A Framework for Monitoring Fisheries Sensitive Watersheds (FSW)

Version 4

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

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

March 16, 2009

Introduction

Fisheries Sensitive Watersheds (FSWs)[1] have been designated in British Columbia (BC) under the *Forest and Range Practices Act* (FRPA) and the *Government Actions Regulations* (GAR)[2]. To qualify as a FSW, watersheds must meet two criteria: they must have significant fisheries values and watershed sensitivity. Watersheds which have been designated as FSWs by the Minister require Forest Act agreement holders to establish results and strategies in their Forest Stewardship Plans consistent with the objective(s) set by the Minister. FSW designation has been undertaken for two reasons. First, designation is intended to conserve natural hydrological conditions, bed dynamics and channel integrity, as well as the quality, quantity, and timing of water flow. Second, designation is intended to prevent cumulative effects that would have adverse effects on fish habitat. Ultimately, the goal of FSW designation is to conserve fish habitat and the natural functions and processes required to maintain fish habitats now and in the future, while forest management activities proceed. For a description of the process for designating a watershed as a FSW refer to Reese-Hansen and Parkinson (2006). Effectiveness monitoring is required to determine if FSW designation has achieved this goal.

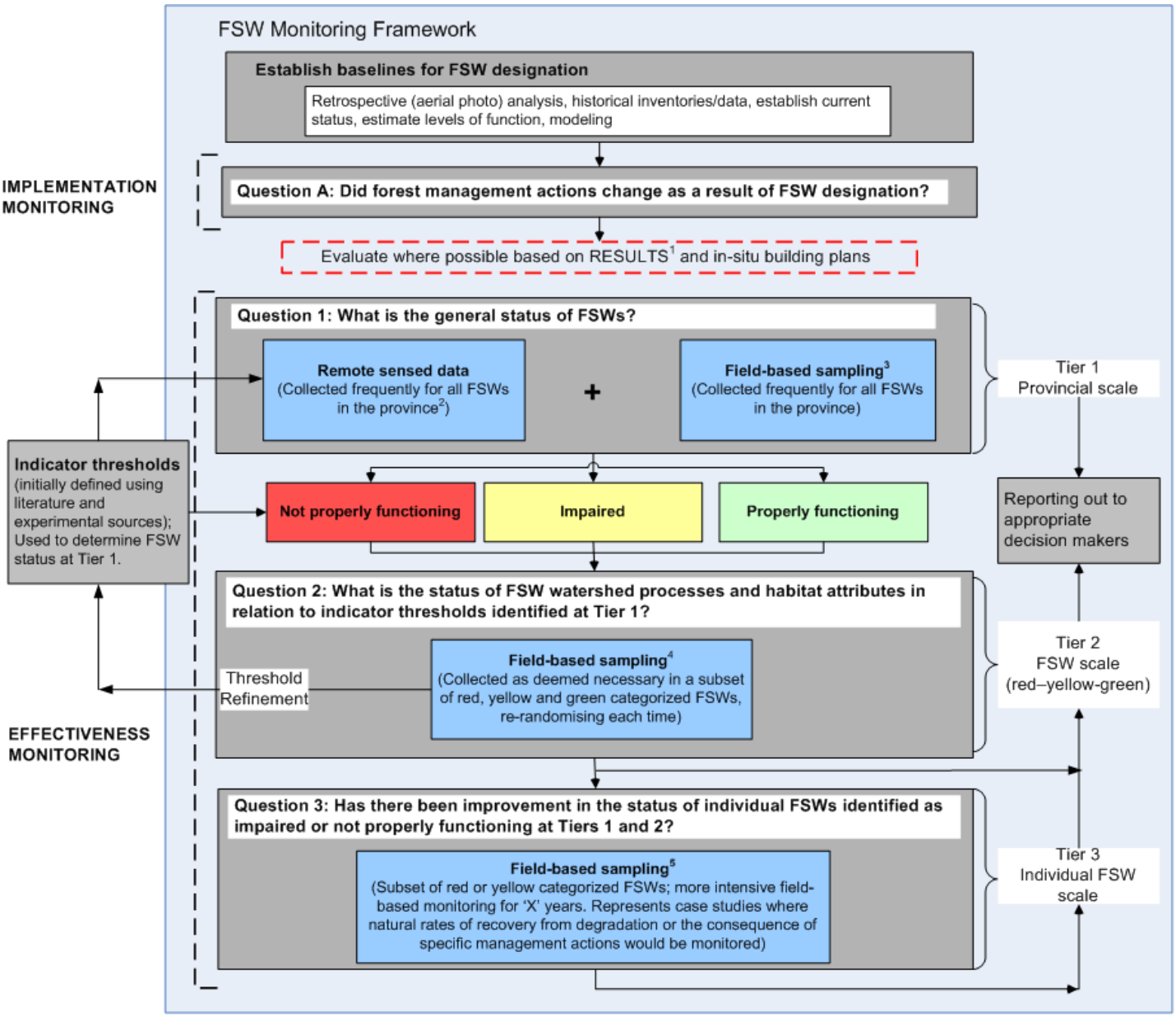
MacDonald et al. (1991) describe effectiveness monitoring as "an assessment of whether specified activities had the desired effect". They suggest that effectiveness monitoring should include both the overall plan and the individual components of a project in order to assess management actions and to learn from the results. For example, monitoring effectiveness solely at the broad scale of the overall plan may reveal if a plan is succeeding or failing, but if failing, will not to reveal the reason for failure. Additionally, monitoring the overall effectiveness of large-scale land management plans, while useful for supporting long term conservation strategies, will generally provide interpretable results only after extended periods of time (Reid and Furniss 1998). A protracted learning period provides limited support to shorter-term adaptive management decisions, reducing the benefits to managers. Reid and Furniss (1998) suggest that all "state-of-the-art' management plans incorporate an element of adaptive management, such that outcomes of actions are monitored in a manner that allows future decisions to benefit from experience. Consequently, the effectiveness of individual components of a larger management plan that have shorter temporal and smaller spatial scales should also be evaluated so as to provide direct feedback on the effectiveness of specific management actions. This subset of monitoring must provide managers with information that identifies necessary adaptations, and it must do so as quickly as possible so that changes can be made before degradation proceeds too far (Reid and Furniss 1998). If monitoring does not provide interpretable results over the time-span needed for adaptive management, the success of broad-scale programs will be compromised.

