NIWA
Taihoro Nukurangi

Monitoring and Evaluation to Support Adaptive Management of River Flows

Prepared for, and in collaboration with, regional councils and unitary authorities

November 2022

Prepared by:
Rick Stoffels
Doug Booker
Paul Franklin
Robin Holmes (Cawthron)
In collaboration with regional councils and unitary authorities

For any information regarding this report please contact:

Rick Stoffels
Group Manager
Freshwater Modelling
+64-3-343 7832
rick.stoffels@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 8602
Riccarton
Christchurch 8440

Phone +64 3 348 8987

NIWA CLIENT REPORT No:	2022366CH
Report date:	November 2022
NIWA Project:	ELF20504

Revision	Description	Date
Version 1.0	Final Report	15 December 2022

Quality Assurance Statement		
	Reviewed by:	Cathy Kilroy
	Formatting checked by:	Nic McNeil
	Approved for release by:	Neale Hudson

Contents

Tables

Figures

Brief summary

Objectives and strategies

We present a plan to guide effective and efficient monitoring and evaluation (*M&E*) of instream ecological response to river flow management by regional councils and unitary authorities (*councils*), in accordance with the National Policy Statement for Freshwater Management 2020 (*NPSFM*). Our aim was to develop an M&E plan supporting:

A. Scientifically defensible evaluation of the extent to which application of water take limits (*take limits*) enable councils to meet objectives in regional plans;

B. adaptive management of take limits, which involves using data collected as part of the M&E program to reduce uncertainty about how take limits have affected, and are likely to affect, environmental outcomes.

The need for such an M&E plan was prompted by regional councils who noted that current monitoring programs were not sufficient to meet the requirements of the NPSFM with respect to the management of water takes and environmental flow regimes. Development of this plan was funded by an Envirolink Tool grant.

This M&E plan is consistent with the latest recommendations in the resource management literature and best practice in structured decision making. M&E objectives were defined in collaboration with councils to meet the requirements of the NPSFM as well as any additional requirements and aspirations of councils.

We argue that if we are to meet the requirements of the NPSFM, then it may be necessary for M&E design to depart from the status quo—*surveillance monitoring* (e.g., State of the Environment monitoring)—which generates data useful for several purposes, but often does not efficiently generate the information and tools required to understand and forecast how policy and management decisions affect environmental outcomes. The M&E program described here should better yield the understanding and tools to determine how decisions affect outcomes.

We recommend strategies for meeting M&E objectives that differ from those employed in surveillance monitoring in several key ways, including:

- M&E should yield a cause-effect understanding of how flow management decisions affect environmental outcomes. This involves:
  - designing metrics and methods in light of clear conceptual models that link water take decisions to flow regimes, then flow regimes to environmental outcomes; and
  - selecting metrics that, as far as practicable, reduce uncertainty about the mechanisms by which flow regimes directly and indirectly affect hydrogeomorphic and ecological outcomes.
- Rather than spreading monitoring effort across many sites, with incomplete co-occurrence of metrics—reducing our potential for relating stressors and responses—we argue for more concentrated monitoring of informative, interrelated metrics at fewer sites.

- To minimise the trade-off between the need to generate a cause-effect understanding through concentrated monitoring at fewer sites, and the need to evaluate and forecast outcomes across numerous sites, we recommend a *hub-and-spoke* design:
  - Hub sites are designed to reduce major uncertainties about how decisions affect outcomes, as well as generate the information for development of evaluation- and decision-support tools. Monitoring is more intensive at hubs.
  - Spoke sites are primarily designed for spatial extrapolation of likely outcomes based on the information and tools generated at hub sites. Monitoring is much less intensive at spokes.
- Monitoring protocols are designed such that we may develop quantitative relationships between metrics representing drivers (e.g., wetted area) and response (e.g., benthic invertebrate standing crop). This requires less focus on sampling protocols rigidly aimed at building a time series *per se*, and more focus on structuring our sampling to cover the domains of the driver metrics that affect response.

Spatial design

This M&E plan was designed to be implemented across two broad types of rivers:

1. *Runoff-fed rivers*: Rivers whose flow is primarily fed by runoff, and whose benthic sediment may be dominated by either fine (silt, sand) or coarse substrates (sand, gravel, cobbles, boulders, etc.).
2. *Spring-fed rivers:* Spring-fed rivers consisting of fine and/or coarse substrates.

Instead of recommending specific locations or statistical properties of site distributions, we recommend eight practical criteria for locating monitoring sites within the recommended hub-spoke design.

Metrics

We recommend *metrics* to be monitored, not *attributes* in the NPSFM sense. In the context of the NPSFM, attributes require specification of a target state, a baseline state and timeframes over which targets will be met (in addition to further requirements of attributes outlined in Clauses 3.10 and 3.11 of the NPSFM). Therefore, attributes are not just metrics for assessing the extent to which values are provided for, but are also regulatory instruments. We recommend metrics that best facilitate adaptive management of river flows. It is not necessary to define quantitative targets, nor baselines, nor timeframes over which targets will be met for the metrics we recommend. We distinguish between metrics and attributes to reduce the regulatory burden on councils, without compromising M&E objectives.

Some of the river flow metrics we recommend could be converted to NPSFM hydrological attributes in the future. The metrics we recommend would inform transparent and defensible adaptive management of any such hydrological attributes in the long term.

Ten potential metrics are proposed for runoff-fed hubs. Five of these are proposed—but monitored much less intensively—for runoff-fed spokes. Six metrics are proposed for spring-fed hubs. At this stage we do not recommend establishing spring-fed spokes.

Metrics were designed to support adaptive management of two forms of river flow modification:

1. Run-of-river water takes taken during dry periods, resulting in *low flows*.
2. *Flow harvesting*; water takes during periods of medium to high flows, to be stored for irrigation during dry periods, resulting in reduced frequency of flushing flows and “freshes”.

Where to from here? Implementing this plan

We have presented a plan for M&E to support the adaptive management of river flows under the NPSFM. Irrespective of the requirements of the NPSFM, the plan we have presented ‘stands on its own’ to support credible, relevant and legitimate adaptive management of river flows throughout New Zealand.

Further work is required before this M&E plan can be implemented:

1. A per-metric cost-benefit analysis needs to be conducted.
2. The metrics and methods need to be refined following careful consideration by council core representatives and feedback from those representatives to the science team.
3. If this M&E plan is implemented, the science team and council core representatives should scope the evaluation- and decision-support tools to be developed.
4. Funding for the plan’s implementation must be obtained, following refinement in light of the cost-benefit analysis.
5. Site selection should be coordinated across the representatives from regional and unitary councils and the science team who jointly developed this M&E plan.
6. Data management and storage standards need to be developed.

Detailed executive summary

Overarching aim and scope

We present a plan to guide effective and efficient monitoring and evaluation (*M&E*) of instream ecological response to riverine flow management by regional councils and unitary authorities (*councils*), in accordance with the National Policy Statement for Freshwater Management 2020 (*NPSFM*). Our overarching aim was to develop an M&E plan supporting:

A. scientifically-defensible evaluation of the extent to which application of water take limits (*take limits*) enable councils to meet objectives in regional plans;

B. adaptive management of take limits, which involves using data collected as part of the M&E program to reduce uncertainty about how take limits have affected, and are likely to affect, environmental outcomes.

The need for such an M&E plan was prompted by regional councils who noted that current monitoring programs were not sufficient to meet the requirements of the NPSFM with respect to the management of water takes and environmental flow regimes. Development of this plan was funded by an Envirolink Tool grant.

In accordance with the latest recommendations in the resource management literature and best-practice in structured decision making:

- We move away from *surveillance monitoring*, which has been identified as an approach to M&E that does not support adaptive management of natural resources particularly well. The aim of surveillance monitoring is to document state and trends in indicators of ecosystem health. Surveillance monitoring generates data useful for several purposes, but it often does not efficiently generate the information and tools to understand and forecast how policy and management decisions affect environmental outcomes. We present an M&E program that could better yield the understanding and tools to determine how decisions affect outcomes.

- We place a strong emphasis on setting M&E objectives. Failure to follow best-practice in objective-setting has been identified as a cause of M&E programs yielding data that is not fit for adaptive management.

- We have taken a participatory approach to M&E design. This participatory approach involved a *science team* (scientists from NIWA and Cawthron) working closely with numerous *council representatives* from regions spanning Aotearoa-New Zealand to set objectives. This participatory approach was adopted to improve the relevance and legitimacy of the plan.

We recommend *metrics* to be monitored that are distinct from *attributes* in the NPSFM sense. In the context of the NPSFM, attributes require specification of a target state, a baseline state and timeframes over which targets will be met (in addition to further requirements of attributes outlined in Clauses 3.10 and 3.11 of the NPSFM). Therefore, attributes are not just metrics to be monitored such that we may assess the extent to which values are provided for, but are also regulatory instruments. We recommend metrics that best facilitate adaptive management of river flows. It is not necessary to define quantitative targets, nor baselines, nor timeframes over which targets will be met for the metrics we recommend. We distinguish between metrics and attributes to avoid increasing the regulatory burden on councils, without compromising our overarching aim. Some of the river flow metrics we recommend could be converted to NPSFM hydrological attributes in the future. The metrics we recommend would inform transparent and defensible adaptive management of any such hydrological attributes in the long term.

Primary outputs and design process

A process involving five tasks was used for generating three primary outputs that when joined together comprise a functional M&E plan:

1. a set of M&E objectives to guide design of an M&E program;
2. spatial design guidelines; and
3. a set of metrics for a nationally-consistent M&E program to support adaptive management of river flows.

The five tasks were:

Task A: Documenting the advantages and disadvantages of nationally-consistent monitoring and evaluation

Using desktop analysis, surveys and workshops, we documented the advantages and disadvantages of nationally-consistent M&E, and noted the trade-offs between a centrally-coordinated, nationally-consistent M&E program and one that is region-specific. Documenting these trade-offs is necessary to define scope and manage stakeholder expectations of this M&E plan. Failure to consider advantages and disadvantages of implementing a nationally-consistent M&E plan may prompt the criticism that the full set of options was not considered, lowering the legitimacy and credibility of the plan.

Task B: Flow monitoring and evaluation objectives of the NPSFM 2020

Through desktop analysis and workshops, we identified what actions are required by councils to give effect to the NPSFM with respect to the M&E of ecological responses to river flows (referred to as *flow M&E objectives*)

Task C: Objectives and constraints of councils

The decisions councils make about flow M&E and management will be in response to the long-term visions of communities and mana whenua, as well as the requirements of central government. Council decisions concerning flow M&E and management will also be shaped by resource constraints. To ensure this M&E plan satisfies both the requirements of central and regional government we used surveys and workshops to capture and understand the flow M&E objectives of councils, as well as physical, resource, and infrastructural constraints on meeting M&E objectives.

Task D: Defining river types within which to arrange a network of M&E sites

Different river types may exhibit contrasting ecological and physical responses to changes in the flow regime and may support different freshwater values. As such, we may require different metrics and/or sampling methods for different river types. The science team and council representatives identified types by which metric design could be structured.

Task E: Defining metrics

Metrics were defined and nominated based on flow M&E objectives as well as articulation of:

1. which NPSFM values this M&E plan should address;
2. key hydrological stressors that should be the focus of adaptive management, hence the requirement for M&E; and
3. conceptual mechanistic links between hydrological stressors and potential metrics to be monitored.

Advantages and disadvantages of nationally-consistent monitoring and evaluation

There is no clear policy mandate for nationally-consistent M&E of river flow regimes.

The NPSFM does not direct councils to implement nationally-consistent M&E. The Parliamentary Commissioner for the Environment report *Aotearoa New Zealand's environmental reporting system* called for nationally-consistent M&E.

The Action for Healthy Waterways document included proposals for nationally-consistent M&E of freshwater resources but did not justify such proposals.

Priority 3 of the Regional Councils' Research, Science and Technology Strategy highlights the need for consistent application of decision- and evaluation-support tools across local authorities.

The international, peer-reviewed literature identifies both advantages and disadvantages of nationally-consistent approaches to policy-driven M&E. Consistent with the literature, council representatives and the science team in this project identified numerous advantages and disadvantages to implementing nationally-consistent M&E.

We document the trade-offs that come with both nationally-consistent and region-specific approaches to M&E. These trade-offs highlight that nationally-consistent M&E is not a panacea—it will not meet all the requirements of all stakeholders. This plan for nationally-consistent M&E of ecological response to river flow management may have to be supplemented with region-specific M&E driven by the needs of individual councils.

Flow monitoring and evaluation objectives

By completing Tasks B and C we identified 18 flow M&E objectives, the eight most fundamental of which are outlined in the following table.

Objective (short)	Objective (long)
Determine how current water takes are affecting metric state and trends	Select and monitor sites and metrics to help determine how changes in the flow regime affect metrics within river reaches of freshwater management units (FMUs). This environmental monitoring must be undertaken alongside monitoring of water takes and other forms of hydrological alteration as is required for water accounting purposes.
Forecast how future water takes are likely to affect metric state and trends	Sites and metrics must be selected with a view for how data will support forecasts of how metrics will respond to future environmental changes, including future water management scenarios

Objective (short)	Objective (long)
Disentangle the influence of multiple drivers	Monitor sufficient metrics within river reaches of an FMU such that councils may, as much as practicable, partition the influence of multiple drivers—both anthropogenic and natural—on metrics, and determine the effects of the flow regime on metrics relative to other drivers
Monitoring sites must be representative of the types of rivers most affected by water takes and flow management	Locate sites within FMUs such that the sites are, as much as practicable, representative of the types of rivers within the FMU that: are subject to water takes; and support compulsory values, as specified in Appendices 2A and 2B of the NPSFM 2020
Flow M&E must support adaptive management of water takes	Monitoring must aim to reduce uncertainty about flow-metric relationships, and yield information to facilitate forecasts of how metrics will respond to alternative flow management decisions
Support decision-making and evaluation at multiple scales and locations within regions	Collect data with a plan for how it may be used to facilitate decision-making and evaluation at multiple scales and at several locations throughout FMUs
Minimise costs of data collection and processing	Aim to minimize costs while still meeting other monitoring objectives
Flow M&E must support effective communication of outcomes to all stakeholders	The information collected must be relevant to, and easily understood by, as many stakeholder groups as practicable.

Ten additional objectives are described in the body of this plan. They outline the means by which we may achieve these fundamental M&E objectives.

Challenges and trade-offs among M&E objectives, and potential solutions

The inherent complexity of the systems councils wish to understand and manage, as well as limitations of the scientific processes and tools available to us, make meeting our fundamental M&E objectives challenging. The challenges we face are not insurmountable and we recommend the following M&E strategies to overcome those challenges:

Monitor for a cause-effect understanding between hydrological stressors and metrics within an adaptive management context. Design metrics using conceptual models that, first, capture the effects of the predominant forms of water takes on hydrology—thus identifying key hydrological stressors—and, second, capture the effects of hydrological stressors on geomorphological and ecological metrics. Conceptual models should identify metrics that reduce uncertainty about how flow regimes directly and indirectly affect high-value species like fishes.

Metrics should be designed to increase sensitivity to hydrological stressors. Species' abundances offer a more sensitive indicator of environmental effects than species' presences. When monitoring presence-absence we lose all information about how changes in flow affected species abundances prior to extirpation.

Monitor metrics with different rates of response to flow, and with clear conceptual links to each other, and communicate the likely timeframes of response. Evaluation could be strengthened by including metrics with different rates of response to changes in river flows and that, at least conceptually, are likely to be causally related to each other. Doing so enables narratives in evaluations that communicate how short-term responses are contributing to long-term objectives. Causal relationships between fast and slow response metrics should be identified when developing conceptual models.

Timeframes at which we expect metrics to respond to flow management should be explicitly documented, such that we may manage stakeholder expectations concerning how quickly expected environmental outcomes can be met.

Monitor all metrics of primary concern to river flow management at the same long-term sites. If we are to disentangle the influence of multiple stressors on metrics, then we must monitor all metrics of primary management concern at the same sites. This includes both stressor metrics (e.g., river flows; fine sediment; nutrients) and response metrics (e.g., macroinvertebrates).

Use data from the M&E program to develop quantitative decision- and evaluation-support tools and embrace partnerships for developing those tools. Long-term monitoring programs can play a pivotal role in providing the data required to develop evaluation- and decision-support tools, as long as we give careful thought to what we monitor and how we take measurements. This requires defining the temporal and spatial structure of data required to determine how hydrological stressors affect metric state in light of clear conceptual models defining hypothesized or observed cause-effect relationships between metrics. At a minimum, we should anticipate the functional form of relationships between, say, stressor and response metrics, and aim to structure sampling in space/time to ensure samples are well distributed across the stressor and response gradients, such that we may develop mathematical relationships between stressor and response.

Councils may lack the capacity and/or capability to develop the tools required to meet our fundamental M&E objectives. A solution to this problem is to develop and maintain partnerships among councils and other science providers that could help deliver the required data and tools.

Resources for implementing flow M&E are limited, causing major trade-offs between M&E objectives. We contend that the following statements hold given finite resources for flow M&E implementation:

- The requirement for high spatial coverage of information across rivers is at odds with the requirement for more intensive monitoring strategies within rivers to support adaptive management.
- The trade-off between the need to minimize costs and generate the knowledge and tools required to determine how past and future flow management decisions affect riverine values.
- Councils expressed a requirement for information that is relatively simple and easy to communicate with stakeholders. However, fundamental objectives concerning defensible evaluations and forecasts and, more broadly, adaptive management may require intensive monitoring of numerous metrics.

We recommend the following strategies to minimize these trade-offs:

Locate sites and select metrics according to a stratified design, where strata are defined by both river type and the need for river flow management. M&E resources to support adaptive management of flow regimes should be prioritized around types of rivers whose flow regimes are most impacted by water takes, hence where there is the greatest need to meet our fundamental objectives. Stratifying by river types as well as need for flow management will ensure that selection and monitoring of metrics is more efficiently tailored to rivers that likely support different ecological values, and/or exhibit contrasting ecological and geomorphological responses to flow alteration and adaptive management.

Implement a *hub and spoke* spatial design to minimize the trade-off between, on one hand, the need for high spatial coverage of evaluations and forecasts and, on the other hand, the need for intensive M&E to facilitate a cause-effect understanding and adaptive management. A hub-spoke design should subdivide sites into at least two types:

1. *Hub sites*. Long-term sites (i.e., maintained for decades) for intensive monitoring that support development of a cause-effect understanding of how river flow interacts with other variables to affect metrics. Hub sites yield data required for development of evaluation- and decision-support tools. Hub sites would be relatively expensive to operate and so there would likely be very few hub sites per type, per region.

2. *Spoke sites.* Long-term sites where, at a minimum, continuous monitoring of hydrology occurs. The cause-effect understanding and/or tools generated at hub sites could be used to extrapolate the likely environmental outcomes from flow management decisions at long-term spokes. Monitoring of metrics at spokes is less intensive and therefore cheaper, but carefully designed such that we have a strong foundation for application of the tools and knowledge generated at hubs.

Adopt national consistency and partnerships to overcome resource limitations of individual councils. Each council is unlikely to fund numerous hub sites per type. If councils implement hubs in a consistent way, then our ecological understanding and tool development will be leveraged by data collected across:

- more rivers within each type;
- a greater range of climate and geomorphologic contexts within and across types; and
- a larger and more diverse set of flow experiments within types and contexts.

This flow M&E program should be supplemented by strategic research to improve our understanding of how flow affects ecological processes, and to develop decision- and evaluation-support tools. Provision of this supplementary knowledge and tools may be best achieved through partnerships between councils and other research organisations.

One could argue that intensive M&E aimed at supporting adaptive management is at odds with the need for simple forms of information for communicating outcomes to non-scientific stakeholders. This is not necessarily the case. We contend that information obtained for adaptive management can be simplified and effectively communicated to non-scientific stakeholders. By contrast, information obtained through surveillance monitoring programs cannot be easily transformed into forms that support adaptive management.

Spatial design

The science team and council representatives identified two river types most relevant to this M&E plan:

1. ***Runoff-fed rivers*****:** Rivers whose flow is primarily fed by runoff, and whose benthic sediment may be dominated by either fine (silt, sand) or coarse substrates (sand, gravel, cobbles, boulders, etc.). In the context of this plan *runoff-fed rivers* are small- to medium-sized (mostly wadable) rivers.

2. ***Spring-fed rivers:*** Spring-fed rivers consisting of both fine (generally < 2 mm grain diameter) and coarse substrates (sand, gravel, cobbles, boulders, etc.). In the context of this plan spring-fed rivers are small- to medium-sized (mostly wadable) rivers.

We note that, given our definitions of runoff- and spring-fed rivers above, **this M&E plan is limited in scope to small- and medium-sized (mostly wadable) rivers**.

Statisticians have in the past recommended randomizing site locations throughout entire river networks, towards ensuring the data obtained is representative of catchments. Such approaches, however, are often impractical as they do not take into account constraints of any kind (resource constraints, socio-political constraints on site locations, logistical constraints [e.g., ease of access]), and often are recommended to fulfil the objective of surveillance monitoring, not adaptive management of natural resources. Instead of recommending specific locations or statistical properties of site distributions, **we recommend eight practical criteria for locating monitoring hubs** (not detailed in the Executive Summary).

Reference sites are not necessary for implementation of this M&E program, but can strengthen evaluation of the effects of flow management. We define reference sites as sites whose flow regime and physical characteristics are not influenced by local human activities. We explain why reference sites are not necessary in light of a plan for statistical analysis of the outcomes of flow management decisions.

Councils should aim for between one and three hubs per type, per region. Three hubs per type enables within-type replication of hub M&E within each region. Intensive monitoring of several metrics at hubs will likely be expensive, so we appreciate that establishing and maintain three hubs for both runoff- and spring-fed rivers (six hub sites) per region may not be achievable. Some consolation should be taken from the fact that this M&E program will be nationally-consistent, allowing individual councils to better meet our M&E objectives through sharing of data and tools across regions.

Metrics

Metric development focused on the NPSFM compulsory value of *ecosystem health*, specifically components:

- water quantity;
- habitat;
- aquatic life;
- ecological processes.

These components are known to have metrics that are particularly sensitive to river flow. Healthy ecosystems are necessary to support other compulsory values listed within Appendix 1A of the NPSFM, such as *mahinga kai* and *threated species*.

In accordance with our M&E objectives we must also consider how flow regimes affect hydrogeomorphic metrics that shape recreational, aesthetic and cultural values, some of which are represented in the NPSFM other values *natural form and character*, *transport and tauranga waka*, and *fishing* (Appendix 1B of NPSFM 2020). These are also focal values of this M&E plan.

The science team and council representatives agreed that **two broad forms of flow modification were of national concern**:

1. Run-of-river water takes taken during dry periods for immediate use, resulting in reduction in magnitude and increase in duration of **low flows**.
2. **Flow harvesting**; water takes during periods of medium to high flows, to be stored for irrigation during dry periods, resulting in reduction frequency of flushing flows and “freshes”.

Conceptual models were developed linking these forms of flow modification to potential metrics in both runoff- and spring-fed rivers. We present a list of potential metrics for runoff- and spring-fed rivers to meet our M&E objectives and the overarching objective of this plan. The list of metrics may appear resource-intensive, but councils should note that:

- We do not recommend that all metrics be monitored at both runoff- and spring-fed rivers.
- We do not recommend that all metrics be monitored at both hubs and spokes.
- Most metrics are high-priority at hubs only. There are very few hub sites per region.
- Most metrics are not monitored annually. In most cases metrics are sampled during two *sampling years* (see below for explanation) every five-year period, noting the aim for most metrics is to develop functions defining how discharge (flow rate) affects metric state, not to document state and trends alone.
- Several metrics are monitored using automated remote loggers.

For the metrics we recommend at **hub sites**, we present the metric name, then describe the metric using the following structure:

Description: Brief written description of the metric.

Units: Measurement units of the metric.

Values/components: The NPSFM values and value components to which the metric is relevant.

Why? A brief justification of why the metric should be included.

Targets flow modification: The form of flow modification that the metric is most relevant to; low flows, flow harvesting or both.

Time-scales of response detected: The most relevant time-scales at which the metric can be expected to respond to changes in the flow regime. Can be some subset of: Hours. Days. Weeks. Months. Years. Decades.

Priority at hub sites: The level of priority assigned to the metric, or how important the metric is to meeting monitoring objectives. Note that priority levels are not an indication of cost. Priority levels are provided as a guide only—councils may assign higher priority to metrics if they have the resources to do so. One of three levels of priority may be assigned: *essential*; *high*; and *low* (explained in detail in the body of this plan).

Sampling frequency/intensity: Overview of the temporal and spatial intensity of sampling, to give councils a basic understanding of the level of effort—hence resources—required to monitor the metric.

Method, in brief: A brief point-from overview of the method. Presented in this plan primarily to demonstrate the approach to monitoring metrics encouraged in this plan, which departs strongly from surveillance monitoring in a manner that is designed to achieve monitoring objectives for minimal cost.

Minimum data, in brief: Very brief description of the data obtained.

Only metrics deemed *essential* at hub sites are recommended for monitoring at **spoke sites**—councils should view all metrics with a priority level of *low* or *high* at hubs as "nice-to-have" at spokes. Metrics deemed *essential* at hubs are monitored at spokes but with reduced effort. With the exception of fish relative abundance, metrics recommended at spokes define the dynamics of physical habitat—temporal variation in hydrology and how that variation affects physical habitat of a reach. The primary purposes of physical metrics at spokes are:

- To provide the necessary physical foundation for spatial extrapolation of the ecological knowledge obtained through the metrics monitored at hubs.
- Physical metrics monitored at spokes serve as important indicators of riverine values in their own right, particularly with respect to NPSFM values Ecosystem health – habitat, natural form and character, and transport and tauranga waka.

We recommend including fish relative abundance once every five years at spokes for the following reasons:

- Measurement of fish relative abundance is critical for meeting certain of our M&E objectives.
- Fish, being at the top of the riverine food web, integrate biophysical responses to flow regimes at numerous trophic levels below them.

For the metrics we recommend at spokes we present the metric name then describe the metric using the following structure:

Sampling frequency/intensity: Definition as described for hubs.

Method – notable departures from method at hubs: Notes on the key differences between the method as applied at spokes, compared to hubs.

Ten metrics are recommended for runoff-fed hubs. Five of these are recommended—but monitored much less intensively—for runoff-fed spokes.

Six metrics are recommended for spring-fed hubs. At this stage we do not recommend establishing spring-fed spokes. Justification for not having spring-fed spokes is presented within the body of this plan.

All metrics and details of site design are explained in full in this plan. We only present tabular summaries here.

In the tables below a *sampling year* comprises the period Jan-Mar during any specific year of each five-year reporting period. The specific years selected as sampling years is left to the discretion/convenience of councils. During a sampling year, sampling is distributed over three reaches per hub site, and across a range of flow levels, which we assume will vary over the summer-autumn (Jan-Mar) period. During a sampling year, the magnitude of sampling within each reach depends on the metric.

Summary of the metrics to be monitored at runoff-fed hub sites. Form of flow modification (*Flow mod*) that is most relevant to each metric is indicated. The time-scales at which the metric best detects response to changes in the flow regime are indicated. Levels of monitoring priority are presented (Essential, High or Low).

	Flow mod.		Time-scales of response detected							
Metric	**Low flow**	**Flow harvest**	**Hours**	**Days**	**Weeks**	**Months**	**Years**	**Decades**	**Priority**	**Sampling intensity**
Daily discharge	■	■	■	■	■	■	■	■	E	Continuous automated
Depth-velocity distribution	■	■				■	■	■	E	Two sampling years every five years.
Wetted area	■	■				■	■	■	E	Two sampling years every five years.
Water temperature	■		■	■	■	■	■	■	H	Continuous automated
Sediment size distribution		■				■	■	■	E	Two sampling years every five years.
Macroinvertebrate benthic density	■	■				■	■	■	H	Two sampling years every five years.
Macroinvertebrate drift density	■	■				■	■	■	L	Three sampling years every five years.
Fish relative abundance	■	■					■	■	E	Annually
Fish size composition	■	■					■	■	H	Annually
Aerial reach photograph	■	■					■	■	H	Annually

Summary of metrics to be monitored at runoff-fed spoke sites. Form of flow modification (*Flow mod*) that is most relevant to each metric is indicated. The time-scales at which the metric best detects response to changes in the flow regime are indicated.

	Flow mod		Time-scales of response detected							
Metric	Low flow	Flow harvest	Hours	Days	Weeks	Months	Years	Decades	Priority	Sampling intensity
Daily discharge	■	■	■	■	■	■	■	■	E	Continuous automated
Depth-velocity distribution	■	■				■	■	■	E	Two sampling years every five years, but with 1/3 effort of hub sites
Wetted area	■	■				■	■	■	E	Two sampling years every five years.
Sediment size distribution		■				■	■	■	E	Two sampling years every five years, but with 1/3 effort of hub sites
Fish relative abundance	■	■					■	■	E	Once every five years; 1/5 effort of hub sites

Summary of the metrics to be monitored at spring-fed hub sites. Form of flow modification (*Flow mod*) that is most relevant to each metric is indicated. The time-scales at which the metric best detects response to changes in the flow regime are indicated. Levels of monitoring priority are presented (Essential, High or Low).

	Flow mod		Time-scales of response detected							
Metric	Low flow	Flow harvest	Hours	Days	Weeks	Months	Years	Decades	Priority	Sampling intensity
Daily discharge	■		■	■	■	■	■	■	NA	Continuous automated
Discharge spot measurement	■					■	■	■	E	Two sampling years every five years
Cross-sectional area and velocity field	■					■	■	■	E	Two sampling years every five years
Macrophyte cross-sectional area	■					■	■	■	E	Two sampling years every five years
Macroinvertebrate drift density	■					■	■	■	H	Three sampling years every five years
Dissolved oxygen dynamics	■		■	■	■	■	■	■	H	Continuous automated over three 10-day periods, twice every five years

Where to from here? Implementing this plan

We have presented a plan for M&E to support the adaptive management of river flows under the NPSFM 2020. Irrespective of the requirements of the NPSFM, the plan we have presented 'stands on its own' to support credible, relevant and legitimate adaptive management of river flows throughout New Zealand.

Further work is required before this M&E plan can be implemented:

1. **A per-metric cost-benefit analysis needs to be conducted.** This analysis would serve as a foundation for making decisions about which components of this plan should be implemented under various levels of resource constraint.

2. **The metrics and methods need to be refined following careful consideration by council hub representatives and feedback from those representatives to the science team**. Refinement of metrics and methods is an iterative process and this plan represents the first step of that process.

3. **If this M&E plan is implemented, the science team and council hub representatives should scope the evaluation- and decision-support tools to be developed**. This plan has been designed to facilitate sharing among councils of new knowledge, data, and the evaluation- and decision-support tools that can be more rapidly developed when M&E is nationally-consistent. If this plan is implemented, an important next step will be to identify with councils which support tools should be developed and how they are developed.

4. **Funding for the plan's implementation must be obtained, following refinement in light of the cost-benefit analysis**. It is currently not clear how this M&E plan would be funded. The plan itself, we hope, serves as a strong foundation for obtaining the required funding.

5. **Site selection should be coordinated across the council core representatives and the science team**. We have only presented broad guidelines for site selection in Section 6 of this plan. To ensure sites are selected to meet M&E objectives, site selection needs to be workshopped among the science team and hub representatives.

6. **Data management and storage standards need to be developed.** The advantages that result from a nationally-consistent M&E program cannot be met if data are not managed and stored carefully and consistently. We have already provided a solid foundation for best practice in data management and storage by describing broad data requirements in Section 8 of this plan, but a data standards manual would need to be developed.

1 Introduction

1.1 Objectives of this plan

Our objective is to develop a plan to guide effective and efficient monitoring and evaluation (*M&E*) of instream ecological response to riverine flow management by Regional Councils and Unitary Authorities (*councils*), in accordance with the National Policy Statement for Freshwater Management 2020 (NPSFM). We aim to develop an M&E plan supporting:

A. scientifically-defensible evaluation of the extent to which application of water take limits (*take limits*) enable councils to meet objectives in regional plans;

B. adaptive management of take limits, which involves using data collected as part of the M&E program to reduce uncertainty about how take limits have affected, and are likely to affect, environmental outcomes.

1.2 Motivation for this plan

Development of this M&E plan was prompted by regional councils who noted that current monitoring programs did not meet the requirements of the NPSFM with respect to the management of water takes and environmental flow regimes. Development of this plan was funded by an Envirolink Tool grant.

