified;
- the selected evaluation variables did not adequately address the restoration objectives; or
- the evaluation assumptions may not have been valid.

Any such shortcomings should be identified to guide future evaluation.

Efficient collection, compilation and dissemination of data and results in a timely manner is critical for successful completion of future restoration work. This information will have adaptive management benefits to watershed restoration with respect to future treatment design and application.

The reader is referred to the **most recent and current Schedule A** for Effectiveness Evaluation for details and specifications regarding the planning and implementation of effectiveness evaluation projects.