Monitoring Framework

The framework described in **Figure 1** presents a nested, multi-tiered approach to FSW implementation and effectiveness monitoring. Each monitoring tier presented in **Figure 1** refers to the spatial scale (i.e., provincial, across FSWs, within individual FSWs) at which particular monitoring questions should be addressed. These questions will in turn dictate the monitoring components required. For the framework to be

[1] Ministry of Environment. Fisheries Sensitive Watersheds. Available at: http://www.env.gov.bc.ca/wld/frpa/fsw/index.html
[2] Ministry of Forests and Range. Government Action Regulations. Available at: http://www.for.gov.bc.ca/tasb/legsregs/frpa/frparegs/govact/gar.htm

effective the monitoring components must be carefully designed to answer well-defined questions at different scales in an efficient manner.



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).

Figure 1. Conceptual diagram of a monitoring framework protocol incorporating multiple tiers of evaluation for assessing the effectiveness of Fisheries Sensitive Watersheds (FSWs).

Management Questions

There is one "management question" relating to implementation monitoring for FWS. This is:

Have forest management actions changed as a result of FSW designation?

Determining forest management actions that may have changed specifically in response to FSW designation will likely be very difficult (given the large suite of concurrent variables that will affect logging rates and activities). Information from the Ministry of Forest's Silviculture and Land Status Tracking System (RESULTS) possibly in combination with in-situ building plans may, however, allow some assessment of whether forestry changes have occurred that are directly attributable to FSW designation.

There are three tiers of "management questions" relating to effectiveness monitoring that could inform the decision making process for FSWs. The **first tier** of management questions should evaluate the ongoing success of FSW designation as a management action:

What is the overall status of FSWs across the province (i.e., proportions of properly functioning, not properly functioning, and intermediate category watersheds)?