1.3 Approach to achieving a credible, relevant and legitimate plan

Monitoring and evaluation are key activities operating at the boundary between science, policy and management (Westgate et al. 2013). Work at this boundary is most effective when it is credible, relevant and legitimate (Cash et al. 2003). A *credible* M&E plan is one that is developed in light of the latest scientific and technological developments. A *relevant* M&E plan is developed with a clear understanding of how it will serve the needs of managers, iwi, communities, scientists and other stakeholders. *Legitimacy* is a factor that is critical for effective and efficient implementation of natural resource management, but in the context of M&E is often inadequately considered. A *legitimate* M&E plan is one that is respectful of end-users' requirements and is fair in its treatment of the views and interests of end-users.

1.3.1 Achieving credibility by moving away from surveillance monitoring

Environmental M&E has been extensively criticized for failing to provide the information and tools to make better policy and management decisions (Failing and Gregory 2003, Lindenmayer et al. 2013, Waylen et al. 2019). A cause of the problem is what Nichols and Williams (2006) call *surveillance monitoring*[1]. Surveillance monitoring has been the norm for environmental M&E programs. Surveillance monitoring may be necessary when environmental policies, consents or plans set indicator states as regulatory targets and require compliance with those targets. It often involves monitoring indicators that are broadly suggestive of ecosystem health, often with the objective of estimating state and trends in ecosystem health (e.g., some of New Zealand's State of the Environment monitoring). Surveillance monitoring tells us when ecosystem health may be declining, but is often not designed to tell us about the causes of such decline, or to facilitate the anticipation of future changes in the health of ecosystems. The data from surveillance monitoring is useful for

[1] Surveillance monitoring involves monitoring indicators that are broadly suggestive of ecosystem health, with the objective of estimating state and trends in ecosystem health (e.g., some of New Zealand's State of the Environment monitoring).

several purposes, but rarely does a good job of building our understanding of the mechanisms by which past and future policy and management decisions affect outcomes (Nichols and Williams 2006, Lindenmayer and Likens 2010).

We aim to achieve a credible M&E plan by following recommendations for improved M&E (Lovett et al. 2007, Lindenmayer and Likens 2009, Westgate et al. 2013). Most notably, we carefully define M&E objectives in light of policy and management requirements (Olsen et al. 1999, Mace and Baillie 2007). In the case of environmental M&E, building relevance also builds credibility.

1.3.2 Achieving relevance and legitimacy

Large-scale M&E programs are expensive to implement and maintain, so their success hinges on the long-term support of multiple agencies and stakeholders. M&E plans that aim to achieve high legitimacy at the design stage establish a strong foundation for long-term support (Stoffels et al. 2018). Participatory approaches to M&E design can enhance legitimacy and relevance (Parr et al. 2002, Cundill and Fabricius 2009, Bennett et al. 2017, Waylen and Blackstock 2017, Burgman et al. 2022). Groups of individuals with diverse roles and capabilities outperform small specialized groups when it comes to developing solutions to complex problems (Stirling and Burgman 2021).

We take a participatory approach here, and work closely with councils to design this M&E plan. The participants developing this plan comprise a collaboration among NIWA, the Cawthron Institute and 13 regional councils/unitary authorities spanning Aotearoa-New Zealand (Appendix A). This participatory approach to the development of environmental solutions is consistent with the Regional Council's Research, Science and Technology Strategy (RCSAG 2020).

The participants in the design of this M&E plan were limited to councils who are responsible for implementing the NPSFM, which includes engagement with mana whenua (Clause 3.4 of NPSFM 2020). Resources available for preparation of this plan did not permit broadening the participation process beyond councils, who we depended on to represent the interests of additional stakeholders within their region, including mana whenua. We note, however, that this M&E plan will itself be an adaptive one (*sensu* Lindemayer and Likens 2009), and may be modified over time to better reflect additional interests and values of mana whenua, as councils continue to develop stronger partnerships with iwi, hapū and whānau.

The *participants* in the development of this M&E plan consist of a *science team*, *core representatives* of the councils, and *additional representatives* of councils. The science team collaborated with one core representative from each participating council (Figure 1-1). This core representative was a member of the council's scientific team. The core representative contributed to the project by:

- completing surveys;
- participating in workshops;
- reading draft plan sections to ensure accuracy of the content.

Core representatives identified three additional representatives to complete surveys. This group of three additional representatives comprised one additional member of the council's science team, one member of the council's policy/planning team; and one member of the council's consenting team. (Figure 1-1). Thus, we obtained a larger sample of responses that reflected opinions and experiences of diverse roles within councils.

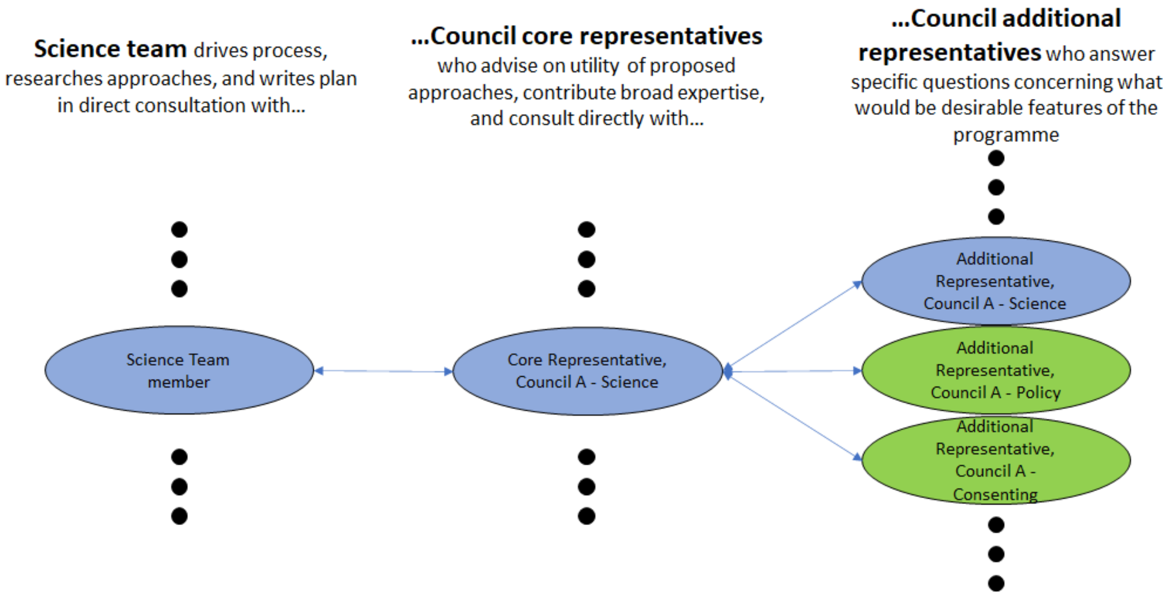

Figure 1-1: Relationships among participants in the design of this plan.

1.3.3 Documenting M&E design

This document serves as a record of the design of the M&E plan. It outlines the processes by which key decisions were made and presents a rationale for the decisions made. When faced with decisions concerning approach, we explored as many solutions as possible, to avoid designs reflecting the personal biases and/or experiences of the participants (Gregory and Keeney 2002). A key factor hindering improvement in environmental M&E has been failure to document program design, such that there is often no record or justification of the decisions made (Parr et al. 2002, Waylen et al. 2019).

2 Scope

2.1 Three most noteworthy elements of scope

2.1.1 We recommend *metrics* to be monitored, not *attributes* in the NPSFM sense

The word *attribute* has a specific meaning in the context of NPSFM implementation. According to Clause 1.4 of the NPSFM, an attribute is "any measurable characteristic that can be used to assess the extent to which a particular value is provided for." But Clause 3.11 of the NPSFM states that in order to achieve planned environmental outcomes "every council must set a target attribute state for every attribute identified for a value." In addition to the specification of a target, councils must also specify baseline attribute states and timeframes over which attribute targets will be met, in addition to further requirements of attributes outlined in Clauses 3.10 and 3.11 of the NPSFM.

Therefore, in the NPSFM sense, attributes are not just metrics to be monitored such that we may assess the extent to which values are provided for, but are also regulatory instruments. Throughout this M&E plan, any use of the word attribute is used in the NPSFM sense, acknowledging the full set of requirements of attributes specified in Clauses 3.10 and 3.11 of the NPSFM.

In Section 7 of this plan we recommend several *metrics*[2] to be monitored for the adaptive management of flow regimes. Metrics are not to be confused with attributes. We recommend metrics that we consider best facilitate adaptive management of water takes—reducing uncertainty about flow-ecology relationships, such that it is possible to more effectively and efficiently manage water takes and achieve planned environmental outcomes (refer to Section 5.3). For most of the metrics we recommend, it is not necessary to define quantitative targets, nor baselines, nor timeframes over which targets will be met.

2.1.2 Metrics will inform the adaptive management of hydrological attributes

Councils need to set water take limits in accordance with the NPSFM and they will need to do so despite substantial uncertainty about how those water take limits affect environmental outcomes (see Clause 1.6, Best Information, of the NPSFM). According to Clause 3.17 of the NPSFM, take limits will need to be expressed as a flow rate, a total volume, or both[3] (a fuller discussion can be found in Section 5.1 of this plan).

The metrics recommended in this plan—and the contents of this M&E plan as a whole—are designed to:

- reduce uncertainty about how planned environmental outcomes are likely to respond to different water take rules or decisions;
- facilitate the development of tools for transparent and defensible
  - evaluation of the outcomes of water take decisions that have been made (assessment of what happened as a result of previous water take decisions); and

[2] We define a metric as any measurable characteristic of a freshwater ecosystem.
[3] Take limits will also need to be set in order to achieve an environmental flow regime, and so must also consider temporal variability in take (flow rates and/or total volumes) among seasons and years (Clause 3.16 of the NPSFM).

- forecasts of the likely outcomes of alternative decisions that could be made in the future (assessment of what might happen in the future, if water take rules are changed); and
- ultimately refine water take limits.

Some of the metrics we recommend may in due course be converted to attributes, serving the functions of a regulatory instrument under the NPSFM as well as a metric to facilitate adaptive management. Identifying which metrics might make suitable flow-sensitive attributes is, however, outside the scope of this plan. Converting metrics into attributes is not necessary for successful implementation of this plan. Converting metrics into attributes will increase the regulatory burden on councils. In differentiating metrics from attributes our aim was to minimise the regulatory burden on councils, without compromising the management of water takes by councils.

The relationship between the metrics recommended in this M&E plan and water take limits is summarised in Figure 2-1, which provides a conceptual model that outlines links between river flow rate observed within a reach and components of ecosystem health affected by flow rate: water quality, physical habitat, and aquatic life and ecological processes. In the long term, the outcomes of this plan, if implemented, will be more effective and efficient water take rules in regional plans.

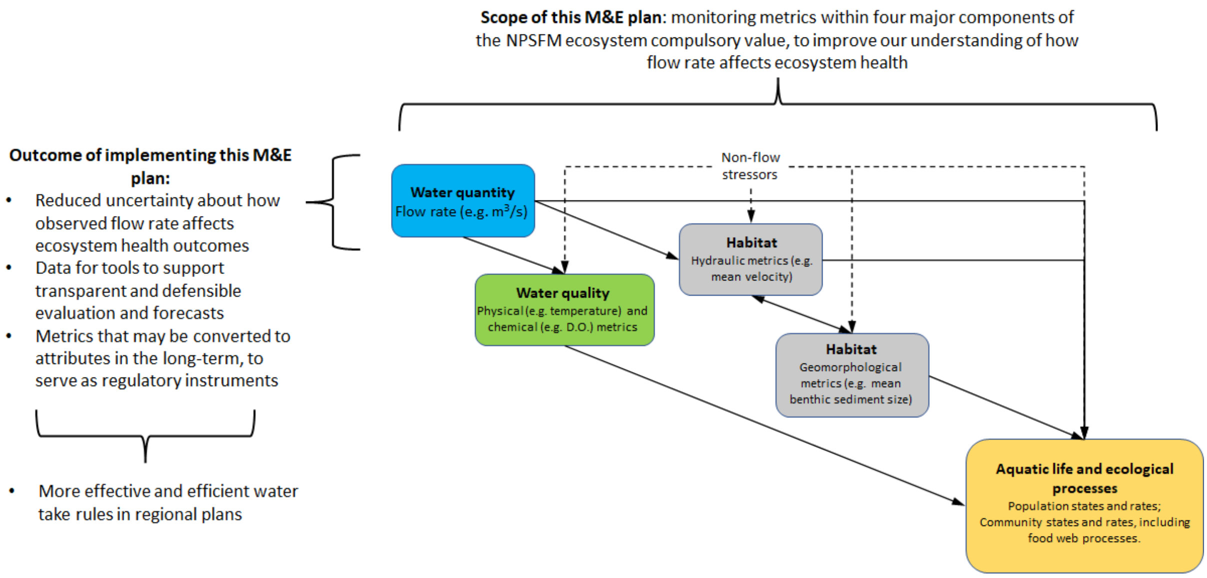

Figure 2-1: Overview of the relationship between the metrics that we recommend for monitoring in this M&E plan and water take rules. Boxes organised by the NPSFM components of ecosystem health.

2.1.3 We do not recommend specific water take limits in this plan

We appreciate that councils need to set water take limits for their 2024 regional plans. Defining water take limits is outside the scope of this M&E plan.

2.2 Additional elements of scope, in brief

The following activities are **within the scope** of this M&E plan:

- Review of the NPSFM 2020 to identify and interpret its requirements with respect to the M&E of river flow regimes and water takes.
- Conducting and analysing surveys of project participants, with the aims of:
  - refining the scope of a nationally consistent M&E plan;
  - building legitimacy through a shared understanding of objectives and limitations of a nationally-consistent M&E plan; and
  - building credibility and relevance through careful definition of objectives.
- Synthesis of the literature on good practice in environmental M&E.
- Use of the basic tools of structured decision making (Gregory et al. 2012, Conroy and Peterson 2013) to define M&E objectives.
- Holding workshops with council core representatives to build legitimacy and increase relevance of the plan, as well as define and identify:
  - river types within which monitoring sites could be distributed;
  - *hydrological stressors*[4] that we wish to manage in an adaptive fashion;
  - conceptual linkages between hydrological stressors and NPSFM values and, in turn, potential attributes.
- Recommendation of strategies to meet M&E objectives and manage trade-offs among those objectives.
- Recommendation of key river types within which to focus a nationally-consistent M&E program.
- Recommendation of metrics.

The following activities are **outside the scope** of this M&E plan:

- Water accounting, including any matters pertaining to the M&E of how water takes affect river flow regimes. This plan focuses on M&E of the links between river flow regimes and riverine ecology.
- Design of a national database for any or all information that might arise from the implementation of this plan.
- Recommendations concerning data curation and data management and storage standards.

[4] Here we define *hydrological stressors* as those properties of the flow regime that negatively affect riverine values. Examples include prolonged low flows that may cause mortality of fishes, and reduced frequency of floods that flush fine sediment from the benthic habitats of rivers. Accordingly, hydrological stressors are not necessarily anthropogenic—events experienced as part of a natural flow regime may have negative effects on riverine values. Water takes may increase the frequency and/or magnitude of hydrological stressors.

- Evaluation of measures of mātauranga Māori. Clause 3.18 of the NPSFM states that monitoring methods must include "measures of mātauranga Māori", identification of which must be undertaken in collaboration with mana whenua and which is outside the scope of the current work.

3 Overview of the design process

A process involving five tasks was used for generating three primary outputs:

1. a set of M&E objectives to guide design;
2. spatial design guidelines; and
3. a set of metrics for a nationally-consistent M&E program to support adaptive management of river flows (Figure 3-1).

The five tasks (rectangles in Figure 3-1) are very briefly explained below. Further details are provided in subsequent sections of this plan.

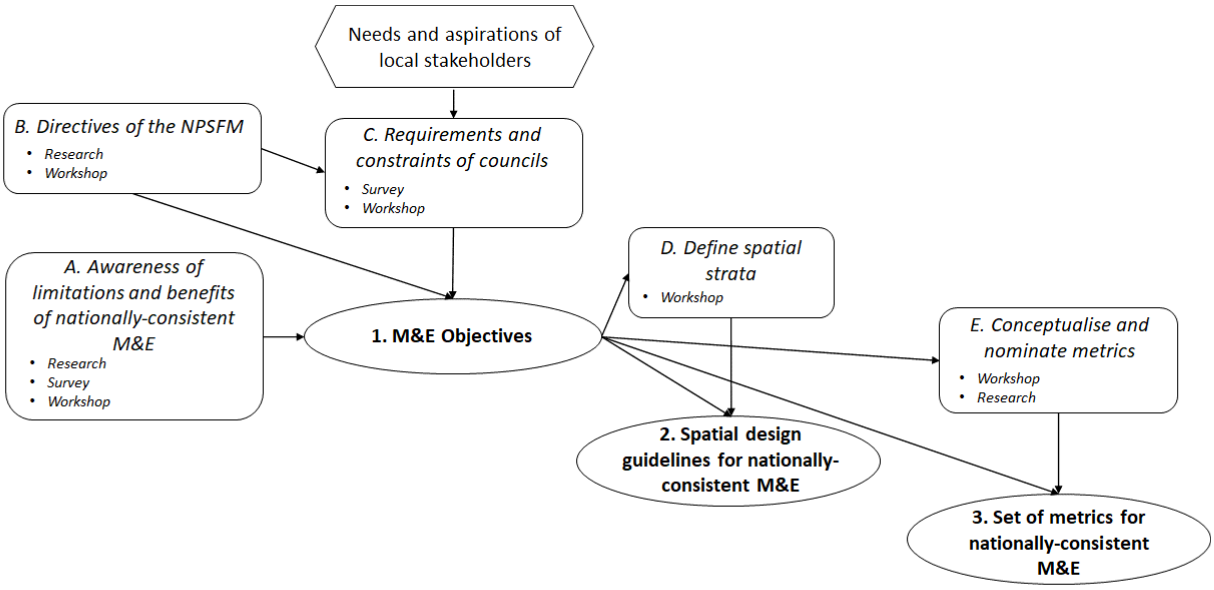

Figure 3-1: Process for defining M&E objectives, spatial design, and metrics of an M&E plan to support adaptive management of river flows. We worked through five tasks (rounded rectangles) to yield three outputs (ovals). Needs and aspirations of iwi and local stakeholders (hexagon) were not directly considered, but indirectly influenced our plan through their influence on the requirements of councils (e.g., as prescribed in regional plans).

3.1 Task A: Deciphering the advantages and disadvantages of nationally-consistent monitoring and evaluation

Failure to consider advantages and disadvantages of implementing a nationally-consistent M&E plan would leave this plan open to the criticism that the full set of options was not considered, thus lowering the legitimacy and credibility of the plan. We summarised the requirements for nationally-consistent M&E presented in policies and plans within New Zealand. We then reviewed the scientific and management literature and summarised the perspectives of M&E experts concerning advantages and disadvantages of nationally-consistent M&E. Finally, we surveyed council core representatives to obtain their perspectives on the advantages and disadvantages of nationally-consistent M&E. The results of the survey were discussed in a workshop.

Further details are provided in Section 4.

3.2 Task B: Flow monitoring and evaluation objectives of the NPSFM 2020

The science team studied the NPSFM 2020 to identify what it requires of councils with respect to the M&E of ecological responses to river flow regimes (hereafter referred to as *flow M&E objectives* for convenience). The science team prepared a summary of the flow M&E objectives of the NPSFM and presented those objectives at a workshop, where the science team and core representatives discussed and achieved a shared understanding of those objectives.

Further details are provided in Section 5.1.

3.3 Task C: Objectives and constraints of councils

The decisions councils make about flow M&E and management will be in response to the needs and aspirations of iwi and specific regional stakeholder groups (Table 3-1), as well as the requirements of central government. Council decisions concerning flow M&E and management will also be shaped by resource constraints, to which the NPSFM (considered in Task B) may give little regard. To ensure this M&E plan satisfies the requirements of both central and regional government, we used a survey and workshops to capture and understand the flow M&E objectives of councils, as well as the resource and infrastructural constraints on meeting M&E objectives.

Further details are provided in Section 5.2.

Table 3-1: Stakeholder groups (not including central government) that may influence decisions of councils. For the purposes of this M&E plan, we define *non-scientific stakeholder groups* as all groups in this table, except scientists. This classification of stakeholder groups follows standard conventions used in structured decision making (Conroy and Peterson 2013; Gregory 2012).

Stakeholder group	Description
Consumers	Members of the public that use freshwater resources either consumptively or non-consumptively. The financial position of individuals in this group is not directly affected by flow management decisions. E.g., anglers/fisherpeople; boaters/canoers/kayakers; those that camp/swim on/in rivers; birdwatchers; trampers.
Non-governmental organisations (NGOs)	NGOs advocate for the conservation or wise use of rivers, but are not part of a government entity legally mandated to manage rivers and/or implement environmental policies/plans. E.g., NZ Forest & Bird.
Economic	Businesses and landholders whose financial position is affected by flow management decisions. Examples include power-generation companies, irrigators, floodplain landholders, regional councils, riverine tourism businesses, and businesses that supply goods and services to these groups.
Resource management agencies (other than councils)	Organisations charged with managing rivers within their legally mandated jurisdiction. E.g., some iwi/hapū; Fish and Game.
Iwi/hapū	Māori have legitimate and legal rights and interests in water management, and there is an expectation from Central government for co-management and full participation in water management decisions.
Scientists	Individuals that have technical knowledge of, or interest in, the flow management decisions, but who are not directly affected by the decision in their technical capacity.

3.4 Task D: Defining spatial strata within which to arrange a network of M&E sites

Our aim was to deliver an M&E plan that is relevant to the major types of rivers subject to significant water takes that are well represented both within regions and among regions, at the national scale. Different river types may respond differently to changes in the flow regime and may support different freshwater values. Therefore, we may require different metrics and/or sampling methods for different river types. To identify major river types that could be incorporated into the M&E, councils nominated river types subject to significant water takes within their region, and that might be differentiated on the basis of their hydrology, geomorphology and ecology. Against each river type councils also submitted assessments of the need for flow management and prevalence within their regions. We then used council nominations of type, need for flow management, and prevalence to define strata within which to arrange a network of M&E sites.

Further details are provided in Section 6.

3.5 Task E: Defining metrics

We defined metrics (see Section 2.1) using a four-step process. The first three steps primarily took place in workshops, allowing council core representatives and scientists to discuss and agree on:

1. which NPSFM values this M&E plan should address directly;
2. key hydrological stressors that should be the focus of adaptive management, hence the requirement for M&E; and
3. conceptual mechanistic links between hydrological stressors and potential metrics to be monitored.

The fourth step involved the science team using information arising from the workshops to nominate flow-affected metrics to support adaptive management under the NPSFM.

Further details are provided in Section 7.

4 Advantages and disadvantages of nationally-consistent monitoring and evaluation

In this section we briefly summarise:

1. the requirement for nationally-consistent M&E as specified in the NPSFM and other relevant policy and planning documents;
2. recommendations on M&E from relevant international literature; and
3. the results of a questionnaire and workshop amongst the science team and core representatives (Appendix A), aimed at identifying advantages and disadvantages of taking a nationally-consistent approach to M&E within New Zealand.

4.1 The requirement for nationally-consistent M&E in policies and plans

Although the NPSFM directs local authorities to undertake monitoring, evaluation and—more broadly—adaptive management of river flow regimes (Section 5.1), it does not state that local authorities must implement a nationally-consistent approach. The Parliamentary Commissioner for the Environment (PCE) report *Aotearoa New Zealand's environmental reporting system* was critical of a localised, authority-specific approach to M&E, and called for (PCE 2019):

- "*the development of a dedicated set of core environmental indicators for the purposes of national environmental reporting, along with*
- *the design of a national-level monitoring network, and*
- *the development, specification and mandating of consistent data collection standards.*

This will ensure New Zealand has a comprehensive and representative national monitoring system with a standardised and consistent approach to collecting, managing and analysing data."

The PCE report addressed the quantity and quality of data available for national environmental reporting by central government. It identified inconsistent approaches across authorities as a barrier to national reporting, and nationally-consistent approaches to monitoring as a way of overcoming that barrier.

Although the Action for Healthy Waterways document included proposals for nationally-consistent M&E of freshwater resources (Ministry for the Environment 2019), the document did not present justifications for such proposals, nor a discussion of the advantages and disadvantages of taking a nationally-consistent approach to M&E.

Priority 3 of the Regional Councils' Research, Science and Technology (RS&T) Strategy highlights the need for consistent application of decision- and evaluation-support tools across local authorities (RCSAG 2020). This RS&T strategy does not, however, state that a nationally-consistent approach to all freshwater M&E is a priority for local authorities.

It follows that there is no clear policy mandate for nationally-consistent M&E of river flow take limits.

4.2 Research on the advantages and disadvantages of nationally-consistent M&E

From our review of the international, peer-reviewed literature, it appears there has been little scholarship aimed at understanding the advantages and disadvantages of nationally-consistent approaches to policy-driven M&E and adaptive management of natural resources. Of the research that does exist, several papers highlight the advantages of taking a coordinated and consistent approach to monitoring. These advantages include (Bricker and Ruggiero 1998, Beard et al. 1999, Parr et al. 2002, Peters et al. 2014):

- Increased efficiency in national evaluation and reporting due to:
  - standard methods of data collection and storage, making it easier to collate, tidy and analyse data;
  - higher coordination of priorities across regions, resulting in less duplication of effort.
- Increased effectiveness of national reporting due to:
  - standard methods of data collection resulting in less method-induced noise in trends, states and responses to policy-driven interventions;
  - increased volumes of data, improving precision of models used to support evaluation.

The disadvantages of nationally-consistent M&E are less commonly considered. An emerging body of literature questions the effectiveness of prescriptive, top-down, policy-driven environmental management practices, albeit not often in the specific context of M&E. This literature acknowledges that socio-ecological systems are *complex adaptive systems* characterized by *wicked problems* (Funtowicz and Ravetz 1993, Game et al. 2014). Solutions to wicked problems are characterized by very high uncertainties; each problem is linked to other problems, such that implementing solutions to one problem often creates surprising socio-ecological responses and potentially new problems elsewhere in the system (Game et al. 2014). Disadvantages to nationally-consistent M&E may include (Danielsen et al. 2009, Lebel and Daniel 2009, Hermans et al. 2013, Chaffin et al. 2014, Game et al. 2014):

- A prescriptive approach may result in effort spent on M&E that is not relevant in some regions.
- Loss of creativity and a diverse range of approaches among regions, from which we may more rapidly learn what works and what does not.
- Less agility of the M&E program, meaning that M&E is less able to adapt to regional surprises and changing, local socio-ecological forces.
- Less engagement with local stakeholders, hence lowered legitimacy and relevance.

4.3 Our perspectives on the advantages and disadvantages of nationally-consistent M&E

The science team and core representatives were asked to submit up to three advantages and three disadvantages of implementing nationally-consistent M&E in response to the NPSFM requirements. A total of 90 statements (50 advantages; 40 disadvantages) were submitted and assigned to one of six advantage classes or six disadvantage classes (Figure 4-1). See Appendix B for detailed presentation of the results of this exercise, and methods applied. The amount of information within statements varied considerably among respondents, such that individual statements were usually relevant to multiple classes. In our analysis, therefore, individual statements could be assigned to up to three classes. Furthermore, separate statements from individual respondents could refer to the same class. This many-to-many mapping of classes to statements meant it was possible for the maximum frequency within classes to exceed the maximum number of statements submitted for either advantages (*N* = 50) or disadvantages (*N* = 40; Figure 4-1).

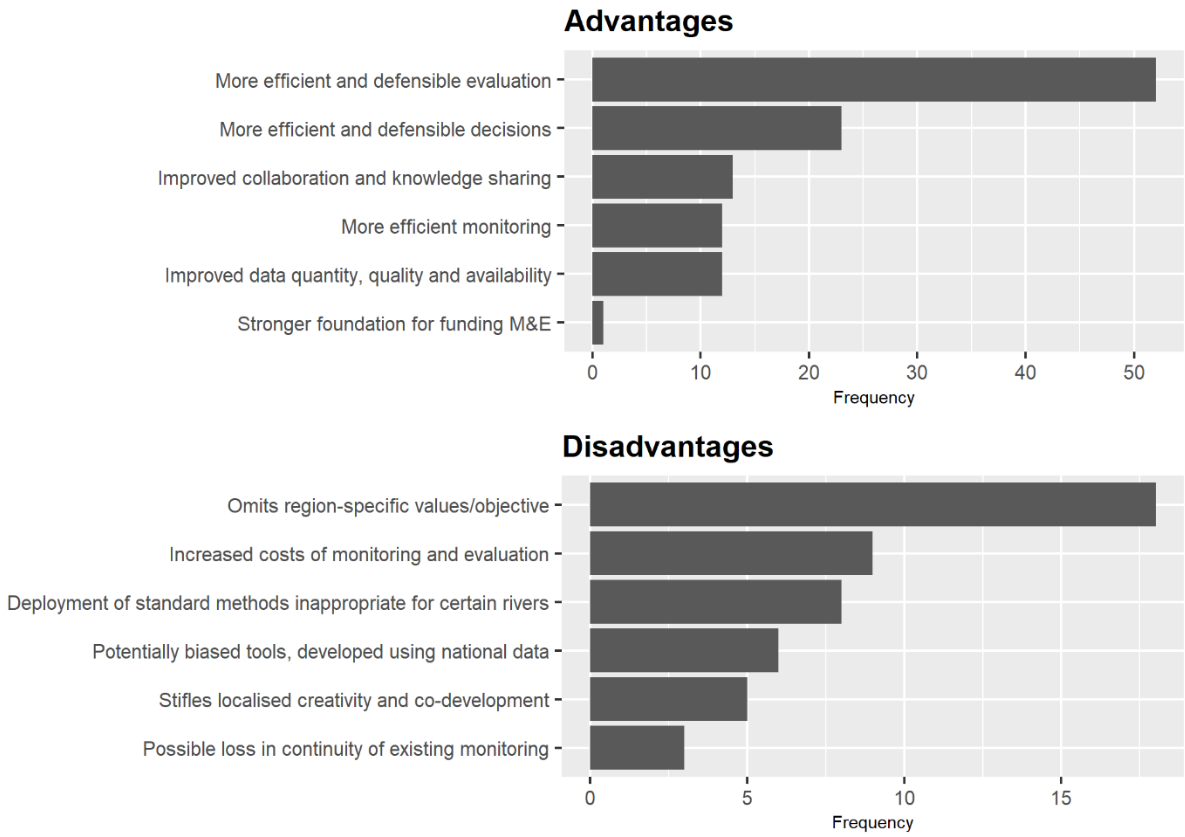

Figure 4-1: Ranking of the advantages and disadvantages of implementing nationally-consistent M&E. Frequency of statements assigned to six advantage/disadvantage themes.

The frequencies with which respondents cited classes of advantages and disadvantages are presented in Figure 4-1, and the explanations for each class are presented in a condensed form in Table 4-1. It is worth noting that respondents may have conflated standardised M&E with co-development of M&E and that this conflation may have increased the respondents' perceptions of advantages arising from consistent M&E. It is possible to have nationally-consistent M&E that is not co-developed, but is prescribed in a top-down fashion by central government.

Particular disadvantages of nationally-consistent M&E may be overcome with good design. For example, instead of implementing a single standardised sampling procedure for a particular metric—one that is likely only going to be effective in certain conditions (see "Deployment of standard methods inappropriate for certain rivers"; Table 4-1; Figure 4-1)—one could implement a family of standardised sampling procedures, covering most of the conditions local practitioners are likely to encounter. Moreover, evaluation and decision-support tools[5] developed using data spanning the nation's catchments will not necessarily yield biased inferences and/or predictions in localised contexts. The magnitude of bias depends on how those models are developed and applied at different spatial scales and locations.

The advantages and disadvantages identified in this analysis are similar to those identified in the literature (but noting the literature cited in the previous section was often not solely referring to M&E, but to ecosystem management/governance in general). One disadvantage we missed was that region-specific approaches may be more agile, hence better able to respond to changing socio-ecological drivers/surprises.

The analysis presented here demonstrates the likely trade-offs among policy-driven M&E approaches. Documenting these trade-offs is important as it demonstrates that **nationally-consistent M&E is not a panacea—it will not meet all the M&E objectives that arise out of policy and council requirements** (see Section 5). Depending on council objectives, this plan for nationally-consistent M&E of ecological response to river flow management may have to be supplemented with region-specific M&E. With respect to this plan, we must identify those M&E objectives, riverine values and metrics that are best suited for nationally-consistent M&E.

Table 4-1: Explanation of the advantages and disadvantages of implementing nationally-consistent M&E.

