TECHNICAL PROTOCOL
FOR THE
FRESHWATER HEALTH ASSESSMENT

WWF-CANADA

June 2015

CONTENTS

INTRODUCTION

WWF-Canada's Freshwater Health Assessment provides a set of core metrics with associated indicators that can be applied across Canada, using currently available monitoring data, to assess ecological health in a scientifically credible manner. This framework for evaluating freshwater health was developed based on international examples of best practice in this field, including drawing from work in South Africa, Australia, the European Union, and the United States (Boulton, 1999; Roux, 1999; Norris and Thoms, 1999; Birk *et al.*, 2012; U.S. Environmental Protection Agency, 2013). Further, we have drawn upon state of watershed reporting initiatives underway in Canada, including work in Ontario, Alberta and New Brunswick (Kidd *et al*., 2011; Alberta Environment and Sustainable Resource Development, 2012; Maaskant and Quinlan, 2012).

There are four metrics in the assessment framework: hydrology, water quality, benthic macro-invertebrates, and fish. These metrics were chosen to provide representation of key elements of the aquatic ecosystems that are also commonly monitored in most Canadian jurisdictions. Additional components of river ecosystems, such as fluvial geomorphology, riparian vegetation, and connectivity may be considered in future iterations of this assessment framework; however, at this time, these components are not comprehensively monitored or reported for many watersheds across Canada.

For each of the four freshwater health metrics, we have developed one or more indicators (i.e., quantitative tests) to assess their current status and, where multiple years of monitoring are available, trends over time. This document describes the four metrics and their associated indicators, including the rationale for choosing them. Also described are the calculations and underlying statistical tests for indicators, and how indicators “roll-up” to the four health metrics.

Based on the health metrics and their associated indicators, we have developed a rubric for scoring freshwater health, i.e., a method for rolling up the metrics to a single overall health score for a watershed or sub-watershed. This rubric is intended to provide a standard for evaluating and improving river health in Canada. Freshwater health scores are calculated at the sub-watershed (i.e., Water Survey of Canada, "Sub-Drainage" areas) and watershed scale (the Pearse watersheds) (Natural Resources Canada, 2010; Pearse, Bertrand and MacLaren, 1985)

Each freshwater health score is accompanied by a measure of data sufficiency, reflecting the availability and accessibility of monitoring data, which is intended to provide an associated level of confidence for each river health score. Further, the data sufficiency criteria are in place to prevent inappropriate extrapolation of limited monitoring (e.g., few sites or minimal duration of sampling) to freshwater health scores at the sub-watershed or watershed scale.

The current version of the Freshwater Health Assessments focuses on river systems and analysis is limited to monitoring data for streams and rivers. Future iterations of the analysis are expected to expand in scope to capture other water body types, including lakes and wetlands. In addition, the assessment is national in scope and intended to provide a consistent framework for evaluating freshwater health at a broad scale. It is not intended to evaluate the impacts of individual projects or to assess the health of water bodies at a site or local scale.

1. HYDROLOGY

The hydrological component of the Freshwater Health Assessment is included to account for changes in a river's flow, specifically the extent of alteration from its natural flow regime. The indicators for this metric are based on the Indicators of Hydrological Alteration (IHA) approach (Richter *et al*., 1996), but focus on changes in monthly and annual flow. Specifically, the hydrology analysis includes measures of trends in monthly and annual flows over time. In addition, in river systems that include dams greater than 10 metres in height (Canadian Dam Association, 2003), an analysis of monthly flows is completed to compare pre-and post-dam conditions. For watersheds without the presence of large dams, an analogous test is completed to compare monthly flows for a benchmark, or historical time period relative to recent conditions.

This analysis assumes that the 'natural flow regime' is captured in the hydrologic record available from flow monitoring. In Canada, this holds true for many sub-watersheds and watersheds, however there are cases where hydrologic alteration may have preceded or occurred simultaneously with hydrologic monitoring. This limitation is acknowledged and addressed, at least in part, through the data sufficiency criteria for the hydrology metric.

For the Freshwater Health Assessment, six indicators of hydrology are used, rolling up to an overall score for the hydrology metric.

1.1. LONG-TERM TRENDS IN MONTHLY FLOW

This indicator is intended to determine directional changes in monthly flow for stations with continuous, long-term, flow data. Significant changes in monthly flows over time suggest changes in magnitude of flow, as well as in the seasonality of important hydrology events, such as floods and low flows. Changes in monthly flow may be attributed to water consumption, alteration of flows by in-stream structures like dams, or changes in precipitation and temperature, including those attributed to climate change.

The analysis of long-term trends is designed to identify directional changes in monthly flow, as well as measure the relative magnitude of that change. Trends are identified using the Mann-Kendall non-parametric tests for detecting monotonic trends in time series data (e.g., Burn and Hag Elnur, 2002; Monk, *et al*., 2011).

The period of study is dependent on the availability of continuous daily flow data, extending from the first year of continuous flow monitoring to present, and excluding monitoring stations with large gaps (i.e., greater than five years). The period of analysis is consistent across all sites in a sub-watershed, but may vary between sub-watersheds. Minimum duration of monitoring for inclusion in this trend analysis is 20 years, however the long-term datasets used here typically had at least 50 years of monitoring without interruptions greater than five years.