Tier 1 monitoring is targeted at the provincial scale and would include all designated FSWs. A subset of non-designated watersheds with high watershed sensitivities comparable to FSWs (but with low fisheries values thus precluding their FSW designation) could also be evaluated. These non-designated watersheds would be paired with similar FSWs to allow comparison of trends within FSW vs. non-FSW watersheds (i.e., a BACIP[3] experimental design, Stewart-Oaten et al. 1986; Underwood 1991). This additional element of evaluation (i.e., FSWs vs. non-FSWs) would be informative but is not considered a core component of the FSW monitoring framework, which may focus solely on FSWs. Since fish are effective at seeking out their preferred habitats, it may not always be possible to find paired watersheds with similarly high sensitivities (and presumably similar habitats) but very different fisheries values. The presence or absence of barriers might help to create watersheds with similarly high sensitivities but contrasting fisheries values.

Tier 1 is designed to evaluate the long-term success of the overall program of FSW designation (i.e., what is the general success of FSW designation as a broad management action). The analysis of remote sensed data and supplementary coarse-scale field data at the Tier 1 level will provide a point of comparison for evaluating long term trends in the status of FSWs. Selected indicators will also be used to define a baseline classification of monitored FSWs (i.e., considered to be properly functioning (green) or not properly functioning (red)[4,5], or some intermediate category (yellow; signs of deteriorating condition but still considered functioning). From here on, properly functioning and not properly functioning will be referred to in this document as PF and NPF, respectively. The assignment of these categories will involve scientific judgments on what constitutes PF or NPF for a specific indicator, and how to appropriately combine indicators given varying conditions and differing levels of uncertainty in the assigned thresholds. It is anticipated that remote-sensed habitat indicators historically used for the province's Level 1 Watershed Assessment Procedure (WAP) (BC MOF 2001) will provide the foundation for categorizing FSW condition at the Tier 1 level.

[3] Before After Control Impact Paired sampling design (BACIP)

[4] The terms properly functioning (PF), impaired, and not properly functioning (NPF) have not yet been defined for use within the FSW monitoring framework. Within the context of this report the terms are used to refer to categories of watersheds that have been demarcated using a set of defined criteria. The task of defining PF, impaired and NPF, as well as the refinement of criteria, is beyond the scope of this document, and will be developed within planned pilot studies. It is possible that a range of watershed categories will ultimately be developed, covering a gradient of condition (e.g., good, moderate, bad, etc.).

[5] An initial starting point for a definition of properly functioning could be a watershed that safely stores and releases water, has vegetative cover that adequately balances exposed sediment sources, and has normal levels of connectivity between all aquatic components that are present in the watershed.

The **second tier** of management questions would examine the condition of a subset of FSWs that are considered to be PF (i.e., green categorization at the Tier 1 scale) versus those that are considered to be NPF or impaired (i.e., red or yellow categorization at the Tier 1 scale):

> **What is the status of watershed processes in PF versus NPF or impaired FSWs (i.e., what is causing particular FSWs to be classified as NPF or impaired at Tier 1)?**

Tier 2 monitoring is designed to provide shorter-term information on the effectiveness of the program in maintaining or improving individual ecosystem processes across FSWs. Tier 2 monitoring serves two functions. First, more detailed field-based monitoring will provide insight into the cause of general problems (initially identified as red or yellow categorizations for the FSW at the Tier 1 level) and continued monitoring will provide information on the ongoing status of FSWs in regards to a suite of key ecosystem processes and habitat attributes. Second, data collected from detailed monitoring of a varied set of FSWs can be used to validate, refine, or revise initial threshold criteria used at the Tier 1 level for defining NPF or impaired watersheds. For example, Tier 2 monitoring may provide information suggesting that the magnitude/extent of a specific type of activity (e.g., road construction) results in observed environmental degradation of magnitude/extent 'X' (e.g., sediment loads). This information can then be used to refine the initial threshold for road density used at the Tier 1 level. It is important to emphasize the need for random sampling at the Tier 2 level to ensure that the data collected in selected watersheds is representative of the FSW watershed category as a whole (i.e., NPF, impaired, and PF).

The **third tier** of management questions asks how damaged areas of NPF or impaired FSWs are recovering over time (either passively or a result of directed management actions):

> **Has there been improvement in watershed processes within individual FSWs designated as NPF or impaired?**

Tier 3 monitoring focuses on a further subset of FSWs selected from those that were identified as NPF or impaired at the Tier 1 level and subsequently monitored in Tier 2. The intention of Tier 3 management questions is to determine whether improvement (or at least stabilization) in the status of individual FSWs has taken place. Tier 3 monitoring would concentrate on areas within the selected watersheds where the greatest problems have been identified. For these focal areas it would be expected that sampling frequency and intensity would be increased beyond that undertaken as part of Tier 2 monitoring. Tier 3 monitoring would provide direct feedback to licensees on the outcomes of any specific management actions they might undertake, (i.e., inform shorter-term adaptive management decisions), would help government enforcement staff determine where best to focus their compliance efforts, and could ultimately inform changes to regulations. Any particular action to be evaluated would be those undertaken by operators in response to discussions with and recommendations from government officials within the FSW program. The nature of the monitoring questions at this scale will be dependent on which watershed processes were originally identified as failing in response to the Tier 2 management question. In addition, depending on the nature of the management question at this Tier, it may not be necessary to select watersheds at random (i.e., there may only be an interest in interpreting the results for the watershed being monitored and not extrapolating up to a larger group of watersheds.)

Appendix A provides a hypothetical example of how the monitoring tiers would feed into each other, i.e., how information from watersheds would be tracked and passed from one tier to another.

Monitoring Questions

For each of the three broad "management questions" listed above there will be specific "monitoring questions" that will need to be asked. For example, at the Tier 1 scale it will be important to know the

condition of FSWs with respect to a specific indicator, or suite of indicators, that will inform the general designations of PF, NPF, or impaired. Managers may need the answers to monitoring questions such as:

What is the status of roads, mass wasting events, or vegetation cover in FSWs?

What is the status of water quality in the FSWs?

What is the overall status of fish habitat in FSWs?

Etc.

Indicator Development

Development of monitoring questions will require that specific indicators be chosen which will act as the basis for comparisons. Indicators to address these questions must be carefully selected to avoid ambiguous results (Mulder et al. 1999; Reeves et al. 2004). Numerous authors (e.g., Noss 1990; MacDonald 1991; Barber 1994) discuss the properties of useful monitoring indicators. Reid and Furniss (1998) summarize these discussions and suggest five key indicator properties:

1. An indicator must respond quickly enough to provide results in the time-frame desired;
2. The cause-effect relationships that control the indicator response must be well understood;
3. Changes in indicator values must be interpretable in terms of the objectives of the program;
4. The indicator signal must be separable from environmental noise; and
5. Indicator measurements must be cost-effective at the required level of precision and accuracy.

The overall plan for FSW monitoring will depend on information relating to hydrological conditions, streambed dynamics and biophysical conditions. Selection of useable indicators will likely come from remote sensed information on channel disturbance, and upslope soil and vegetation condition (for Tier 1 monitoring), as well as field-based inputs required to fully inform the full suite of ecosystem components, processes, and spatial relationships at all monitoring tiers. Current advances in and increased ease of access to information from remote sensing technologies certainly suggest that broad landscape based Tier 1 monitoring of FSWs is feasible. The selection of field-based indicators across the tiers can be derived primarily from the rich set of monitoring methodologies developed for BC's existing Forest and Range Evaluation Program (FREP) (Tripp et al. 2008). It is likely that a large suite of indicators will need to be evaluated to fully answer management questions at different tiers. In this case, data aggregation strategies that pool multiple indicators (e.g., FREP, EPA, AREMP programs) to create composite scores for each tier may have to be developed. At the Tier 1 scale for instance, combining remote-sensed indicators of roads, vegetation cover, and land use into an overall score could provide the basis for a useful high level index for determining watershed condition (e.g., red vs. yellow vs. green).

Sampling Questions / Statistical Issues

The desired level of significance (alpha) to detect whether a difference exists and/or whether a threshold has been exceeded (i.e., whether or not we should reject the null hypothesis) for any particular indicator or suite of indicators remains to be defined. Framing this element requires development of "sampling questions". A sampling question that takes into account the desired level of significance to detect a statistical difference at the Tier 1 monitoring level might be structured as follows:

We will accept only a 10% chance of saying that the status of a FSW indicator is changed when in fact it has not changed (i.e., 0.1 probability of having a false positive– Type I error)

A second sampling issue relates to statistical power, i.e., the ability to actually detect change in the condition of a FSW when it occurs. Statistical power is affected by the following: i) measurement

variability; ii) the probability of Type I error (alpha); iii) sampling intensity; and iv) the effect size (i.e., the larger the impact, the more likely it is to be detected) (Underwood and Chapman 2003). If the FSW monitoring program intends to take a precautionary approach, then the sampling design should have considerable power. The structure for a sampling question addressing statistical power might be:

> **We want to be 80% certain of detecting a change in a FSW indicator when it has occurred (i.e., power of 0.8). This means that we are willing to accept a 20% chance of saying there has not been a change when one has occurred (i.e., 0.2 probability of having a false negative – Type II error).**

The necessary sampling design/effort required to effectively address particular sampling questions with desired levels of both significance and power will have to account for several factors such as the sensitivity of the indicators employed, the resources available to the FSW monitoring program, and the actual degree of certainty needed for the associated management decisions.

Literature Cited

Barber, M.C. 1994. Indicator development strategy. US Environmental Protection Agency. EPA/620/R-9, Washington, D.C., USA.

B.C. Ministry of Forests (B.C. MOF). 2001. Watershed assessment procedure guidebook 2nd ed., Version 2.1. Forest Practices Branch, Ministry of Forests, Victoria, B.C. Forest Practices Code of British Columbia Guidebook.

MacDonald, L.H., A.W. Smart, and R.C. Wissmar. 1991. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska. US Environmental Protection Agency, Region 10, EPA/910/9-91-001, Seattle, Washington, USA.

Mulder, B.S., B.R Noon, T.A. Spies, M.G. Raphael, C.J. Palmer, A.R. Olsen, G.H. Reeves, and H.H. Welsh, Jr. 1999. The strategy and design of the effectiveness monitoring program for the Northwest Forest Plan. Gen. Tech. Rep. PNW-GTR-437. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 138 p.

Noss, R.F. Indicators for monitoring biodiversity: a hierarchical approach. Conservation Biology 4(4): 355-364.

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. Available at: http://www.env.gov.bc.ca/wld/frpa/fsw/index.html

Reeves, G.H., D.B. Hohler, D.P. Larsen, D.E. Busch, K. Kratz, K. Reynolds, K.F. Stein, T. Atzet, P. Hays, and M. Tehan. 2004. Effectiveness monitoring for the aquatic and riparian component of the Northwest Forest Plan: conceptual framework and options. Gen. Tech. Rep. PNW-GTR-57. Portland. OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 71 p.

Reid, L.M. and M.J. Furniss. 1998. On the use of regional channel-based indicators for monitoring. U.S. Forest Service. Arcata, CA. 24pp.

Stewart-Oaten., A., W.W. Murdoch, and K.R. Parker. 1986. Environmental impact assessment: pseudoreplication over time? Ecology 67: 929-940.

Tripp, D.B., P.J. Tshchaplinski, S.A. Bird and D.L. Hogan. 2008. Protocol for evaluating the condition of streams and riparian management areas (riparian management routine effectiveness evaluation). Forest and Range Evaluation Program, B.C. Ministry of Forests and Range, and B.C. Ministry of Environment. Victoria, BC.

Underwood, A.J. 1991. Beyond BACI: experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian Journal of Marine and Freshwater Research 42: 569-587.

Underwood, A.J. and M.G. Chapman. 2003. Power, precaution, Type II error and sampling design in assessment of environmental impacts. Journal of Experimental and Marine Biology and Ecology 296: 49-70.

Appendix A: Hypothetical example of nested Tiers 1 through 3 of the FSW monitoring protocol for a subset of FSW watersheds in southeastern BC

NOTE: All numbers used in this example are illustrative in nature and are not representative of reality

Watersheds are identified as either properly functioning (PF), not properly functioning (NPF), or impaired at Tier 1 using a set of agreed upon criteria and associated thresholds.

e.g., 100 and 50 FSWs are categorized as PF and NPF, respectively.

NOTE: A third (impaired) category could also be incorporated (or more, depending on how many condition categories it was ultimately felt necessary to delineate) but we have confined this simple example to the PF and NPF categories for illustration

NOTE: Comparison of the status and trends of paired FSWs and non-FSWs (both PF and NPF) could also be made at the Tier 1 scale if this was seen as a useful component within an overall evaluation of the FSW program.

Moving forward to Tier 2, a random sample of FSWs would be taken from each category (i.e., PF or NPF) to undergo more detailed field monitoring (the exact nature of Tier 2 monitoring remains to be determined, but will be based ultimately on the management questions of interest, the indicators selected, the sampling design, and the monitoring protocols chosen). In addition, sampling within a watershed should be done using a randomized approach for site selection to ensure that sites in the watershed are representative of the watershed as a whole.

10 of the 100 FSWs categorized as PF and 20 of the 50 FSWs categorized as NPF[6] are randomly selected. These 30 watersheds would then become the focus of Tier 2 monitoring[7].

Based on Tier 2 sampling, the data collected may suggest that 50% of NPF watersheds (i.e., 10 of 20) have been classified NPF at the Tier 1 level because they have sediment loads greater than what would be considered to be within the natural range. Government (i.e., scientists and managers) may then decide that this is an issue worth exploring further in order to better inform forest management actions.

One watershed is selected at random from these 10 NPF FSWs to move on to Tier 3 monitoring[8].

Tier 3 monitoring would focus on the particular problem identified at Tier 2. The intention of Tier 3 monitoring is to get a better sense of the temporal variability in the watershed, to determine whether improvement (or at least stabilization) in the status of the individual FSW is occurring, and hopefully to learn how different management actions might improve the problem. Progress in the latter element would be achieved through a collaborative dialogue between government and forestry operators, where some kind of specific management action could be agreed upon once a particular problem has been identified (e.g., try leaving an extra 5m riparian buffer in some areas to reduce fine sediment inputs, etc.). **Figure 2** provides a map-based illustration of this design approach for a subset of FSW watersheds located in southeastern BC.

[6] The desired amount of sampling effort to occur within each category has not yet been determined. The numbers presented here are simply for illustrative purposes. Because the sample from each category was taken at random it is representative of the category as a whole.

[7] The relative weighting of effort in PF and NPF will depend on the type of question being asked. The weighting of effort in this example is for illustrative purposes only.

[8] The number of watersheds that move on to Tier 3 will depend on the extent of resources available. Also, depending on the nature of the management question at this Tier, it may not be necessary to select watersheds at random (i.e., there may only be an interest in interpreting the results for the watershed being monitored and not extrapolating up to a larger group of watersheds.)

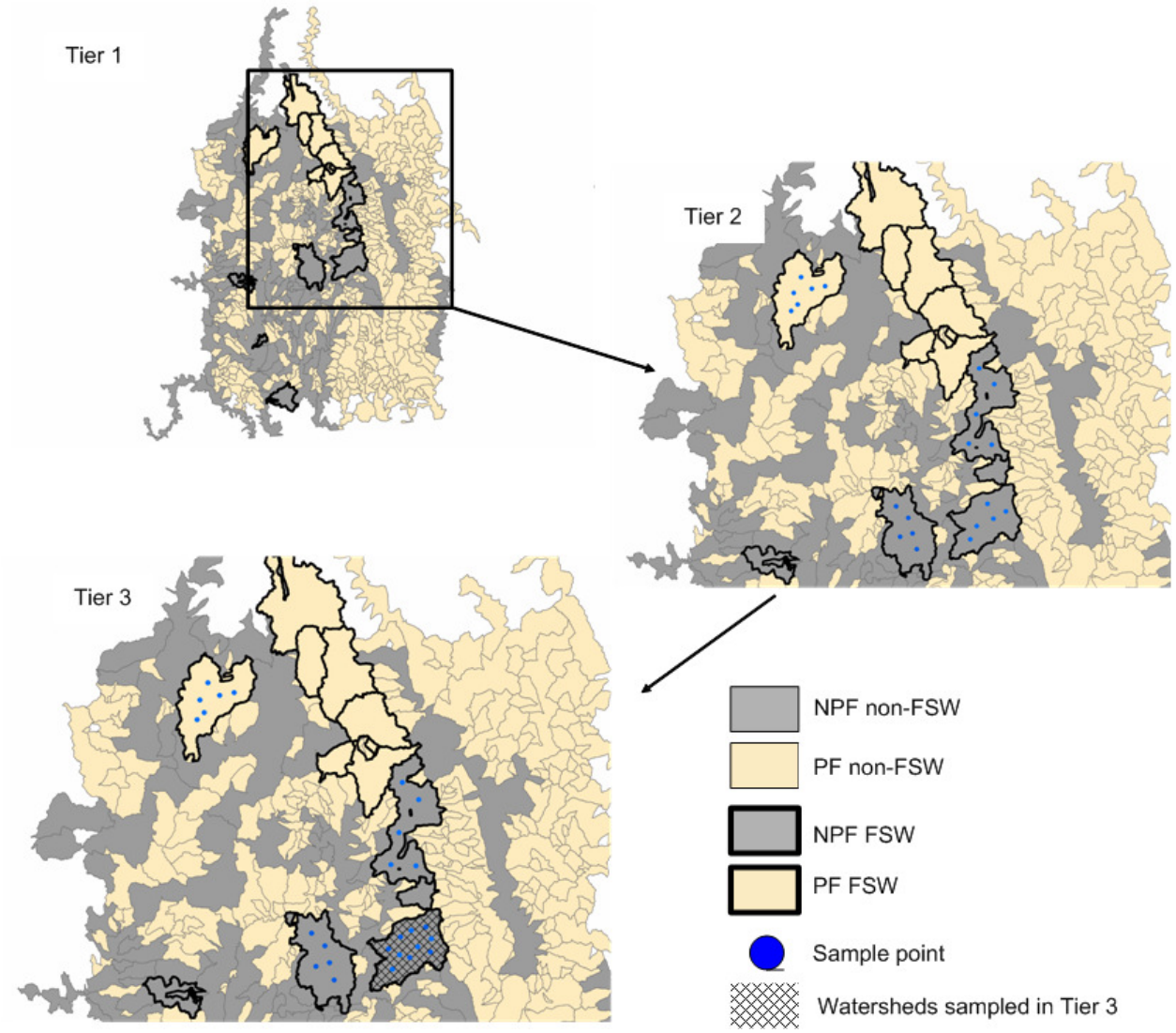


Figure 2. Hypothetical illustration of a nested multi-tiered effectiveness monitoring framework to evaluate the status of BC's FSWs. The maps above are of a subset of FSWs currently designated in an area of southeastern BC. Thicker and darker watershed boundaries indicate FSWs and thinner lighter watershed boundaries indicate non-FSWs. Gray coloured watersheds are those classified through Tier 1 monitoring as not properly functioning (NPF), light coloured watersheds are those classified as properly functioning (PF). The second panel shows a zoom to a subset of FSWs selected for more intensive field-based Tier 2 monitoring (blue points represent potential randomized sampling locations within a watershed). The third panel shows a selected NPF FSW (illustrated by cross –hatching) in which multi-year intensive monitoring will be undertaken to evaluate whether the watershed has stabilized or improved (either passively or in response to directed management actions). Management actions that could be evaluated may vary from status quo (i.e., no change in ongoing practices) to directed restoration depending on operator's willingness to follow recommendations put forth by government scientists and managers monitoring the watershed.