Advantages	
Class	**Explanation**
More efficient and defensible evaluation	Greater efficiency in evaluating effects of interventions as a result of more data collected in a consistent, standardised fashion. Higher precision of evaluation models achieved more rapidly when multiple regional governments have contributed data towards development of evaluation models. Development of evaluation models shared by multiple regional governments shares/reduces costs. More defensible evaluation as a result of co-development of evaluation models among multiple experts/stakeholders. Higher legitimacy, resulting from co-development, means higher defensibility. More precise evaluation models—resulting from more data—increasing defensibility.
More efficient and defensible decisions	Gains in efficiency of decision-making as a result of more data being incorporated into decision-support models. More data may result in decision-support models being fit for purpose, sooner. Faster reduction of uncertainties. Decisions may be more defensible when made under a nationally-consistent adaptive management framework, co-developed among multiple regional governments. Easier to defend decisions with high legitimacy. Reduced uncertainties in decision-support models as a result of models incorporating more data across replicated management contexts/interventions.

[5] Such as models of how ecological attributes respond to flow regimes.

Improved collaboration and knowledge sharing	Nationally-consistent approach that is co-developed will improve collaboration amongst multiple actors, and may result in more rapid improvement of approaches and reduction of uncertainties as a result of knowledge sharing. Increased legitimacy and relevance of adaptive management, resulting from collaboration.
More efficient monitoring	Reduced duplication of effort in developing methods and designs, including quality assurance procedures and, potentially, data pipelines. A common approach means technical methodological training can be shared across regional governments, and monitoring skills are more easily transferred across regions.
Improved data quality, quantity and availability.	A standardised approach will result in more comparable data. Co-development of sampling methodologies among numerous stakeholders ensures protocols pass several "acid-tests," improving accuracy of data. If a consistent M&E program also results in consistent data management pipelines and storage, then data may become more available. Increased data availability then feeds back into other classes of advantages.
Stronger foundation for funding M&E	Co-development of all approaches results in high legitimacy, relevance and credibility, improving the proposal to central/national government to fund M&E to support policy implementation.

Disadvantages

Class	Explanation
Omits region-specific values/objectives	Certain localised values (e.g., culturally-significant species) and management objectives (e.g., how to set allocation limits in intermittent streams) may not be covered well by a consistent M&E/adaptive management plan. Possible that no single set of indicators/metrics will be appropriate for detecting responses to allocation decisions in all stream types, which vary considerably among regions.
Increased costs of M&E	Existing region-specific M&E commitments may not necessarily be covered by a new national plan, but may need to be continued to fulfil regional plans. Hence implementing a new consistent M&E plan may increase costs of M&E. If the consistent M&E covers values/objectives that are not very relevant to a specific region, then those M&E components may represent superfluous expenses for that regional government. A new, consistent M&E plan may contain methods, processes and/or infrastructures that are new to specific regions, increasing training, capability and infrastructural costs.
Deployment of standard methods inappropriate for certain rivers	If standard sampling methods are implemented, they may not be appropriate for certain regions or rivers; a localised, bespoke sampling approach may be more effective.
Potentially biased tools developed using national data	Evaluation and decision-support models developed using national data may yield inferences/predictions that are "pulled" towards the national average response, hence may be biased with respect to localised responses.
Stifles localised creativity and co-development	A nationally-consistent approach may reduce the involvement of local stakeholders in M&E, reducing legitimacy amongst certain local stakeholders.

	Designing M&E to be nationally-consistent limits the scope of regional governments to design their own approaches, resulting in less creativity and diversity in design at a national scale, across regions. A diversity of approaches at a national scale may facilitate learning which M&E approach yields best outcomes (regionalised designs serve as M&E experiments).
Possible loss in continuity of existing monitoring	Resources are limited, so implementing a policy-driven, nationally-consistent approach may draw resources away from regionalised programs, possibly causing cessation of regionalised programs. It is possible existing M&E activities could be absorbed into a new nationally-consistent program, but if so, that may result in a change of methods, leading to inconsistencies in methods through time, within regions.

5 Flow monitoring and evaluation objectives

Defining explicit objectives is one of the hardest tasks in natural resource management (Conroy and Peterson 2013)[6]. Environmental M&E has been extensively criticized for failing to yield data and tools to effectively support decision-making and decision-evaluation (Lindenmayer et al. 2013, Waylen et al. 2019), instead yielding a focus on 'documenting the decline' without pointing to solutions. A major cause of the problem is poorly considered M&E objectives. Scientists and managers responsible for designing environmental M&E have been reluctant to move away from traditional M&E objectives (Waylen and Blackstock 2017) and, when formulating objectives, have been constrained by "what is" instead of considering "what could be" (Cundill and Fabricius 2009). In other words, M&E designers frequently fall into the following decision traps:

- the *status quo trap*—the situation where the decisions we make are biased towards maintenance of the status quo; and
- the *sunk-cost trap*—which occurs when we make choices that justify past, flawed choices.

The status quo and sunk-cost traps are major contributors to poor decision-making in general (Hammond et al. 1998). In the context of natural resource management, these cognitive biases frequently lead to what Conroy and Peterson (2013) call *pre-emptive rejection of objectives*, which in turn contributes to poor environmental management.

Some simple strategies can help avoid the status quo and sunk-cost traps (Hammond et al. 1998, Gregory et al. 2012, Conroy and Peterson 2013) and we employed them here to formulate objectives. When formulating M&E objectives we deliberately downplayed effort and/or costs of objectives. We translated NPSFM policies and clauses to M&E objectives as faithfully as we could, without questioning the feasibility of policy requirements. We did not let specification of an objective be constrained by other objectives. Questions within the online survey (see Section 5.2.1) were often framed to elicit M&E ideals from councils. Questions pertaining to resource constraints were presented in the last section of the survey; we did this to avoid having council responses to the majority of questions anchored[7] by mental models based on resource limitations. These strategies were an attempt to discourage "failure-fearing" (Hammond et al. 1998) and to encourage reimagining "what could be" rather than "what is" (Cundill and Fabricius 2009, Conroy and Peterson 2013).

Our approach was likely to yield trade-offs among objectives. We saw this as an advantage as it allowed us to explicitly document those trade-offs, achieving two ends: First, it allowed the science team and the core representatives to achieve a shared understanding of the unavoidable limitations of flow M&E. Second, explicit documentation of the unavoidable trade-offs among M&E objectives increased the transparency of the M&E decisions made (consistent with the recommendation of Waylen et al. 2019).

[6] The following excerpt from Conroy and Peterson (2013) is relevant: 'In short, it is not an understatement to say that everything, and we mean everything, is largely determined by the objectives. Right now you might be thinking, "this should be easy, I already know my objectives." Unfortunately, it isn't easy. In our experience, we find that identifying objectives is the most difficult step in the structured decision-making process.'

[7] 'Anchoring' occurs when the decisions we make (responding to a survey question, in this case) are biased by information we receive prior to making the decision.

5.1 Objectives of the NPSFM 2020

The National Objectives Framework (NOF) comprises Subpart 2 of the implementation section (Part 3) of the NPSFM. The NOF presents clauses that describe the process by which councils must manage freshwater resources. Certain clauses of the NPSFM require fuller explanations in order to advance implementation.

5.1.1 A challenging shift in how we prioritise freshwater values

Te Mana o te Wai is a fundamental concept underpinning the NPSFM, and presents a challenging shift in how we prioritise freshwater ecosystem values within New Zealand. Te Mana o te Wai states that there is a hierarchy of obligations such that managers must prioritise:

A. first and foremost, the health and well-being of water bodies and freshwater ecosystems;

B. second, the health needs of people (such as drinking water);

C. third, the ability of people and communities to provide for their social, economic, and cultural well-being, now and in the future.

With respect to the management of river flows, Te Mana o te Wai requires us to ensure a river's flow regime comprises the necessary hydrological characteristics to maintain and/or restore critical ecological processes before we consider managing river flows to meet the needs of humans.

NPSFM Policy 1—"Freshwater is managed in a way that gives effect to Te Mana o te Wai"—is an overarching policy that we were mindful of when developing this M&E plan. Te Mana o te Wai acknowledges Māori *tikanga* (indigenous laws and values) in central government policy, reiterates that river health is a national priority for New Zealand, and underscores the importance of an M&E plan that facilitates successful restoration and maintenance of New Zealand's riverine values.

5.1.2 The requirement to manage water takes

NPSFM Policy 11 requires that "Freshwater is allocated and used efficiently, all existing over-allocation is phased out, and future over-allocation is avoided." The NPSFM 2020 defines "*over-allocation*, in relation to both the quantity and quality of freshwater, as the situation where (a) resource use exceeds a *limit*; or (b) if limits have not been set, a *freshwater management unit* (FMU) or part of an FMU is degraded or degrading" (NPSFM 2020 Section 1.4). Limits to resource use are set in order to meet *attribute target states*, which are in turn set to meet expected *environmental outcomes* stated in regional plans.

NPSFM Policy 11 therefore directs councils to set water take limits to ensure the flow regimes of rivers are such that, at a minimum, the freshwater values listed in Appendices 2A and 2B of the NPSFM are maintained or enhanced.

In addition to Policy 11, other NPSFM policies are relevant to river flows. Given that flow is a "master variable" driving river ecology (Power et al. 1995, Walker et al. 1995), NPSFM Policies 7 – 10 implicitly require improved management of water takes:

Policy 7: The loss of river extent and values is avoided to the extent practicable.

Policy 8: The significant values of outstanding water bodies are protected.

Policy 9: The habitats of indigenous freshwater species are protected.

Policy 10: The habitat of trout and salmon is protected, insofar as this is consistent with Policy 9.

Clause 3.16 of the NPSFM—Setting environmental flows and levels—states that councils must set *environmental flows* for FMUs. Environmental flows "must be expressed in terms of the water level and flow rate, and may include variability of flow (as appropriate to the water body) at which, for flows and levels in rivers, any taking, damming, diversion, or discharge of water meets the environmental outcomes for the river, any connected water body, and receiving environments."

Clause 3.16 of the NPSFM directs councils to define "water level(s) and flow rate(s)" of rivers such that their ecological values (first and foremost, following Te Mana o te Wai) are maintained or enhanced. We note that defining an environmental flow as a single static water level or flow rate is unlikely to result in ecological outcomes being met. Ecological processes in rivers have adapted to—and are therefore driven by—a flow regime (Lytle and Poff 2004). From a statistical perspective, a river's flow regime is defined by many variables (e.g., frequency of events where discharge exceeds/drops below some threshold; mean duration of such events; seasonal patterns) (Poff et al. 1997). Broadly, the frequency, duration, magnitude, timing and rates of change of specific discharge values affect riverine processes (Poff et al. 1997). For the purposes of this M&E plan, we define a *flow regime* as:

- a *quantifiable representation of the main characteristics of a time series of discharges, calculated over a period spanning many years*. The flow regime may represent variability at several temporal resolutions.

In this M&E plan we view the challenge of setting environmental flows (NPSFM 2020, Clause 3.16) as *specifying a flow regime that supports the expected environmental outcomes of regional plans*.

Within the above definition there is deliberate flexibility in the temporal resolution used to characterise different aspects of a flow regime. Some aspects of a flow regime are best characterised using flow data with a fine temporal resolution. For example, 15-minute resolution flow data are required to characterise instantaneous peak flows and sub-daily rates of change. It may be necessary to manage these aspects of a flow regime because they are relevant to sediment movement and hydropeaking respectively. Other aspects of a flow regime are best characterised using coarser resolution flow data. For example, monthly mean flow data are often used to characterise seasonality, which may need to be managed because it is relevant to seasonal ecological processes such as fish spawning.

Clause 3.17 of the NPSFM—Identifying take limits—states that in order to protect a river's environmental flow regime (defined in response to Clause 3.16) every council:

A. "must identify take limits for each FMU; and

B. must include the take limits as rules in its regional plan(s); and

C. must state in its regional plan(s) whether (and if so, when and which) existing water permits will be reviewed to comply with environmental flows and levels; and

D. may impose conditions on resource consents" (Clause 3.17(1); NPSFM 2020).

"Take limits must be identified that:

A. provide for flow or level variability that meets the needs of the relevant water body and connected water bodies, and their associated ecosystems; and

B. safeguard ecosystem health from the effects of the take limit on the frequency and duration of lowered flows or levels; and

C. provide for the life cycle needs of aquatic life; and

D. take into account the environmental outcomes applying to relevant water bodies and any connected water bodies (such as aquifers and downstream surface water bodies), whether in the same region or another region" (Clause 3.17(4); NPSFM 20202).

"Take limits must be expressed as a total volume, a total rate, or both a total volume and a total rate, at which water may be:

A. taken or diverted from an FMU or part of an FMU; or

B. dammed in an FMU or part of an FMU" (Clause 3.17(2); NPSFM 2020).

Taken together, Clauses 3.16 and 3.17 of the NOF require:

1. an understanding of how take limits and imposed conditions within resource consents (e.g., restrictions on when water can be taken) within FMUs affect flow regimes;

2. an understanding of how river flow regimes within FMUs affect expected environmental outcomes.

To obtain an understanding of how take limits and imposed conditions within resource consents affect flow regimes (requirement 1, presented immediately above), councils must design and implement effective and efficient *water accounting systems* that (a) monitor the volumes of water being taken from a river; and (b) have the systems and/or tools to determine how observed[8] water takes affect the flow regimes of rivers within an FMU. **Guidance on design and implementation of water accounting systems is not within the scope of this M&E plan**. This M&E plan focuses on how hydrological and ecological metrics might be monitored and used by councils to develop an understanding of how river flow regimes affect environmental outcomes (requirement 2 of Clauses 3.16 and 3.17).

5.1.3 The requirement for monitoring, evaluation and reporting

NPSFM Policy 13 requires "the condition of water bodies and freshwater ecosystems is systematically monitored over time, and action is taken where freshwater is degraded, and to reverse deteriorating trends."

NPSFM Policy 14 requires **"**information (including monitoring data) about the state of water bodies and freshwater ecosystems, and the challenges to their health and well-being, is regularly reported on and published."

Several clauses including 3.8, 3.18, 3.19 and 3.30 of the NOF contain specific guidance to give effect to Policies 13 and 14. These clauses do, however, require some interpretation to facilitate their

[8] Observed water take from a river may be defined as the monitored or modelled volume of water taken from a river by consented takes upstream of a specific river reach during a specific time period. Observed water take should not be conflated with consented water take.

translation into M&E objectives. To that end, we highlight the key statements pertaining to monitoring and evaluation in Clauses 3.8, 3.18, 3.19 and 3.30 then present our interpretation of these statements as M&E objectives.

Clause 3.8 of the NPSFM—Identifying FMUs and special sites and features—states that "monitoring sites for an FMU must be located at sites that are either or both of the following: (a) representative of the FMU or relevant part of the FMU; (b) representative of one or more primary contact sites in the FMU" (Clause 3.8 (4)).

Clause 3.18 of the NPSFM—Monitoring—states "monitoring methods must recognise the importance of long-term trends, and the relationship between results and their contribution to evaluating progress towards achieving long-term visions and environmental outcomes for FMUs and parts of FMUs."

Clause 3.18 also states that "methods must include measures of:

A. mātauranga Māori; and

B. the health of indigenous flora and fauna."

Measures of mātauranga Māori must be set in collaboration with mana whenua and are outside the scope of the current work. Metrics pertaining to the health of indigenous flora and fauna will be considered in Section 7.

Clause 3.19 of the NPSFM—Assessing trends—states "in order to assess trends in attribute states (that is, whether improving or deteriorating), every regional council must:

1. (a) determine the appropriate period for assessment (which must be the period specified in the relevant attribute table in Appendix 2A or 2B, if given); and (b) determine the minimum sampling frequency and distribution of sampling dates (which must be the frequency and distribution specified in the relevant attribute table in Appendix 2A or 2B, if given); and (c) specify the likelihood of any trend.

2. If a deteriorating trend is more likely than not, the regional council must: (a) investigate the cause of the trend; and (b) consider the likelihood of the deteriorating trend, the magnitude of the trend, and the risk of adverse effects on the environment.

3. If a deteriorating trend that is the result of something other than a naturally occurring process is detected, any part of an FMU to which the attribute applies is degrading and Clause 3.20 applies.

4. If a trend assessment cannot identify a trend because of insufficient monitoring, the regional council must make any practicable changes to the monitoring regime that will or are likely to help detect trends in that attribute state."

Clause 3.30 of the NPSFM—Assessing and reporting—presents numerous challenging requirements, including but not limited to:

At least every five years every regional council must prepare and publish:

- "a comparison of the current state of attributes as compared with target attribute states;

- an assessment of whether the target attribute states and environmental outcomes for each FMU or part of an FMU in the region are being achieved and, if not, whether and when they are likely to be;
- if monitoring shows that an FMU or part of an FMU is degraded or degrading, information on the known or likely causes;
- a description of the environmental pressures on each FMU (such as water takes, sources of contaminants, or water body modification) as indicated by information from the freshwater accounting systems referred to in Clause 3.29;
- an assessment of the cumulative effect of changes across multiple sites within an FMU and multiple attributes during the period covered by the assessment;
- predictions of changes, including the foreseeable effects of climate change, that are likely to affect water bodies and freshwater ecosystems in the region."

Our interpretation of these clauses within the context of river flow management yields the following objectives (Os) of our plan for flow M&E:

O1.—Determine how current water takes are affecting metric state and trends. *Select and monitor sites and metrics to help determine how changes in the flow regime affect metrics within river reaches throughout FMUs. This environmental monitoring must be undertaken alongside monitoring of water takes and other forms of hydrological alteration as is required for water accounting purposes.*

O2.—Forecast how future water takes are likely to affect metric state and trends. *Sites and metrics must be selected with a view for how data will be used to forecast the response of metrics to future environmental changes, including future water management scenarios.*

An effective water accounting program is critical to meeting O1 and O2 but is outside the scope of the current M&E plan. Objectives 1 and 2 have been formulated under the assumption that councils have a water accounting program that determines relationships between water takes and flow regimes of rivers within FMUs.

O3.—Disentangle the influence of multiple drivers. *Monitor sufficient environmental metrics within rivers of an FMU such that councils may, as much as practicable, partition the influence of multiple drivers—both anthropogenic and natural—on metrics, and determine the effects of the flow regime on metrics relative to other drivers.*

The NPSFM does not offer guidance on the tools and/or processes by which councils might achieve Clauses 3.18, 3.19 and 3.30, hence O1, O2 and O3. Later, in Section 5.2, we discuss tools/processes that may facilitate meeting O1, O2 and O3, following the input from councils during this project.

O4.—Monitoring sites must be representative of the types of rivers most affected by water takes and flow management*. Locate sites within FMUs such that the sites are, as much as practicable, representative of the types of rivers within the FMU that:*

- *are subject to water takes; and*
- *support compulsory or other values, as specified in Appendices 2A and 2B of the NPSFM 2020.*

***O5.—Implement adaptive monitoring*[9]*.** Once every five years, this M&E plan should be reviewed to determine whether it is fulfilling the objectives stated at the beginning of this plan. Such a review would be aimed at identifying which elements of this plan were adopted, and which elements should be retained, modified, or dropped from the plan. If the M&E plan is not fulfilling M&E objectives it should be amended to do so, in a manner that, as much as practicable, preserves the integrity of previous monitoring time series data.*

Objective 5 is consistent with good practice in monitoring and evaluation (Lindenmayer and Likens 2009), as well as Clause 3.19(4).

5.1.4 The requirement for adaptive management

Adaptive management is an iterative process involving:

- making decisions despite high uncertainty;
- monitoring the outcomes to decisions;
- evaluating how the decision(s) affected the outcomes;
- updating our knowledge/models of the system based on those evaluations; and
- using updated knowledge/models to predict the likely outcomes of alternative decisions in the future (Walters 1997, Gregory et al. 2006).

Well-designed M&E programs play a keystone role in adaptive management (Westgate et al. 2013).

Clause 1.6 of the NPSM—Best information—requires councils to make decisions to give effect to the NPSFM despite uncertainty about the outcomes. Clause 1.6 (2) states that, where uncertainty exists “local authorities must: (a) prefer sources of information that provide the greatest level of certainty; and (b) take all practicable steps to reduce uncertainty.” Clause 1.6 (3) states that “A person who is required to use the best information available at the time… must not delay making decisions solely because of uncertainty about the quality or quantity of the information available.”

Clause 3.16 2(b) of the NPSFM states that environmental flows (flow regimes) may be adapted over time to ensure proposed environmental outcomes are met in the long term.

These clauses—along with the policies and clauses presented in the preceding section on the requirement for monitoring, evaluation and reporting—imply the need for adaptive management of river flow regimes. We note, however, that the NPSFM 2020 makes no explicit mention of ‘adaptive management’. We also recognise that execution of adaptive management would require alterations to water allocation and consenting processes that are beyond the scope of our current work.

Consistent with Clauses 1.6 and 3.16 of the NPSFM, we add the following M&E objective:

***O6.—Flow M&E must support adaptive management of water takes**. Monitoring must aim to reduce uncertainty about flow-metric relationships, and yield information to facilitate transparent and defensible evaluations and forecasts of how metrics have and will respond to flow management decisions.*

[9] “Adaptive monitoring” is a term coined by Lindenmayer and Likens (2009), now part of the environmental M&E vernacular.

Objective 6 states the overarching purpose of objectives O1, O2 and O3, and ties together those objectives in the context of an iterative decision-making cycle.

5.2 Objectives of councils

5.2.1 Survey questions

Having established the policy requirements of central government, we then obtained council perspectives on:

- what metrics should be monitored; and
- how monitoring data should be used to inform decision-making and evaluation.

The perspectives of council core and additional representatives were obtained through an online survey. Answers to survey questions were used to formulate flow M&E objectives additional to those arising from the NPSFM. Responses to survey questions and M&E objectives were echoed back to core representatives at workshops, with the objective of collaboratively finalizing M&E objectives. This process was employed to:

- increase the relevance of M&E objectives to council requirements; and
- increase the legitimacy of the M&E plan as a result of the science team and core representatives working together to achieve a shared understanding of M&E objectives.

Questions and workshop discussions focused on identifying and understanding the dual requirements of policy and non-scientific stakeholders (Table 3-1), as well as resource constraints of councils. Acknowledging that the survey was to inform workshop discussion, we made it short and simple, comprising 19 multi-choice questions divided into five sections. A comments box was included at the end of each section, allowing council representatives to share additional perspectives and context. The full survey is presented in Appendix C and we briefly outline its structure and objectives here.

After obtaining information about the respondent's role, how long they had been working in the resource management sector and their workplace, questions were organized into the following sections:

1. Alignment of metrics with values of stakeholders and partners.
2. Ease of reporting and communicating with stakeholders and partners.
3. Sensitivity of metrics to flow management decisions.
4. Tools and processes to support decisions and evaluations.
5. Resource and logistical constraints.

Alignment of metrics with values of stakeholders and partners. Despite calls to more closely align environmental M&E with stakeholder values (Turnhout et al. 2007, Waylen and Blackstock 2017), many—perhaps most—freshwater M&E programs focus on metrics that have several degrees of separation from the values of non-scientific stakeholders (Mace and Baillie 2007, Friberg et al. 2011, Kuehne et al. 2017). A focus on metrics that have little direct relevance to the values and concerns of stakeholders may make it more difficult for councils to communicate the desired outcomes of flow interventions, and demonstrate the importance of environmental flows to society (Chess et al. 2005).

The first section of our survey contained questions designed to elicit council perspectives on the importance of aligning metrics with stakeholder values. Answers to questions within this section informed which metrics might be measured as well as how they should be measured (e.g., presence/absence cf. abundance of species). At the beginning of this section we highlighted to respondents that questions within this section referred to non-scientific stakeholders that do not receive financial, social and/or political benefits from water takes. That is, with respect to economic stakeholder group members (Table 3-1), respondents were requested to consider only those economic stakeholder group members whose financial position may be negatively affected by water takes.

Ease of reporting and communicating with stakeholders and partners. Questions in this section were about what forms of information would best facilitate reporting and communicating environmental outcomes to non-scientific stakeholder groups, including economic stakeholders that benefit from water takes (e.g., irrigators and agribusiness). Some questions in this section were about how a metric's units of measurement and/or degree of variable reduction might affect communication with stakeholders. By *degree of variable reduction* we mean the extent to which responses of numerous individual variables (e.g., species) are integrated, then reduced down to single multivariate indices. Metrics that are abstract summaries of specific structures (e.g., invertebrate community indices) or processes (e.g., metabolism) may be deemed most interesting or relevant by scientists, but their relevance to the values of non-scientific stakeholders is often not obvious (Chess et al. 2005, Seeteram et al. 2018, Stainback et al. 2020). If the relationships between abstract metrics and the properties of the rivers most valued by stakeholders are not (a) well understood; and (b) carefully communicated to stakeholders, then the relevance of M&E to non-scientific stakeholders may be reduced (Schiller et al. 2001, Chess et al. 2005). Loss of relevance to stakeholders undermines the legitimacy of the M&E plan.

Other questions in this section were presented to elicit council perspectives on how place- and scale-specificity of information might affect reporting to non-scientific stakeholders. The values of non-scientific stakeholders are very often place-specific (Verbrugge et al. 2017). This creates a challenge for councils—and for environmental M&E in general—because it is unlikely resources will permit M&E of all rivers within a region. Capturing expectations that stakeholders may have concerning the place- and scale-specificity of evaluations allows us to explicitly account for possible tensions between ideals on one hand, and what we can realistically achieve on the other.

Sensitivity of metrics to flow management decisions. The speed with which a metric exhibits a response to flow events and regimes will vary among metrics. Some metrics (e.g., depth) may respond almost instantaneously to changes in river flow, while others (e.g., presence/absence of a long-lived fish species) may not exhibit a rapid response, and only respond after many years of sustained changes to the flow regime (Thompson et al. 2018). Questions in this section of the survey were designed to prompt thinking about duration of council reporting cycles and, by extension, the speeds of metric response to flow alteration that would best facilitate reporting to stakeholders, including central government.

Which metrics we measure will also affect how well we can attribute ecological response to specific changes in the flow regime and, hence, specific decisions. While reach hydrology is directly affected by changes in flow, macroinvertebrate community structure may be indirectly affected by changes in flow as well as other stressors. Consequently, evaluation of flow impacts on hydrology will be associated with higher confidence than evaluation of flow impacts on macroinvertebrates and other metrics that are indirectly affected by flow and multiple interacting stressors. Some questions in this section of the survey were presented to elicit council opinions on the need to monitor metrics whose responses can be confidently linked to flow management decisions.

Tools and processes to support decisions and evaluations. Flow management decisions and evaluations can be contentious and are closely scrutinized by stakeholders (Poff et al. 2003, Ryan et al. 2021, Stewardson et al. 2021). M&E programs can yield critical data for the development of decision- and evaluation-support tools to support more defensible and transparent decisions and evaluations, but such outputs need to be explicitly stated as objectives at the design phase (Stoffels et al. 2018).

Answers to questions posed within this section informed workshop discussions on the extent to which data collected by an M&E plan should be used to develop quantitative decision- and evaluation-support tools. Some questions within this section of the survey were posed to increase our understanding of the degree to which flow management decisions will be scrutinized by stakeholders within regions. To provide further context to the answers provided, we determined the level of experience council representatives have had with water resource disputes within their regions. Other questions within this section were designed to obtain council perspectives on the extent to which quantitative tools have historically been used to support water management decisions or evaluations within their region.

Resource and logistical constraints. Questions within this section of the survey were presented to ensure we explicitly captured council views about resource (e.g., funding) and logistical (e.g., constraints on site locations) constraints. A shared understanding of such constraints helps frame subsequent discussion at workshops and helps to ensure the M&E plan is fit for purpose.

5.2.2 Survey results

A total of 41 people responded to the survey, representing 13 councils (Table A-1). A total of 19 respondents worked within councils in a policy/planning and/or consenting role (hereafter, *PPC experts*), while 22 respondents were scientists (Table 5-1). Fifty-six percent of respondents had over 10 years of experience in the natural resource management sector, while 93% of respondents had more than five years of experience in the sector (Table 5-1).

Table 5-1: Composition of respondents by role and experience.

Role	Experience	*N*
Policy/planning and/or consenting (PPC experts)	> 10 years	11
	5 – 10 years	7
	< 5 years	1
Scientist	> 10 years	12
	5 – 10 years	8
	< 5 years	2

Alignment of metrics with values of stakeholders and partners

Scientists and PPC experts generally shared similar views on the alignment of metrics with stakeholder values (Figure 5-1). Council representatives *agreed* or *strongly agreed* that metrics should directly show how stakeholder values respond to changes in river flow (Figure 5-1, Q1). The majority (>80%) of council representatives either *disagreed* or *strongly disagreed* that presence/absence of species was sufficient for communicating outcomes to stakeholders (Figure 5-1, Q2). Subsequent workshop discussions showed that, taken together, responses to Q1 and Q2 yield the following objective:

***O7.—Determine how flow management affects the abundance of fishes and other high-value species**. Data must be collected with a clear plan for how it will be used to evaluate effects of the flow regime on the abundance of fishes and potentially other species of direct value to stakeholders (e.g., other mahinga kai species like kōura and kākahi).*

All scientists and PPC experts *agreed* or *strongly agreed* that the amount of water in a river *per se* is important to stakeholders within their regions (Figure 5-1, Q3). These responses, as well as subsequent discussion during the workshops demonstrated that councils are aware that the amount of water *per se* is important to numerous recreational services that rivers provide, such as kayaking, boating and swimming (Mosley 1983, Brown et al. 1991, Cocklin et al. 1994).

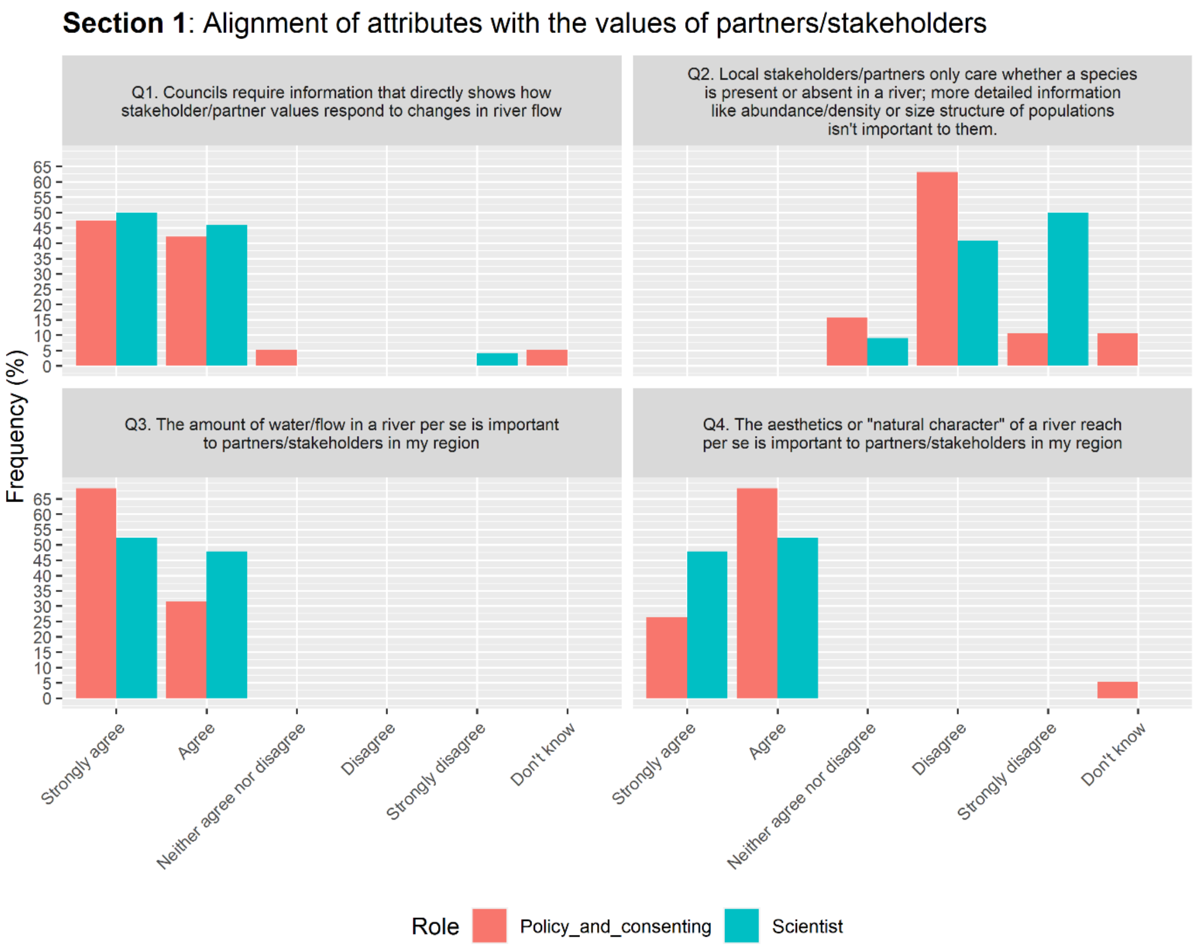

Figure 5-1: Council perspectives concerning the alignment of flow-dependent metrics with the values of stakeholders.

Rivers that are dry or have very little flow may have aesthetics that are undesirable to stakeholders (Pflüger et al. 2010). Indeed, all survey respondents *agreed* or *strongly agreed* that the aesthetics and/or natural character of a river are important to stakeholders (Figure 5-1, Q4). This response may reflect councils' awareness of several things.

First, an awareness of Te Mana o te Wai, a fundamental overarching concept of the NPSFM (Clause 1.3 of NPSFM 2020), reflecting Māori *tikanga* (indigenous laws and values), and which prioritises the health/wellbeing of the river in and of itself, over the services that the river provides to humans (Te Aho 2019). Māori do not espouse an anthropocentric view of a river as a resource to be used or a hazard to be controlled, but instead as a living being with its own *mana* (authority/prestige) and *mauri* (life-force) (Te Aho 2010, Harmsworth et al. 2016). According to Māori tikanga and *kaitiakitanga* (guardianship[10]), a river's natural hydrogeomorphic processes must be understood and respected (Brierley et al. 2019).

Second, council responses may reflect an awareness that the values of all stakeholders—irrespective of ethnicity—are linked to natural hydrogeomorphic processes in rivers (Mosley 1983, Brierley et al. 2019). The importance of natural hydrogeomorphic processes in the maintenance of values is reflected in the NPSFM values presented in Appendix 1B(1)—Natural form and character, Appendix 1B(3)—Wai tapu and Appendix 1B(1)—Transport and tauranga waka.

Third, council responses to Q4 may also reflect an awareness of New Zealand's Resource Management Act (RMA) which states that the "*natural character*" of rivers and "their margins" is a "matter of national importance" and something to be preserved by councils[11]. Although *natural character* is not defined within the RMA, the RMA's directive to preserve it is reflected in regional plans and policies[12], where natural character is broadly interpreted as the extent to which natural riverine processes have been preserved (Boffa Miskell 2018).

Responses to Q3 and Q4 lead to the following objective:

O8.—Determine how flow management affects key hydrogeomorphic metrics of a river reach*. Key hydrogeomorphic metrics must be identified and monitored at gauged sites[13]. Prioritise hydrogeomorphological variables that most strongly affect valued species (O7).*

Ease of reporting and communicating with stakeholders and partners

Scientists and PPC experts shared similar responses to questions presented in Section 2 of the survey (Figure 5-2). The majority (>80%) of respondents thought that it was *very important* or *extremely important* that metrics are easily interpreted by stakeholders in units that align with their understanding and values (Figure 5-2, Q5). When compared with abstract ecological indices (e.g., an MCI score), metrics that align with stakeholder values and have units that are easily understood (e.g., number of fish within a reach above a minimum size limit) may be easier to communicate to the

---

[10] More than guardianship, *kaitiakitanga* is the ethic of protecting the environment for its own sake, as well as for present and future generations to use and enjoy (Te Aho 2019).

[11] Resource Management Act 1991 No 69 (as at 12 April 2022), Public Act Contents – New Zealand Legislation. Within Section 6 of the RMA (Matters of national importance), it states that "all persons exercising functions and powers under it, in relation to managing the use, development, and protection of natural and physical resources, shall recognise and provide for the following matters of national importance: (a) the preservation of the natural character of the coastal environment (including the coastal marine area), wetlands, and lakes and rivers and their margins, and the protection of them from inappropriate subdivision, use, and development."

[12] E.g.,: [Natural character of braided rivers | Environment Canterbury (ecan.govt.nz)]

[13] Sites at which discharge is monitored.

public and may increase legitimacy of M&E and river management more generally (Schiller et al. 2001, Chess et al. 2005).

The response to Q6 somewhat contradicts the response to Q5. Approximately 60% of scientists and 40% of PPC experts *agreed* or *strongly agreed* that multivariate indices are an effective way to communicate ecological outcomes to stakeholders (Figure 5-2, Q6). This may reflect the experience of council representatives in communicating the meaning of commonly used ecological indicators (e.g., MCI) to stakeholder groups. However, approximately 40% of council staff and PPC experts *neither agreed nor disagreed* or *disagreed* that multivariate indices are an effective way of communicating ecological outcomes. There is, therefore, some ambivalence amongst council representatives concerning how well multivariate indices communicate ecological outcomes to stakeholders.

When asked if we require metrics that summarise ecosystem state and trends down to as few numbers as possible, the majority of council representatives *neither agreed nor disagreed* (Figure 5-2, Q7). Just under 50% of scientists and approximately 35% of PPC experts *agreed* or *strongly agreed* that metrics should summarise ecosystem state and trends down to as few numbers as possible. In asking Q7 we reframed Q6 with the aim of further eliciting the consistency of council opinions on the extent to which ecological information should be simplified, reduced and/or represented as abstract indices. As was the case with Q6, there was some ambivalence among council representatives concerning the extent to which ecological information should be reduced or simplified (Figure 5-2, Q6-Q7).

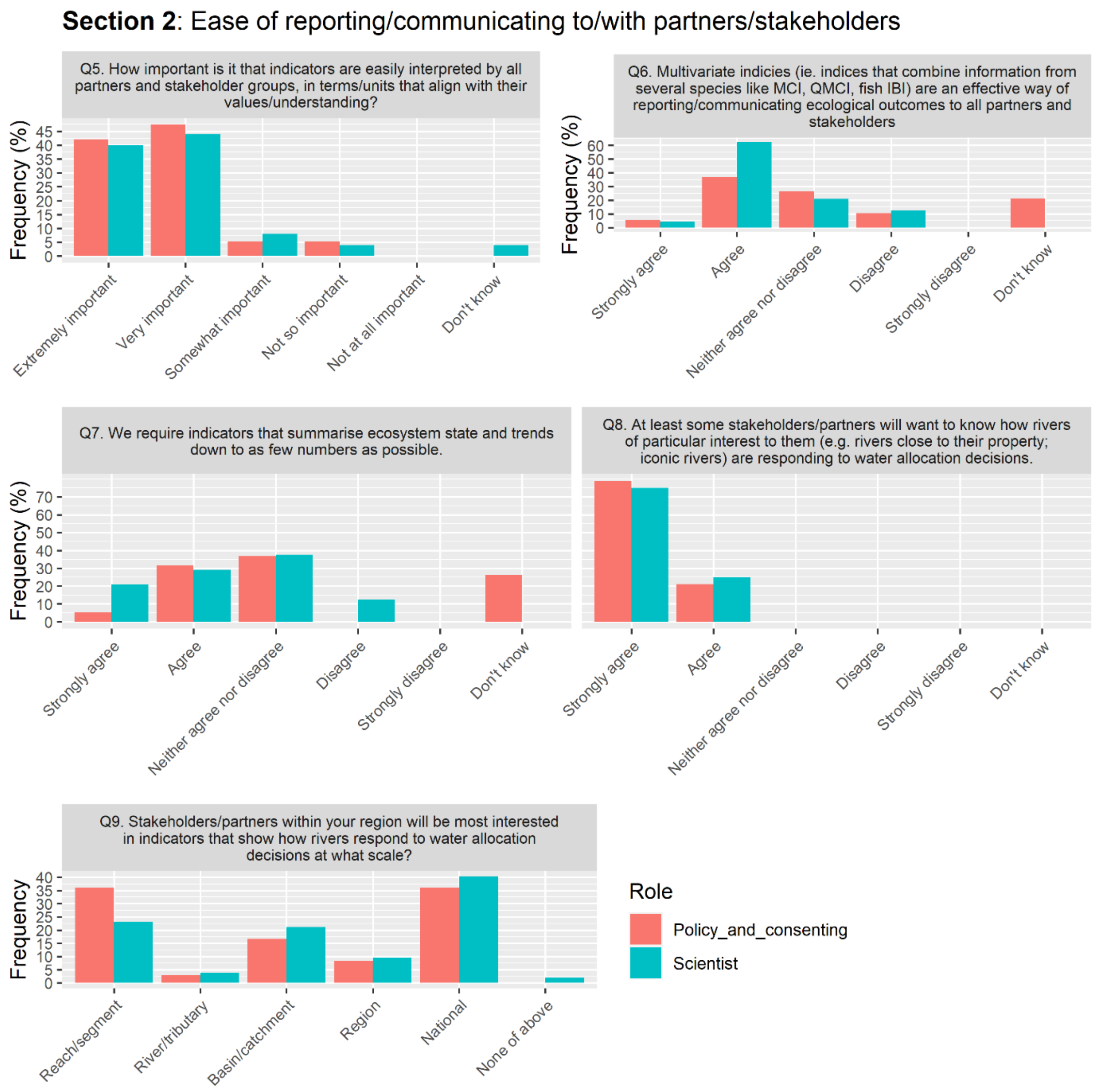

Figure 5-2: Council perspectives on how what we monitor (metrics) and where we monitor may affect reporting/communicating with partners/stakeholders.

Ambivalence concerning the extent to which ecological information should be reduced and/or summarized as abstract indices is also reflected in the written responses of council representatives. Several very experienced (> 10 years) council representatives noted the advantages and disadvantages that come with using multivariate metrics:

"Simple, single number metrics are definitely good reporting tools but there is a balance to be struck between simplicity and meaning...so shouldn't be the goal above all others."

"High level/holistic measures are always well received, but often people want to know the why of these, so being able to drill down through the metrics is important."

"I find some of the more holistic measures are used inappropriately because of the way they are represented... I'm not sure if this is in an effort to simplify a range of metrics to make a situation easier to discuss, but I don't think it does anyone any favours in the long run, and leads to situations

where people use these metrics inappropriately (like saying water abstraction isn't having an impact, because the MCI is high, a pretty typical situation in resource consent applications)."

"*Indicators need to be understood and directly meaningful to stakeholders' conceptual understanding of their resource, and their understanding of pressures on their resource.*"

"*Indicators at present (in my view) are deficient for reporting/meeting objectives for reporting river health generally and more importantly identifying key driving pressures to focus on/address.*"

"*I think it's more important to make sure that attributes are determined in accordance with the NOF process than it is to select them because they're more easily understood or able to be communicated."*

"Disconnects between an indicator and value/understanding may highlight opportunity for education around indicators and how they represent values and/or highlight poor understanding."

"Multivariate indicators are poorly understood by lay stakeholders."

Although difficult to synthesise, on balance responses to Q5-Q7 led to the following M&E objectives:

O9.—Metrics should be easily understood by non-scientific stakeholders. *Metric units should, as much as practicable, be intuitive to non-scientific stakeholders.*

O10.—Keep the number of metrics to a minimum. *To facilitate communication of simple messages to non-non-scientific stakeholders, choose metrics carefully and sparingly.*

Objectives 8 and 9 are consistent with the recommendations of Mace and Bailey (2007).

At least 70% of council representatives *strongly agreed* that some stakeholders will want to know how rivers of particular interest to them are responding to water take decisions (Figure 5-2, Q8). This response reflects awareness that stakeholder interests in natural resource management are often very localized (Waylen and Blackstock 2017). Council response to Q9 presents a great challenge to the design of this M&E plan as it shows that, ideally, the M&E plan should yield information to facilitate decision-making and evaluation at multiple spatial resolutions; from individual river reaches/segments through to the national scale (Figure 5-2, Q9):

O11.—Support decision-making and evaluation at multiple scales and locations*. Collect data with a plan for how it may be used to facilitate decision-making and evaluation at multiple scales and at several locations throughout FMUs.*

Sensitivity of metrics to flow management decisions

At least 65% of council scientists thought that metrics would ideally exhibit significant change (improvement or deterioration) to changes in a rivers' flow regime in five years or less (Figure 5-3, Q10). That is, we require some metrics that are sensitive to changes in the flow regime and do not exhibit long (>5 years) temporal lags in response. The requirement of councils for metrics to exhibit responses within five years aligns with the five-yearly reporting cycle of the NPSFM (Clause 3.30 of NPSFM).

O12.—Include metrics that exhibit rapid responses to changes in the flow regime*. At least some of the metrics monitored should exhibit measurable responses to changes in the flow regime within five years. Metrics that respond to changes in flow within one and three years are also desirable.*

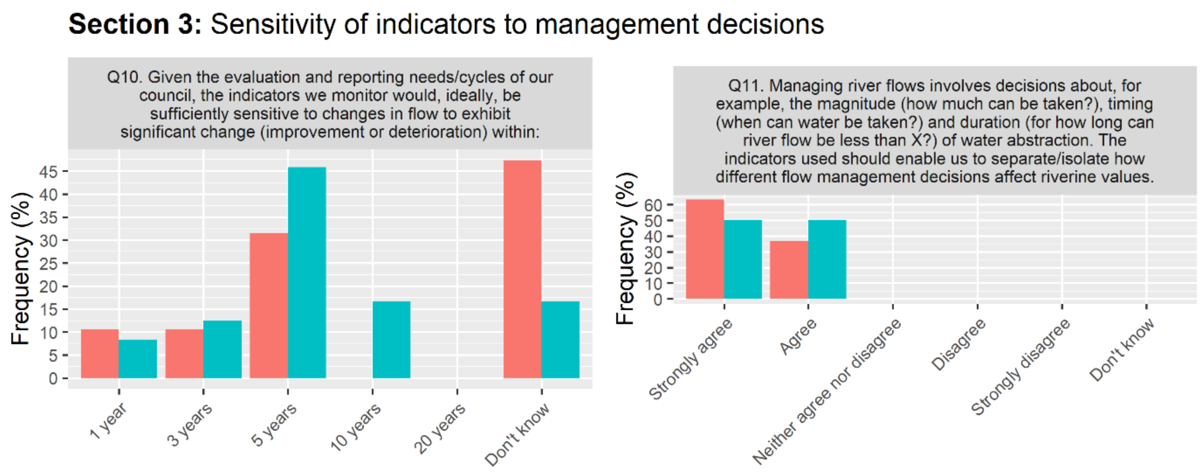

Figure 5-3: Council perspectives on the sensitivity of metrics to flow management decisions.

All council representatives either *agreed* or *strongly agreed* that the metrics selected should facilitate separation of the effects of different hydrological components of the flow regimes (e.g., duration of a specific discharge; frequency of specific discharges) (Figure 5-3, Q11). Responses to Q11 add further weight to O1 and O2—arising from the NPSFM and captured in Section 5.1.3—which point to the need for monitoring and evaluation to separate effects of water takes from background variability not driven by water takes *per se* (O1) and other stressors (O3; e.g., nutrients).

O13.—Select metrics that are sensitive to the dominant hydrological stressors resulting from water takes*. Identify the hydrological components of the flow regime most affected by water takes and select metrics that are most sensitive to those components, based on a conceptual model of the system.*

Tools and processes to support decisions and evaluations

As context to the responses to Section 4, council representatives were asked a multi-choice question to ascertain the level of experience they have had with water resource disputes. The responses indicate that a large majority of representatives have extensive—often first-hand—experience with water resource disputes and their resolution:

- 80% of respondents - *have first-hand experience, e.g., through environment court, local court hearings, community/stakeholder meetings*.
- 15% of respondents - *don't have first-hand experience, but have been in the industry long enough to be well aware of how freshwater resource disputes arise, and how they are resolved*.
- 5% of respondents - *have little first-hand experience or awareness of how freshwater resource disputes arise, or how they are resolved*.

Responses to Q12 and Q13 show that a strong majority of council representatives *agree* or *strongly agree* that flow management decisions are going to be contentious and scrutinized by stakeholders (Figure 5-4). These responses reflect the strong competition between the environment and consumer/economic stakeholders (Table 3-1) for freshwater in New Zealand, and echo similar observations made in other countries (Naiman et al. 2002, Acreman et al. 2014, Ruan et al. 2021, Ryan et al. 2021).

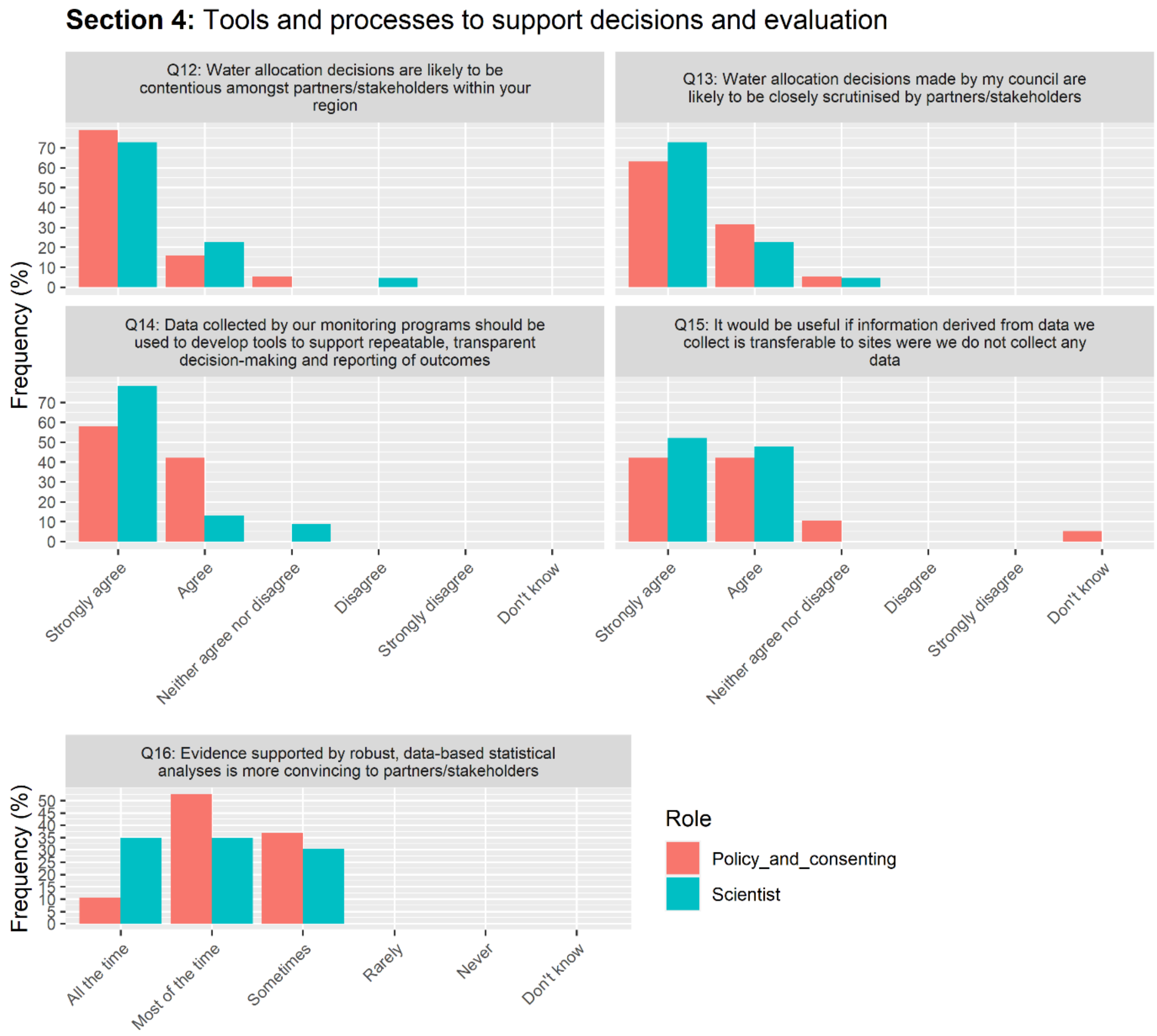

Figure 5-4: Council responses to questions about tools and processes that may support decisions and evaluation.

Responses to Q14 reflect a desire of council representatives to ensure that data collected as part of this M&E plan supports development of tools to support evaluation of flow management decisions (hindcasting; looking backward), as well as tools to support decision-making (forecasting; looking ahead) (Figure 5-4, Q14). Council representatives also indicated strong support for any tools or processes that would facilitate the extrapolation of knowledge gained from monitoring sites to unmonitored sites (Figure 5-4, Q15). Discussion at the workshops following this survey confirmed the following requirement of council representatives:

O14.—Design monitoring to support decision- and evaluation-support tools*. Monitoring should be designed such that the data obtained support development of models that facilitate transparent and defensible:*

- *evaluation of how past flow management decisions affected ecological outcomes;*

- *forecasts of the potential future outcomes of flow management choices;*
- *extrapolation of general ecological responses to flow management decisions in space, towards inferring how flow likely affected/affects responses at places where there is little/no data.*

Support by councils for O14 is further reflected in their experience with water resource disputes, where they have found that evidence supported by quantitative tools/analyses is, at least *most of the time*, more convincing to stakeholders (Figure 5-4, Q16). Objective 14 provides a means of achieving O1, O2 and O3 arising from policy, and also provides some guidance on what forms of evidence "provide the greatest level of certainty" (Clause 1.6 of the NPSFM) and so are preferred by councils.

Resource and logistical constraints

All council scientists and at least 60% of PPC experts *agreed* or *strongly agreed* that costs associated with monitoring were a concern (Figure 5-5, Q17). Given the many and varied demands that the NPSFM places on councils[14], this result is unsurprising yet worth explicitly capturing to help frame subsequent discussion and design of the M&E plan. Responses to Q17, followed by workshop discussion among the science team and core representatives led to the following objective:

O15.—Minimise costs of data collection and processing*. Aim to minimize costs while still meeting other monitoring objectives.*

At least 75% of scientists and 60% of PPC experts *agreed* or *strongly agreed* that new monitoring infrastructural requirements (Figure 5-5, Q18) and site locations (Figure 5-5, Q19) should leverage off existing monitoring infrastructure and sites. There was some ambivalence among councils in their response to these questions, with approximately 20% of scientists *neither agreeing nor disagreeing* or *disagreeing* to Q18 and Q19. With respect to the response of PPC experts to the statements in Q18 and Q19, at least 30% *neither agreed nor disagreed* or had no opinion (*don't know*; Figure 5-5). Written comments in Section 5 of the survey highlight the challenge of leveraging new site and infrastructural requirements off existing sites and infrastructure:

"*My answers to Q18 and 19 are basically neutral, as again, the current monitoring network was developed for specific purposes (e.g., flow monitoring is traditionally focused on flood protection, and far less on low flow monitoring), so in this case, new monitoring requirements designed to monitor low flows will probably need a very different set of waterway types, and should focus on areas of high water demand. Obviously, where possible we need to leverage as much value as we can from existing sites, but I would not like to be in a situation where we were required to only monitor our historic sites just because that is where they are. They may not be appropriate for assessing impacts of water allocation*."

"*We aim to utilise existing monitoring where possible but understand this is not always possible and new indicators and sites are required which will have new direct costs*."

[14] The NPSFM requires monitoring of 23 attributes at a minimum (those attributes specifically listed in Appendices 2A and 2B of the NPSFM). The list of 23 attributes in Appendix 2 of the NPSFM does not include other attributes that councils are required to define in response to all the compulsory and other values that councils must consider (Appendix 1 of the NPSFM), including flow-sensitive attributes.

"Some … water quality FMUs are under-represented by the existing monitoring network, which is being addressed by the monitoring network review process. This is to implement NPS-FM 2020 but involves a huge resource cost."

"If the network is not fit-for-purpose than a redesign is probably necessary."

"Councils can't monitor everything at all places, the cost to do this to an appropriate standard is prohibitive. The NPS allows for modelling/relationships to be established between sites. Work by (scientists) are (sic) increasingly showing the driving effect long term weather patterns can have on water quality trends, therefore long term existing sites are highly valued. For these reasons it is important to build off existing networks wherever possible (acknowledging this isn't always possible)."

"Regardless of existing data, existing/historical sites may not be representative of parts of the riverscape where information on flows is most important from a regional biological effects point of view."

"New monitoring should be targeted and specific to achieve required outcome. while alignment/leveraging existing monitoring sites can provide efficiency, in the first instance I believe it's more appropriate to aim for targeted outcomes before compromising by aligning with an existing location."

The written responses above acknowledge the value of the long-term time series collected at existing sites and the cost-savings that come with using existing infrastructure and sites. The responses also acknowledge, however, that many existing sites may not be located within catchments most affected by water takes and/or may not support freshwater values of greatest concern.

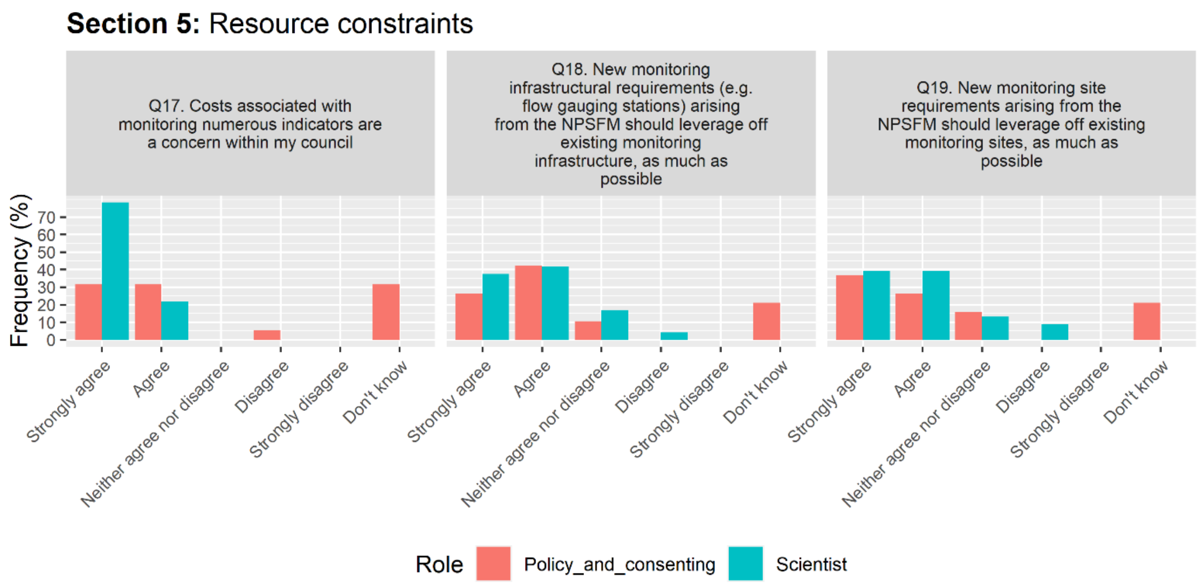

Figure 5-5: Council perspectives on the extent to which resource and logistical constraints affect monitoring and evaluation.

Responses to Q18 and Q19, as well as comments in Section 5 of the survey and subsequent workshop discussion lead to the final monitoring objective:

O16.—Leverage off existing sites and infrastructure where possible*. Use existing sites and infrastructure in cases where doing so does not undermine our ability to meet other monitoring objectives.*

5.3 Challenges and trade-offs among objectives, and potential solutions

5.3.1 Objectives network

Consideration of the requirements of the NPSFM and of councils yielded 16 M&E objectives (collated in Table 5-2). Consistent with the recommendation of Conroy and Peterson (2013), we organized our objectives into an objectives network. Objectives networks help to differentiate means and fundamental objectives.

- *Fundamental objectives* are those objectives that matter most. Success of a program is best measured by the extent to which fundamental objectives have been met.
- *Means objectives* are useful only in as much as they help us achieve fundamental objectives; they are our means of achieving more fundamental objectives.

Fundamental objectives are presented towards the top of our objectives network, while means objectives occur towards the bottom (Figure 5-6). Differentiating fundamental and means objectives is deemed an essential part of structured decision making, as it helps us identify what matters most, and the means of achieving what matters most (Gregory et al. 2012, Conroy and Peterson 2013). Objectives networks can also help us identify hidden and stranded objectives. *Hidden objectives* are objectives that we have failed to explicitly identify, but are implied by certain means objectives. *Stranded objectives* are fundamental objectives that have no explicit links with means objectives, and hence are at risk of not being met.

Table 5-2: Monitoring and evaluation objectives arising from central government policy and council requirements. The eight most fundamental objectives occur in grey shaded rows.

Number	Objective (short)	Objective (long)
1	Determine how current water takes are affecting metric state and trends	Select and monitor sites and metrics to help determine how changes in the flow regime affect metrics within river reaches throughout freshwater management units (FMUs). This environmental monitoring must be undertaken alongside monitoring of water takes and other forms of hydrological alteration as is required for water accounting purposes.
2	Forecast how future water takes are likely to affect metric state and trends	Sites and metrics must be selected with a view for how data will be used to forecast the response of metrics to future environmental changes, including future water management scenarios.
3	Disentangle the influence of multiple drivers	Monitor sufficient environmental variables within river reaches of an FMU such that councils may, as much as practicable, partition the influence of multiple drivers—both anthropogenic and natural—on metrics, and determine the effects of the flow regime on metrics relative to other drivers
4	Monitoring sites must be representative of rivers most affected by water takes and flow management	Locate sites within FMUs such that the sites are, as much as practicable, representative of the rivers within the FMU that: are subject to water takes; and support compulsory values, as specified in Appendices 2A and 2B of the NPSFM 2020

Number	Objective (short)	Objective (long)
5	Implement adaptive monitoring	Once every five years, this M&E plan should be reviewed to determine whether it is fulfilling the objectives stated at the beginning of this plan. Such a review would be aimed at identifying which elements of this plan were adopted, and which elements should be retained, modified, or dropped from the plan. If the M&E plan is not fulfilling M&E objectives it should be amended to do so, in a manner that, as much as practicable, preserves the integrity of previous monitoring data time series
6	Flow M&E must support adaptive management of water takes	Monitoring must aim to reduce uncertainty about flow-metric relationships, and yield information to facilitate transparent and defensible evaluations and forecasts of how metrics have and will respond to flow management decisions
7	Determine how flow management affects the abundance of fishes and other high-value species	Data must be collected with a clear plan for how it will be used to evaluate effects of the flow regime on the abundance of fishes and potentially other species of direct value to stakeholders (e.g., other mahinga kai species like kōura and kākahi)
8	Determine how flow management affects key hydrogeomorphic metrics of a river reach	Key hydrogeomorphic metrics must be identified and monitored at gauged sites. Prioritise hydrogeomorphological variables that most strongly affect values (O7)
9	Metrics should be easily understood by non-scientific stakeholders	Metric units should, as much as practicable, be intuitive to non-scientific stakeholders
10	Keep the number of metrics to a minimum	To facilitate communication of simple messages to non-scientific stakeholders, choose metrics carefully and sparingly
11	Support decision-making and evaluation at multiple scales and locations within regions	Collect data with a plan for how it may be used to facilitate decision-making and evaluation at multiple scales and at several locations throughout FMUs
12	Include metrics that exhibit rapid responses to changes in the flow regime	At least some of the metrics monitored should exhibit measurable responses to changes in the flow regime within five years. Metrics that respond to changes in flow within one and three years are also desirable
13	Select metrics that are sensitive to the dominant hydrological stressors resulting from water takes	Identify the hydrological components of the flow regime most affected by water takes and select metrics that are most sensitive to those components, based on a conceptual model of the system
14	Design monitoring to support decision- and evaluation-support tools	Monitoring should be designed such that the data obtained supported development of models that facilitate transparent and defensible: evaluation of how past flow management decisions affected ecological outcomes; forecasts of the potential future outcomes of flow management choices; extrapolation of general ecological responses to flow management decisions in space, towards inferring how flow likely affected/affects responses at places where there is little/no data
15	Minimise costs of data collection and processing	Aim to minimize costs while still meeting other monitoring objectives

Number	Objective (short)	Objective (long)
16	Leverage off existing sites and infrastructure where possible	Use existing sites and infrastructure in cases where doing so does not undermine our ability to meet other monitoring objectives
Additional objectives originally hidden; added after development of the objectives network:		
17	Flow M&E must support effective communication of outcomes to all stakeholders	The information collected must be relevant to, and easily understood by, as many stakeholder groups as practicable. [*This fundamental objective was added after drafting an objectives network, to explicitly capture in the network why other means objectives were important (e.g., O8; O9)*]
18	Monitor all metrics of management concern at the same sites	Monitor all metrics of primary concern to management and policy at the same long-term sites. This includes metrics deemed to be stressors (e.g., hydrological stressors, fine sediment, nutrients) and response (e.g., macroinvertebrates fish). [*This means objective was added after drafting an objectives network and discussing strategies to achieve fundamental objective O2*].

A network of the 16 initial objectives revealed no stranded objectives (Figure 5-6). Objective 5—implement adaptive monitoring—is a means of improving the efficacy and efficiency with which we meet all other objectives in the long term. It appears stranded in the network (Figure 5-6) only because it was impractical to draw links between it and all other objectives.

The network did highlight one hidden objective: Objective 17—Flow M&E must support effective communication of outcomes to all stakeholders—this is required to clarify the fundamental objective underpinning means objectives pertaining to ease of communicating with non-scientific stakeholders (e.g., O8, O6, O9). Of the eight most fundamental objectives (shaded boxes of Figure 5-6), five were driven by the NPSFM (Figure 5-6). Objective 6 is the most fundamental objective and points to the requirement for M&E to support adaptive management of river flow regimes. The means of achieving O6 are primarily by meeting objectives O1, O2 and O3. This objectives network is complex, but is typical of those characterizing problems in natural resource management (Conroy and Peterson 2013).

In the sections below we describe the challenge of meeting the fundamental objective of adaptive management and identify some major trade-offs among objectives. Describing the challenge of meeting objectives helps to build a shared understanding of the challenges involved with adaptive management. Identifying major trade-offs increases transparency of the M&E design process and helps manage stakeholder expectations around what we can realistically achieve. We outline some strategies used in this plan to overcome challenges and manage trade-offs. By identifying strategies to overcome challenges we also identify any major hidden objectives.

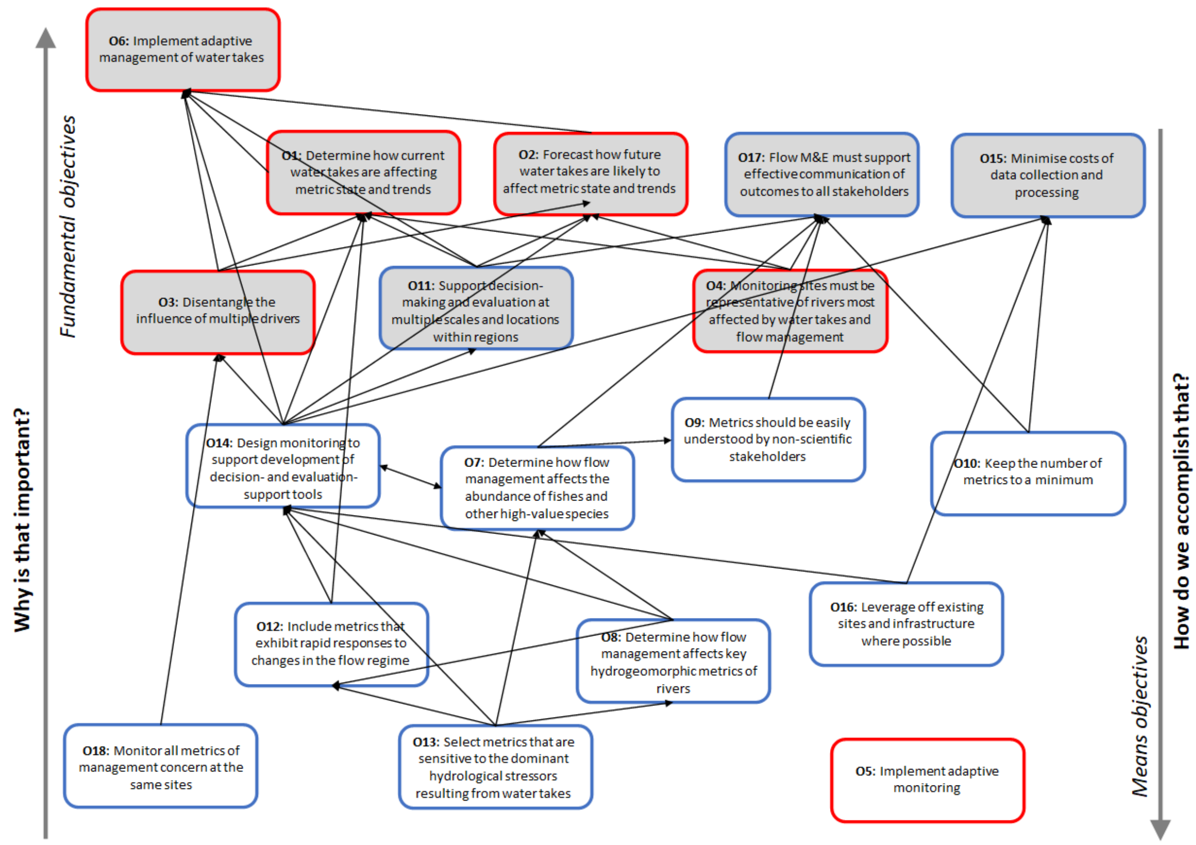

Figure 5-6: Monitoring objectives network arising from the requirements of central government and councils. Objectives arising from NPSFM requirements occur in red boxes, while those arising from additional council requirements occur in blue boxes. See text for further explanation.

5.3.2 Determining how flow management decisions affect riverine values

Fundamental objectives arising from the NPSFM (O1, O2 and O3) require determination of the extent to which past flow management decisions have influenced instream values. Those objectives also require forecasts of the likely impacts of future management decisions—in addition to other factors—on instream values. Fundamental objectives O1, O2 and O3 are the means by which we achieve adaptive management of river flows (O6; Figure 5-6). Attribution of observed ecosystem states to management decisions, and forecasting the response of ecosystems to future decisions, comprise challenges on the frontier of environmental science (Ferraro and Pattanayak 2006, Polasky et al. 2011, Adams et al. 2019). Meeting fundamental objectives O1, O2 and O3 is challenging for several reasons, including:

- Ecological metrics are affected by numerous anthropogenic and natural drivers, which often interact with each other to produce complex non-linear and context-dependent effects (Levin et al. 1997, Blaustein and Kiesecker 2002, Naiman 2013). Scientists often do not have sufficient information to determine the full set of drivers of ecological metrics. Even when the most important drivers of a metric are known, determining the relative effects of drivers is complicated by:
  - our limited understanding of the components of a system and how those components interact (epistemic[15] or structural uncertainty), which in turn limits

[15] *Epistemic uncertainty* refers to our imperfect knowledge of the basic cause-effect pathways that drive system state and dynamics

our ability to parametrize models that accurately partition variance among drivers;

- data that are imprecise, biased or too few, lowering confidence in inferences and forecasts of specific effects (ontological[16] or statistical uncertainty); and
- basic limitations to what mathematical models can do (McElreath 2016).

- River flow interventions—take limits in the case of the NPSFM—are not often administered within a classical experimental design consisting of replicated controls and treatments (Konrad et al. 2011). This makes analysing effects of past decisions and generalizing forecasts of future decisions more difficult (Stewart-Oaten and Bence 2001, Stoffels et al. 2018).
- Lagged biophysical responses to flow events and regimes further complicate attribution of ecological response to flow management decisions (Thompson et al. 2018). Discrete hydrological events within a year, like a summer low flow or a drying event, can have lasting effects on populations that persist for years after normal hydrology resumes (Boulton 2003, Shenton et al. 2012).
- The biophysical state of riverine ecosystems is the result of many years of the cumulative effects of the prevailing and preceding flow regimes. Consequently, when we attempt to restore a more natural flow regime it may take many years of sustained, improved flow management before we see shifts in biophysical state (Graf 2006, Hamilton 2012).

Thus, the inherent complexity of the systems we wish to understand and manage, as well as limitations of the scientific processes and tools available to us, make meeting certain requirements of the NPSFM difficult. Nevertheless, long-term monitoring plays a pivotal role in unravelling such complexity (Fraser et al. 2013, Peters et al. 2014). In the material below we discuss five M&E strategies that include our means objectives (Table 5-2) and which may facilitate overcoming the above challenges.

Monitor for a cause-effect understanding between hydrological stressors and metrics within an adaptive management context. Metrics whose states are driven by hydrological stressors should be selected. We do this in Section 7.3 using conceptual models, following means objective O13 (Table 5-2). Conceptual models of Sections 7.2 and 7.3 will, first, capture the effects of the predominant forms of water takes on hydrology—thus identifying key hydrological stressors—and, second, capture the effects of hydrological stressors on ecological metrics.

To ensure flow M&E is relevant to and understood by stakeholders (fundamental objective O17), we must monitor high-value species like fishes (means objective O7) whose relationships with flow may be highly uncertain. We may reduce that uncertainty through adaptive management (fundamental objective O6). To achieve this, conceptual modelling in Section 7.3 will identify the most likely causal pathways by which flow affects species, moving us towards monitoring approaches that help us better understand the processes by which flow affects values. This may in turn lead to identification of hidden means objectives. For example, we may hypothesise that a key causal pathway linking fishes and low flow is reduced carrying capacity as a result of reduced habitat quantity (e.g., less

[16] *Ontological uncertainty* refers to statistical uncertainties arising from biased or imprecise data, as well as the inherent randomness of the processes we study

area) and quality (e.g., changed hydraulics) for macroinvertebrate production. We may include macroinvertebrate density, coupled with wetted area, as metrics to reduce uncertainty about how low flows affect fish carrying capacity. More generally, if conceptual modelling shows that high-value metrics are indirectly linked to flow via numerous intermediate cause-effect links, then including metrics that help us better understand that chain of cause and effect should improve adaptive management of river flow regimes.

Metrics should be designed to increase sensitivity to hydrological stressors. Means objectives O7 and O9 point to the need to move away from metrics based on species' presence-absence and instead monitor species' abundances. Objectives O7 and O9 arose from the need to ensure information is relevant to stakeholders, but meeting those objectives will also facilitate an improved understanding of how flow management affects values. Species' abundances offer a more sensitive indicator of environmental effects than species' presences. Metrics based on the presence-absence of species only change score when a species is locally extirpated. When monitoring presence-absence, we lose all information about how changes in flow affected species abundances prior to extirpation, by which stage reversal of the drivers of this change may not result in a return of the species.

Monitor metrics with different rates of response to flow, and with clear conceptual links to each other, and communicate the likely timeframes of response. To ensure flow M&E is relevant to stakeholders (O17; O7), we must monitor metrics likely to exhibit slow responses to flow. However, to meet means objective O12 and facilitate 5-yearly reporting required by the NPSFM (fundamental objective O1), metrics that exhibit measurable responses to changes in flow within 1-5 years are also desirable. Evaluation could be strengthened by including metrics with different rates of response to flow and that, at least conceptually, are likely to be causally related to each other. Doing so enables narratives in evaluations that communicate how short-term responses are contributing to long-term objectives. Causal relationships between fast and slow metrics will be identified in the conceptual modelling exercise (Section 7.3).

Timeframes at which we expect metrics to respond to flow management should be explicitly documented, such that we may manage stakeholder expectations concerning how quickly expected environmental outcomes can be met. In Section 8 we include anticipated timeframes of response against metrics.

Monitor all metrics of primary concern to management at the same long-term sites. Fundamental objective O3 reflects the NPSFM requirement to, as much as practicable, disentangle the influence of multiple stressors on metrics, towards identifying how specific management interventions affect outcomes. To achieve O3 we must monitor all metrics of primary management concern at the same sites. This includes both stressor metrics (e.g., hydrology; fine sediment; nutrients) and response metrics (e.g., macroinvertebrates). This may seem obvious and, therefore, not worthy of explicit mention. However, it is common to see river monitoring programs emerge in a piecemeal manner, each of which targets specific stressors or values and exhibits limited spatial overlap with other programs (Bricker and Ruggiero 1998, Stoffels et al. 2018). Given the importance of this strategy for meeting fundamental objectives, yet the frequency with which it is not used, we include it as an additional means objective in our objective network (O18; Figure 5-6; Table 5-2).

Use data from the M&E program to develop quantitative decision- and evaluation-support tools and embrace partnerships for developing those tools. Quantitative models can facilitate (a) partitioning of multiple drivers and attribution of environmental response to flow management decisions; and (b) forecasting the effects of alternative management decisions (Stewart-Oaten and Bence 2001, Ferraro and Hanauer 2014, Webb et al. 2018). Consequently, such models would facilitate meeting fundamental objectives O1, O2, O3 and O5. A major factor limiting development of evaluation- and decision-support models is availability of data (Walters 1997). Long-term monitoring programs can play a pivotal role in providing those data, as long as we give careful thought to what we monitor and how we take measurements (Stoffels et al. 2018). This requires defining the temporal and spatial structure of data required to determine how hydrological stressors affect metric state. Again, this needs to be done in light of clear conceptual models defining hypothesized or observed cause-effect relationships between metrics (Section 7.3). Ideally, the spatial and temporal structure of sampling would be designed with preferred modelling applications in mind. This would lead to specification of the parameters to be estimated, including how samples need to be taken in order to estimate those parameters. We acknowledge, however, that environmental modelling is a rapidly evolving discipline, and that there is no consensus among scientists on which models are optimal for adaptive management of natural resources (Jakeman et al. 2006, Robson et al. 2008, Robson 2014). At a minimum, we should anticipate the functional form of relationships between, say, stressor and response metrics, and aim to structure sampling in space/time to ensure samples are well distributed across the stressor and response gradients, such that we may develop mathematical relationships between stressor and response.

Councils may lack the capacity and/or capability to develop the tools required to meet O14. A solution to this problem is to develop and maintain partnerships among councils and other science providers that could deliver the required tools.

5.3.3 Managing trade-offs among monitoring objectives

Resources for implementing flow M&E are limited, causing major trade-offs between fundamental objective O15—minimise costs of data collection and processing—and all other fundamental objectives. Given our objectives network (Figure 5-6), notable trade-offs include:

- To meet fundamental objectives O11 and O4, M&E must generate information to support evaluation and forecasting of flow outcomes across numerous rivers within regions; rivers relevant to stakeholder interests, and representative of the types requiring adaptive management of flows. This requirement for high spatial coverage of information across rivers may be at odds with the requirement for more intensive monitoring strategies within rivers to support adaptive management (O1, O2, O3).
- The trade-off between the need to minimize costs (fundamental objective O14) and generate the knowledge and tools required to determine how past and future flow management decisions affect riverine values (many other objectives).
- Councils expressed a requirement for information that is relatively simple and easy to communicate with stakeholders (e.g., O9, O10 and O17). However, fundamental objectives concerning defensible evaluations and forecasts (O1, O2, O3) and, more broadly, adaptive management (O6), may require intensive monitoring of numerous metrics, which may not be amenable to simple formats of communication.
- Locations of existing sites may not well represent rivers or catchments most affected by water takes and flow management. The potential exists, therefore, for a trade-off between Objectives 4 and 16, depending on where current monitoring sites are located within regions.

Some strategies to manage these trade-offs are presented below.

Locate sites and select metrics according to a stratified design, where strata are defined by both river type and the need for river flow management. M&E resources required to support adaptive management of flow regimes should be prioritized around types of rivers whose flow regimes are most impacted by water takes[17], hence where there is the greatest need to meet our fundamental objectives. Stratifying by river types as well as need for flow management will ensure that selection and measurement of metrics is more efficiently tailored to rivers that likely support different ecological values, and/or exhibit contrasting ecological and geomorphological responses to flow alteration and adaptive management. Given limited resources, it is possible that not all river types can be covered by this approach. In that case, criteria such as the instream values that specific rivers support, or prevalence of river types, may be used to prioritise site locations.

Implement a *hub and spoke* spatial design to minimize the trade-off between, on one hand, the need for high spatial coverage of evaluations and forecasts and, on the other hand, the need for intensive M&E to facilitate a cause-effect understanding and adaptive management. A hub-spoke design would also minimize the trade-off between the need to reduce M&E costs and the need to generate the data to support adaptive management of flows. A hub-spoke design might subdivide sites into three types:

1. Hub sites. Long-term sites (i.e., maintained for decades) for intensive monitoring that supports development of a cause-effect understanding of how flow interacts with other variables to affect metrics. Hub sites would include continuous monitoring of discharge, as well as monitoring of other stressors with which discharge may interact (e.g., fine sediment, nutrients) and high-value metrics likely to be affected by flow (e.g., abundance of specific species and some geomorphic metrics). Monitoring at hub sites would include other metrics that enable us to reduce uncertainty about the chain of cause and effect between hydrology and those high-value metrics (following the recommendations in Section 5.3.2). Hub sites would also generate the data required for development of evaluation and forecasting tools. To that end, sampling design at hub sites needs to facilitate more than evaluation of the states and trends of individual metrics *per se*, and instead needs to facilitate development of mathematical relationships between metrics. Hub sites should be situated on rivers deemed high-priority for adaptive management of river flows (following the stratified design, above). Hub sites would be expensive to operate and so there would likely be very few hub sites per type, per region.
2. Long-term spoke sites Long-term sites where, at a minimum, continuous monitoring of hydrology occurs. Several long-term spokes would be located within each river type, in each region and/or on rivers of high socio-economic importance without hub sites. The cause-effect understanding and/or tools generated at hub sites could be used to extrapolate the likely environmental outcomes from flow management decisions among long-term spokes. Thus, certain outcomes at long-term spokes would be predicted—using either qualitative or quantitative models—not necessarily observed. Some stakeholders may view predictive evaluation of outcomes as an evaluation method with low legitimacy, but it is almost certainly necessary given resource

[17] This is not necessarily where water takes are greatest. For example, irrigators may abstract large quantities of water from a river with very high mean discharge (e.g., Clutha River), such that the ratio of take:discharge volumes is low and there is little impact on the flow regime.

constraints and the need to evaluate outcomes across many rivers that are representative of those types we wish to manage, and the rivers of concern to stakeholders. If resources permit, predictive inference at long-term spokes should be supplemented by low-intensity surveillance monitoring (less metrics, perhaps measured less frequently) to corroborate (or challenge) predictive evaluations of management decisions, and to support evaluation of trends in metric states across broad spatial extents.

3. *Short-term spoke sites*. These are temporary sites for short-term (e.g., < 5 years) M&E of specific river management interventions. To add value to the program, these short-term M&E projects might include a subset of the metrics monitored at hub and long-term spoke sites, as well as the same monitoring methods for those metrics.

A similar design was recommended by a panel of experts in the USA, to better facilitate adaptive management (Bricker and Ruggiero 1998).

Adopt national consistency and partnerships to overcome resource limitations of individual councils. Each council is unlikely to fund numerous hub sites per river type. A low number of hub sites per region reduces the ability of individual councils to meet fundamental objectives. However, a nationally-consistent approach to implementation of hub sites across numerous councils would increase the potential for effective adaptive management as a result of understanding and tools being leveraged by more data collected across:

- more rivers within each type;
- a greater range of climate and geomorphologic contexts within and across types; and
- a larger and more diverse set of flow experiments within types and contexts.

Other reasons to adopt a nationally-consistent approach have been covered in Section 4. Section 4 also presents numerous reasons why multi-organisation partnerships can improve efficacy and efficiency of flow M&E. We reiterate that a nationally-consistent program will not be able to meet all M&E objectives of individual councils (Section 4), so councils should set aside resources for region-specific programs.

This flow M&E program should be supplemented by strategic research to improve our understanding of how flow affects ecological processes, and to develop decision- and evaluation-support tools. Provision of this supplementary knowledge and tools may be best achieved through partnerships between councils and other research organisations (e.g., Cawthron and NIWA).

There are **no easy ways to minimise the trade-off between Objectives 4 and 16**. If existing long-term monitoring sites (e.g., SoE sites) are not located on river segments most relevant to adaptive management of the flow regime, then they will not be suitable hub sites for implementing this M&E program. If existing long-term monitoring sites are not located on high-priority rivers, but are located on rivers still relevant to flow adaptive management (e.g., where refining limits on water takes is required), then they may make suitable long-term spoke sites. As much as practicable, councils should avoid falling into the sunk-cost and status quo traps (see Section 5) by tacking this M&E program onto existing inappropriate sites, but instead establish new hub sites for flow M&E.

One could argue that intensive M&E aimed at supporting adaptive management is at odds with the need for simple forms of information for communicating outcomes to non-scientific stakeholders (e.g., Objectives 17 and 10). This is not necessarily the case. We see no reason why the information obtained for adaptive management cannot be simplified and effectively communicated to non-scientific stakeholders. By contrast, information obtained through surveillance monitoring programs cannot be easily transformed into forms that support adaptive management.

6 Spatial design

The purpose of this section is to present a framework for selecting and grouping M&E sites in terms of their number, location and purpose.

6.1 Defining and prioritising river types

We suggest a stratified, hub-spoke monitoring network, with strata defined by river type and the need for adaptive management of river flow (*need for flow management*). Rivers with a high *need for flow management* are those rivers:

- whose natural flow regimes have been strongly modified by water takes; and
- requiring management of water takes in order to meet environmental outcomes stated in regional plans, in accordance with the NPSFM.

For the purposes of M&E design it is useful to define the major *river types*, across which responses to water takes and monitoring methods are likely to vary. Rivers have different natural flow regimes, geomorphologies and ecological communities. As such, response to altered flow regimes may vary across rivers, and monitoring methods suitable for one type of river may not be suitable for another.

Approaches to the development of river typologies vary considerably, with the approaches employed reflecting the experiences and values of the scientists developing them (Tadaki et al. 2014). For the purposes of this plan, we could have employed a statistical approach that uses hydrological, geomorphological and ecological data (among additional forms of data) to objectively classify rivers into a typology for the adaptive management of flows (e.g., Kennard et al. 2010, Olden et al. 2012). Such analysis was outside the scope of the current project.

Instead of a statistical approach to defining a river typology, we utilized the collective expert knowledge of council core representatives and the science team to subjectively define types, in order to obtain a very simple stratification for flow M&E that was relevant to council needs. We deemed a more elaborate statistical approach unnecessary as:

- we anticipate that resourcing hub sites will be difficult and that, therefore, the number of river types should be minimized to only those most relevant to flow adaptive management throughout New Zealand; and
- given the need for very few types, it was likely that a more elaborate, statistical approach would have a low benefit:cost ratio, in that it would likely tell core representatives and the science team what they already knew (e.g., that for the intents and purposes of this M&E plan, spring-fed creeks should be treated differently to large braided rivers).

The process for defining types consisted of three steps. **Step 1** took place in a workshop setting and involved council core representatives accessing an online spreadsheet and, based on their expert knowledge, listing key river types within their regions that are subject to significant flow modification, and might be differentiated on the basis of their hydrology, geomorphology and ecology. Alongside each river type councils were asked to also provide a subjective, categorical assessment of:

- the need for flow management (high, medium, low);

- examples of rivers within their type, so that the science team was better able to aggregate council-defined types after the workshop (Step 2);
- any additional comments concerning the significance of the river type to adaptive management of flows (again to assist with Step 2); and
- prevalence within their region (high, medium, low), included to help prioritization of M&E resources (assisting with Step 3).

Step 1 yielded river types that were common to numerous councils whilst accepting that written descriptions and subjective judgements about labels (e.g., "spring") and degrees (e.g., "medium") would vary across councils (Table D-1).

Step 2 involved the science team using the council descriptions of types, alongside their examples, and assigning council-defined types to a typology with a consistent nomenclature across councils/regions (Table D-1). Table 6-1 summarises the results of the council nominations of types. In most cases councils differentiated types based on the source of the flow (ie. whether the river's discharge was primarily sourced from spring/groundwater flows or rainfall runoff), whether or not rivers were braided, and the primary substrate (soft-bottom/fine sediment versus coarse sediment) (Table 6-1; Table D-1).

Source of flow is a common characteristic used to differentiate riverine environments in New Zealand (Snelder and Biggs 2002) and elsewhere (Whiting and Stamm 1995, Whiting and Moog 2001, Lusardi et al. 2021). A river's flow regime is very strongly influenced by its source of flow (Whiting and Stamm 1995). For example, relative to rivers primarily fed by runoff, flow in streams primarily fed by springs and/or groundwater usually has a narrow range and is less variable within and among years (Whiting and Stamm 1995, Griffiths et al. 2008, Snelder and Booker 2013). Compared with runoff-fed rivers, mean daily flow of spring-fed streams is less influenced by rainfall events. The geomorphology of spring-fed rivers is unique, and more static than that of runoff-fed rivers (Lusardi et al. 2021). Further, spring-fed rivers often support different ecological values to those found in runoff-fed rivers (e.g., macrophyte-based faunal assemblages in spring-fed rivers) (Lusardi et al. 2021).

Braided rivers have unique geomorphological dynamics that are maintained by a balance between the natural flow regime, sediment supply, and landscape setting in relation to channel gradient. They may also support unique values (e.g., waterfowl that nest on braided river islands), so the tendency of councils to separate braided rivers as a unique type is logical.

Rivers that naturally have contrasting dominant substrate compositions (e.g., soft- versus hard-bottomed) experience different flow regimes and support different faunal assemblages. Consequently, rivers with contrasting substrates may require different monitoring methods and/or exhibit unique responses to the flow regime. For example, methods for sampling benthic invertebrates differ between hard- and soft-bottomed rivers. Furthermore, the contrasting invertebrate assemblages of hard- and soft-bottomed rivers may have different responses to changes in hydraulics.

Table 6-1: A set of river types for a hub-spoke network, nominated by council representatives, following the first two steps of defining types. Mean need for flow management and mean prevalence is provided. These are means of categorical variables (1 = high; 2 = medium; 3 = low) and so only indicate subjective, relative magnitudes of water take and prevalence within regions.

Type	Regions that nominated type	N (regions)	Need for flow management	Prevalence
Runoff_CoarseSub Runoff-fed rivers primarily comprised of coarse substrates (sand, gravel, cobbles, boulders, etc)	Canterbury; Otago; Wellington; Northland; Bay of Plenty; Hawke's Bay; Nelson; Taranaki; Waikato	9	1.9	1.4
Spring_FineCoarseSub Spring-fed rivers consisting of both fine (generally < 2 mm grain diameter) and coarse substrates (sand, gravel, cobbles, boulders, etc)	Canterbury; Northland; Auckland; Bay of Plenty; Waikato; Taranaki; Gisborne; Otago; Nelson	9	1.8	2
Runoff_FineSub Runoff-fed rivers primarily comprised of fine substrates (generally < 2 mm grain diameter)	Waikato; Hawke's Bay; Gisborne; Northland; Taranaki; Nelson; Wellington; Otago; Canterbury	9	2.2	2.1
Runoff_Braided_CoarseSub Runoff-fed braided rivers, primarily composed of coarse substrates (sand, gravel, cobbles, boulders, etc)	Otago; Canterbury; Hawke's Bay; Bay of Plenty; Gisborne	5	2.2	2.2
Runoff_FineCoarseSub Runoff-fed rivers comprised of either fine or coarse substrates, depending on segment	Auckland; Gisborne; Northland; Waikato	4	1.6	2.2
LakeFed_CoarseSub Rivers with a lake-fed source of flow/hydrology, primarily consisting of coarse substrates (sand, gravel, cobbles, boulders, etc)	Auckland; Canterbury; Otago	3	1.7	3
Spring_FineSub Spring-fed rivers consisting of fine (generally < 2 mm grain diameter) substrates	Wellington; Hawke's Bay	2	1.5	2
Spring_CoarseSub Spring-fed rivers consisting of coarse substrates (sand, gravel, cobbles, boulders, etc)	Hawke's Bay; Wellington	2	2	2.5
Runoff_Urban_CoarseSub Rivers in urbanized catchments consisting of coarse substrates (sand, gravel, cobbles, boulders, etc)	Auckland; Canterbury	2	3	2.5

Type	Regions that nominated type	N (regions)	Need for flow management	Prevalence
Runoff_Urban_FineSub Rivers in urbanised settings consisting of fine substrates (generally < 2 mm grain diameter)	Auckland	1	1	2
LakeFed_FineSub Rivers with a lake-fed source of flow/hydrology, primarily consisting of fine substrates (generally < 2 mm grain diameter)	Northland	1	3	3

Step 3 was to prioritise the types. Given this is a program for nationally-consistent flow M&E we assigned highest priority to types that are common to numerous regions. The types presented in Table 6-1 have been sorted in descending order by the number of regions/councils that nominated them. Table 6-1 also includes mean level of need for flow management and mean prevalence, both of which were calculated across subjective, categorical values submitted by councils (Appendix D). Mean need for flow management and prevalence within regions can be interpreted as follows:

- 1 = high;
- 1-2 = medium-high;
- 2 = medium;
- 2-3 = medium-low; and
- 3 = low.

Three types are conspicuous by their representation across numerous regions (Table 6-1):

- Runoff-fed rivers primarily comprised of coarse substrates (sand, gravel, cobbles, boulders, etc.; *Runoff_CoarseSub*).
- Spring-fed rivers consisting of both fine (generally < 2 mm grain diameter) and coarse substrates (sand, gravel, cobbles, boulders, etc.; *Spring_FineCoarseSub*).
- Runoff-fed rivers primarily comprised of fine substrates (generally < 2 mm grain diameter; *Runoff_FineSub*).

Rivers that councils nominated as *Runoff_FineCoarseSub* are essentially rivers that have the characteristics of two other types: *Runoff_CoarseSub* and *Runoff_FineSub*. We therefore combined these three types into runoff-fed rivers comprised of fine and/or coarse benthic sediments. Combining these three types had a beneficial impact on conceptualization in Section 7.3, as there was no indication in the literature to suggest that the effects of hydrological stressors (Section 7.2) were qualitatively different among these three types. For the purposes of this plan, it was sufficient to know that metrics and methods for rivers belonging to the types *Runoff_FineCoarseSub* must be suitable for fine-/soft- and coarse-/hard-bottomed rivers.

We therefore defined **two focal river types for this M&E plan**:

1. ***Runoff-fed rivers*: ** Rivers whose flow is primarily fed by runoff, and whose benthic sediment may be dominated by either fine (silt, sand) or coarse substrates (sand, gravel, cobbles, boulders, etc.). In the context of this plan *runoff-fed rivers* are small- to medium-sized (mostly wadable) rivers. Rivers within this type may be braided or non-braided.

2. ***Spring-fed rivers:*** Spring-fed rivers consisting of both fine (generally < 2 mm grain diameter) and coarse substrates (sand, gravel, cobbles, boulders, etc.). In the context of this plan spring-fed rivers are small- to medium-sized (mostly wadable) rivers.

These two river types had mean management needs between medium-high and medium-low, and mean prevalence within regions between medium-high and medium-low (Table 6-1).

We note that, given our definitions of runoff- and spring-fed rivers above, **this M&E plan is limited in scope to small- and medium-sized (mostly wadable) rivers**.

6.2 Locating sites

Statisticians have in the past recommended randomizing site locations throughout entire river networks, towards ensuring the data we obtain is representative of catchments. Such approaches, however, are often impractical as they do not take into account constraints of any kind (resource constraints, socio-political constraints on site locations, logistical constraints [e.g., ease of access]), and often are recommended to fulfil the objective of surveillance monitoring, not adaptive management of natural resources. Instead of recommending specific locations or statistical properties of site distributions, we recommend the following practical criteria for locating hub sites:

- Sites should be generally representative of the rivers and the reaches of the type (fundamental objective O4). Note that sites should be located on representative reaches as well as representative rivers to ensure that evaluations and forecasts based on hub sites are generally applicable to as many other reaches and rivers within the type as possible.

- Sites should support NPSFM and stakeholder values to ensure M&E is relevant to both policy and stakeholders (fundamental objective O17).

- Locate sites on rivers and reaches subject to significant flow modification, to ensure flow M&E is targeted to rivers that most require adaptive management (fundamental objective O4). By *significant flow modification* we mean rivers/reaches where the ratio of total volume of permitted water take to mean annual discharge is relatively high.

- Sites need to be within reaches suitable for long-term automated flow gauging, or have an alternative method for constructing a river flow time series such as correlated upstream gauging station, while still satisfying the other criteria presented here.

- Hub sites are pivotal to meeting our most fundamental objective (adaptive management; O6) and so require long-term commitment. As such, hub sites need to be at locations conducive to long-term maintenance of monitoring access and infrastructure.

- Hub sites should have good access. Numerous metrics may be monitored at hub sites, which may require the installation of infrastructure (e.g., for remote automated monitoring). Good access may also facilitate use of the site by non-council research agencies, which adds value to the data collected as part of this flow M&E program.
- Consistent with Objective 16 (Table 5-2), if existing long-term monitoring sites meet the above criteria, then they should be used. This will preserve and enhance the value of existing data time series.
- Fundamental Objective 3 requires us to disentangle the influence of multiple stressors. To meet this objective, hub sites would ideally be located in rivers/reaches most relevant to flow management, but also in rivers/reaches relevant to management of other key stressors/metrics (e.g., nutrients, fine sediment) of the FMU within which types are located. Selection of sites suitable for meeting multiple requirements of the NPSFM will increase the efficiency of all freshwater M&E under the NPSFM.

The above criteria also apply when locating spoke sites. Spoke sites will be used for spatial extrapolation of evaluations and forecasts generated (primarily) by hub sites, so councils should locate spokes within rivers and reaches where extrapolations are most required (e.g., rivers of high stakeholder interest, but not containing hub sites).

6.3 Analysis issues. Do we require reference sites?

We define reference sites as sites whose flow regime and physical characteristics are not influenced by local human activities. Reference sites are not necessary for a functioning M&E program, but can strengthen evaluation of the effects of management interventions—flow interventions in our case (Stewart-Oaten and Bence 2001). Ecologists have argued that several reference or control sites are necessary for defensible evaluation of interventions (Underwood 1992), but expert statisticians have demonstrated that such assertions are false (Stewart-Oaten and Bence 2001, Stewart-Oaten 2008).

Locating reference sites specifically for evaluation of flow effects is challenging. Fundamental objective O3 requires us to disentangle the effects of multiple drivers that will almost certainly be interacting with flows to produce complex effects at all hub and spoke sites we may choose. When selecting reference sites the question therefore arises: which stressors do we hope to control for? The answer may be “all of them”, in which case councils may consider choosing reference sites for M&E under the entire NPSFM, not specifically in the context of this plan.

We reiterate, however, that reference sites are not necessary for defensible evaluation of flow management decisions under this plan. If data are collected according to this plan, then it will be suitable for intervention analysis. *Intervention analysis* was first described by Box and Tiao (Box and Tiao 1965, Box and Tiao 1975), and has since been applied in ecological contexts (Stewart-Oaten et al. 1986, Stewart-Oaten and Bence 2001). Intervention analysis requires time series of metrics and their drivers before and after a particular management decision. It involves developing a model of metric temporal dynamics as a function of drivers at a site, including stressors, and using that model to simulate how metrics would have responded, had the intervention not occurred. The simulated outcome in the absence of the intervention is often referred to as the counterfactual scenario (Gelman and Hill 2007, Stewart-Oaten 2008). Intervention analysis has been used for evaluation of the outcomes of environmental flow decisions (Stoffels et al. 2018). Intervention analysis offers many strengths for defensible evaluation, but its success hinges on the careful collection and maintenance of quality time series and modelling techniques that are more sophisticated than the classical ANOVA approaches employed in spatially-replicated control-impact designs (Stewart-Oaten and Bence 2001). In the context of this M&E plan, national consistency in the collection of time series will facilitate more rapid development of more robust models for intervention analysis.

6.4 How many hub sites per river type?

Answering this question requires knowledge of the availability of resources within councils and how councils wish to distribute those resources across all aspects of NPSFM implementation (not just implementation of this M&E plan). Councils must, therefore, answer this question themselves. We nevertheless offer some perspectives below.

Councils should aim for between one and three hub sites per type, per region. Three hub sites per type enables within-type replication of hub M&E within each region. This would in turn facilitate a better understanding of how spatial variability and context within types and regions affects the outcomes of flow adaptive management. Intensive monitoring of several metrics at hub sites will likely be expensive, so we appreciate that establishing and maintaining three hub sites for both runoff- and spring-fed rivers (six hub sites) per region may not be achievable. Some consolation should be taken from the fact that this M&E program will be nationally-consistent, allowing individual councils to better meet our fundamental and means objectives (Table 5-2) through sharing of data and tools across regions.

7 Metrics – scoping and conceptual foundation

7.1 Focal NPSFM values

The NPSFM covers numerous socio-ecological *values*, which are decomposed into measurable *attributes*. The NPSFM contains four *compulsory* freshwater values (Table 7-1; Appendix 1A of NPSFM 2020). Compulsory values—and the requirements for their management under the NOF—apply to *every FMU*. Compulsory values are differentiated from *other values* (Appendix 1B of NPSFM 2020) in that local authorities must consider whether other values apply to the resource limits or management plans designed to achieve outcomes within some—not necessarily all—FMUs (Clause 3.9.2 NPSFM 2020).

Table 7-1: NPSFM compulsory values. The ecosystem health value is composed of five components. "Limit or plan?" indicates those attributes requiring a limit on resource use (e.g., water abstraction) or an action plan in order to achieve attribute targets (Clause 3.12 NPSFM 2020).

Value	Contributing biophysical component	NPSFM descriptors	Attributes (rivers)	Limit or plan?
Ecosystem health	Water quality	Phys-chem; e.g.,: temperature; DO; pH; fine sediment; nutrient; toxicants	DO (mg/L)	Plan
			DRP (mg/L)	Plan
			Ammonia (toxicity)	Limit
			Nitrate (toxicity)	Limit
			Suspended fine sediment (m)	Limit
	Water quantity	"Extent and variability in water level or flow"		
	Habitat	"The physical form, structure and extent of the water body, its bed, banks and margins; its riparian vegetation; and its connections to the floodplain and groundwater.	% deposited fine sediment cover	Plan
	Aquatic life	"The abundance and diversity of biota and including microbes, invertebrates, plants, fish and birds."	Fish IBI	Plan
			Macroinvertebrate MCI and QMCI	Plan
			Macroinvertebrate ASPM	Plan
			Periphyton concentration	Limit
	Ecological Processes	"The interactions among biota and their physical and chemical environment such as primary production, decomposition, nutrient cycling and trophic connectivity."	Metabolism (g O_2 / sq m / day)	Plan
Human Contact		Extent to which an FMU support recreation on the water. Issues: pathogens; water clarity; deposited sediment; plant growth; cyanobacteria; other toxicants	*E. coli* (primary contact sites during bathing season)	Plan
			E. coli (general)	Limit

Threatened Species	Extent to which an FMU supports threatened species. All life stages and associated stage-specific habitat dependencies need to be protected/restored.
Mahinga Kai	Kai must be safe to eat and plentiful. The site must also be in good condition (its mauri must be intact).

This M&E plan does not present guidance on the monitoring and evaluation of all NPSFM attributes. Not all NOF attributes are appropriate metrics of how NPSFM values respond to river flows and some attributes have not been defined for all values/components most likely to respond to flow (e.g., water quantity; Table 7-1). Furthermore, Section 4 highlighted trade-offs associated with implementing a nationally-consistent approach. In light of those trade-offs, one could argue that a wise approach to M&E may be to "hedge one's bets" and limit nationally-consistent M&E to a subset of the NOF values and metrics, allowing region-specific approaches to others.

It follows that, in order to develop this plan, we had to:

- identify which NPSFM values this plan was to focus on;
- define flow-sensitive metrics to serve as effective indicators of how flow regimes affect focal values.

These tasks accord with Clauses 3.9 and 3.10 of the NPSFM 2020.

To determine the *focal values* of this plan, a workshop was held involving the science team and the core representatives of councils. The consensus during that workshop was to focus the plan on the compulsory value of *ecosystem health*, specifically components (Table 7-1):

- water quantity;
- habitat;
- aquatic life;
- ecological processes.

These components are known to have metrics that are particularly sensitive to flow. Healthy ecosystems are necessary to support other compulsory values listed within Appendix 1A of the NPSFM, such as *mahinga kai* and *threated species*.

In accordance with objectives O8 and O17 flow M&E must also consider how flow regimes affect hydrogeomorphic metrics that shape recreational, aesthetic and cultural values (see Sections 5.2.2 and 5.3.1), some of which are represented in the NPSFM other values *natural form and character*, *transport and tauranga waka*, and *fishing* (Appendix 1B of NPSFM 2020). These are also focal values of this M&E plan.

7.2 Hydrological stressors to be managed adaptively

Here we define the key hydrological stressors resulting from water takes. By doing so we ensure that the metrics we select, and the methods used to measure them, will be sensitive to hydrological stressors, towards improving our cause-effect understanding of how flow management decisions affect values (following the strategy for meeting fundamental objectives outlined in Section 5.3.2). The science team and council core representatives achieved a shared understanding of hydrological stressors through discussion at a workshop. **Two forms of flow modification were of national concern**:

1. Water takes during dry periods, resulting in **low flows**.
2. **Flow harvesting**; water takes during periods of medium to high flows, to be stored for irrigation during dry periods.

Conceptual links between these two major forms of flow modification and hydrological stressors were then visualised in a workshop and summarised in graphical form, as presented in Figure 7-1 and Figure 7-2. We list key hydrological stressors below. Potential effects of hydrological stressors on riverine ecosystem health are then discussed in Section 7.3.

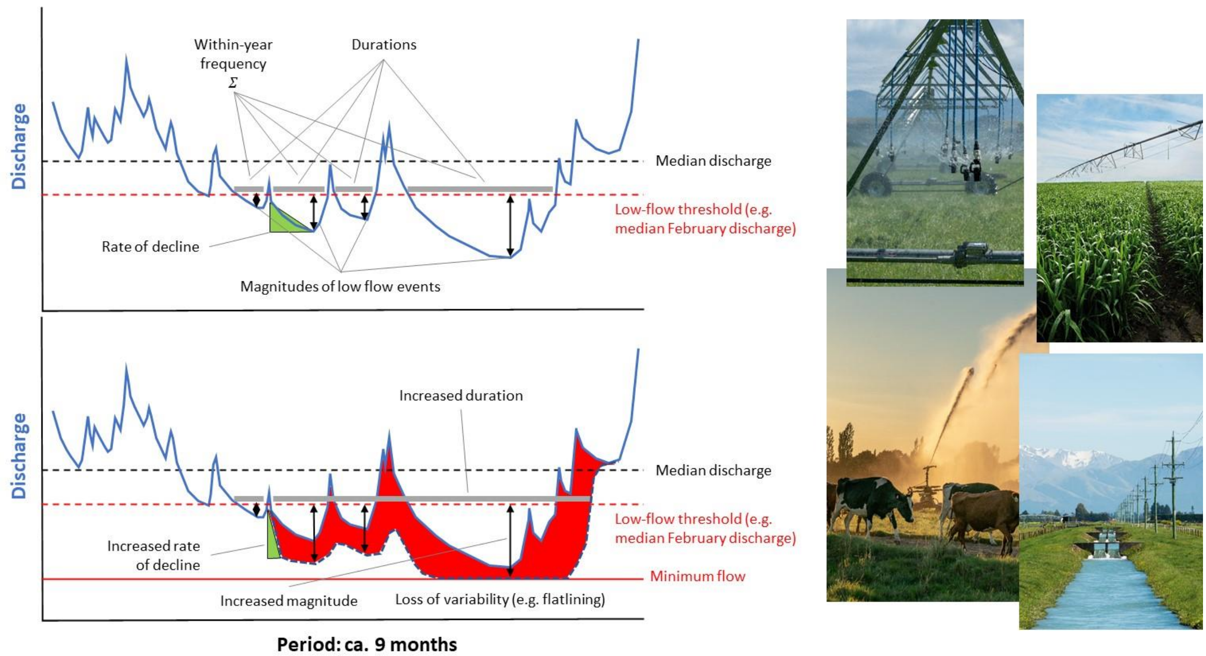

Figure 7-1: Conceptualisation of the impact of water takes during dry periods on a hydrograph and resultant hydrological stressors. Solid and dashed lines indicate hydrographs without and with water takes, respectively. Red filled region indicates volume of water taken away from natural hydrograph.

Hydrological stressors resulting from **low flows** are well studied (Smakhtin 2001) and include the following:

1. *Duration of a low flow event*. The number of days during which discharge is below the *low flow threshold*[18] (Figure 7-1).

[18] It is not necessary to define a low flow threshold for this M&E plan to be effective. Although the statistical properties of low flow duration—and other hydrological stressors—may be implicit within this plan, their identification through simple conceptual models allows us to determine how hydrology needs to be monitored at core and spoke sites.

2. The *within-year frequency of low flow events*. The frequency with which discharge drops below the low flow threshold (Figure 7-1).

3. The *magnitude of a low flow event*. The difference between the low flow threshold and the minimum discharge observed during a low flow event (Figure 7-1).

4. The *rate of decline in discharge* during a low flow event. The per-day rate at which discharge declines during a low flow event (Figure 7-1).

Water takes during dry periods exacerbate all four of the above hydrological stressors (Figure 7-1).

Hydrological stressors resulting from flow harvesting are not as well studied as those resulting from water takes during low flows. Accordingly, our conceptualisation of the links between flow harvesting and hydrological stressors (Figure 7-2) was not as detailed as that pertaining to low flows (Figure 7-1). Very broadly, one hydrological stressor resulting from **flow harvesting** was noted during the workshop attended by the science team and the core representatives:

5. *Flood frequency*. The frequency of flows above various high-flow/flood thresholds within a year (Figure 7-2). The family of FRE-X statistics (e.g., FRE3; frequency of flows within a year exceeding three times the long-term median) is an example of flood frequency statistics (Booker 2013). High-flow harvesting will reduce the frequency of high flows in the flow regime.

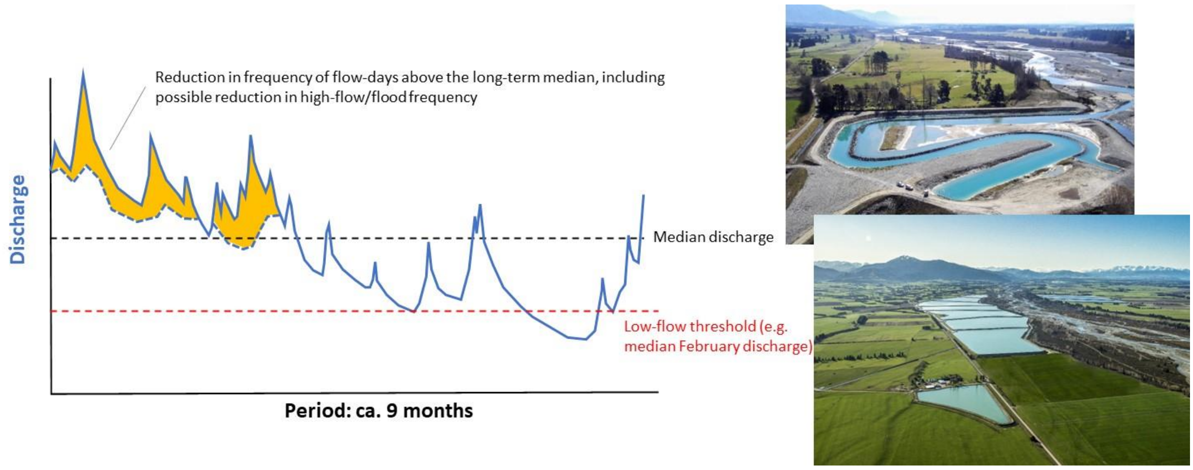

Figure 7-2: Conceptualisation of the impact of flow harvesting on a hydrograph and resultant hydrological stressors. Solid and dashed lines indicate hydrographs without and with water takes, respectively. Filled region indicates volume of water taken.

7.3 Conceptual links between flow modification and potential metrics

Our approach to developing conceptual models of links between the two major forms of flow modification—low flows and flow harvesting—and potential metrics was as follows:

- To ensure flow M&E would be relevant to the NPSFM value ecosystem health, conceptual models were structured by the components of ecosystem health (Figure 7-3).

- Working with the basic structure of Figure 7-3, we started by including metrics that would best support effective communication of outcomes and decisions to stakeholders (Fundamental Objective 17), then worked our way back from those metrics to identify additional key metrics that would support an understanding of how flow management decisions affect outcomes (Fundamental Objective 6). To this end, conceptual models began with fish abundance and population composition (Means Objectives 7 and 9) under the *aquatic life and ecological processes* component, and hydrogeomorphic properties supporting key values (Means Objectives 8 and 9) under the *habitat* components. This approach ensured our conceptualisations remained grounded by our objectives (Section 5.3.1) and helped to ensure that only the most relevant details were considered.
- The conceptual models are mostly self-explanatory and only briefly explained in point-form with key supporting literature.
- We develop conceptual models outlining the effects of low flows on both runoff- and spring-fed rivers. A conceptual model outlining the effects of flow harvesting is only developed for runoff-fed rivers, as the hydrological stressors associated with flow harvesting are not relevant to spring-fed rivers.
- After a brief workshop discussion on the key hydrological stressors within the context of Figure 7-1 and Figure 7-2, a 'brainstorming' session was held, during which council core representatives and the science team suggested flow-affected metrics. The aim of this brainstorming session was to increase the relevance and legitimacy of the M&E plan to council requirements. The brainstorming session ensured the science team did not miss any potential metrics of concern to councils. Further details and results of this exercise are presented in Appendix E.

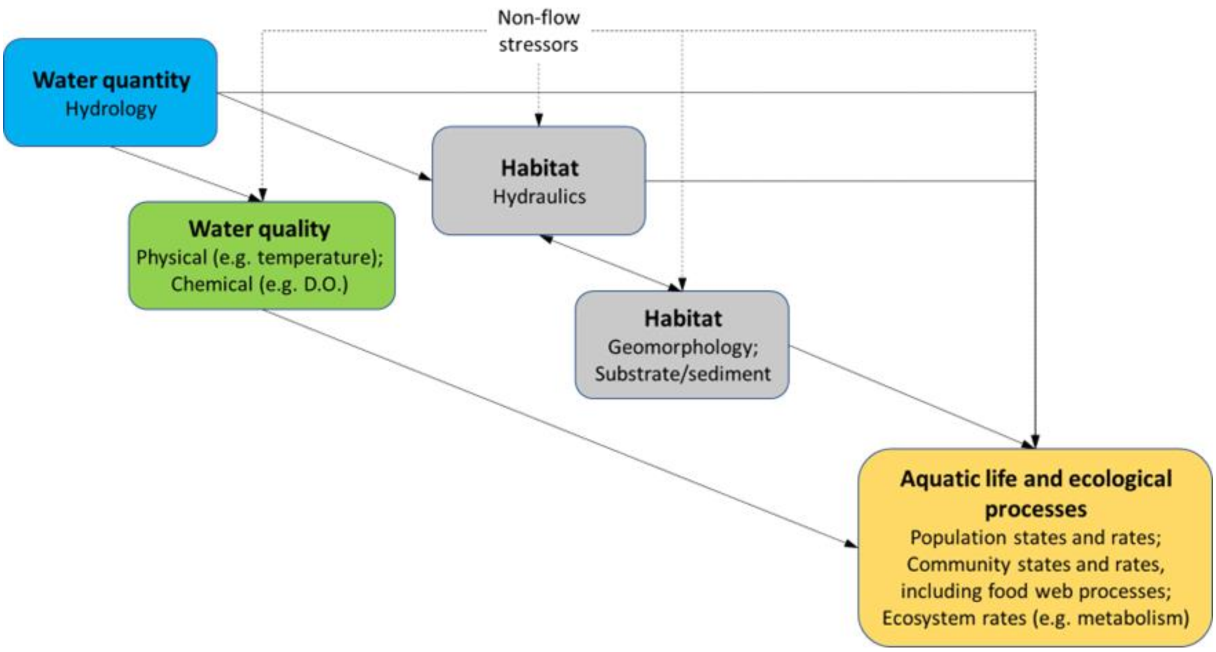

Figure 7-3: Flow-affected interactions between the NPSFM components of ecosystem health. The ecosystem health components are: water quantity, water quality, habitat, aquatic life and ecological processes.

7.3.1 Low flows

Runoff-fed rivers

A conceptualisation of the effects of low flows on metrics within runoff-fed rivers is presented in Figure 7-4. Based on our understanding of the literature, there was no reason to prepare different conceptual models for hard- and soft-bottomed runoff-fed rivers as the basic, qualitative ecological relationships are the same (Rolls et al. 2012). Figure 7-4 is structured by the NPSFM components of ecosystem health (water quantity, water quality, habitat, aquatic life and ecological processes). The conceptual model was based on research summarised in Smakhtin (2001), Rolls et al. (2012), Dewson et al. (2007b) and King et al. (2015). Additional supporting literature is occasionally provided in the explanation of the conceptual model, below.

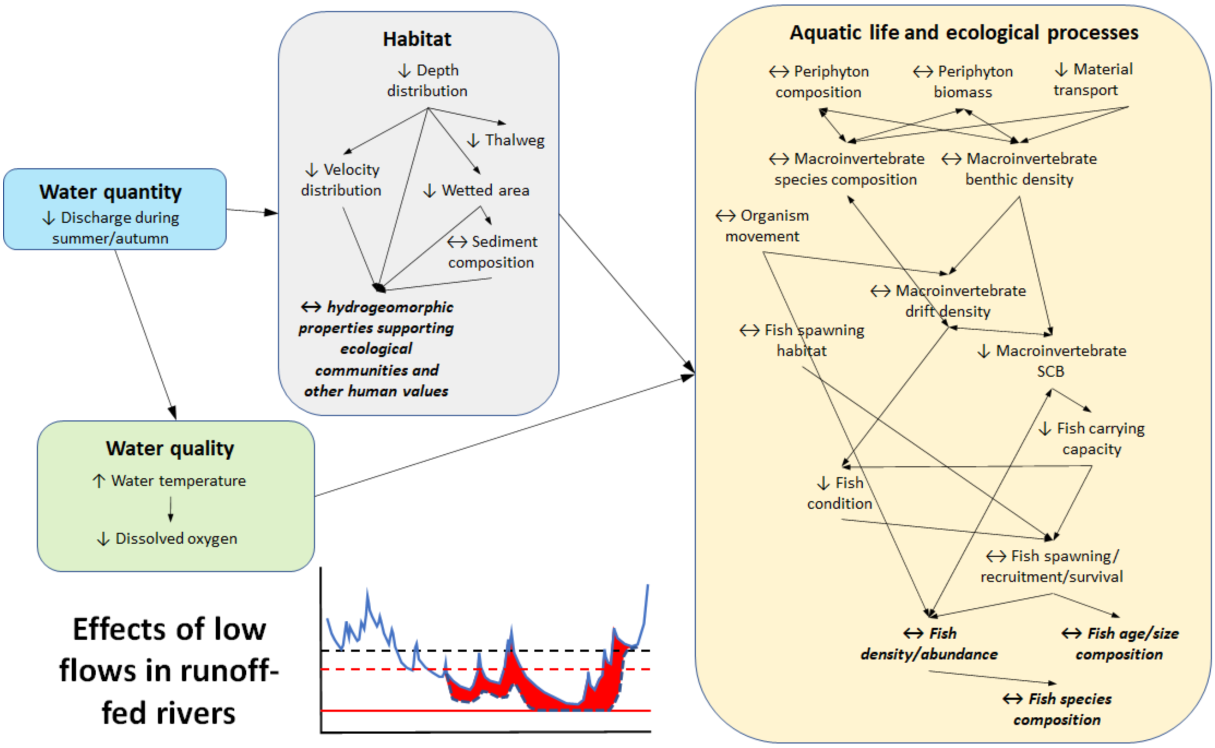

Figure 7-4: Conceptual model of the effects of low river flows on metrics within runoff-fed rivers grouped by water quality, physical habitat, and aquatic life and ecological processes. Up- and down-arrows denote increasing and decreasing respectively. Bidirectional lateral arrows denote change; either increase or decrease, depending on the specific low-flow hydrological stressor considered, its timing, duration and magnitude. Arrows between metrics in different components (boxes) were not included; including them resulted in a very high density of arrows throughout the diagram that did more to obfuscate than clarify.

Low flows may have the following effects on **physical habitat,** resulting in changes in the hydrogeomorphic properties of a river reach that directly (e.g., swimming; boating) and indirectly (support of aquatic life and ecological processes) support focal values (Figure 7-4):

- Reduced discharge will reduce mean depth of the water column throughout a river reach, and may change other statistical properties of the depth distribution of a reach (maximum depth, range, variance, etc.).

- Reduced mean depth will change the thalweg[19] of a river reach.
- Reduced depth will generally reduce reach wetted area. As flow decreases, the magnitude of reduction in wetted area will depend on reach morphometry; the shallower the gradient of the stream bottom, the greater the reduction in wetted area per unit change in depth.
- Mean water velocity of a river reach generally decreases as discharge is reduced.
- As mean water depth and wetted area decline, the sediment size composition of the benthos may also change (Hakala and Hartman 2004).

Low flows may have the following effects on **water quality** (Figure 7-4):

- Reduced river discharge can increase water temperature (Caissie 2006, Booker and Whitehead 2022).
- Higher water temperatures and reduced mixing of the water column may decrease dissolved oxygen in specific rivers where/when ecosystem respiration is high (e.g., in rivers with lots of organic matter, hence high rates of microbial decomposition) (Diaz and Breitburg 2009).

Flow-mediated effects on physical habitat and water quality will interact to affect **aquatic life and ecological processes** supporting fish populations (Figure 7-4):

- Broad periphyton types (such as thin films or long filaments) tend to be associated with specific depths and velocities (Biggs and Hickey 1994, Biggs and Stockseth 1996), so changed hydraulics during low flows may also change the periphyton composition of a river reach, as well as the biomass (mass per unit area) of periphyton as measured by Chl-*a* concentration (Suren et al. 2003b). Increased water temperature during low flows will also interact with changed hydraulics to affect periphyton composition and biomass (Miller et al. 2007).
- Reduced velocity and discharge will reduce rates of organic matter transport downstream, increasing retention of organic matter (Boulton and Lake 1992, Dewson et al. 2007a).
- Changes in periphyton composition and biomass, organic matter retention and water quality will affect macroinvertebrate species composition and biomass (per unit area) (Suren et al. 2003a, Haxton and Findlay 2008, Brooks et al. 2011).
- Reductions in wetted area may change the composition of benthic sediment/ substrata. Macroinvertebrates and types of periphyton have specific substrate preferences, hence a change in substrate composition is likely to change the composition of the benthic community (Quinn and Hickey 1990, Biggs et al. 1999, Shearer et al. 2015, Hoyle et al. 2017). A change in the size composition of benthic sediment may also affect the availability of spawning and refuge habitat of fishes, in turn affecting population survival rates and, ultimately, population size (Magoulick and Kobza 2003, Davey et al. 2006).

[19] The thalweg of a river reach is the longitudinal profile of maximum depth along a river reach.

- Changes in the macroinvertebrate community, as well as reduced velocity and discharge will also change the composition and density of drifting macroinvertebrates (Sotiropoulos et al. 2006).
- Although we know that low flows affect periphyton and macroinvertebrate species composition and biomass, the direction and magnitude of effects depends on the spatial context of the river reach (land-use, riparian habitat, etc), and the states of the four low flow hydrological stressors identified in Section 7.2. The states of the four low flow stressors will vary in time, throughout the summer-autumn period. It follows that the directions and magnitudes of effects on periphyton and macroinvertebrates will be dynamic during low flow events (Rolls et al. 2012).
- Despite the dynamic effects of low flows on macroinvertebrate density and species composition at relatively small scales (e.g., within particular channel units (riffles, runs, etc.) and microhabitats (e.g., patches within riffles) (Fausch et al. 2002)), at larger spatial scales we can expect a reduction reach-wide, total standing crop biomass of macroinvertebrates, as a consequence of the reduction in wetted area of the reach (Walters and Post 2011).
- The effects outlined above combine to reduce fish carrying capacity at the reach scale (Hakala and Hartman 2004), with the greatest reductions in carrying capacity occurring for large-bodied fishes at higher trophic levels (McCann et al. 2005). Reduced fish carrying capacity will lower condition of individuals in fish populations and in turn lead to reduced survival and recruitment (Cowx et al. 1984), with the end result being reduced fish abundance and changed fish population/community structure (Figure 7-4).
- Increased water temperatures during low flows can affect fish populations via direct and indirect mechanisms, with the direction of the effects (positive/negative) dependent on the magnitude of heating relative to the species' thermal tolerances.

Spring-fed rivers

Scientists know less about the ecology of spring-fed rivers than that of runoff-fed rivers (Lusardi et al. 2021). Scientists' understanding of how water abstraction affects the ecology of spring-fed rivers is particularly poor. The hydrology, geomorphology and ecology of spring-fed rivers strongly contrasts with that of runoff-fed rivers. Relative to runoff-fed rivers, spring-fed rivers are characterised by (Whiting and Stamm 1995, Whiting and Moog 2001, Griffiths et al. 2008, Lusardi et al. 2021):

- Low temporal variation in discharge within and among years—less variable flow regimes.
- Static channel morphology.
- Benthic substrates characterised by fine sediments (silts, sands, gravels).
- Rectangular cross-sections and large width:depth ratios.
- Less pronounced pool-riffle sequences with less longitudinal heterogeneity of channel units.

- Low bedform complexity, with habitat complexity being dominated by the contribution of macrophytes (biotic habitat complexity rather than abiotic habitat complexity).
- Less variable and cooler thermal regimes.
- Ecology driven by macrophytes and the organisms they support.
- Highly complex interactions between flow and macrophytes. For example, reduced discharge may increase macrophyte growth, such that a reduction in depth and wetted cross-sectional area that may have resulted from reduced discharge may be compensated for by macrophytes filling more of the river's cross-section. This may result in the somewhat surprising outcome of no change in depth and cross-sectional area with reduced discharge, depending on the time of year (Champion and Tanner 2000, Willis et al. 2017).
- Higher levels of macroinvertebrate production and macroinvertebrate communities comprised of different species with different life histories—notably, life-histories that are less dependent on flow seasonality.

Noting the above, we present a simple conceptual model (Figure 7-5) focusing on the relationships between discharge, macrophytes, water quality, macroinvertebrates and the fish populations they support.

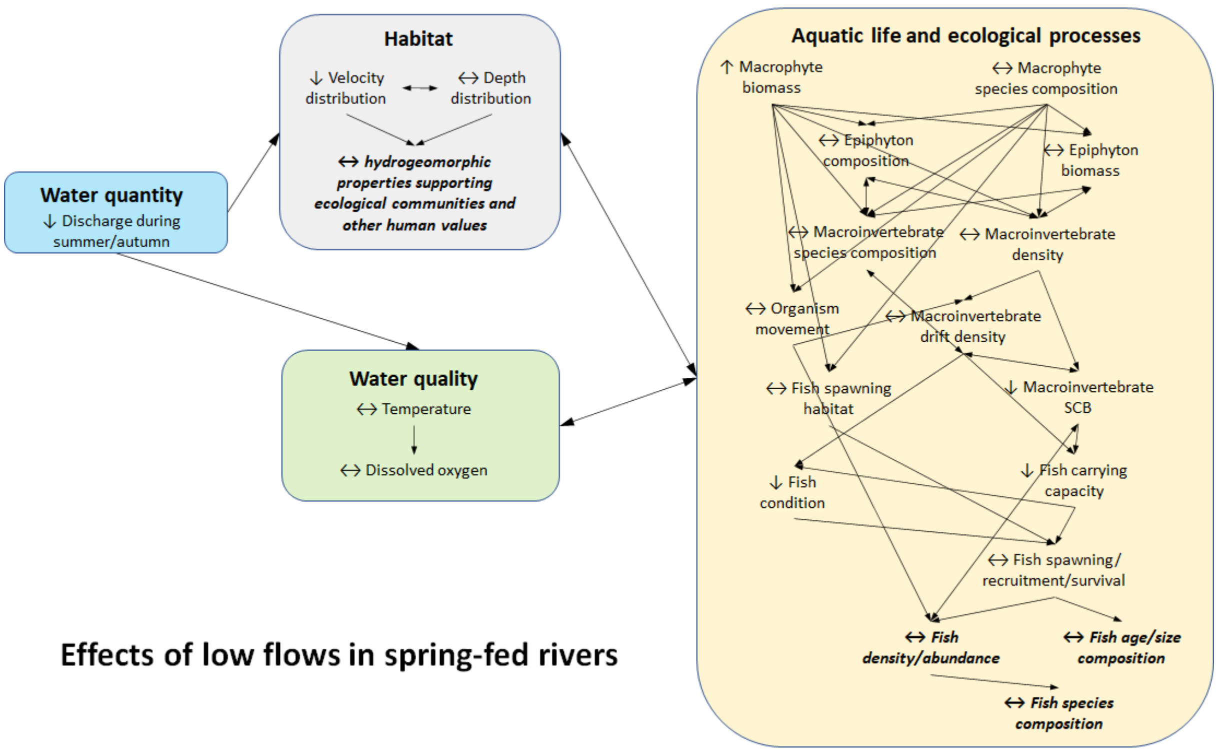

Figure 7-5: Conceptual model of the effects of low flows on metrics within spring-fed rivers, grouped by components of ecosystem health. Up- and down-arrows denote increasing and decreasing respectively. Bidirectional lateral arrows denote change; either increase or decrease, depending on the specific low-flow hydrological stressor considered, its timing, duration and magnitude. Arrows between metrics in different components (boxes) were not included; including them resulted in a very high density of arrows throughout the diagram that did more to obfuscate than clarify.

Low flows may have the following key impacts on **physical habitat** in spring-fed rivers, resulting in changes in the hydrogeomorphic properties of a river reach that directly (e.g., canoeing) and indirectly (support of aquatic life and ecological processes) support focal values (Figure 7-5):

- Reduced velocity (Biggs 1996, Riis and Biggs 2003).
- Depth may decrease, depending on the extent to which reduced velocity increases macrophyte growth; increased macrophyte biomass can compensate for reductions in depth associated with lowered discharge (Champion and Tanner 2000).

Low flows may have the following effects on **water quality** in spring-fed rivers (Figure 7-5):

- Reduced discharge may affect water temperatures, but the magnitude of any effect in spring-fed rivers is likely to be much smaller than that of effects in runoff-fed rivers. There is much uncertainty about the impact of flow on water temperature in spring-fed rivers as (a) spring-fed rivers have a more stable (and cooler) thermal regime than runoff-fed rivers; and (b) macrophyte growth—which may increase during low flows—can shade the water column hence cool water temperatures (Lusardi et al. 2021).
- Increased macrophyte growth during low flows may increase diel variability in dissolved oxygen, leading to potential hypoxia at night when respiration exceeds photosynthesis, and hyperoxia during the daytime when photosynthesis outpaces respiration (Kaenel et al. 2000, Wetzel 2001).

Flow-mediated effects on physical habitat and water quality will interact to affect **aquatic life and ecological processes** supporting fish populations (Figure 7-5):

- Reduced velocity may increase macrophyte growth, hence macrophyte biomass. Changed hydraulic conditions may also affect macrophyte species composition (Riis and Biggs 2003, Franklin et al. 2008, Riis et al. 2008). These changes to the macrophyte community have feedback effects on water quality and physical habitat (Franklin et al. 2008).
- Living macrophytes are generally not a preferred food of macroinvertebrates. Macrophytes do, however, support epiphyton communities that in turn support a rich and diverse macroinvertebrate community. Increased macrophyte biomass and/or a change in the macrophyte community will likely change the epiphyton community supporting higher trophic levels in the food web (Lusardi et al. 2021).
- Changes in the biomass and composition of epiphyton as well as the macrophyte community itself can increase the biomass of macroinvertebrates in reaches of spring-fed rivers, which in turn will affect the density and composition of drifting invertebrates (Gregg and Rose 1985, Lusardi et al. 2018).
- The interaction between epiphyton and macroinvertebrates will be characterised by both bottom-up and top-down effects. That is, an increase in epiphyton biomass may increase macroinvertebrate biomass, but increased macroinvertebrate biomass may also suppress epiphyton biomass.
- Dense stands of macrophytes may limit longitudinal movement of organisms.

- A change in the density and composition of macrophytes may affect the quality and quantity of fish spawning and recruitment habitat (Lusardi et al. 2018).
- We propose that the dominant mechanisms by which low flows affect fish populations of spring-fed rivers is via macrophyte-mediated impacts on (a) the availability of macroinvertebrate prey; and (b) dissolved oxygen, particularly low dissolved oxygen at night.

7.3.2 Flow harvesting

A conceptualisation of the effects of flow harvesting on metrics within runoff-fed rivers is presented in Figure 7-6. As was the case for our low flow conceptualisation (Figure 7-4), there was no reason to prepare different conceptual models for hard- and soft-bottomed runoff-fed rivers as the basic, qualitative ecological relationships are the same (Junk et al. 1989, Puckridge et al. 1998, Winemiller 2004, Humphries et al. 2014). Figure 7-6 is structured by the NPSFM components of ecosystem health (water quantity, habitat, aquatic life and ecological processes). Water quality is not included as the impacts of flow harvesting on water quality (e.g., temperature, dissolved oxygen; Figure 7-4) are likely minor. The conceptual model was based on several syntheses of the extensive body of research addressing the role played by floods in the structuring and functioning of rivers (Junk et al. 1989, Poff et al. 1997, Puckridge et al. 1998, Bunn and Arthington 2002, Winemiller 2004, Humphries et al. 2014, Stoffels et al. 2022). Additional supporting literature is occasionally provided in the explanation of the conceptual model.

Flow harvesting may have the following effects on **physical habitat,** resulting in changes in the hydrogeomorphic properties of a river reach that directly (e.g., swimming; boating) and indirectly (support of aquatic life and ecological processes) support focal values (Figure 7-6).

- Reducing the frequency of high flows will in turn reduce the frequency with which specific velocity thresholds are exceeded. Reducing the frequency of high velocities may in turn result in less frequent mobilisation and transport of fine sediment (sands and silts) hence, conversely, higher retention of fine sediment within river reaches (Hoyle et al. 2017).
- Floods are the physical force shaping the geomorphic structure of rivers—they interact with underlying geology to shape bed form and gradient, hence physical habitat diversity in lateral, longitudinal and vertical dimensions of the river (Chessman et al. 2006, Graf 2006).
- Floods are a—perhaps *the*—critical determinant of the natural character of a river (Brierley et al. 2019). Loss of a river's natural character impacts upon the capacity of the river to support aquatic life and ecosystem services to humans (including recreational values).

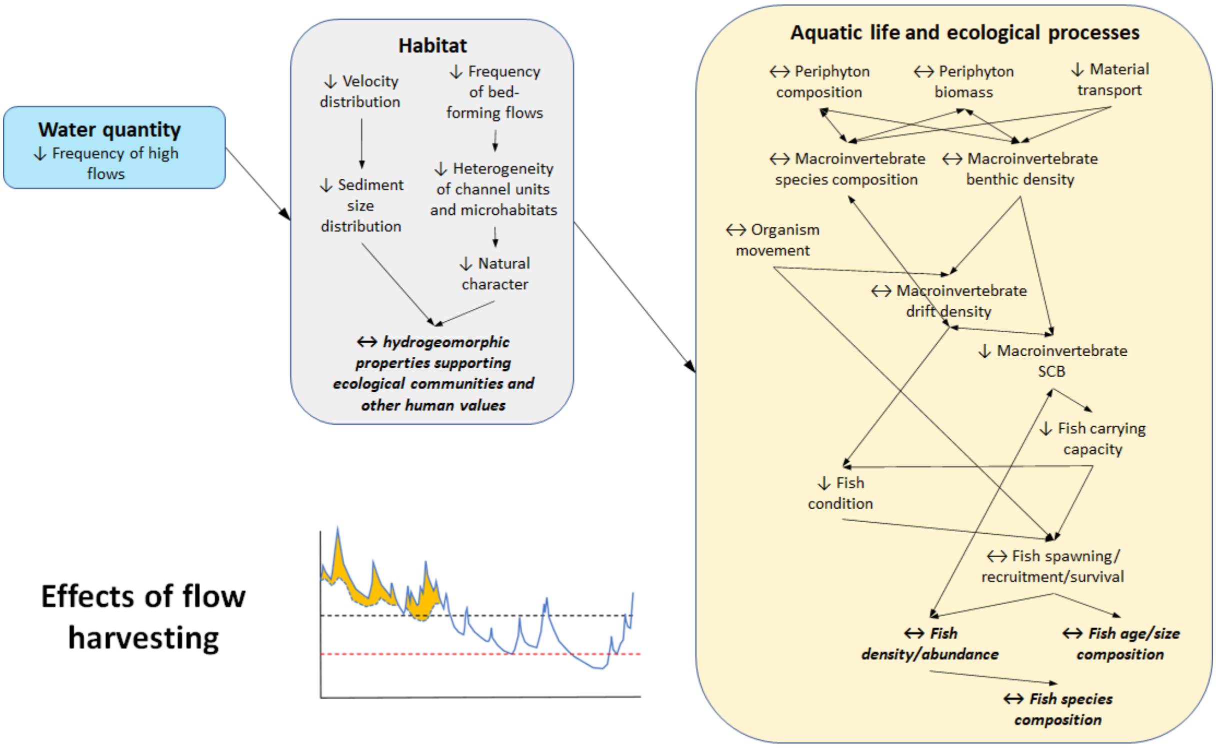

Figure 7-6: Conceptual model of effects of flow harvesting on metrics grouped by physical habitat and aquatic life and ecological processes. Up- and down-arrows denote increasing and decreasing respectively. Bidirectional lateral arrows denote change; either increase or decrease. Arrows between metrics in different components (boxes) were not included; including them resulted in a very high density of arrows throughout the diagram that did more to obfuscate than clarify.

A reduction in the frequency of floods may affect **aquatic life and ecological processes** indirectly, via effects on physical habitat, and directly (Figure 7-6). Many indirect effects are (broadly) qualitatively similar to those captured for low flows (Section 7.3.1), so we do not repeat those descriptions here. Instead we highlight some important differences between flow harvesting and low flows with respect to the mechanisms by which river flow may indirectly affect aquatic life and ecological processes.

- Sediment composition exerts a strong influence over periphyton and macroinvertebrate communities, hence the community of predators at higher trophic levels that depend on them for nutrition (fishes, koura, etc.; discussed above). Low flows may change the sediment composition of a river reach instantaneously as a consequence of a change in the benthic habitats that are inundated. Assuming other components of the flow regime are preserved (i.e., water takes only affect low flows, not floods etc.), low flows should have a relatively minor influence over the sediment composition of river reaches as a whole (inundated and dry areas of the river bed) over the long term. By contrast, a reduction in flood frequency may affect the sediment composition of an entire river reach (inundated and dry river bed) over the long term, as a result of chronic increases in the retention of fine sediment and reduced bed mobility. One could therefore suggest that reduced frequency of high flows poses a more serious threat to riverine food webs in the long term, in that the effects may be less easily reversed than those of low flows (Boulton et al. 1998, Shearer et al. 2015, Hoyle et al. 2017).

- Reducing the frequency of bed-forming flows may in turn change the composition of physical habitat at numerous scales. Long-term effects may include reduction in diversity of hydraulic habitat and the abundance of slackwater habitats (Vietz et al. 2013, Vietz and Finlayson 2017, Grams et al. 2022). Although little is known about the physical forces acting upon riverine fish recruitment in New Zealand, the international literature has shown that slackwater habitat—and physical habitat diversity in general—has a positive effect on rates of fish recruitment (Humphries et al. 2020, Stoffels et al. 2022). It is possible, therefore, that a reduction in the diversity of channel units and microhabitats may reduce fish recruitment, hence population size in New Zealand rivers.

Reducing the frequency of high flows may affect aquatic life and ecological processes in rivers directly:

- The life-histories of riverine fauna and flora are adapted to a natural flow regime (Lytle and Poff 2004). Freshwater scientists have demonstrated how many and varied critical population processes are triggered by high flow events—from the timing of insect emergence and oviposition (Lytle 2008), to fish spawning (King et al. 2010) and movement (David and Closs 2002), and riparian plant recruitment (Bunn et al. 2006, Tonkin et al. 2018). We have a poor understanding of how high flow events affect the life histories of New Zealand's riverine organisms. However, given the large body of evidence for adaptation to natural flow regimes in other countries, we may expect such adaptations in our flora and fauna as well.

8 Metrics – descriptions and recommendations

8.1 Overview

In this section we recommend metrics based on the M&E objectives and strategies identified in Section 5, and the conceptualisations presented in Sections 7.2 and 7.3. The list of metrics may appear resource-intensive, but the reader should note that:

- We do not recommend that all metrics be monitored at both runoff- and spring-fed rivers.
- We do not recommend monitoring of all metrics at both hub and spoke sites.
- Most metrics are high-priority at hub sites only. There are very few hub sites per region.
- Most metrics are not monitored annually. In most cases metrics are sampled twice during every five-year period, noting the aim for most metrics is to develop functions defining how discharge (flow rate) affects metric state.
- Several metrics are monitored using automated remote loggers.

We present metrics using consistent structures of explanation (explained below) for both hub and spoke sites for runoff- and spring-fed rivers.

8.1.1 Metrics and methods described in this plan should be considered first drafts

Designing metrics and methods that best meet M&E objectives is a great challenge. It is unlikely that the metric methods described below will be ready for implementation in their current form. They will require consideration by, and feedback from, councils such that they can be further refined to best meet our M&E objectives. A lot of careful thought has been put into the design of the metrics and methods presented below, but they should nevertheless be viewed as **draft metrics and methods**.

8.1.2 Structure of metric descriptions at hub sites

For the metrics we recommend at **hub sites**, we present the metric name then describe the metric using the following structure:

Description: Brief written description of the metric.

Units: Measurement units of the metric.

Values/components: The NPSFM values and value components to which the metric is relevant.

Why? A brief justification of why the metric should be included.

Targets flow modification: The form of flow modification that the metric is most relevant to: low flows, flow harvesting or both.

Time-scales of response detected: The most relevant time-scales at which the metric can be expected to respond to changes in the flow regime. Can be some subset of: hours, days, weeks, months, years, decades.

Priority at hub sites: The level of priority assigned to the metric, or how important the metric is to meeting monitoring objectives (Section 5.3.1). Note that priority levels are not an indication of cost—for example, *essential* metrics have draft methods that we have designed with monitoring Objectives 10 and 15 as constraints (Table 5-2; Figure 5-6). Priority levels are provided as a guide only—councils may assign higher priority to metrics if they have the resources to do so. One of three levels of priority may be assigned:

- *Essential*: Metric must be monitored at the hub site by all councils participating in this nationally-consistent program as the metric as described is essential to meeting monitoring objectives. The information supplied by this metric for evaluation/forecasting at a site cannot be assumed or approximated based on information from other sites/regions or the literature.
- *High*: Metric should be monitored at hub sites across 5 or more councils, but not necessarily all councils. The information supplied by this metric for evaluation/forecasting at a site may be assumed or approximated based on information from other sites/regions. But the information is still required from a subset of regions within the first five years of the program.
- *Low*: Metric not critical to meeting fundamental objectives but still would add significant value. Metric should be monitored at hub sites across 3-5 councils, but not necessarily all councils. The information supplied by this metric for evaluation/forecasting at a site may be assumed or approximated based on information from other sites/regions.

Sampling frequency/intensity: Overview of the temporal and spatial intensity of sampling, to give councils a basic understanding of the level of effort—hence resources—required to monitor the metric.

Method, in brief: A brief point-form overview of the method is presented in this plan primarily to demonstrate the approach to selection of monitoring metrics encouraged in this plan. , The approach departs strongly from surveillance monitoring (annual sampling of state for trends) in a manner that is designed to achieve monitoring objectives (Section 5.3.1) for minimal cost.

Minimum data, in brief: Very brief description of the data obtained.

8.1.3 Structure of metric description at spoke sites

Only metrics deemed *essential* at hub sites (see Section 8.1.2) are recommended for monitoring at spoke sites—councils should view all metrics with a priority level of *low* or *high* at hub sites as a "nice-to-have" at spoke sites. Metrics deemed *essential* at hub sites are monitored at spoke sites but with reduced effort. With the exception of fish relative abundance, metrics recommended at spoke sites define the dynamics of physical habitat—temporal variation in hydrology and how that variation affects physical habitat of a reach. The primary purposes of physical metrics at spoke sites are:

- To provide the necessary physical foundation for spatial extrapolation of the ecological knowledge obtained through the metrics monitored at hub sites. Following the strategies presented in Section 5.3.3, data obtained at hub sites will be used to develop evaluation- and decision-support tools. Those tools may be applied at spoke sites if we monitor the physical metrics at spoke sites that serve as predictors in those tools.

- Physical metrics monitored at spoke sites serve as important indicators of riverine values in their own right, particularly with respect to NPSFM values Ecosystem health – habitat, natural form and character, and transport and tauranga waka.

We recommend including fish relative abundance once every five years at spoke sites for the following reasons:

- Measurement of fish relative abundance is critical for meeting Fundamental Objective 17 via Means Objectives 7 and 9.
- Fish, being at the top of the riverine food web, integrate biophysical responses to flow regimes at numerous trophic levels below them (Poff and Zimmerman 2010).

For the metrics we recommend at **spoke sites**, we present the metric name then describe the metric using the following structure:

Sampling frequency/intensity: Definition as described for hub sites.

Method – notable departures from method at hub sites: Notes on the key differences between the method as applied at spoke sites, compared to hub sites.

8.2 Tabular summary of metrics at runoff-fed rivers

To facilitate an overview of the metrics suggested, tabular summaries are presented below.

In the tables below a *sampling year* comprises the period Jan-Mar during any specific year of each five-year reporting period. The specific years selected as sampling years is left to the discretion/convenience of councils. During a sampling year, sampling is distributed over three reaches per hub site, and across a range of flow levels, which we assume will vary over the summer-autumn (Jan-Mar) period. During a sampling year, the magnitude of sampling within each reach depends on the metric.

8.2.1 Hub sites

Table 8-1: Summary of the metrics to be monitored at runoff-fed hub sites. Form of flow modification that is most relevant to each metric is indicated ("Flow mod."). The time-scales at which the metric best detects response to changes in the flow regime are indicated. Levels of monitoring priority are presented (Essential, High or Low; see text for explanation). Grey cells are applicable to the respective metric.

	Flow mod.		Time-scales of response detected							
Metric	**Low flow**	**Flow harvest**	**Hours**	**Days**	**Weeks**	**Months**	**Years**	**Decades**	**Priority**	**Sampling intensity**
Daily discharge	■	■	■	■	■	■	■	■	E	Continuous automated
Depth-velocity distribution	■	■				■	■	■	E	Two sampling years every five years.
Wetted area	■	■				■	■	■	E	Two sampling years every five years.
Water temperature	■		■	■	■	■	■	■	H	Continuous automated
Sediment size distribution		■				■	■	■	E	Two sampling years every five years.
Macroinvertebrate benthic density	■	■				■	■	■	H	Two sampling years every five years.
Macroinvertebrate drift density	■	■				■	■	■	L	Three sampling years every five years.
Fish relative abundance	■	■					■	■	E	Annually
Fish size composition	■	■					■	■	H	Annually
Aerial reach photograph	■	■					■	■	H	Annually

Table 8-2: Summary of metrics to be monitored at runoff-fed spoke sites. Form of flow modification that is most relevant to each metric is indicated ("Flow mod."). The time-scales at which the metric best detects response to changes in the flow regime are indicated. Levels of monitoring priority are presented (Essential, High or Low; see text for explanation). Grey cells are applicable to the respective metric.

	Flow mod		Time-scales of response detected							
Metric	**Low flow**	**Flow harvest**	**Hours**	**Days**	**Weeks**	**Months**	**Years**	**Decades**	**Priority**	**Sampling intensity**
Daily discharge	■	■	■	■	■	■	■	■	E	Continuous automated
Depth-velocity distribution	■	■				■	■	■	E	Two sampling years every five years, but with 1/3 effort of hub sites
Wetted area	■	■				■	■	■	E	Two sampling years every five years.
Sediment size distribution		■				■	■	■	E	Two sampling years every five years, but with 1/3 effort of hub sites
Fish relative abundance	■	■					■	■	E	Once every five years; 1/5 effort of hub sites

8.3 Tabular summary of metrics for spring-fed rivers

8.3.1 Hub sites

Table 8-3: Summary of the metrics to be monitored at spring-fed hub sites. Form of flow modification that is most relevant to each metric is indicated (“Flow mod."). The time-scales at which the metric best detects response to changes in the flow regime are indicated. Levels of monitoring priority are presented (Essential, High or Low; see text for explanation). Grey cells are applicable to the respective metric.

	Flow mod		Time-scales of response detected							
Metric	Low flow	Flow harvest	Hours	Days	Weeks	Months	Years	Decades	Priority	Sampling intensity
Daily discharge	■		■	■	■	■	■	■	NA	Continuous automated
Discharge spot measurement	■					■	■	■	E	Two sampling years every five years
Cross-sectional area and velocity field	■					■	■	■	E	Two sampling years every five years
Macrophyte cross-sectional area	■					■	■	■	E	Two sampling years every five years
Macroinvertebrate drift density	■					■	■	■	H	Three sampling years every five years
Dissolved oxygen dynamics	■		■	■	■	■	■	■	H	Continuous automated over three 10-day periods, twice every five years

8.4 Site structure – runoff-fed rivers

When describing metrics and methods we frequently refer to *sites* and *reaches* within sites. The site description below focuses on hub sites, but spoke sites also should have the same basic structure. As specified in Section 6.4, councils implementing this plan should have between one and three hub sites for runoff-fed rivers and for spring-fed rivers. Hub sites within runoff-fed streams have the following basic setup (Figure 8-1):

- A site has the characteristics described in Section 6.2.
- A site is divided into three *reaches*.
- If the (approximate) mean width of the site is *w*, then a reach has length ca. 8*w*.
- Councils must clearly specify *w* and total actual length of each reach, as metric calculation (e.g., wetted width) is dependent on reach length.
- The *river right* and *left* are, respectively, the right and left banks of the river when facing the same direction as the flow.
- Various metrics are sampled along transects, of which there are nine per reach. Sampling along equidistant transects ensures that sample location is unbiased with

respect to longitudinal habitat heterogeneity[20], such that our metrics are representative of the reach as a whole.

- Each transect has ID X.Y where X is reach number (1, 2 or 3) and Y is transect number (1, 2,…, 9).
- Transects are comprised of *sampling points*, the specific locations of metric measurements.
- Sampling points should be spatially-explicit, given the need to align numerous measurements in space and time such that we may develop quantitative relationships between metrics. This means that sampling points need to have a consistent ID nomenclature and be labelled R1, R2,…, R5, where R denotes *right bank* and the integer denotes a sequence from the right bank. When paired with columns denoting date, transect ID, and site ID, we have unique sample IDs for every sample taken.

Further details drawn in Figure 8-1 are referred to in the metric descriptions, below.

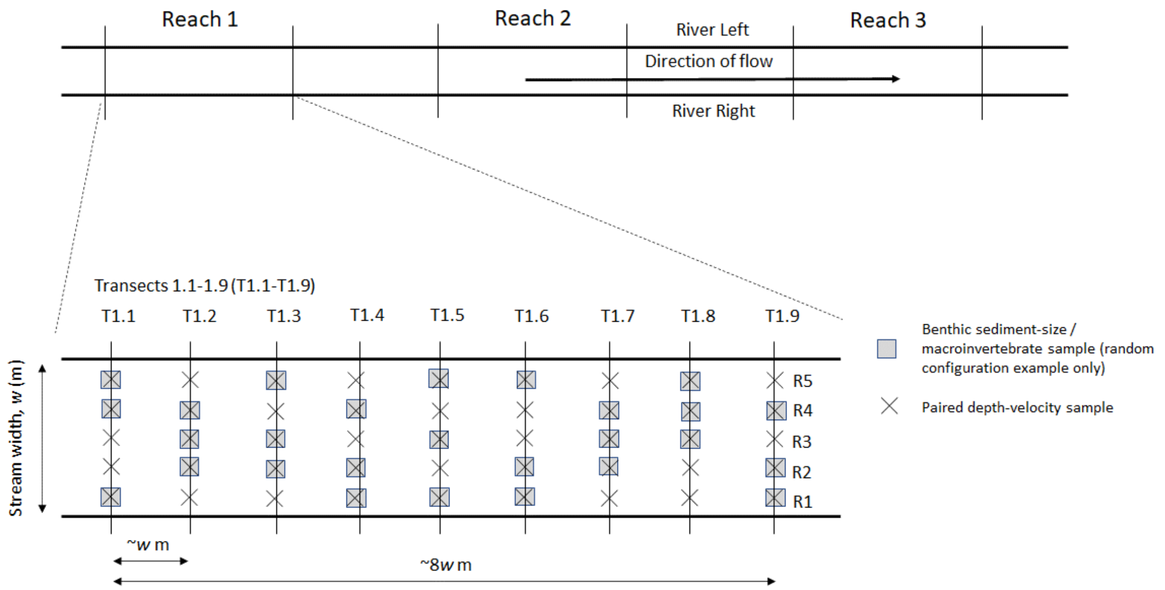

Figure 8-1: Basic setup of hub sites within runoff-fed rivers. Average river width at a site denoted *w*.

[20] E.g., ensures we don't just sample at the easiest crossings.

8.5 Metrics for runoff-fed rivers – hub sites

8.5.1 Stressors that interact with flow regimes to affect outcomes

Following our recommendations in Section 5.3, stressors that interact with flow to affect outcomes should be monitored at the same sites. At a minimum we recommend monitoring the following NPSFM attributes and metrics at hub sites:

1. Suspended fine sediment.
2. Deposited fine sediment.
3. Dissolved reactive phosphorus.
4. Dissolved inorganic nitrogen (not technically an attribute, but see Clause 3.13).

8.5.2 Mean daily discharge

Description: Time series of mean daily flow rate.

Units: $m^3 s^{-1}$.

Values/components: Ecosystem health – water quantity.

Why? Following Clause 3.17 of the NPSFM, take limits may be defined as flow rate (synonymous with rate of discharge). Mean daily discharge is the key variable that will link environmental responses to water take rules.

Targets flow modification: Low flows and flow harvesting.

Time-scales of response detected: Hours. Days. Weeks. Months. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Discharge sampled hourly, throughout every day of year.

Method, in brief:

- Automated, continuous, remote monitoring at gauged stations.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, mean discharge. Rows correspond to individual days.

8.5.3 Paired depth-velocity distribution

Description: Paired frequency distributions of water depths and velocities at a site, obtained by sampling water depth and velocities at the same points throughout a site.

Units: cm (depth); $m s^{-1}$ (velocity).

Values/components: Ecosystem health – habitat. Natural form and character. Transport and tauranga waka.

Why? Both depth and velocity are critical determinants of hydrogeomorphic properties supporting ecological communities and other human values. Depth and velocity are particularly important drivers of community composition and abundance of periphyton and macroinvertebrates, hence fish carrying capacity. Functions defining the relationship between mean daily discharge and depth-velocity distributions could be developed. Paired depth-velocity distributions would serve as a predictor of responses at higher trophic levels. At hub sites, paired depth-velocity measurements would be sampled at the same points used for periphyton and/or macroinvertebrate sampling, such that we may develop national functions defining the relationships between depth, velocity and periphyton/macroinvertebrates.

Targets flow modification: Low flows (primarily).

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Two *sampling years* every five years[21].
- A total of 5 (sampling points) x 9 (transects) = 45 paired observations for each reach (given three reaches: 45 x 3 = 135 paired observations) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 135 x 2 = 270 paired depth-velocity measurements per hub site, every five-year period (this equates to only 54 observations per year, per hub site).

Method, in brief:

- Sampling to take place during two years within every five-year period. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge and depth-velocity distributions. Timing of sampling years to be determined by councils.
- A sampling year defined by three contiguous months January, February, March. This late summer-autumn period is when instream productivity is at its highest, and when low flows are most likely to occur. It is also during a time most conducive to field work.
- Three transects from each of three reaches will be sampled per month (Jan, Feb, Mar) of the sampling year. E.g., during January transects TX.1, TX.4 and TX.7 would be sampled from Reach X=1, X=2, X=3; and during February transects TX.2, TX.5 and TX.8 would be sampled from Reach X=1, X=2, X=3, and so on (Figure 8-1). Sampling this way across three months helps us to obtain discharge-depth/velocity relationships across a range of discharges (assuming discharge varies across the three sampling months in a sampling year). Sampling across three months within a year also helps us to obtain data representative of the low flow period as a whole, not just a narrow period within the late summer-autumn period.
- During each month (Jan, Feb, Mar) paired depth-velocity distributions will be sampled from five points across the transect (Figure 8-1). The positioning of sampling points

[21] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

laterally, across the transect should be random to ensure depth-velocity observations are representative of the transect.

- At each point, depth measured with a ruler to nearest cm.
- At each point velocity measured in middle of water column and ca. 25% of column depth above the bottom and below the surface (three velocity readings per point). Velocimeter should take time-averaged observations over at ten seconds.
- Paired depth-velocity observations must be taken at the same time.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID, sampling point ID, depth, velocity. Rows correspond to individual measurements of depth and velocity at points along transects.

8.5.4 Wetted area

Description: Wetted area of site, obtained as a simple function of mean wetted width over the known length of the site.

Units: m^2

Values/components: Ecosystem health – habitat. Natural form and character. Transport and tauranga waka.

Why? Critical determinant of hydrogeomorphic properties supporting ecological communities and other human values. Wetted area determines the surface area available for macroinvertebrate and fish production. Functions defining the relationship between mean daily discharge and wetted area could be developed. When coupled with macroinvertebrate density (see metric below) wetted area can be used to predict macroinvertebrate standing crop biomass at a site and, in turn, fish carrying capacity.

Targets flow modification: Low flows.

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- At least two *sampling years* every five years[22].
- A total of 9 transects along which width is measured per reach (given three reaches: 9 x 3 = 27 observations of width) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 27 x 2 = 54 paired depth-velocity measurements per hub site, every five-year period (this equates to only ca. 11 observations per year, per hub site).

[22] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

Method, in brief:

- Sampling to take place during two years within every five-year period. Timing of sampling years to coincide with that of depth-velocity sampling. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge and depth-velocity distributions.
- Temporal-spatial structure of sampling to follow that described for depth-velocity distributions.
- Width measured along each transect using an appropriate method[23].

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID, width. Although the metric of interest is wetted area (m^2), data should be stored as numerous wetted widths (m), from which calculation of area can be easily automated. Rows correspond to individual measurements of widths along transects. Numerous rows per site, per date/discharge.

8.5.5 Water temperature

Description: Continuously-logged water temperature during summer-autumn.

Units: °C.

Values/components: Ecosystem health – Water quality.

Why? Critical determinant of aquatic life and ecological processes, and increases during low flows. Aquatic plants and animals of New Zealand rivers are ectothermic, so their physiological rates—hence growth, reproduction and survival—are strongly determined by water temperature. Functions defining the relationships between mean daily discharge and water temperature would be useful for adaptive management.

Targets flow modification: Low flows.

Time-scales of response detected: Hours. Days. Weeks. Months. Years. Decades.

Priority at hub sites: High.

Sampling frequency-intensity:

- Temperature measured every 30 minutes by logger, every day of Dec-May period each year.

Method, in brief:

- Three water temperature loggers should be securely installed within a deep pool within each reach of the site (Figure 8-1).
- Loggers should be set to log water temperature every 30 min throughout the months of Dec – May, when low flows may occur. Loggers should be removed for data download at the end of each year.

[23] Easiest with a digital range finder.

- Logged data would be used to build a function describing the relationship between discharge and water temperature during summer-autumn. The aim is to determine water temperature over the domain of discharges experienced during a low flow period.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, unique logger ID, date, time, temperature. Numerous rows per site, per date/discharge, per logger—each row a unique measurement on the half-hour. Number of rows corresponds to the number of 30 min intervals during Dec-May each year.

8.5.6 Sediment size distribution

Description: Size frequency distribution of benthic sediment within a site.

Units: mm for individual measurements, but entered as frequencies under size categories (see Minimum data, in brief).

Values/components: Ecosystem health – Habitat. Natural form and character.

Why? Critical determinant of benthic ecological processes, including macroinvertebrate—hence fish—productivity. We hypothesise that one of the main mechanisms by which flow harvesting will influence ecosystem health is by changing benthic sediment size distributions, and the ecological processes affected by sediment size. Functions defining the relationship between flood frequency and sediment size distribution could be developed. When coupled with macroinvertebrate density (see metric below) sediment size statistics can be used to predict macroinvertebrate standing crop biomass at a site and, in turn, fish carrying capacity.

Targets flow modification: Flow harvesting (primarily) and low flows.

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Two *sampling years* every five years[24].
- A total of 3 (sampling points) x 9 (transects) = 27 individual observations for each reach (given three reaches: 27 x 3 = 81 observations) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 81 x 2 = 162 individual Wolman counts per hub site, every five-year period (this equates to only 32 Wolman counts per year).

Method, in brief:

- Sediment size distributions to be sampled using Wolman counts. Wolman counts, while being more labour intensive, are superior to visual assessments as they offer a more objective and informative (greater number of sediment size categories) metric of benthic sediment distribution.
- Sampling to take place during two years within every five-year period. Annual sampling not necessary as the aim is to build functions defining the relationship between

[24] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

discharge and sediment distributions. Timing of sampling years to coincide with those of depth-velocity sampling.

- Temporal-spatial structure of sampling within sampling years follows that outlined for depth-velocity sampling, with the following exception:
  - Three (cf. five) points along each transect are to be sampled (Figure 8-1).
  - The three sampling points along the transect are to be a random subset of the five sampling points used for paired depth-velocity observations.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID, sampling point ID and 16 columns corresponding to the 16 Wolman sediment size categories. Rows correspond to individual Wolman counts along transects—spreadsheet cells under each Wolman size category column should contain the frequency with which that sediment category was encountered for each Wolman count. Numerous rows per site, per date/discharge.

8.5.7 Macroinvertebrate benthic density

Description: Taxon-specific density of benthic macroinvertebrates.

Units: individuals m^2 (for each taxon).

Values/components: Ecosystem health – Aquatic life.

Why? The macroinvertebrate community comprises the majority of aquatic animal diversity within New Zealand's rivers. The macroinvertebrate community drives most fish production in New Zealand rivers. Density of macroinvertebrates is required to understand and predict the consequences of changes in the macroinvertebrate community to fishes, hence high-value species of the ecosystem. Taxon-specific densities would be useful, as not all macroinvertebrates are favoured prey items of fishes. Further, taxon-specific densities can shed light on which functional groups / guilds are impacted by changes in the flow regime. Functions defining the relationship between depth, velocity, sediment and macroinvertebrate densities could be developed, facilitating a predictive understanding of the indirect impacts of flow regimes on macroinvertebrates. In turn, when coupled with reach-scale measurements of wetted area and available habitat, such functions allow estimation of reach-wide macroinvertebrate standing crop biomass and fish carrying capacity.

Targets flow modification: Low flows and flow harvesting (via sediment impacts).

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Two *sampling years* every five years[25].
- A total of 3 (sampling points) x 9 (transects) = 27 individual observations for each reach (given three reaches: 27 x 3 = 81 observations) each sampling year (Figure 8-1).

[25] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

- Given two sampling years per reporting cycle: 81 x 2 = 162 individual samples per hub site, every five-year period (this equates to only 32 samples per year).

Method, in brief:

- Macroinvertebrate densities to be sampled using Hess or surber samplers[26].
- Sampling to take place during two years within every five-year period. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge, depth, velocity, sediment size and macroinvertebrate density. Timing of sampling years to coincide with those of depth-velocity sampling.
- Temporal-spatial structure of sampling within sampling years follows that outlined for sediment size distributions. Macroinvertebrate sampling points should be the same as those used for Wolman counts (conduct Wolman counts after macroinvertebrate sampling).
- Depth and velocity should be recorded at the point and time of macroinvertebrate sampling.
- Macroinvertebrate sample sorting in the laboratory should follow usual procedures with taxon-specific counts at the lowest taxonomic resolution practicable.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID code, sampling point ID (same as that used for corresponding Wolman count) and macroinvertebrate taxon. Rows correspond to individual Hess/surber samples along transects.

8.5.8 Macroinvertebrate drift density

Description: Taxon-specific density of drifting macroinvertebrates.

Units: individuals/m^3 (for each taxon).

Values/components: Ecosystem health – Aquatic life.

Why? The macroinvertebrate community comprises the majority of aquatic animal diversity within New Zealand's rivers. The macroinvertebrate community drives most fish production in New Zealand rivers. While benthic density allows an assessment of how flow-induced changes to physical habitat affect standing crop biomass of macroinvertebrates, drift density allows assessment of how flow affects the subset of benthic macroinvertebrate entering the drift, hence the density of macroinvertebrate available to drift-feeding fishes. Drift density is the result of both benthic food web (e.g., productivity) and behavioural responses to flow. Density of macroinvertebrates is required to understand and predict the consequences of changes in the macroinvertebrate community to fishes, hence high-value species of the ecosystem. Taxon-specific densities would be useful, as not all macroinvertebrates are favoured prey items of fishes. Further, taxon-specific densities can shed light on which functional groups / guilds are impacted by changes in the flow regime. Functions could be developed, facilitating a predictive understanding of the direct and indirect impacts of flow regimes on drifting macroinvertebrates. In turn, such functions shed light on fish carrying capacity.

[26] Either Hess or surber samples are appropriate – the key thing is to obtain quantitative samples for density. Once a particular sampling apparatus is chosen (Hess or surber), keep using that apparatus to maintain consistency.

Targets flow modification: Low flows and flow harvesting (via sediment and velocity impacts).

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: Low.

Sampling frequency/intensity:

- Three *sampling years* every five years.
- A total of 3 sampling points for each reach, each sampling year. Given three reaches: 3 x 3 = 9 observations per hub site, each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 9 x 3 = 24 individual drift samples per hub site, every five-year period.

Method, in brief:

- Macroinvertebrate drift densities to be sampled using drift nets fitted with a calibrated velocimeter such that we may determine volume of water sampled.
- Sampling to take place during three years within every five-year period. Two of the three years should coincide with those used for sampling depth-velocity (and numerous other) metrics. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge, velocity and macroinvertebrate drift density.
- Temporal structure of sampling within sampling years follows that outlined for macroinvertebrate benthic density. Drift sampling points at the discretion of councils, but should be recorded relative to site layout (Figure 8-1).
- Depth should be recorded at the point and time of macroinvertebrate sampling.
- Macroinvertebrate sample sorting in the laboratory should follow usual procedures with taxon-specific counts at the lowest taxonomic resolution practicable.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID code, unique sample code and macroinvertebrate taxon. Rows correspond to individual drift samples.

8.5.9 Fish relative abundance

Description: Species-specific catch-per-unit-effort.

Units: Number of individuals per unit sampling effort (catch-per-unit-effort, CPUE).

Values/components: Ecosystem health – Aquatic life.

Why? The macroinvertebrate and physical habitat metrics described earlier provide an indirect assessment of the fish-supporting capacity of river reaches, but they do not provide a direct measure of which fish species are actually within a reach or the relative abundance of those fishes. Given the high-value nature of fishes—and Objectives 7, 9 and 17 of this plan (Figure 5-6; Table 5-2)—it is necessary to monitor fish abundance. Flow regimes affect fishes through a suite of direct and indirect mechanisms, so deciphering the drivers of changes in species' CPUE will be challenging. Monitoring the metrics described above at the same reaches where fish monitoring takes place will improve our ability to decipher the effects of the flow regime of fish CPUE. Following Objective 7, monitoring of fish presence-absence is not sufficient. Relative abundance is a more sensitive metric of change than presence-absence.

Targets flow modification: Low flows and flow harvesting.

Time-scales of response detected: Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Annual sampling.

Method, in brief:

- Follow the method of Joy et al. (2013), with the 150 m length sample reach—as required by the Joy et al. (2013) method—being within each of the three reaches we describe in Section .

Minimum data, in brief: See Joy et al. (2013).

8.5.10 Fish size composition

Description: Species-specific total lengths, indicating changes in population structure.

Units: Total length (mm).

Values/components: Ecosystem health – Aquatic life.

Why? Fish size structure is a good indicator of the age- and stage-structure of a fish population. Stage-structure of a population strongly affects population dynamics—how resistant and resilient the population is to environmental change. Further, the size structure of a population may change before CPUE does, so offers a particularly sensitive metric of fish population response.

Targets flow modification: Low flows and flow harvesting.

Time-scales of response detected: Years. Decades.

Priority at hub sites: High.

Sampling frequency/intensity:

- Annual sampling.

Method, in brief:

- Follow the method of Joy et al. (2013), with the 150 m length sample reach—as required by the Joy et al. (2013) method—being within each of the three reaches we describe in Section 8.4.

Minimum data, in brief: See Joy et al. (2013).

8.5.11 Aerial photograph

Description: Aerial/drone image of each study reach within the site.

Units: NA

Values/components: Ecosystem health – habitat. Natural form and character. Transport and tauranga waka.

Why? A large amount of information can be extracted from multi-year series of drone images, such as degree of braiding, encroachment of tree species / riparian cover, channel evolution, wetted area. This richness of information, coupled with the very low cost of obtaining such images, makes this a "no-brainer" metric.

Targets flow modification: Low flows and flow harvesting.

Time-scales of response detected: Years. Decades.

Priority at hub sites: High.

Sampling frequency/intensity:

- Annual sampling.
- Three images per site (one for each of three reaches).

Method, in brief:

- Aerial photographs of each reach, taken within the same month (April) each year.

Minimum data, in brief: Annual series of digital images.

8.6 Metrics for runoff-fed rivers – spoke sites

8.6.1 Mean daily discharge

Sampling frequency/intensity:

- Discharge sampled hourly, throughout every day of year.

Method – notable departures from method at hub sites:

- NA. Install automated, continuous, remote monitoring at gauged stations.

8.6.2 Paired depth-velocity distribution

Sampling frequency/intensity:

- 1/3 the sampling effort at hub sites, on a per-site basis.
- Two *sampling years* every five years.
- A total of 5 (sampling points) x 3 (transects) = 15 paired observations for each reach (given three reaches: 15 x 3 = 45 paired observations) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 45 x 2 = 90 paired depth-velocity measurements per spoke site, every five-year period.

Method – notable departures from method at hub sites:

- 1/3 the number of transects sampled.
- Sampling should take place over the same period (Jan-Mar) outlined for hub sites, but the timing of transect sampling within that period more flexible at spoke sites. Councils should aim to sample depth-velocity distributions (total of 45 at a site, within a sampling year) at spoke sites across a range of different discharge/flow levels.

8.6.3 Wetted area

Sampling frequency/intensity:

- At least two *sampling years* every five years.
- A total of 9 transects along which width is measured per reach (given three reaches: 9 x 3 = 27 observations of width) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 27 x 2 = 54 paired depth-velocity measurements per spoke site, every five-year period.

Method – notable departures from method at hub sites:

- None. Widths very easy to obtain so method follows that for hub sites. Timing more flexible and may be best done at the same time spoke sites are sampled for depth-velocity distributions.

8.6.4 Sediment size distribution

Sampling frequency/intensity:

- 1/3 the sampling effort at hub sites, on a per-site basis.
- Two *sampling years* every five years.
- A total of 3 (sampling points) x 3 (transects) = 9 individual observations for each reach (given three reaches: 9 x 3 = 27 observations) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 27 x 2 = 54 individual Wolman counts per spoke site, every five-year period.

Method, in brief:

- 1/3 the number of transects sampled.
- Sampling should take place over the same period (Jan-Mar) outlined for hub sites, but the timing of transect sampling within that period more flexible at spoke sites. Councils should aim to sample sediment distributions (total of 27 Wolman counts at a site, within a sampling year) at spoke sites across a range of different discharge/flow levels.

8.6.5 Fish relative abundance

Sampling frequency/intensity:

- 1/5 the sampling effort at hub sites, on a per-site basis.
- One sampling year every five years, on a per-site basis.

Method – notable departures from method at hub sites:

- None. Follow the same method as described for hub sites, but rotate sampling at a spoke sites on a 5-year schedule.

8.7 Site structure – spring-fed rivers

As was the case for runoff-fed rivers, when describing metrics and methods for spring-fed rivers we frequently refer to *sites* and *reaches* within sites. Here we provide a site description for hub sites on spring-fed rivers. Spoke sites would have the same basic structure, but **at this stage we are not recommending spoke sites for spring-fed rivers**, due to (a) the difficulty of discharge gauging in spring-fed, macrophyte-dominated rivers, and (b) the resource-intensive nature of obtaining useful data from macrophyte-dominated rivers (see metric descriptions below). As specified in Section 6.4, councils implementing this plan should have between one and three hub sites for runoff-fed rivers and for spring-fed rivers. Hub sites within spring-fed streams have the following basic setup (Figure 8-2):

- A site has the characteristics described in Section 6.2.
- A site is divided into three *reaches*.
- If the (approximate) mean width of the site is w, then a reach has length ca. $8w$.
- Councils must clearly specify w and total actual length of each reach, as metric calculation (e.g., wetted width) is dependent on reach length.
- The *river right* and *left* are, respectively, the right and left banks of the river when facing the same direction as the flow.
- Various metrics are sampled along transects, of which there are nine per reach. Sampling along equidistant transects ensures that sample location is unbiased with respect to longitudinal habitat heterogeneity[27], such that our metrics are representative of the reach as a whole.
- Each transect has ID X.Y where X is reach number (1, 2 or 3) and Y is transect number (1,2,…,9).
- Transects are comprised of *sampling points*, the specific locations of metric measurements.
- Sampling points should be spatially-explicit, given the need to align numerous measurements in space and time such that we may develop quantitative relationships between metrics. This means that sampling points need to have a consistent ID nomenclature and be labelled R1, R2,…, R7, where R denotes *right bank* and the

[27] E.g., ensures we don't just sample at the easiest crossings.

integer denotes a sequence from the right bank. When paired with columns denoting date, transect ID, and site ID, we have unique sample IDs for every sample taken.

Further details drawn in Figure 8-2 are referred to in the metric descriptions, below.

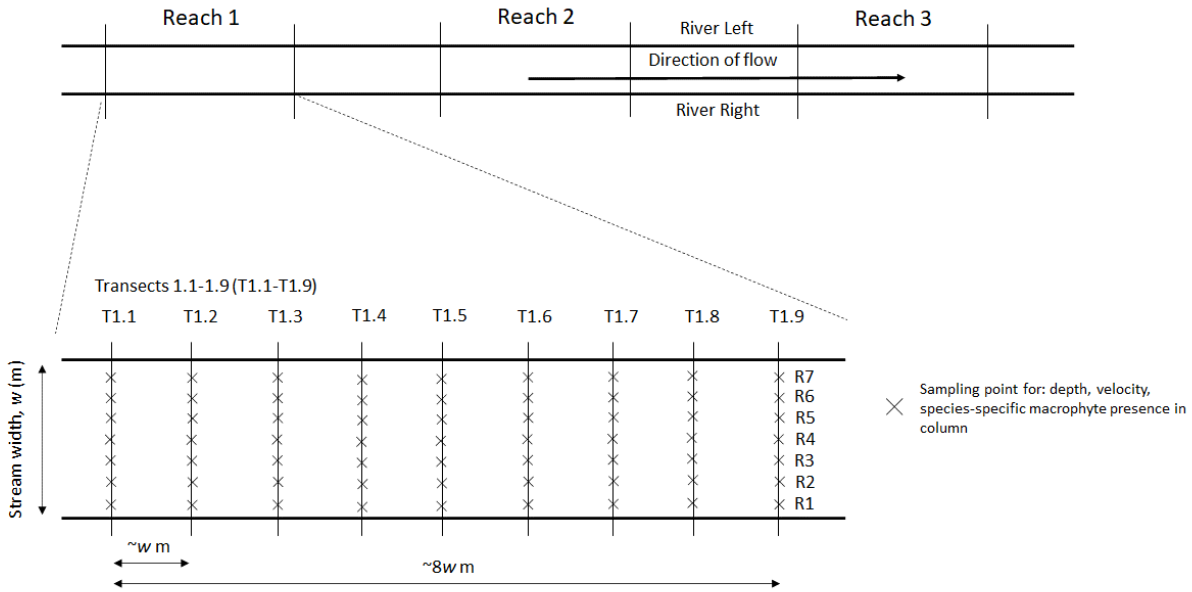

Figure 8-2: Schematic outlining the basic structure of a monitoring site on a spring-fed river. A site consists of three reaches. Each reach is comprised of nine transects. The river right and left are, respectively, the right and left banks of the river when facing the same direction as the flow. Notation for transect and sampling point IDs is given.

8.8 Metrics for spring-fed rivers – hub sites

8.8.1 Stressors that interact with flow regimes to affect outcomes

Following our recommendations in Section 5.3, stressors that interact with flow to affect outcomes should be monitored at the same sites. At a minimum we recommend monitoring the following NPSFM attributes and metrics at hub sites:

1. Suspended fine sediment.
2. Deposited fine sediment.
3. Dissolved reactive phosphorus.
4. Dissolved inorganic nitrogen (not technically an attribute, but see Clause 3.13).

8.8.2 Mean daily discharge

If a suitable site can be found for installation of an automated discharge gauging station then councils should do so. However, automated monitoring of discharge in spring-fed rivers is very difficult. In most cases, discharge gauges are calibrated using ratings curves, which assume that the relationship between wetted cross-sectional area and the velocity field through that wetted cross-section is constant in time. Macrophytes block the wetted cross-section. If there is variation in the extent to which macrophytes block flow[28], then gauges based on traditional ratings curves yield inaccurate discharges.

Installation of flow gauging stations in spring-fed rivers with macrophytes requires further investigation if this M&E plan were to be implemented.

8.8.3 Discharge spot measurement

Description: The discharge through a river reach as measured manually on a specific day.

Units: $m^3 s^{-1}$.

Values/components: Ecosystem health – water quantity.

Why? We aim to develop quantitative relationships between discharge and flow-dependent metrics. Noting the difficulty of automated flow gauging in spring-fed, macrophyte dominated rivers, there remains a need to obtain estimates of discharge during the times at which other flow-dependent metrics are sampled. It is essential to have estimates of discharge on the days we sample other metrics, described below.

Targets flow modification: Low flows.

Time-scales of response detected: Days. Months. Years.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Two *sampling years* every five years[29].
- A total of 5 (sampling points) x 9 (transects) = 45 paired observations for each reach (given three reaches: 45 x 3 = 135 paired observations) each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 135 x 2 = 270 paired depth-velocity measurements per hub site, every five-year period (this equates to only 54 observations per year, per hub site).

Method, in brief:

- We note that metric *cross-sectional area and velocity field* also yields data suitable for estimating discharge at a point in time. However, the metric *cross-sectional area and velocity field* is not specifically aimed at yielding a discharge estimate—transects are located randomly with respect to environmental conditions, such that we may obtain data on, for example, the extent to which macrophyte stands affect the cross-sectional velocity field. Such transects are not best suited for estimating discharge, so we recommend an additional estimate of cross-sectional area and velocity at a location with minimal macrophyte coverage for estimation of discharge on the day the metrics below are measured.

[28] Due to, for example, temporal variation in macrophyte species composition, variation in space and time in the amount of macrophyte growth, etc. Also note that macrophyte growth either side of the cross-section from which a ratings curve is obtained can strongly influence the relationship between cross-sectional area and discharge.
[29] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

- Discharge spot measurements are to be taken on the same days that metrics *cross-sectional area and velocity field* (Section 8.8.4) and *macrophyte cross-sectional area* (Section 8.8.5) are sampled. Consequently, three discharge spot measurements (one in each of three reaches at a site) will be taken during each of three months (Jan, Feb, Mar).
- Locate transect within each reach such that the channel cross-section has relatively homogeneous flow throughout and is relatively macrophyte-free.
- For each transect record:
  - wetted width;
  - at ten randomly-positioned point along the transect, record depth and velocity in the middle of the water column and ca. 25% of column depth above the bottom and below the surface (three velocity readings per point). The velocimeter should take time-averaged observations over at least ten seconds.

8.8.4 Cross-sectional area and velocity field

Description: The gross cross-sectional wetted area of the river, including macrophyte-free and macrophyte-filled areas. Also a cross-sectional grid of water velocities throughout the cross-section.

This metric also yields the data required for estimating wetted widths, hence wetted area.

Units: m^2; $m\ s^{-1}$.

Values/components: Ecosystem health – habitat. Natural form and character. Transport and tauranga waka.

Why? Total cross-sectional area is required in order to estimate the proportion of total cross-sectional area occupied by macrophytes (see *Macrophyte cross-sectional area*), hence the flow blockage factor (see *Macrophyte cross-sectional area* for further justification) (Green 2005). As noted in our conceptualisation (Section 7.3), macrophyte biomass and species composition is affected by low flows. Changes in macrophyte stand characteristics during low flows in turn alters the distribution of water velocities through the water column. Water velocity distributions are a critical determinant of biological processes in spring-fed rivers, such as epiphyton growth and macroinvertebrate production and drift. This metric also supplies wetted width, hence wetted area, as long as councils adhere to the site plan and methods described.

Targets flow modification: Low flows.

Time-scales of response detected: Month. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Sampling intensity high, but based on recommendations in Green (2005), noting the very high spatial heterogeneity and the need to obtain a relatively large sample size to obtain sufficient accuracy.

- Two *sampling years* every five years[30].
- A total of 7 (sampling points) x 9 (transects) = 63 paired observations for each reach (given three reaches: 63 x 3 = 189 paired observations) each sampling year (Figure 8-2).
- Given two sampling years per reporting cycle: 189 x 2 = 378 paired observations per hub site, every five-year period (this equates to ca. 66 observations per year, per hub site).

Method, in brief:

- Sampling to take place during two years within every five-year period. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge cross-sectional area and velocity distributions. Timing of sampling years to be determined by councils.
- A sampling year is defined by three contiguous months: January, February, March. This late summer-autumn period is when instream productivity is at its highest, and when low flows are most likely to occur. It is also during a time most conducive to field work.
- Three transects from each of three reaches will be sampled per month (Jan, Feb, Mar) of the sampling year. E.g., during January transects TX.1, TX.4 and TX.7 would be sampled from Reach X=1, X=2, X=3; and during February transects TX.2, TX.5 and TX.8 would be sampled from Reach X=1, X=2, X=3, and so on (Figure 8-2). Sampling this way across three months helps us to obtain discharge-metric relationships across a range of discharges (assuming discharge varies across the three sampling months in a sampling year). Sampling across three months within a year also helps us to obtain data representative of the low flow period as a whole, not just a narrow period within the late summer-autumn period.
- During each month (Jan, Feb, Mar) wetted width of the transect (m) should be measured. On the same day, paired depth-velocity distributions will be sampled from seven sampling points across the transect (Figure 8-2). The positioning of sampling points laterally, across the transect, should be random to ensure depth-velocity observations are representative of the transect.
- At each sampling point, depth measured with a ruler to nearest cm.
- At each sampling point velocity should be measured in the middle of the water column and ca. 25% of column depth above the bottom and below the surface (three velocity readings per point). Velocimeter should take time-averaged observations over at least ten seconds. Velocity to be taken at a point in the water column where, as much as practicable, the meter is not in physical contact with the macrophyte.
- Paired depth-velocity observations must be taken at the same time.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID code, sampling-point ID, stream width (m), depth (cm), velocity (m s^{-1}).

[30] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

Rows correspond to individual measurements of depth and velocity at sampling points along transects.

8.8.5 Macrophyte cross-sectional area

Description: The species-specific macrophyte composition and coverage of the river cross-section (hence reach).

Units: m^2; although data recorded as macrophyte presence-absence within a cross-sectional grid.

Values/components: Ecosystem health – aquatic life.

Why? As noted in our conceptualisation (Section 7.3), macrophytes play a dominant role in the ecological and biogeochemical processes of spring-fed rivers. Macrophyte species composition and coverage is affected by low flows. In turn, macrophyte species composition can affect abundance and composition of macroinvertebrates, among other ecological metrics. Macrophyte blockage may increase during low flows and is a concern to river managers, as it may detrimentally affect other instream values, such as fish movement and feeding, and macroinvertebrate drift. We may use these data to build functional relationships between discharge, macrophyte composition and blockage. In turn, when linked with the metric *macroinvertebrate drift density*, we may improve our understanding of how discharge in spring-fed rivers affects macroinvertebrate drift via effects on macrophytes. Cross-sectional measurements of macrophyte coverage along numerous transects—as specified below—provide estimates of blockage superior to those provided by % coverage of water surface by macrophytes and macrophyte biomass (Green 2005).

Targets flow modification: Low flows.

Time-scales of response detected: Month. Years. Decades.

Priority at hub sites: Essential.

Sampling frequency/intensity:

- Sampling intensity high, but based on recommendations in Green (2005), noting the very high spatial heterogeneity and the need to obtain a relatively large sample size to obtain sufficient accuracy.
- Two *sampling years* every five years[31].
- A total of 7 (sampling points) x 3 (observations per point through the column) x 9 (transects) = 189 observations of macrophyte presence-absence for each reach (given three reaches: 189 x 3 = 567 paired observations) each sampling year (Figure 8-2).
- Given two sampling years per reporting cycle: 567 x 2 = 1134 observations per hub site, every five-year period (this equates to only 54 observations per year, per hub site).

[31] Noting that, under the NPSFM, reporting is carried out on a five-year cycle.

Method, in brief:

- Sampling to take place during two years within every five-year period. Annual sampling not necessary as the aim is to build functions defining the relationship between discharge cross-sectional area and velocity distributions.
- Method largely follows that outlined in Section 8.8.4 for *cross-sectional area and velocity field*, with the following addendum:
  - Wherever velocity is measured when measuring *cross-sectional area and velocity field*, also record the presence of any macrophyte species. This will enable us to obtain a crude approximation of % cross-sectional coverage of macrophytes.
- *Macrophyte cross-sectional area* should be measured while measuring *cross-sectional area and velocity field*.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID code, sampling-point ID, stream width (m), depth (cm), velocity (m s^{-1}) and macrophyte taxa. Rows correspond to individual measurements of depth, velocity and macrophyte presence (1) or absence (0) at sampling points along transects.

8.8.6 Macroinvertebrate drift density

Description: Taxon-specific density of drifting macroinvertebrates.

Units: individuals m^3 (for each taxon).

Values/components: Ecosystem health – Aquatic life.

Why? See justification for the same metric proposed for runoff-fed rivers (Section 8.5.8). In addition: macroinvertebrate production can be significantly higher in spring-fed rivers than in runoff-fed rivers. Sampling and processing macroinvertebrates attached to macrophytes can be very laborious. By contrast, drift samples may be obtained in open patches of spring-fed rivers that yield relatively 'clean' macroinvertebrate samples that are less costly to process.

Targets flow modification: Low flows.

Time-scales of response detected: Months. Years. Decades.

Priority at hub sites: High.

Sampling frequency/intensity:

- Two *sampling years* every five years.
- A total of 3 sampling points for each reach, each sampling year. Given three reaches: 3 x 3 = 9 observations per hub site, each sampling year (Figure 8-1).
- Given two sampling years per reporting cycle: 9 x 2 = 18 individual drift samples per hub site, every five-year period.

Method, in brief:

- As described for runoff-fed rivers, but net positioned in open patches of the reach, where flow is concentrated with minimal interference by macrophytes.
- Sampling to take place within two days either side (before or after) discharge spot measurement, noting we are unlikely to have gauged discharge data.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, transect ID code, unique sample code and macroinvertebrate taxon. Rows correspond to individual drift samples.

8.8.7 Dissolved oxygen dynamics

Description: Diel dynamics of dissolved oxygen concentrations during low flow periods. See NPSFM Appendix 2A table 7.

Units: mg O_2 L^{-1}.

Values/components: Ecosystem health – Water quality.

Why? As described in Section 7.3.1, macrophytes play a dominant role in the ecosystem processes of spring-fed rivers. Macrophytes play a particularly dominant role in the dissolved oxygen dynamics of spring-fed rivers. Excess macrophyte growth can cause hyperoxia during the day and hypoxia at night. Hypoxia is the stressor of greatest concern (McArley et al. 2021), having lethal effects on many instream organisms of high value (Diaz and Breitburg 2009).

Targets flow modification: Low flows.

Time-scales of response detected: Hours. Days. Weeks. Months. Years. Decades.

Priority at hub sites: High.

Sampling frequency/intensity:

- Two *sampling years* every five years.
- Automated logging over three ten-day periods, coinciding the sampling of other metrics above.

Method, in brief:

- Following NPSFM protocol, if available.
- The objective is to obtain continuous automated measurements over 10-day periods following measurement of the other metrics above (notably *discharge spot measurement*, *cross-sectional area and velocity field* and *macrophyte cross-sectional area*), such that we develop a quantitative understanding of how discharge affects DO dynamics through macrophyte-mediated mechanisms.
- Accurate logging of DO dynamics in rivers is non-trivial and it is common to see poor quality (often unusable) DO time series due to the logger not being properly calibrated or poorly maintained. Pilot trials should be carried out before implementation to refine place- and logger-specific protocols.

Minimum data, in brief: Columns specifying unique site name and ID code, site latitude, site longitude, date, reach ID, transect ID code, unique sample code and DO concentration. Rows correspond to individual DO concentration measurements.

8.9 Metrics for spring-fed rivers – spoke sites

At this stage we are not recommending spoke sites for spring-fed rivers, due to (a) the difficulty of discharge gauging in spring-fed, macrophyte-dominated rivers, and (b) the resource-intensive nature of obtaining useful data from macrophyte-dominated rivers (see metric descriptions for spring-fed hub sites).

9 Where to from here? Implementing this plan

We have presented a plan for M&E to support the adaptive management of river flows under the NPSFM 2020. Irrespective of the requirements of the NPSFM, the plan we have presented 'stands on its own' to support credible, relevant and legitimate adaptive management of river flows throughout New Zealand.

This M&E plan comprises a significant departure from the much-criticised surveillance approach to M&E (Section 1.3.1). It has been carefully developed to maximise credibility, relevance and legitimacy, but further work is required before it can be implemented:

1. **A per-metric cost-benefit analysis needs to be conducted.** This analysis would serve as a foundation for making decisions about which components of this plan should be implemented under various levels of resource constraint.

2. **The metrics and methods need to be refined following careful consideration by council hub representatives and feedback from those representatives to the science team**. Refinement of metrics and methods is an iterative process and this plan represents the first step of that process.

3. **If this M&E plan is implemented, the science team and council hub representatives should scope the evaluation- and decision-support tools to be developed**. This plan has been designed to facilitate sharing among councils of new knowledge, data, and the evaluation- and decision-support tools that can be more rapidly developed when M&E is nationally-consistent. If this plan is implemented an important next step will be to workshop with councils which support tools are developed and how they are developed.

4. **Funding for the plan's implementation must be obtained, following refinement in light of the cost-benefit analysis**. It is currently not clear how this M&E plan would be funded. The plan itself, we hope, serves as a strong foundation for obtaining the required funding.

5. **Site selection should be coordinated across the council core representatives and the science team**. We have only presented broad guidelines for site selection in Section 6 of this plan. To ensure sites are selected to meet M&E objectives, site selection needs to be workshopped among the science team and hub representatives.

6. **Data management and storage standards need to be developed.** The advantages that result from a nationally-consistent M&E program (Section 4) cannot be met if data are not managed and stored carefully and consistently. We have already provided a solid foundation for best practice in data management and storage by describing broad data requirements in Section 8, but a data standards manual would need to be developed.

10 Acknowledgements

Development of this M&E plan was funded by an Envirolink Tools grant to RJS. We are greatly indebted to Susie Osbaldiston (Northland Regional Council), our council champion, for her support and patience during the development of this plan. We sincerely thank our council core representatives for their contributions to this M&E plan: Jason Augspurger (Otago Regional Council), Manas Chakraborty (Northland Regional Council), Dan Fake (Hawke's Bay Regional Council), Paul Fisher (Nelson City Council), Roger Hodson (Environment Southland), Adrian Meredith and Duncan Gray (Environment Canterbury), Paul Murphy (Gisborne District Council), Regan Phipps and Darin Sutherland (Taranaki Regional Council), Kolt Johnson (Auckland Council), Mike Thompson (Greater Wellington Regional Council), and Thomas Wilding (Waikato Regional Council).

11 Glossary of abbreviations and terms

Adaptive management	A coupling of management, monitoring, evaluation and forecasting that facilitates reducing uncertainty about how management and policy decisions affect environmental outcomes. Inherent to adaptive management is a philosophy of "learning by doing." Adaptive management can be effective when we wish to manage a system under high levels of uncertainty about how the system will respond.
Environmental outcome	As per NPSFM: in relation to a value that applies to an FMU or part of an FMU, a desired outcome that a regional council identifies and then includes as an objective in its regional plan(s) (see NPSFM clause 3.9).
Epistemic uncertainty	Refers to our imperfect knowledge of the basic cause-effect pathways that drive system state and dynamics
Flow regime	A quantifiable representation of the main characteristics of a time series of discharges, calculated over a period spanning many years. The flow regime may represent variability at several temporal resolutions.
Freshwater Management Unit (FMU)	As per NPSFM: All or any part of a water body or water bodies, and their related catchments, that a regional council determines under (NPSFM) clause 3.8 is an appropriate unit for freshwater management and accounting purposes.
Hydrological stressors	Those properties of the flow regime that negatively affect riverine values. Examples include prolonged low flows that may cause mortality of fishes, and reduced frequency of floods that flush fine sediment from the benthic habitats of rivers.
NPSFM	National Policy Statement for Freshwater Management 2020
Ontological uncertainty	Statistical uncertainties arising from biased or imprecise data, as well as the inherent randomness of the processes we study
Status quo trap	The situation where the decisions we make are biased towards maintenance of the status quo.
Sunk-cost trap	Occurs when we make choices that justify past, flawed choices
Surveillance monitoring	Surveillance monitoring involves monitoring indicators that are broadly suggestive of ecosystem health, with the objective of estimating state and trends in ecosystem health (e.g., some of New Zealand's State of the Environment monitoring). Surveillance monitoring tells us when ecosystem health may be declining, but is not designed to tell us about the causes of such decline, nor is it designed to facilitate anticipation of future changes in the health of ecosystems.

Take limit	Following NPSFM: a limit on the amount of water that can be taken from an FMU or part of an FMU, as set under clause 3.17.
Tikanga	Māori laws and values.

12 References

Acreman, M. C., I. C. Overton, J. King, P. J. Wood, I. G. Cowx, M. J. Dunbar, E. Kendy, and W. J. Young. 2014. The changing role of ecohydrological science in guiding environmental flows. Hydrological Sciences Journal **59**:433-450.

Adams, V. M., M. Barnes, and R. L. Pressey. 2019. Shortfalls in Conservation Evidence: Moving from Ecological Effects of Interventions to Policy Evaluation. One Earth **1**:62-75.

Beard, G. R., W. A. Scott, and J. K. