Workplan to Pilot the Fisheries Sensitive Watersheds (FSW) Monitoring Framework

Final Version

Prepared for:

British Columbia Ministry of Environment
4th floor, 2975 Jutland Rd., Victoria
P.O. Box 9338, Stn Prov Govt
Victoria, BC, V8W 9M1

Prepared by

Darcy Pickard, Marc Porter, Katherine Wieckowski, and Dave Marmorek
ESSA Technologies Ltd.
Suite 300, 1765 West 8th Avenue
Vancouver, BC V6J 5C6

April 24, 2009

Citation: **Pickard, D., M. Porter, K. Wieckowski, and D. Marmorek**. 2009. Workplan to Pilot the Fisheries Sensitive Watershed (FSW) Monitoring Framework. Report prepared by ESSA Technologies Ltd., Vancouver, BC. for BC. Ministry of Environment, Victoria. 16 pp.

Table of Contents

List of Figures

Introduction

Under the *Forest and Range Practices Act* (FRPA) and *Government Actions Regulations* (GAR), the BC Ministry of Environment (MOE) has developed policy and procedures that guide a program for evaluating and designating drainages as Fisheries Sensitive Watersheds (FSWs) (Reese-Hansen and Parkinson 2006). A FSW is a watershed that has both i) significant fish values and ii) watershed sensitivity. Thirty-one FSWs have been designated by the Minister of the Environment and over the course of the next several years there are plans to identify and designate additional watersheds throughout the province as FSWs (L. Reese-Hansen, BC Ministry of Environment, pers. comm.). A comprehensive monitoring plan will be essential for ensuring that critical watershed processes and associated resource values are maintained within designated FSWs. The conceptual background for a FSW monitoring approach and an intended multi-tiered framework for the monitoring plan are described in Wieckowski et al. 2008 & 2009, respectively. The purpose of this report is to develop a flexible work-plan of specific tasks and analyses that must be completed for MOE to effectively move from a conceptual framework to on-the-ground deployment of a statistically reliable and cost-effective monitoring and evaluation program for FSWs.

The FSW monitoring framework described in Wieckowski et al. (2009) (Appendix B) presents a nested, three-tiered approach to FSW implementation and effectiveness monitoring. Each monitoring tier represented in Appendix B operates at a different spatial scale (i.e., provincial, across FSWs, within individual FSWs) and determines the scale at which data will be collected. Key management questions also differ at each of these three scales (Wieckowski et al. 2009). These management questions will in turn dictate the monitoring components required. For the framework to be effective the monitoring components must be carefully designed to answer well-defined questions at different scales in an efficient manner. Each element of the proposed framework for FSW effectiveness monitoring (Tiers 1 to 3 in Figure 1) has existing uncertainties that must be addressed within the work-plan before MOE can proceed with provincial-scale monitoring of FSWs. The workplan represents a modular approach, allowing various individual elements to be dropped temporarily as needed to adjust for variable budget scenarios, without hindering overall progress within the workplan

At the Tier 1 monitoring level there is a need to determine the most versatile, cost-effective remote sensed method(s)/dataset(s) available for capturing watershed information. Criteria need to be established for assigning Tier 1 watershed condition ratings (i.e., red, yellow, green) based on this remote sensed data. There is also a need to determine whether remote sensed data will be sufficient for these Tier 1 condition ratings, or whether there is a need for supplementation with field data collected at this scale. Finally, methods need to be developed for reliably linking and validating watershed condition as defined at Tier 1, to finer interpretations of watershed processes defined at Tier 2. At the Tier 2 monitoring level there is a need to resolve the most appropriate watershed stratifications, and best statistical designs for selecting both the watersheds to monitor, and the sampling sites within watersheds. While Tier 3 monitoring will ultimately be important for evaluating the results of specific actions on maintaining/restoring FSW processes, providing formal empirical evidence to adaptively revise management, uncertainties around Tier 3 monitoring will not be directly addressed within this work-plan. As MOE has not defined the management questions or actions of interest at the Tier 3 scale for FSWs it is not yet possible to evaluate Tier 3 data needs and potential sampling and monitoring approaches.

For Tier 1 and Tier 2 monitoring we propose a general sampling design structure (Figure 1) which we believe should be effective and flexible, while also ultimately easy to implement for broadscale FSW monitoring. We discuss our recommendations and describe the items which need to be resolved in order to implement the monitoring framework. Finally we provide suggestions for how to proceed with implementation of monitoring as funds become available.

Proposed Design Structure for Tier 1 and 2 Monitoring

We based the sampling design structure on the concepts presented in Wieckowski et al. (2008) (e.g., Section 6: Approaches/design options to monitoring watersheds; Appendix D: Developing a sampling design). The nested structure of the FSW framework (Appendix B) is such that Tier 1 monitoring is independent of the other Tiers, but Tier 2 monitoring is dependent on the outcome from Tier 1 monitoring. The first step in any sampling problem is to define the target population.[1] In the first year of the design the target population for both Tier 1 and Tier 2 are all current FSWs in BC. The primary sampling unit[2] for both Tier 1 and Tier 2 is the watershed. The secondary sampling unit for Tier 2 is a stream reach within watersheds. The general integrated design structure for Tier 1 and 2 monitoring is shown in Figure 1.

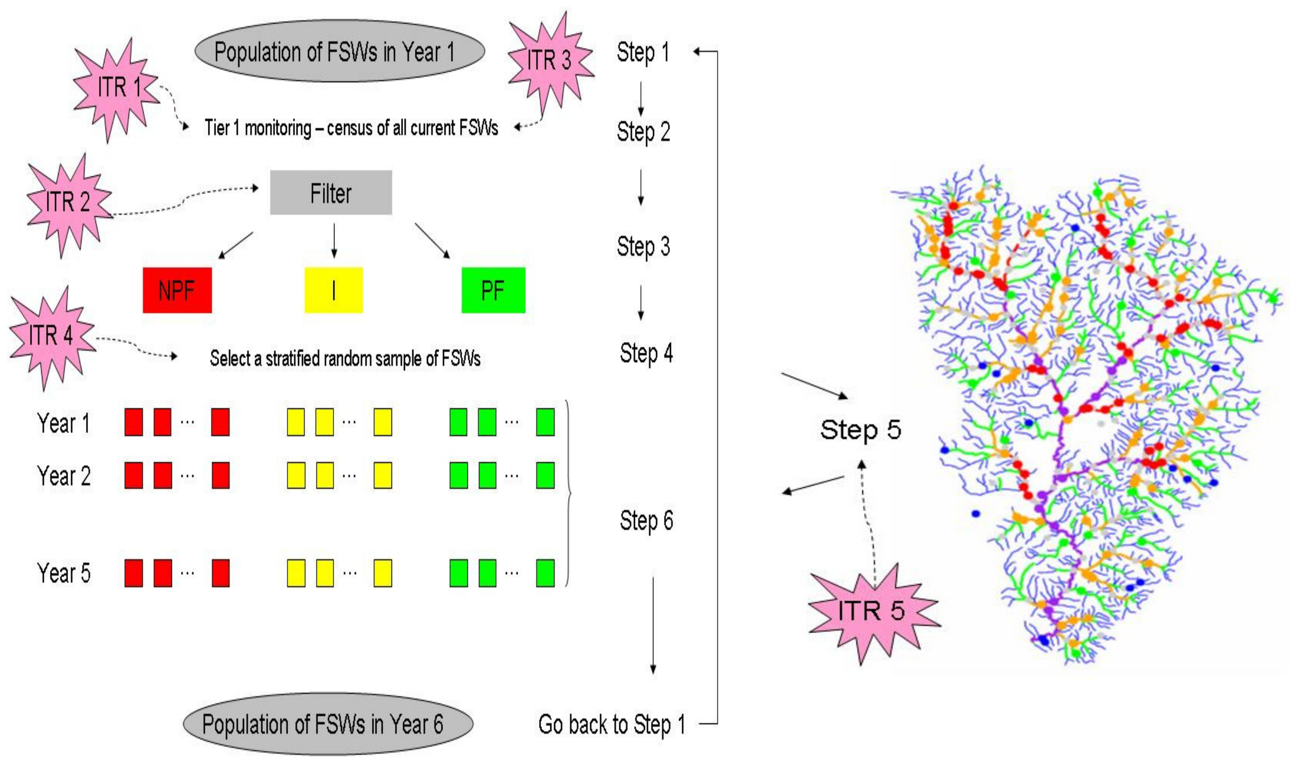


Figure 1. The general structure of the proposed design for Tier 1 and Tier 2 sampling, illustrating "items left to be resolved" (ITR) in order to implement the FSW monitoring framework and proposed sampling design (NP = not properly functioning, PF = properly functioning, I = impaired). Coloured boxes represent FSW watersheds. The zoomed-to basin for Step 5 represents the hydrology network with (hypothetical) selected sampling points for an example FSW.

[1] Target population: The population about which information is wanted (Cochran 1977).
[2] Sampling unit: The *sampling unit* is the actual unit of measurement. The population is divided into many sampling units. When sub-sampling is necessary, the secondary sampling unit is nested within the primary sampling unit (i.e., stream reaches within watersheds).

Sampling design steps illustrated in Figure 1:

Step 1: Identify the target population (i.e., the complete current list of FSWs in the Province)

Step 2: Complete Tier 1 monitoring for all FSWs every 5 years (the actual frequency adopted will depend on the cost of obtaining/interpreting Tier 1 data and the annual variability in the Tier 1 metrics).

Step 3: Assess the status of each FSW using the Tier 1 data;

Step 4: Select a stratified random sample of FSWs (i.e., primary sampling units) from within each functioning class (i.e., not properly functioning, impaired, or properly functioning)

Step 5: Within each selected watershed, select a stratified random sample of stream reaches (i.e., the secondary sampling unit).

Step 6: Complete Tier 2 monitoring using some version of a repeated[3] design across the 5 years

Go back to Step 1: Redefine the target population for both Tier 1 and Tier 2, adding any new FSWs that have been designated since the last time the Tier 1 monitoring occurred. The cycle then repeats itself as illustrated in Figure 1.

Items to resolve (ITRs) indicated in Figure 1:

ITR 1: Develop the Tier 1 monitoring component of the FSW framework (i.e., identify data sources and indicators).

ITR 2: Establish classification criteria and determine how to roll up stream reach scale assessments to the watershed scale.

ITR 3: Resolve Tier 1 monitoring, sampling design questions.

ITR 4: Resolve Tier 2 monitoring, design details for sampling of FSWs (i.e., primary sampling units)

ITR 5: Resolve Tier 2 monitoring, design details for sampling of reaches within FSWs (i.e., secondary sampling units)

Items to be resolved (ITRs)

1) Develop the Tier 1 monitoring component of the FSW framework (ITR 1).

1.1 Evaluate Potential Remote Sensed Methods/Datasets

Remote sensed data are increasingly being used in environmental monitoring. There are a number of potential remote sensed datasets currently available in British Columbia, (see section on remote sensing protocols in Wieckowski et al. (2008)). In the mid-1990s, BC developed rigorous Watershed Assessment Procedures. These procedures consisted of three levels of analysis. Level 1 involved a reconnaissance level analysis intended to act as a coarse filter to identify watersheds that may have impacts from the cumulative effects of past forest harvesting or planned future forest harvesting (IWAP 1995). Level 1 data were intended to be extracted from remote sensed data (e.g., aerial photographs and associated geographic information system (GIS) interpretation). Level 2 and 3 WAP assessments involved increasingly intensive and more thoroughly detailed field-based monitoring.

[3] Repeated Design: A sampling design where at least some sampling units are purposely re-visited over time.

The WAP Level 1 monitoring correlates reasonably well with the FREP Routine Riparian Effectiveness Evaluation (RREE) assessments of condition (e.g., Bowron Lake Study, Nordin et al. 2008). Unfortunately, the time and cost of implementing WAP Level 1 evaluations can be considerable (e.g., with regards to both collecting the necessary remote sensed data and in interpreting it). A workshop of provincial habitat monitoring experts held in Victoria, Nov. 25-26, 2008 proposed developing a streamlined version of the WAP Level 1, a so-called 'WAP-*lite*', with the intention of capturing sufficient information to filter FSW watersheds, but with reduced time and cost so that it could be applied more broadly. In addition the workshop participants agreed that the Tier 2 (i.e., on the ground, within watershed) monitoring would use the RREE protocol.

Based on discussions at the Victoria workshop we propose the following general strategy for Tier 1 monitoring be explored within the work-plan:

- Use the WAP Level 1 as the initial target. The WAP Level 1 collects data to inform 13 impact indicators in one of four categories: Peak flow, Surface Erosion, Riparian Buffer, and Mass Wasting. Begin by reviewing the existing WAP Level 1 protocol and the data sources required to obtain these indicators. Then evaluate alternative remote sensed data sources and compare them to the WAP Level 1 using the following questions:
    - What remote sense data are available (e.g. air photos, LandSat, SPOT, Quickbird etc.)?
    - How frequently are data available (e.g. annually, once every 5 years etc.)?
    - At what spatial scale (e.g. 1:20K, 1:50K, etc.)?
    - How much does it cost to obtain data (e.g., cost of coverage for entire province or parts of province to obtain imagery)?
    - How much does it cost to process data (e.g., how many person hours)?
    - How much processing effort is involved (e.g., are the data manually processed or is the procedure automated)?
    - What interpreted information can the data give us (e.g., road density, vegetation layers, etc.)?
- Evaluate the relative ability of each of the different remote sense methods to obtain data on the 13 impact indicators defined by WAP Level 1 (see Appendix A for a sample form). Also list any additional indicators that may be available from each remote sensed data source.

This review will provide the building blocks for developing the WAP-*lite* approach. As a minimum, WAP-*lite* should address all four categories of impact indicators. The next step will be to take this information and trial different potential indicators and associated criteria (Section 2.1) that can be used to filter FSW watersheds into three different condition levels (i.e., red, yellow, and green). As a first cut at the WAP-*lite*, we recommend choosing 2-3 indicators for each category that are available for less cost/effort than the WAP Level 1 indicators. Preferably they will be indicators that will be available at a minimum frequency of once every five years. The primary objective will be to determine if a WAP-*lite* assessment can be developed that is both quicker and cheaper than the original WAP Level 1 process, yet still correlates well with the RREE condition assessments.

2) Develop watershed scale classification criteria (ITR 2).

2.1 Establish Remote Sensed Criteria for Assigning Functional Status to FSWs.

Validation studies have shown that the RREE protocol (for evaluating the condition of streams and riparian management areas) correlates well with more detailed protocols (Tripp 2007). The aggregate RREE condition score also correlates with the condition of fish habitat (Tripp 2007). The next critical task is to determine how the WAP-*lite* indicators relate to the RREE metrics, to come up with criteria that can be used to assign all watersheds to one of three categories: Not Properly Functioning (NPF), Impaired (I), or Properly Functioning (PF). This process will be iterative and may also require refinement in the selection of WAP-*lite* indicators (Figure 2). If strong linkages cannot be established between the WAP-*lite* indicators and the Tier 2 monitoring indicators, then it may be necessary to include some field monitoring at a sample of sites in the Tier 1 monitoring design (e.g., temperature or water quality). This would be advisable only if Tier 1 field-based monitoring didn't increase costs beyond that of the original WAP level 1 or if it was decided to drop Tier 2 monitoring for cost or other reasons (Section 3.3).

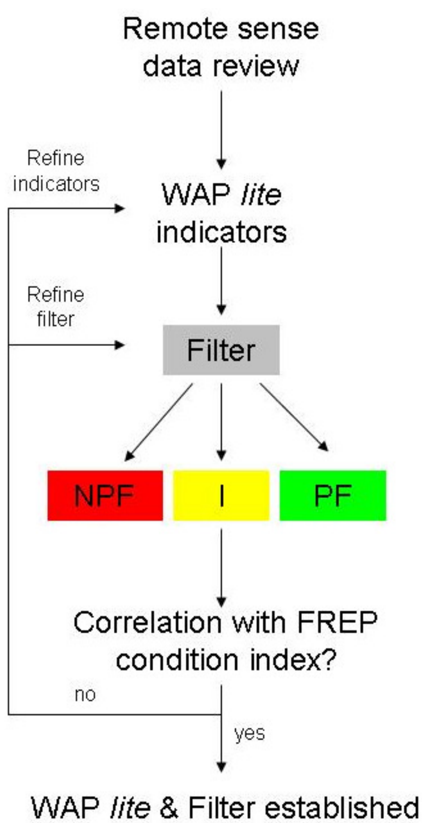


Figure 2. Process of developing a set of WAP-*lite* remote sensed indicators for evaluating the condition category of each FSW.

The original WAP Level 1 methods calculated an index score for each impact category. The index scores are all on the same scale (i.e., 0 to 1; no impact to potential high impact) and were based on real data from a large number of test watersheds (i.e., 40 for the Interior (IWAP) and 20 for Coastal watersheds (CWAP)) (BC Ministry of Forests 1995a and 1995b). The effort in establishing these scoring systems for the Forest Practices Code should be exploited as much as possible. If these are not directly applicable to the new WAP-*lite* indicators, then the same general strategy could be used (i.e., look at the range of plausible responses for the indicator and set the extreme values to 0 and 1 and interpolate between these points). A straight line fit between 0 and 1 is probably a reasonable place to start, but other fits would also be possible. For categorical responses you may or may not want to translate this to a score on the 0-1 scale. The Interior Columbia Technical Recovery Team (ICTRT) uses categorical responses for assessing the spatial structure and diversity of salmonids and may provide some useful examples for how to interpret categorical responses within an overall scoring system (ICTRT 2007).

Given data on a number of WAP-*lite* indicators, and a standardized score for each indicator, the next step is to actually build the filter. The indicators could be grouped first as in the WAP Level 1 generating a score for each category. Alternatively, the raw indicators could be interpreted directly. The goal of the scoring system is to find a set of classification rules that generally classifies the FSWs into the same bins as the RREE protocol. Nordin et al. (2008) employed Principal Component Analyses (PCA) as a statistical method for determining bin thresholds in the Bowron pilot study. Classification and Regression Tree (CART) methods would be a comparable and perhaps more versatile tool for helping to determine the appropriate thresholds for each of the impact indicators (Pickard and Porter 2008). Some subjectivity is required with all of these methods. In general one looks for natural patterns, clusters, and breaks in the distributions of indicators.

2.2 Determine How to Link Reach-Scale Indicators to Watershed Scale Assessments of Condition.

In order to compare WAP-lite (i.e., Tier 1 monitoring data) assessments of condition with RREE (i.e., Tier 2 monitoring data) assessments of condition, we need to first ensure they are both on the same scale. The WAP-lite assessments are at the watershed scale. The RREE assessments are at the stream reach scale. Assuming the program selects a statistically valid sample of stream reaches (Section 5), it is possible to obtain watershed scale estimates from the stream reach data. However, the program needs to decide the best way to combine the raw stream reach data to produce estimates of condition at the watershed scale.

The program could assess the condition of each stream reach using standard RREE methods (i.e., use 15 evenly weighted yes/no questions to interpret the 56 FREP metrics). Then the program could use the sampled reaches to estimate the proportion of each watershed in each of the four condition stages as described by RREE (i.e., 1) the stream and riparian habitat is not functioning properly; 2) the stream is in properly functioning condition, but at high risk; 3) the stream is in properly functioning condition, but at risk; 4) the stream is in properly functioning condition). Finally the program could compare the observed proportion of the watershed in each category against (to be determined) thresholds to determine FSW functional status. For example: if we estimate that > 50% of the watershed is properly functioning and not at risk (i.e., RREE category 4), then the FSW is properly functioning (PF); if we estimate that > 50% of the watershed is not properly functioning (i.e., RREE category 1), then the FSW is not properly functioning (NPF); else Impaired (I). These examples are hypothetical; actual thresholds would require expert input, peer review, examples from the literature, or at a minimum agreement to update as we learn.

The program may also decide to use a different roll up scheme than that used in the RREE. For example, it may suit the purposes of the FSW monitoring program to weight some indicators differently than RREE. Or, the program might want to report the average answer to each of the 15 questions across all sites in a watershed and then roll up to a watershed result. There are many different possibilities; refer to Appendix B in the FSW conceptual model document (Wieckowski et al. 2008) for guidance on alternative roll up approaches and a summary of different strategies used by a variety of monitoring programs.

3) Resolve Tier 1 monitoring, sampling design questions (ITR 3).

Tier 1 sampling design details will depend on the outcome of the tasks described above. Once the first two items have been resolved, it should be relatively easy to answer the following questions.

3.1 Is a spatial census of FSWs feasible if we are only relying on remote sense data?

This will depend on the outcome of the remote sensed data evaluation. If a spatial census of all FSWs is feasible, then there are no further sampling questions to address for Tier 1 remote sensed data.

3.2 How frequently should the Tier 1 monitoring be completed?

This depends on the cost of obtaining and interpreting the remote sense data and the annual variability of the Tier 1 metrics. If a metric changes slowly it doesn't need to be monitored as frequently. However, if a metric changes too slowly it may not be able to provide feedback on changes in watershed health in a useful time-frame.

3.3 Should field-based sampling be incorporated? If so why and when?

It is unrealistic to think that field monitoring could ever be completed for a census of watersheds. Therefore, it would be most desirable if we could establish strong correlations between the WAP-*lite* remote sense indicators and the RREE indicators and not include any field-based sampling in the Tier 1 monitoring. If a strong correlation between the WAP-*lite* and the RREE indicators cannot be found, then it may be necessary to supplement the Tier 1 monitoring with some field-based sampling (e.g., temperature or water quality). We identify two ways that this could be accomplished:

i. Use a double sampling approach (Pollack et al. 2002). Collect supplemental field-based data for a random sample of FSWs. Use the relationship between the field-based data and WAP-*lite* to infer what the field-based data would be at the FSWs where we only completed the WAP-*lite*. This may increase the cost of the program to a point on par with the original WAP Level 1. If this true, it may make more sense to undertake the original WAP Level 1 approach rather than creating a new system.

ii. Rather than completing a census of FSWs during the Tier 1 monitoring stage, we could take a random sample and complete both the WAP-*lite* and the field-based monitoring at a sample of watersheds. The framework would then proceed as planned, but with selection of a subset of watersheds for Tier 2 monitoring from the original sample of watersheds selected for Tier 1 monitoring.

The one other situation where we believe that field-based sampling should be incorporated into the Tier 1 monitoring is if no Tier 2 sampling is planned for that year. It is too risky to entirely trust such important resources to remote sensed data. Having eyes on the ground at some stage is critical to ensuring that managers do not miss something obvious which might indicate a habitat relationship has changed or that something that we hadn't thought to monitor in the first place has changed. In the manufacturing world this concept is known as management by walking around (MBWA), and it is a critical component to effective management. Even in a controlled manufacturing setting unexpected things happen which is why some degree of MBWA is always necessary. Now consider natural resource management where we have far less control over the processes and recognize that we cannot possibly foresee every scenario, making it critical to always keep some eyes on the ground.

4) Resolve Tier 2 monitoring, design details for sampling of FSWs (ITR 4).

4.1 Watershed scale sampling strata (Figure 1, Step 4):

<u>Stratification Across FSWs</u>: These strata would be used for the selection of FSWs (i.e., the primary sampling unit) for sampling from the population of possible FSWs (i.e., the target population).

- Coastal vs. Interior: Use the same classification as the WAP to accommodate physiographic and biogeoclimatic differences (Interior Watershed Assessment Procedure Guidebook (IWAP) 1995). Each watershed in the province permanently belongs to the same stratum.
- Functioning Status (Not Functioning Properly, Impaired, and Functioning Properly): Each FSW is reclassified once per cycle (i.e., each time the Tier 1 monitoring is completed). If the classification rules change it is important to re-assess the historical results according to the new rules.

Subset items to resolve:

1. Decide if the proposed stratification for selecting watersheds is acceptable.

4.2 Selection of Watersheds (Primary Sampling Unit) (Figure 1, Step 4):

Step 4 of the general sampling design (Figure 2) is to select a sample of watersheds for Tier 2 monitoring. In this section we describe the spatial and temporal selection of FSWs within the general sampling design. Section 6.3 of the FSW Conceptual Model Document (Wieckowski et al. 2008) provides a detailed discussion of the sampling concepts discussed here.

4.2.1 Spatial selection

Use a Generalized Random Tessellation Stratified (GRTS) design (Stevens & Olsen 2004) to select a random yet spatially balanced sample of watersheds from the population of watersheds. This design can incorporate strata important at the watershed scale (i.e., region and functioning status).

4.2.2 Temporal selection

We recommend generating a new sampling design (i.e., selecting a new set of FSWs for Tier 2 monitoring) at the start of each cycle rather than generating a single design now that runs far into the future (i.e., specifying the list of FSWs to be sampled for the next 50 years). It will be much easier to implement and analyze a new randomized design each cycle than to try and implement a rotating panel design across many years. Firstly, because a key part of the FSW frame-work and the general sampling design is the use of functioning status as a stratification variable. We expect the functioning status to change as a result of management actions, but we can't predict which FSWs will belong to which strata (i.e., NPF, I, PF) in the future. Secondly, because the population of FSWs (i.e., the target population) will change as new FSWs are identified. This would be less of a concern if FSWs were expected to be added at random, but it seems likely that FSWs will be added in batches according to need, historical designation, regional/local concerns and assessments etc. It is possible that there may be a systematic difference between FSWs added early on and those that were added later. If we were to lock a design in place now, we wouldn't end up including any new FSWs, perhaps resulting in a bias. Re-randomizing the selection of FSWs each cycle is a simple way to ensure that the design is always valid. The simplicity of the method is one of the most important components, as in our experience long-term monitoring programs are more likely to succeed if the people implementing them understand the design.

Within a design cycle, there are several different ways we could select watersheds across years:

a) re-visit the same set of watersheds each year for the full cycle (e.g., 5 years in a row),
b) rotate through a new set of watersheds each year, or
c) use an intermediate design where we return to some watersheds for at least two years in a row in order to maximize the sample size while still providing an estimate of annual variability.

For FSW Tier 2 monitoring we are interested in both change within individual FSWs and the overall status of FSWs. Option (a) would result in the best estimate of individual change (e.g., how does a specific FSW change over time?), but it has the lowest overall sample size for estimating status (e.g., what is the average behaviour of all NPF FSWs?). Option (b) would result in the greatest number of watersheds being sampled and so the best estimate of status, but we would not be able to estimate annual variability, we would not be able to estimate the individual change, and we would only obtain a weak estimate of average gross trend (e.g., how is the population of FSWs changing). Option (c) is an example of what is typically called a rotating panel design and tries to address both questions at the same time. Option (c) would normally be our first choice, but its value is reduced in the case of FSWs by the fact that we recommend re-randomizing at the start of the each cycle. Option (c) is the most complicated design to implement and analyze, and as previously mentioned that is a serious flaw. Overall we think that Option a) is the best choice for this project. There are additional benefits to having a set of watersheds monitored for several years in a row:

- It is easy to understand and implement logistically (e.g., consistent field crews; new FSWs are easy to incorporate).
- It provides a good understanding of annual variability, which is critical because some field metrics may vary quite a bit depending on the environmental conditions for a given year.
- It should provide a good understanding of what watershed functions are working well versus those processes showing impairment. Hopefully this will provide sufficient information to convince forest managers, operators, and other decision makers of problems on the ground.
- The data from a specific watershed might be useful for industry when they are defining their ‘special management plan’. The average behaviour of NPF watersheds compared with PF watersheds (e.g., what were the typical causes of failure) could be used to guide ‘special management plans’ even in watersheds that had not been selected.
- When a specific problem is found in a NPF watershed, this watershed could be selected for Tier 3 monitoring to test the success of a specific management action. The 5 years of Tier 2 data would provide excellent ‘before’ data in a Before-After design.
- There is a disincentive for forest licensees to have their watershed classified in the ‘Not Properly Functioning’ status category. Since this stratum will probably be weighted the heaviest, such watersheds will be more likely to be monitored in detail, bringing poor forest practices into public scrutiny.

Subset items to resolve:

1. Decide if the proposed spatial and temporal strategy for selecting watersheds is acceptable.

4.3 Determine the Number of Watersheds to Sample (Figure 1, Step 4).

The number of watersheds required to answer Tier 2 monitoring questions depends on the number of strata, the importance of each strata (e.g., we might weight NPF watersheds more heavily than PF watersheds), the variability within each strata, the annual variability, the desired precision, and cost. Refer to MOE Monitoring and Evaluation Toolbox (Pickard & Porter 2008) for details, examples, and references of pre-monitoring analyses including: sample size and power calculations.

Subset items to resolve:

1. Obtain an estimate of variability within and between watersheds within each stratum. That means at least two, but preferably at least three watersheds within each stratum to simply estimate the variability.
2. Determine the required level of precision (refer to Appendix C in the FSW conceptual model document; Wieckowski et al. 2008).
3. Complete sample size and power calculations to determine how many FSWs should be sampled within each strata.

5) Resolve Tier 2 monitoring, design details for sampling of stream reaches within FSWs (ITR 5).

5.1 Within watershed scale sampling strata (Figure 1, Step 5)

<u>Stratification Within FSWs</u>: The intention is to use the RREE protocols to complete the field sampling in the FSWs. However, FSW monitoring will be focused on assessing watershed condition, and not just condition at a cut-block as is the case for the RREE. As a result selection of sample sites will be unique for FSWs, even if RREE field protocols are employed once a site is selected.

- <u>Stream order:</u> Streams of Strahler stream order 1-2 are often not fish bearing but are still important to the health of the watershed (Myer et al. 2003). They may be harder to access but are probably easier to sample once there. Streams of Strahler stream order 3 and higher are typically fish bearing.. They may be easier to access since they are lower in the watershed, but there may be other logistical difficulties when sampling due to the increased size of the streams. The RREE protocols use a different definition of stream class based on channel width, these range from 1.5m to 20m in width. Strahler stream order will be much easier to use as a stratification variable because this information is already available with provincial GIS layers and so does not require field measurements to define the population. Generally speaking, stratification is used to make the sampling design more efficient. It is most useful when the variability between strata is large compared to variability within a stratum (Cochran 1977). However, more strata are not necessarily better (Pickard and Porter 2008). In order to decide on the optimal number and grouping of stream order strata (e.g., two strata, a=1-2, b=3 or higher; or 4 strata, a=1, b=2, c=3, d=4 or higher etc.), consider the tradeoff between precision and cost as more strata are added to the design.
- <u>Cutblock vs. not cutblock:</u> The only reason to consider cut-blocks is to enable better integration of the FREP and FSW monitoring. If we stratified on cut-blocks we would be able to produce estimates for cut-blocks alone as well as for the entire watershed and so FSW sites could be used to supplement the FREP study.

Subset items to resolve:

1. The grouping and relative importance of different stream orders:
   - If a stream is > 20m wide, will the RREE methods still work?
   - Do we want to include all stream orders in the sample?

- How should they be grouped? Using all 6 stream orders separately may not be a very efficient design.
- How many of each stratum? Proportional allocation or weighted?

2. Should we include a cutblock stratum or not?

5.2 Selection of Sample Reaches Within a Watershed (Secondary sampling unit) (Figure 1, Step 5).

5.2.1 Spatial selection:

Use a Generalized Random Tessellation Stratified (GRTS) design (Stevens & Olsen 2004) to select a random but spatially balanced sample of points from a given watershed (see example for Step 5 in Figure 1). The selection of points would incorporate the within-watershed scale strata of importance (i.e., stream order and possibly cut-blocks). The RREE protocol recommends a minimum sampling unit of 30 channel widths, but the GRTS sample will only define a point in the watershed. The randomly selected GRTS point can be interpreted as the mid-point (or an endpoint, decide before you get there) of the sampled reach (Figure 3).

5.2.2 Temporal selection:

For Tier 2 monitoring, we recommend re-randomizing the selection of sites within the watershed each year in order to maximize the sample size for the watershed. The question we are most interested in is the overall status of the watershed and net changes in this overall status over time. We are not interested in following a small part of the watershed over time, although this might be incorporated into a Tier 3 monitoring design depending on the question.

Subset items to resolve:

1. Decide if the proposed spatial and temporal strategy for selecting watersheds is acceptable.

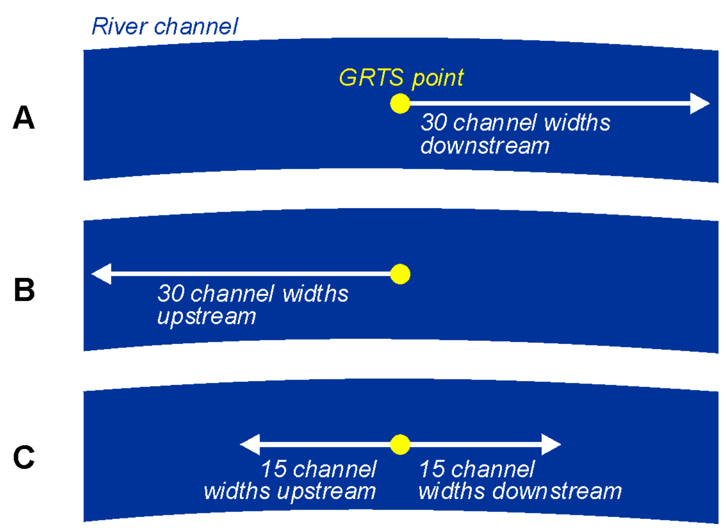


Figure 3. The sampling unit (i.e., 30 channel width section of the stream) can be located by starting at the GRTS point and moving down stream 30 channel widths (A), or moving upstream 30 channel widths (B), or the GRTS point can be the mid-point of the sampling unit (C). All three orientations are valid provided that it is decided prior to arriving at a site.

5.3 Determine the Number of Sites to Sample Within the Watershed (Figure 1, Step 5)

The required number of sites (i.e., sample reaches) within a watershed to answer Tier 2 monitoring questions depends on the same factors described for watershed selection (Section 4.3 above).

Subset items to resolve:

1. Obtain an estimate of variability (of stream reach indicators) within and between strata
    - Review existing studies where RREE monitoring has occurred (Derek Tripp has indicated that he can provide these data for the province)
        - obtain preliminary estimates of variability for different potential groupings of the stream order strata
        - obtain preliminary estimates of variability for different reaches within a watershed
2. Determine Per/Sample Costs for Sampling within a Watershed
    - Evaluate potential costs for sampling based on historical RREE projects and sampling of remote sites that was undertaken for past provincial fisheries inventory exercises within BC.
3. Determine the required level of precision (refer to Appendix C in the FSW conceptual model document (Wieckowski et al. 2008))
4. Complete sample size and power calculations to determine how many stream reaches should be sampled within each strata and FSW.

Overall Approach

We have identified a number of items to resolve to successfully implement the Fisheries Sensitive Watershed Framework. The concept of using a Tiered FSW monitoring framework is the key component that needs to be assessed to determine if the proposed framework is feasible. ITRs 1-2 address this need. If we can't resolve these items, the entire FSW framework will have to be re-visited. As such, the initial focus for the workplan should be ITRs 1-2 until success or failure of the framework has been established. However, many of the later Items can be tackled opportunistically as data is collected for ITRs 1-2. In this section we provide some suggestions to help MOE proceed under a number of different budget scenarios.

In an ideal world a random sample of FSWs would be selected and both the WAP-*lite* and the RREE assessments would be completed for each of the FSWs. Different filters could then be applied to the WAP-*lite* data and compared to the RREE outcomes. In reality there may be limited funding available for monitoring. Here are some strategies that could be used to accommodate varying levels of funding:

i. <u>Availability of data:</u> Select watersheds where RREE, WAP Level 1, and other remote sensed data are already available. Test out the WAP-*lite* indicators in these watersheds first. Examples may include:
    - Bowron River Watershed: RREE and WAP Level 1 have been completed here;
    - Lake Else: currently implementing WAP evaluations; or
    - 53 sensitive watersheds in the Charlottes: not yet established as FSWs, but did have WAPs completed historically.

ii. Proximity of watersheds: Select watersheds that are close to MOE offices in Victoria to enable Ministry staff to complete the field work themselves.

iii. Contrast in watersheds: As a minimum select two watersheds; one that is expected to be 'not properly functioning' and one that is expected to be 'properly functioning'. Select one of each from within either coastal FSWs or else from Interior FSWs. Eventually you will want to evaluate FSWs in both strata, but if budgets are limited choose two watersheds from within the same strata first.

Each of these strategies represents 'convenience sampling' (i.e., their non-random selection would not be considered sufficiently robust to allow inference to the larger population of FSWs). While these strategies are not ideal, the unfortunate reality is that these approaches may be all that are affordable. We can use convenience samples to develop the initial WAP-*lite* indicators (ITR 1) and the initial filter for classifying FSW condition (ITR 2). The program will however have to remain flexible enough to allow the classification filter to be improved as more data is collected. As part of the proposed FSW monitoring framework (Appendix B) the filter can be periodically updated. If new watersheds are added each year, eventually a sufficient number of watersheds will have both WAP-*lite* and RREE monitoring in place to be able to make inferences to the population of FSWs. Until the WAP-*lite* protocols and the associated condition filter have been fully validated it is important to sample all strata (NPF, I, and PF). Once the classification procedure has been established on a reliable sample of all three strata, we may wish to weight the strata differently (i.e., it will likely be more useful to focus principally on the NPF watersheds).

In order to determine the required number of watersheds (ITR 4), we need to monitor multiple watersheds within each of the proposed strata. As described above, if possible we'd select the pilot watersheds randomly, but it is probably unrealistic. Data from multiple watersheds are needed within each stratum to complete the sample size and power calculations. It may not be possible to collect sufficient data to even complete power calculations for several years. However, Tier 2 data have to be collected to address ITR 2, and so the sample size calculations can be addressed simultaneously.

In order to determine the required number of sites within a watershed (ITR 5), we need to monitor multiple stream reaches within each of the proposed strata. For whichever watersheds are selected to address ITR 1-2, we can simultaneously address this question. We recommend that the program initially oversample the watershed and document the effort (time and cost) to complete each sample and to move between reaches. Then use this information to complete power analyses to determine how many samples are required to yield the desired level of power.

As described above, it would be ideal to randomly select the pilot watersheds, but it is probably unrealistic. Estimates of variability from non-randomly selected watersheds won't necessarily be representative of the target population, but they will be a reasonable starting place. Furthermore, if sites are not selected randomly it would be prudent to keep the sample size a bit on the high side until sufficient data have been collected to be confident in the sample size requirements.

Appendix A. Sample Form to Compare Potential Remote Sense Data Sources

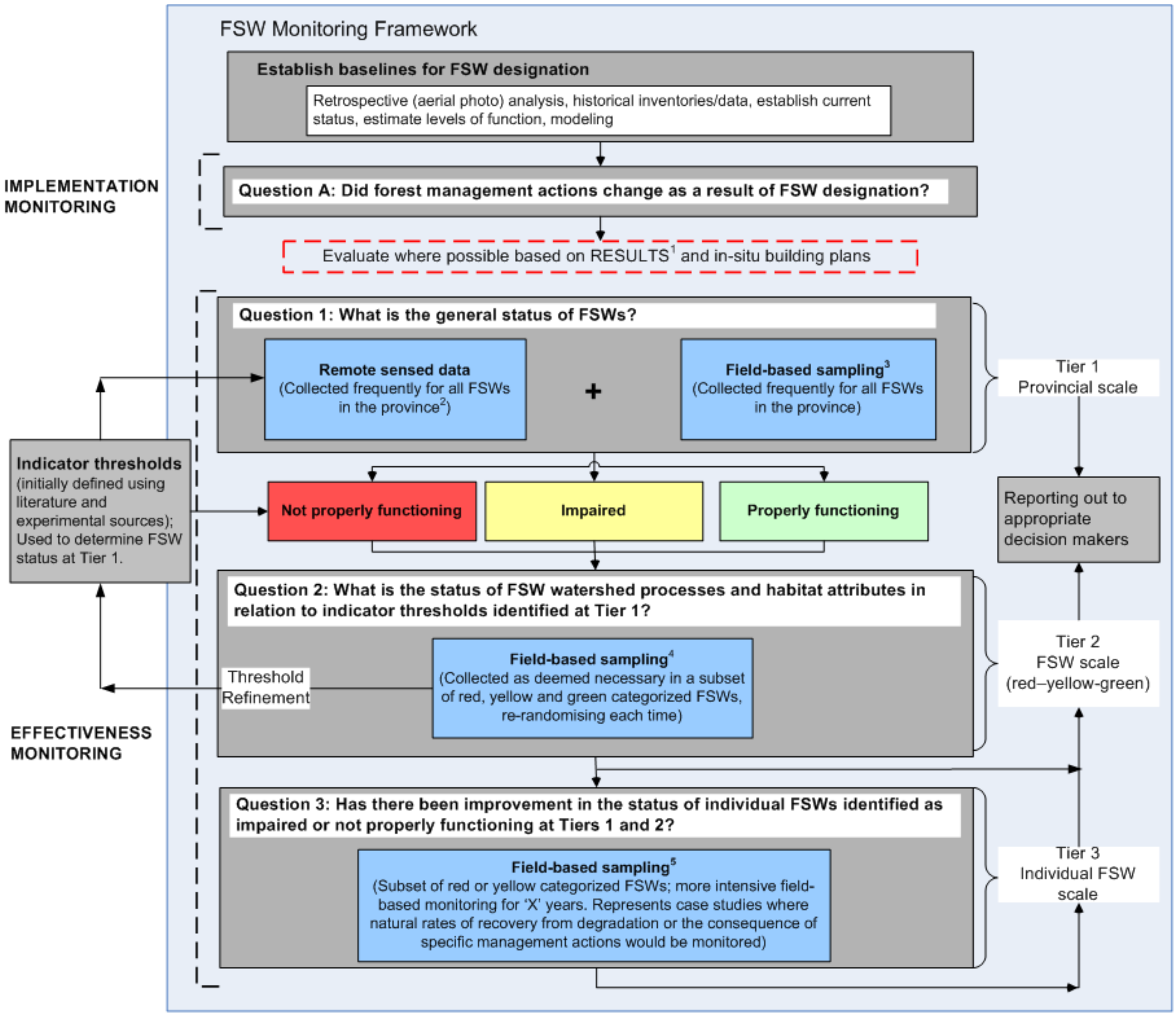
Method:__________________ compared with WAP Level 1

Realistic frequency of sampling:
Spatial scale available (e.g. whole province):
Cost to obtain data:
Cost to process or interpret data:
Time to process or interpret data:

WAP level 1 indicator	Description	stressor	Quality of data (better, worse, about the same)	Processing effort (cost/time)
1	Peak flow index	peak flow		
2	Road density above H60 line	peak flow		
3	Road density (used for assessing peak flow changes)	peak flow		
4	Density of roads on erodible soils	peak flow		
5	Density of roads less than 100m from a stream	surface erosion		
6	Density of roads on erodible soils less than 100m distance from a stream	surface erosion		
7	Density of stream crossings	surface erosion		
8	Road density (used for assessing surface erosion)	surface erosion		
9	portion of streams that have been logged to the stream bank	riparian buffer		
10	portion of fish-bearing streams that have been logged to the stream bank	riparian buffer		
11	Density of landslides in the watershed	mass wasting		
12	density of roads on unstable or potentially unstable terrain	mass wasting		
13	Portion of streambanks that have been logged on slopes >60%	mass wasting		

Alternative indicators available with this method				
Indicator #	Description	stressor	In what way is it more or less useful than any of the WAP Level 1 indicators	Processing effort (cost/time)

Appendix B. FSW Monitoring Framework



1. RESULTS is the MOF's Reporting Silviculture and Land Status Tracking System
2. Monitoring of a subset of non-FSWs could also be used for comparative purposes. Such monitoring would likely only focus on those watersheds that display high sensitivity but have not been designated FSW because of low fisheries value, and could thus be paired with similar FSWs.
3. Field-based sampling at Tier 1 will focus on a few coarse-scale metrics and could be based on a rotating panel design. In a rotating panel design only a certain proportion of all FSWs would be monitored in any particular year with revisits occurring based on a set return cycle. Continuation of field-based sampling at Tier 1 may be dependent on whether strong linkages can be made between field-based and remote sensed sampling. If strong linkages can be established it may be sufficient to use only remote sensed data to evaluate and stratify FSW watersheds at the Tier 1 level.
4. Sampling design, effort, and frequency at Tier 2 to be determined during pilot study. Will rely principally on FREP field protocols for data collection related to selected FSW indicators.
5. Allocation of effort between FSW categories (i.e., red, yellow, green) in Tier 2 is to be determined. For example, may choose to allocate equal effort across all three categories until data is available to inform how sampling effort should be allocated across watershed categories.
6. Tier 2 will provide validation monitoring for Tier 1 to ensure that red/yellow/green classifications are appropriate. In addition, Tier 2 sampling will provide a means to assess causes of indicator failure at the Tier 1 level.
7. Tier 3 sampling design will be dependent on what causal factors were identified in Tier 2. Extrapolation of Tier 3 results will be dependent on the way in which case study watersheds are selected (i.e., random vs. non-random).

Literature Cited

Cochran, William Gemmell. 1977. Sampling Techniques. John Wiley & Sons, Inc., New York, NY.

Interior Columbia Basin Technical Recovery Team. 2005. Viability criteria for application to Interior Columbia Basin Salmonid ESUs. Review Draft, 2007. URL http://www.nwfsc.noaa.gov/trt/trt_documents/ictrt_viability_criteria_reviewdraft_2007_complete.pdf

Nordin, L., D. Maloney, J. Rex, and P. Krauskopt.2008. The Bowron River Watershed: a landscape level assessment of post-beetle change in stream riparian function. Ministry of Forests and Range, Northern Interior Forest Region, Prince George, BC.

Pollack, Kenneth H., James D. Nichols, Theodore R. Simons, George L. Farnsworth, Larissa L. Bailey, and John R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics. 13:105-119.

B.C. Ministry of Forests. 1995a. Interior Watershed Assessment Procedure Guidebook (IWAP), Level 1 Analysis. Forest Practices Branch, Ministry of Forests, Victoria., B.C. Forest practices code of British Columbia Guidebook.

B.C. Ministry of Forests. 1995b. Coastal Watershed Assessment Procedure Guidebook (CWAP), Level 1 Analysis. Forest Practices Branch, Ministry of Forests, Victoria., B.C. Forest practices code of British Columbia Guidebook.

Myer, J.L. et al. 2003. Where rivers are born: the scientific imperative for defending small streams and wetlands. American Rivers and Sierra Club.

Pickard, D. and M. Porter. 2008. Monitoring and Evaluation Toolbox. Draft report prepared by ESSA Technologies Ltd., Vancouver, B.C. for Kathy Paige, Ecosystems Branch, B.C. Ministry of Environment, Victoria. 51 pp.

Reese-Hansen, L. and E. Parkinson. 2006. Evaluating and designating Fisheries Sensitive Watersheds: an overview of BC's new FSW procedure. B.C. Ministry of Environment.

Stevens, D.L., and A.R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99 (465): 262-278.

Tripp, D. 2007 (draft). Development and testing of extensive-level indicators and methods for determining if current forestry practices are sustainably managing riparian, aquatic ecosystem, and fish-habitat values. Prepared for Ministry of Forests Research Branch, Ecology and Earth Sciences Section, Victoria, BC.

Wieckowski, K., D. Pickard, M. Porter, D. Robinson, D. Marmorek, and C. Schwarz. 2008. A conceptual framework for monitoring Fisheries Sensitive Watersheds (FSW).. Report prepared by ESSA Technologies Ltd. for BC Ministry of the Environment (MOE), Victoria, BC. 61 p.

Wieckowski, K., M. Porter, D. Marmorek, and D. Pickard. 2009. A framework for monitoring Fisheries Sensitive Watersheds (FSW). Report prepared by ESSA Technologies Ltd. for BC Ministry of the Environment (MOE), Victoria, BC. 11 p.