Sub-watershed scores are calculated as the average percentage change in median monthly flow, measured as the relative change in median monthly flow per year, reported as an average across studied stations and weighted by the median annual flow per station. Scores are classified using the following categories: 0 - 0.099% = Very Good; 0.10-0.99%= Good; 1-4.99%= Fair; 5-9.99%= Poor; 0-100%= Very Poor.

Watershed scores are calculated using a weighted mean of the sub-watershed scores, based on the number of monitoring stations in each sub-watershed.

1.2. RECENT-TERM TRENDS IN MONTHLY FLOW

This indicator is similar to 1.1 – "Long-Term Trends in Monthly Flow" – and is intended to determine directional change in monthly flow for stations with continuous flow data, but is calculated based on monitoring data from more recent years. Such an approach permits a greater number of monitoring sites to be included in the analysis, and hence to increase representation of river reaches and sub-sub-watersheds. The period of study for this indicator is based on availability of monitoring data, defined as the year of activation of the last monitoring station with continuous daily flow data and extending until the most recent year of available monitoring. The period of analysis is consistent across all sites in a sub-watershed, but may vary between sub-watersheds. Minimum duration of monitoring for inclusion in this trend analysis is 20 years.

Analysis and scoring are the same as described for indicator 1.1.

1.3. TRENDS IN ANNUAL FLOW

This hydrology indicator is intended to determine directional change in annual flow, based on longest-term available continuous flow data. The period of study is dependent on the availability of continuous daily flow data, extending from the first year of continuous flow monitoring to present, and excluding stations with large gaps in monitoring (i.e., greater than five years). The period of analysis is consistent across all sites in a sub-watershed, but may vary between sub-watersheds. Minimum duration of monitoring for inclusion in this trend analysis is 20 years, however the long-term datasets used here typically had at least 50 years of monitoring without interruptions greater than five years.

Trends in median annual flow using linear regression analysis and Mann-Kendal non-parametric trend analysis.

Sub-watershed scores are calculated as the percentage change in median annual flow, for stations where a statistically significant trend in median annual flow is observed, and measured as the relative change in median monthly flow per year, and reported as an average across studied stations. Scores are categorized as follows: 0 - 0.099% = Very Good; 0.10 - 0.99%= Good; 1 - 4.99%= Fair; 5 - 9.99%= Poor; 10 -100%= Very Poor.

Watershed scores are calculated using a weighted mean of the sub-watershed scores, based on the number of monitoring stations in each sub-watershed.

1.4. ALTERATION OF VARIANCE IN MONTHLY FLOWS

This indicator is used to detect specific impacts on variance of monthly flows, for instance reduced variability of flow due to dam operations. In sub-watersheds where large dams (>10 metres in height) exist, this indicator provides a mechanism to assess the relative impact of dams on monthly flows. Where dams do not exist, an equivalent test for changes in median monthly flows is performed between an historical and recent time period. Ideally, monitoring data would be present for 30 years in both time periods (e.g.,30 years of monitoring both pre- and post- dam operation) to ensure that natural variability in the flow regime is accounted for in this test. However, due to limited availability of such long-term monitoring, all monitoring stations with more than 40 years of monitoring are included in the analysis for this indicator.

Alteration of variance in monthly flows is calculated as the number of months with a statistically significant difference in standard deviation of monthly flow between the two time periods (i.e., pre- vs. post-dam operation, or historical vs. recent), based on Fligner-Killeen non-parametric test of homogeneity of variances (Conover, Johnson and Johnson, 1981).

Sub-watershed scores are calculated as the percentage of total months, for all stations analyzed in the sub-watershed with significantly different variances in monthly flow between the two time periods analyzed, and categorized based on percentiles: 0 - 9.9% = Very Good; 10 - 29.9% = Good; 30 - 44.9% = Fair; 45 - 59.9% = Poor; >60% = Very Poor.

1.5. MAGNITUDE OF CHANGE IN MEDIAN MONTHLY FLOWS

This indicator is intended to report on the magnitude change in monthly flows between two time periods. In sub-watersheds where dams (>10 metres in height) exist, this indicator provides a mechanism to assess the relative impact of dams on monthly flows. Where dams do not exist, an equivalent test for changes in median monthly flows is performed between an historical and recent time period. This approach is similar to Richter's (2009) Sustainability Boundary Approach.

Magnitude of change in median monthly flows is calculated as the percentage change in median monthly flow between the two time periods of interest for months in which a statistically significant difference in median monthly flow (i.e., pre- vs. post-dam operation, or historical vs. recent) is observed, as determined by a Mann-Whitney non-parametric test.

Sub-watershed scores are calculated as the average percentage change in median flow across studied stations in each sub-watershed weighted by the median annual flow per station. Scores are classified using the following categories 0 - 9.9% = Very Good; 0 - 29.9% = Good; 30 - 44.9% = Fair; 45 - 59.9% = Poor; >60% = Very Poor.

1.6. OVERALL HYDROLOGY SCORE

A total hydrology score is calculated for each watershed and associated sub-watersheds. Numeric values are assigned to the four scoring classes received for each hydrology indicator, where sufficient data are available, to generate a numeric score (i.e., Very Good= 5; Good =4; Fair = 3; Poor = 2; Very Poor = 1). The numeric values are then summed to generate a total hydrology score based on all indicators.

The maximum available score is determined for each sub-watershed as the sum of the maximum scores for each of the hydrology indicators completed. In cases where the test is not completed due to lack of sufficient data, these indicators are not included in the maximum available score.

Overall hydrology score is calculated as the percentage of maximum available score for each study unit (i.e. watershed and sub-watersheds) as the quotient of the total hydrology score divided by the maximum available score, and then classified into an overall hydrology score based on the following categories: 84 - 100% = Very Good; 72 - 83.9% = Good; 67 – 71.9% = Fair; 60 - 66.9% = Poor; 0 - 59.9% = Very Poor.

The watershed will receive an overall hydrology score of Data Deficient if 50 percent or more of the sub-watersheds are designated as Data Deficient.

1.7. DATA SUFFICIENCY FOR HYDROLOGY

Data sufficiency scores for the hydrology indicators are determined from spatial and temporal criteria related to the availability and accessibility of monitoring data. Ideally, evaluation of hydrology would include long-term monitoring in all sub-sub-watersheds, extending from pre-industrial conditions (or pre-dam operation in those study units that include large dams) until present.

Data sufficiency for hydrology is evaluated for three time periods, defined by three start dates: (1) the year of first available monitoring; (2) the first year of widespread available monitoring; and (3) the year of operation of the last available monitoring station. The overall hydrology data sufficiency score is determined by the highest available score across those three start dates.

For each sub-watershed, spatial criteria for data sufficiency are determined based on Water Survey of Canada "sub-sub-drainage areas" (Natural Resources Canada, 2010).

Scoring categories are:

Sufficient = Monitoring Stations in 90% or more sub-watersheds, including locations downstream of dams, with flow data extending back at least 30 years before dam construction, <u>or</u> in undammed systems until at least 1930, with continuous monitoring (i.e., without gaps totalling more than 5 years) until 2010 or more recent.

Moderately Sufficient = Monitoring Stations in the majority of sub-watersheds, including locations downstream of dams, with flow data extending back to at least 1950 with continuous monitoring, without gaps totalling more than 5 years, until 2010 or more recent.

Partially Sufficient = Monitoring stations in at least one sub-watershed extending back to at least 1980 with continuous data, without gaps totalling more than 5 years, until 2005 or more recent.

Insufficient = Monitoring not meeting the minimum conditions for Partially Sufficient.

2. WATER QUALITY

Water quality is one of the more commonly monitored characteristics of freshwater systems in Canada (Dunn and Bakker, 2011), due to consequences for aquatic life, but as well for drinking water and recreation. Similar to the Water Quality Index endorsed by the Canadian Council of Ministers of the Environment (CCME) (1999), the water quality indicator used in this analysis is focused on chemical, rather than biological components (e.g., *E. coli* or other bacteria, viruses or protozoa). Future iterations of Freshwater Health Assessments may include indicators specifically related to human health and well-being; however, the current version is primarily focussed on aquatic life.

For the Freshwater Health Assessment, one indicator comprising multiple measures of water quality is used, rolling up to an overall score for the water quality metric.

2.1. EXCEEDANCE OF WATER QUALITY GUIDELINES

The indicator used for water quality is the proportion of all measurements of a chosen set of widely monitored parameters that exceed a set of three thresholds. Similar to the CCME Water Quality Index, the first threshold used in the analysis is drawn from water quality objectives and guidelines for aquatic life, as defined by provincial and/or federal governments and compiled by Environment Canada (2011). The value used for this threshold is specific for each water quality parameter, and that threshold value may vary from province to province. It has been shown that in some cases provincial or federal guidelines for water quality parameters may not adequately represent specific water bodies (Khan *et al*., 2005), for instance, where levels of a parameter are elevated due to naturally occurring concentrations in the river substrate. Further, where natural background levels of a parameter are low, provincial or federal guidelines may be overly permissive. To address the limitations of the generic guideline approach, we adopted two additional thresholds that are set based on the historical distribution of observed values of a parameter in a watershed. Specifically, those thresholds are the 75^{th} and 90^{th} percentile of the distribution of observed values.

Measurements of the following water quality parameters and contaminants were assessed relative to the three thresholds described above: arsenic, aluminum, ammonia, cadmium, chlorides, copper, dissolved oxygen, iron, lead, mercury, nickel, nitrates, nitrites, nitrogen, phosphorus, pH, turbidity, uranium and zinc. The following persistent organic pollutants were assessed: aldrin, chlordane, DDT, dieldrin, hexachlorobenzene, hexachlorocyclohexane, polychlorinated biphenyls, toxaphene and U. The following endocrine-disruptive compounds were assessed: bisphenol A, benzyl butyl phthalate, bis(2-ethylhexyl) phthalate, di-n-butyl phthalate, dimethyl phthalate, dionyl phthalate, dinoctyl phthalate, dtrazine, bentazone, bromozyil, clopyralid, dicamba, dimethanamid, glyphosate, 2,4-D, and MCPA. The following polycyclic aromatic hydrocarbons were also assessed: acenaphtene, acridine, anthracene, benz[a]anthracene, benzo[a]pyrene, chrysene, fluorene, fluoranthene, naphthalene, phenanthrene, pyrene and quinoline.

Other water quality parameters that may be of interest in specific river systems, such as hydrocarbons, have not been included at this time, since they are not consistently monitored across the country. Future iterations of the Freshwater Health Assessment methodology may evolve to include additional parameters as more widespread monitoring becomes available.

The relative proportion of measurements that exceeded each of the three thresholds are calculated for the full suite of water quality parameters. An overall index of water quality exceedances is created as a weighted average of the three proportions:

$$\text{Water Quality Score} = \frac{(PGuide * 3) + (P90 * 2) + (P75 * 1)}{6}$$

where,

PGuide = Proportion of water quality measurements that exceed federal or provincial water quality guideline
P90 = Proportion of water quality measurements that exceed 90th percentile of historical distribution
P75 = Proportion of water quality measurements that exceed 75th percentile of historical distribution

Sub-watershed scores are calculated as the mean proportion of exceedances for all monitoring sites in the sub-watershed for the five most recent years of available monitoring, weighted by the number of sites and water quality parameters monitored per year. Scores are categorized into the following classes: 0.000 - 0.069 = Very Good; 0.070 - 0.139 = Good; 0.140 - 0.239 = Fair; 0.240 - 0.499 = Poor; >0.500 = Very Poor.

Watershed scores are calculated as an average of the sub-watershed scores, weighted by the number of sampling sites found in each sub-watershed.

In addition, a Mann-Kendall test is completed to determine if directional trends exist in water quality over time.

2.2. DATA SUFFICIENCY FOR WATER QUALITY

Data sufficiency for the water quality metric is determined based on both spatial and temporal criteria. Ideally, water quality monitoring would be available for each sub-sub-watershed of a sub-watershed study unit, including with long-term monitoring at each station.

Spatial criteria for data sufficiency are determined based on Water Survey of Canada "sub-sub-drainage areas” (Natural Resources Canada, 2010).

Sufficient = Data are available from monitoring that covers 90% of all sub-sub-watersheds or more and encompass at least three years of monitoring within the last ten years including at least one year of monitoring after 2008. In addition, measurements for 10 or more water quality parameters are available for at least 90% of samples.

Moderately Sufficient = Data are available from monitoring that covers 75% of sub-sub-watersheds or more, spans at least three years of monitoring within the last ten years including at least one year of monitoring after 2008, and includes measurements for 10 or more water quality parameters in at least 75% of samples.

Partially Sufficient = Data are available from monitoring that covers 25% of sub-sub-watersheds or more, spans at least two years over the last ten years, including at least one year of monitoring after 2008. In addition, data is available for more than one monitoring stations per sub-basin, and includes measurements for 10 or more parameters in at least 10% of samples.

Insufficient = Data and monitoring do not meet the minimum conditions for Partially Sufficient.

3. FISH

Fish are an integral component of aquatic health assessments due to their role in aquatic food webs, and high value to humans (e.g., for food, recreation and spiritual reasons). Fish health can be assessed based on a wide array of characteristics which range from coarse indicators like number of species, and grow in detail to include population-scale measures like abundance of species, and further extend to measurements of individual fish, for instance size and weight. Community-level indicators that account for changes in abundance of the full assemblage of fish species in a watershed are increasingly becoming acknowledged as a preferred approach for assessing aquatic ecological integrity (Ganasan and Hughes, 1998; Pont, *et al*., 2006). In some instances, it is even possible to apply evaluative models of fish communities that include factors like species life history characteristics and sensitivity to disturbance. Such indices of fish biological integrity have great potential for evaluating freshwater health; however, they are typically dependent on measuring changes in fish species abundance from an expected level, either based on historical monitoring or regional models. Unfortunately, the majority of publically available fish monitoring data in Canada are based on inventories of species presence/absence or are focused on individual species of commercial or recreational value. Data of this type limit the applicability of the more detailed analyses of the fish community, such as the multi-metric indices of fish biotic integrity (Angermeier and Karr, 1986). There are, however, several promising initiatives that could eventually lead to inclusion of such approaches in the Freshwater Health Assessment and other reporting in the future, such as the Index of Native Fish Integrity being developed in Alberta, and through the Ontario Stream Assessment Protocol (Stevens, Council and Sullivan, 2010; Stanfield, 2012).

For the Freshwater Health Assessment, one indicator of native fish species richness is used, rolling up to an overall score for the fish metric.

3.1. SIGNIFICANT DECLINE IN NATIVE FISH SPECIES RICHNESS

Currently in Canada, available fish monitoring data are often restricted to the presence/absence of a number of fish species; hence, it is only possible to report in a consistent matter across the country using this rather simple or “coarse” approach. Based on presence/absence data, it is possible to count the observed fish species richness at a site (i.e., the number of native species found). By itself, this count has limited value as an indicator of fish health; however, when observed over time, trends in native species richness can be detected. Declines in native species richness are strongly indicative of a decline in the overall health of the fish community.

The Mann-Kendall non-parametric test for directional trends in time-series data is used to detect the presence of statistically significant declines in median and total annual fish species richness over time.

Fish metric scores for each sub-watershed and watershed are categorized into the following three classes: No Significant Trend in both median and total annual fish species richness = Good; Significant Trend in either median or annual fish species richness = Fair; Significant Trend in both median and annual fish species richness = Poor.

Future versions of the Freshwater Health Assessment may incorporate more sophisticated measures of fish trends, if sufficient data are available on a broad geographic basis.

3.2. FISH DATA SUFFICIENCY

Data sufficiency for the fish metric is determined based on both spatial and temporal criteria. Comprehensive fish monitoring would include sampling for each sub-sub-watershed, including with long-term monitoring at representative sites for each sub-sub-watershed.

For sub-watersheds, spatial criteria for data sufficiency are determined based on Water Survey of Canada "sub-sub-drainage areas" (Natural Resources Canada, 2010).

Sufficient = Data are available from monitoring that covers at least 90% of sub-sub-watersheds, spans at least three years of monitoring within the last ten years, including one year of monitoring after 2008.

Moderately Sufficient = Data are available from monitoring that covers at least 90% of sub-sub-watersheds, for at least one year of monitoring within the last five years.

Partially Sufficient = Data are available from monitoring that covers at least 25% of sub-sub-watersheds and spans at least two years of monitoring within the last ten years, including one year of monitoring after 2008.

Insufficient = Data and monitoring do not meet the minimum conditions for Partially Sufficient.

4. BENTHIC MACRO-INVERTEBRATES

4.1. HILSENHOFF BIOTIC INDEX

For the Freshwater Health Assessment, one indicator of benthic macro-invertebrates, comprising multiple measurements of invertebrates, is calculated, rolling up to an overall score for this metric.

The Hilsenhoff Biotic Index (HBI) (1988) is an established metric of benthic community health, initially developed for study of disturbance due to organic pollutants, but has also shown to be a robust indicator of other sources of human-caused disturbances. The HBI approach accounts for sensitivity of benthic invertebrates to disturbance and, based on their relative abundance in the sample of the benthic community, indicates the degree of disturbance of the sampling location. For the Freshwater Health Assessment, WWF-Canada calculates HBI values based on family-level sensitivity values. Also, Hilsenhoff's (1988) scoring framework is slightly reclassified to match the WWF-Canada Freshwater Health Assessment categories, as follows:

<table>
<tr><th>HBI Score</th><th>Hilsenhoff's (1988) Categories</th><th>WWF Categories</th></tr>
<tr><td>0.00-3.75</td><td>Excellent</td><td rowspan="2">Very Good</td></tr>
<tr><td>3.76-4.25</td><td>Very Good</td></tr>
<tr><td>4.26-5.00</td><td>Good</td><td>Good</td></tr>
<tr><td>5.01-5.75</td><td>Fair</td><td>Fair</td></tr>
<tr><td>5.76-6.50</td><td>Fairly poor</td><td>Poor</td></tr>
<tr><td>6.51-7.25</td><td>Poor</td><td rowspan="2">Very Poor</td></tr>
<tr><td>7.26-10.00</td><td>Very poor</td></tr>
</table>

The Canadian Ecological Flow Index (CEFI) (Armanini *et al*., 2011) has potential to become a flow-centric index of benthic communities; however, CEFI requires a reference condition approach to be present for application, which is not currently available for most watersheds.

Sub-watershed benthic scores are calculated as the median HBI for all monitoring sites in the sub-watershed from the five most recent years of monitoring, weighted by the number of sites monitored per year, and categorized into the following classes: 0.00 - 4.25 = Very Good; 4.26 - 5.00 = Good; 5.01 - 5.75 = Fair; 5.76 - 6.50 = Poor; >6.51 = Very Poor.

Watershed scores are calculated as the average of the sub-watershed scores, weighted by the number of sampling sites in each sub-watershed.

In addition, to determine directional changes in benthic community response to habitat alteration (i.e., improvement or decline in benthic community health), a non-parametric Mann-Kendal test is applied to determine the presence of a directional trend in HBI scores over time.

4.2. DATA SUFFICIENCY FOR BENTHICS

Data sufficiency for the benthic macro-invertebrate metric is determined based on both spatial and temporal criteria. Ideally, benthic monitoring would be available for each sub-sub-watershed of the river watershed, including with long-term monitoring at each station.

For sub-watersheds, spatial criteria for data sufficiency are determined based on Water Survey of Canada "sub-sub-drainage areas" (Natural Resources Canada, 2010).

Sufficient = Data are available from monitoring that covers at least 90% of sub-sub-watersheds, spans at least three years of monitoring within the last ten years, including one year of monitoring after 2008.

Moderately Sufficient = Data are available from monitoring that covers at least 90% of sub-sub-watersheds, for at least one year of monitoring within the last five years.

Partially Sufficient = Data are available from monitoring that covers at least 25% of sub-sub-watersheds and spans at least two years of monitoring within the last ten years, including one year of monitoring after 2008.

Insufficient = Data and monitoring do not meet the minimum conditions for Partially Sufficient.

5. OVERALL SCORE

5.1. OVERALL RIVER HEALTH SCORE CALCULATION

Categorical scores for each of the four freshwater health metrics are assigned a corresponding numeric value out of 5: Very Good = 5; Good = 4; Fair = 3; Poor = 2; Very Poor = 1; Data deficient = 0.

These numeric values are added to create an overall numeric score.

The overall numeric score is converted to a percentage of the maximum available score (i.e., total score / sum of max score for all metrics that do not score "Data Deficient", expressed as a percentage).

The proportion of maximum score is classified into an overall health category, based on the following classification: Very Good = 83 - 100%; Good = 73.3 – 82.9%; Fair = 60 - 73.29%; Poor = 53.3 – 59.9%; Very Poor = 0 – 53.29%.

5.2. OVERALL DATA SUFFICIENCY CALCULATION

Categorical scores for Data Sufficiency of each metric are assigned a corresponding numeric value out of 3: Sufficient = 3; Moderate = 2; Partial = 1; Insufficient = 0.

These numeric values are added to create an overall Data Sufficiency score.

Overall numeric score is converted to a percentage of maximum Data Sufficiency score (4 metrics X 3 = 12).

The proportion of maximum score is classified into an overall data sufficiency category, based on the following classification: Sufficient = 100 - 75 %; Moderately Sufficient = 74.9 - 50%; Partially Sufficient = 49.9% - 25%; Insufficient = 24.9% - 0%.

If any Data Sufficiency metric scores “Insufficient”, then the corresponding river health score automatically becomes "Data Deficient".

If 50% or more of the freshwater health metrics received Data Sufficiency scores of "Insufficient", the overall freshwater health score automatically becomes "Data Deficient".

If 50% or more of the sub-watersheds received overall Data Sufficiency scores of "Insufficient", the overall freshwater health score for the watershed becomes "Data Deficient".

6. SUMMARY OF METRICS, INDICATORS AND SCORING

Metric	Indicator		Value Ranges Minimum	Value Ranges Maximum	Health Category	Score
Hydrology	Long-Term Trends in Monthly Flow	Average percentage change in median monthly flow per year, measured as the relative change in median monthly flow per year, reported as an average across studied stations and weighted by the median annual flow per station.	0.00	0.099	Very Good	5
			0.10	0.99	Good	4
			1.00	4.99	Fair	3
			5.00	9.99	Poor	2
			10.00	100.00	Very Poor	1
	Trends in Annual Flow	Average percentage change in median annual flow,reported as an average across studied stations and weighted by the median annual flow per station.	0.000%	0.099%	Very Good	5
			0.101%	0.990%	Good	4
			1.000%	4.990%	Fair	3
			5.000%	9.990%	Poor	2
			10.000%	100.000%	Very Poor	1
	Hydrological Alteration due to dams	Percentage of total months, for all stations analyzed, with significantly different variance in monthly flow pre- vs. post-dam operation.	0.0%	9.9%	Very Good	5
			10.0%	29.9%	Good	4
			30.0%	44.9%	Fair	3
			45.0%	59.9%	Poor	2
			60.0%	100.0%	Very Poor	1
		Percentage change in mean monthly flow pre-and post-dam, averaged across studied stations.	0.0%	9.9%	Very Good	5
			10.0%	29.9%	Good	4
			30.0%	44.9%	Fair	3
			45.0%	59.9%	Poor	2
			60.0%	100.0%	Very Poor	1
	Overall Hydrology Score	Overall Hydrology Score = Total of Hydrology Scores / Maximum Possible Hydrology Score, expressed as a percentage	84.0%	100.0%	Very Good	5
			72.0%	84.0%	Good	4
			67.0%	71.9%	Fair	3
			60.0%	66.9%	Poor	2
			0.0%	59.9%	Very Poor	1

Metric	Indicator	Value Ranges Minimum	Value Ranges Maximum	Health Category	Score
Water Quality	Exceedance of water quality thresholds. Weighted average of exceedances of three thresholds: water quality guidelines, 90th percentile and 77th percentile. Expressed as a proportion of total measurements. Reported for the most recent year of monitoring.	0.000	0.069	Very Good	5
		0.070	0.139	Good	4
		0.140	0.235	Fair	3
		0.250	0.490	Poor	2
		0.500	1.000	Very Poor	1

Metric	Indicator	Value Ranges Minimum	Value Ranges Maximum	Health Category	Score
Benthic Macro-Invertebrates	Median Hilsenhoff Biotic Index (HBI) score for the basin, based on the most recent year of monitoring.	0.00	4.25	Very Good	5
		4.26	5.00	Good	4
		5.01	5.75	Fair	3
		5.76	6.50	Poor	2
		6.51	10.00	Very Poor	1

Metric	Indicator	Value Ranges	Health Category	Score
Fish	Presence of statistically significant decline in median and/or total fish species richness.	None	Good	4
		One of Two	Fair	3
		Two of Two	Poor	2

Metric	Indicator	Value Ranges Minimum	Value Ranges Maximum	Health Category
Overall Score	River health score overall, expressed as the total of the four river health metrics as a percentage of the maximum possible river health score. * If two or more river health metrics receive a Data Sufficiency score of "Inadequate", the overall river health score will be "Data Deficient"	83.0%	100.0%	Very Good
		73.3%	82.90%	Good
		60.0%	73.29%	Fair
		53.3%	59.99%	Poor
		0.0%	53.29%	Very Poor

REFERENCES

Alberta Environment and Sustainable Resource Development. "Guide to Reporting on Common Indicators Used in State of the Watershed Reports." 2012. http://environment.gov.ab.ca/info/library/8713.pdf (accessed 2013).

Angermeier, Paul L., and James R. Karr. "Applying an Index of Biotic Integrity Based on Stream-Fish Communities: Considerations in Sampling and Interpretation." *North American Journal of Fisheries Management*, 1986: 418-429.

Armanini, D.G., N. Horrigan, W.A. Monk, D.L. Peters, and D.J. Baird. "Development of a benthic macroinvertebrate flow sensitivity index for Canadian Rivers." *River Research and Applications*, 2011: 723-737.

Birk, Sebastien, et al. "Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive." *Ecological Indicators*, 2012: 31-41.

Boulton, Andrew J. "An overview of river health assessment: philosophies, practice, problems and prognosis." *Freshwater Biology*, 1999: 469-479.

Burn, Donald H., and Mohamed A. Hag Elnur. "Detection of hydrological trends and variability." *Journal of Hydrology*, 2002: 107-122.

Canadian Council of Ministers of the Environment. "Canadian water quality guidelines for the protection of aquatic life: CCME Water Quality Index 1.0, Users Manual." In *Canadian environmental quality guidelines*. Winnipeg, 1999.

Canadian Dam Association. "Register of Dams in Canada." 2003.

Conover, William J., Mark E. Johnson, and Myrle M. Johnson. "A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data." *Technometrics*, 1981: 351-361.

Dunn, Gemma, and Karen Bakker. "Fresh Water-Related Indicators in Canada: An Inventory and Analysis." *Canadian Water Resources Journal*, 2011: 135-148.

Environment Canada. "Freshwater Quality Indicator, Data Sources and Methods." 2011.

Ganasan, V., and Robert M. Hughes. "Applicaton of an index of biological integrity (IBI) to fish assemblages of the rivers Khan and Kshipra (Madhya Pradesh), India." *Freshwater Biology*, 1998: 367-383.

Hilsenhoff, William H. "Rapid Field Assessment of Organic Pollution with a Family-Level Biotic Index." *Journal of the North American Benthological Society*, 1988: 65-68.

Khan, Amir Ali, Annette Tobin, Renee Paterson, Haseen Khan, and Richard Warren. "Applications of CCME Procedures for Deriving Site-Specific Water Quality Guidelines for the CCME Water Quality Index." *Water Quality Research Journal of Canada*, 2005: 448-456.

Kidd, Scott D., Curry Allen, and Kelly R. Munkittrick. "The Saint John River: A State of the Environment Report." 2011. http://www.unb.ca/research/institutes/cri/_resources/pdfs/criday2011/cri_sjr_soe_final.pdf (accessed 2011).

Maaskant, Karen, and Cathy Quinlan. "Upper Thames River Watershed Report Cards." 2012. http://www.thamesriver.on.ca/Watershed_Report_Cards/Watershed_Report_Cards-2012.htm.

Monk, Wendy A., Daniel L. Peters, Allen Curry, and Donald J. Baird. "Quantifying trends in indicator hydroecological variables for regime-based groups of Canadian Rivers." *Hydrological Processes*, 2011: 3086-3100.

Natural Resources Canada. *National Hydro Network: Data Product Specifications Distribution Profile, Edition 1.1.* Sherbrooke: Natural Resources Canada, 2010.

Norris, Richard H., and Martin C. Thoms. "What is river health." *Freshwater Biology*, 1999: 197-209.

Pearse, Peter H., Françoise Bertrand, and James W. MacLaren. T. *Currents of Change: Final Report of the Inquiry on Federal Water Policy*. Environment Canada, 1985.

Pont, D., et al. "Assessing river biotic condition at a continental scale: a European approach using functional metrics and fish assemblages." *Journal of Applied Ecology*, 2006: 70-80.

Richter, Brian D. "Re-Thinking Environmental Flows: From Allocations and Reserves to Sustainability Boundaries." *River Research and Applications*, 2009.

Richter, Brian D., Jeffrey V. Baumgartner, Jennifer Powell, and David P. Braun. "A Method for Assessing Hydrologic Alteration within Ecosystems." *Conservation Biology*, 1996: 1163-1174.

Roux, D. J. "Design of a National Programme for Monitoring and Assessing the Health of Aquatic Ecosystems, with Specific Reference to the South African River Health Programme." *Environmental Science Forum*, 1999: 13-32.

Stanfield, Les. "Reporting on the condition of stream fish communities in the Canadian tributaries of Lake Ontario, at various spatial scales." *Journal of Great Lakes Research*, 2012: 196-205.

Stevens, Cameron E., Trevor Council, and Michael G. Sullivan. "Influences of Human Stressors on Fish-Based Metrics for Assessing River Condition in Central Alberta." *Water Quality Research Journal of Canada*, 2010: 35-46.

U.S. Environmental Protection Agency. *National Rivers and Streams Assessment 2008-2009: A Collaborative Survey.* Draft, Washington: U.S. Environmental Protection Agency, 2013.

APPENDIX 1. DATA SOURCES

Watershed	Sources
Peace-Athabasca	Alberta Biodiversity Monitoring Institute (ABMI); Alberta Environment & Sustainable Resource Development; Alberta Regional Aquatic Monitoring Program (RAMP); Canadian Dam Association; Environment Canada; Fish and Wildlife Management Information System; HYDAT, Water Survey of Canada (Environment Canada); Alberta Fish and Wildlife Information System; British Columbia Environmental Monitoring System; British Columbia Fisheries Information System; British Columbia Ministry of the Environment; Canadian Aquatic Biomonitoring Network (CABIN), Environment Canada; Pacific Yukon Water Quality Monitoring Program; British Columbia Environmental Monitoring Service (BCEMS)
Fraser	BC Fish Information Summary System (FISS); Canadian Aquatic Biomonitoring Network (CABIN), Environment Canada; Canadian Dam Association; HYDAT, Water Survey of Canada (Environment Canada); British Columbia Environmental Monitoring Service (BCEMS)
Ottawa	Banque de données sur la qualité du milieu aquatique (BQMA); Ministère du Développement Durable, Environnement et Lutte contre les changements climatiques (MDDELCC); Ministère des Forêts, Faune et Parcs (MFFP); Canadian Dam Association; Centre d'expertise hydrique Quebec (CEHQ); Flowing Waters Information System (FWIS); The Centre for Community Mapping; HYDAT, Water Survey of Canada (Environment Canada); Ontario Ministry of Natural Resources; Ontario Benthos Biomonitoring Network (OBBN); Ontario Provincial Water Quality Monitoring Network
St. John-St.Croix	Canadian Dam Association; Department of the Environment, Government of New Brunswick; Environment Canada; HYDAT, Water Survey of Canada (Environment Canada), UNET
Skeena	BC Fish Information Summary System (FISS); Canadian Aquatic Biomonitoring Network (CABIN), Environment Canada; Canadian Dam Association; HYDAT, Water Survey of Canada (Environment Canada); British Columbia Environmental Monitoring Service (BCEMS)
South Saskatchewan	Alberta Environment & Sustainable Resource Development; Canadian Dam Association; Fish and Wildlife Management Information System; HYDAT, Water Survey of Canada (Environment Canada); Water Security Agency of Saskatchewan (WSAS)
Great Lakes	Ontario Ministry of the Environment, Canadian Aquatic Biomonitoring Network (CABIN), Lower Simcoe Conservation Authority (LSCA), Credit Valley Conservation Authority (CVC), Toronto Region Conservation Authority (TRCA), Water Survey of Canada (Environment Canada), Canadian Dam Association; HYDAT, Ontario Provincial Water Quality Monitoring Network; Upper Thames River Conservation Authority (UTRCA), EcoSpark, Ontario Benthos Biomonitoring Network (OBBN), North Bay-Mattawa Conservation Authority (NBMCA)
St.Lawrence	Banque de données sur la faune aquatique et son environnement (BDFAE); Banque de données sur la qualité du milieu aquatique (BQMA); Ministère du Développement Durable, Environnent et Lutte contre les changements climatiques (MDDELCC); Ministère des Forêts, Faune et Parcs (MFFP); Canadian Dam Association; Centre d'expertise hydrique Quebec (CEHQ); HYDAT, Water Survey of Canada (Environment Canada); Canadian Aquatic Biomonitoring Network (CABIN); Groupe d'éducation et d'écosurveillance de l'eau (G3E); Parc National du Mont-Tremblant; Organismes de bassins versants (OBV); Canards Illimités/Ducks Unlimited; Ontario Ministry of the Environment
North Shore-Gaspé	Banque de données sur la qualité du milieu aquatique (BQMA); Canadian Dam Association; Centre d'expertise hydrique Quebec (CEHQ); HYDAT, Water Survey of Canada (Environment Canada)
Maritime Coastal	Bluenose Coastal Action Foundation, Canadian Aquatic Biomonitoring Network (CABIN), Fisheries and Oceans Canada, Ecology Action Centre, Environment Canada, Government of Nova Scotia, Water Survey of Canada (Environment Canada)

Lower Mackenzie	British Columbia Environmental Monitoring System (BCEMS); British Columbia Fisheries Information System; Canadian Aquatic Biomonitoring Network (CABIN), Environment Canada; Fisheries and Oceans Canada; Geoscience BC, Government of Yukon; Triton Environmental Consultants Ltd.; Water Survey of Canada (Environment Canada); Aboriginal Affairs and Northern Development Canada (AANDC), Benthic Information System for the Yukon, Environment Yukon, Government of Yukon; Gwich'in Renewable Resources Board; Alberta Biodiversity Monitoring Institute (ABMI); Alberta Fisheries and Wildlife Management Information System, British Columbia Fisheries Information System; AENS; Alberta Environment & Sustainable Resource Development; Canada-Alberta Oil Sands Environmental Monitoring System
North Saskachewan	Alberta Biodiversity Monitoring Institute (ABMI); Alberta Environment & Sustainable Resource Development; Canadian Dam Association; Environment Canada; Fish and Wildlife Management Information System; HYDAT, Water Survey of Canada (Environment Canada); Saskatchewan Water Security Agency (SWSA); Government of Alberta; Fisheries & Wildlife Management Information System (FWMIS)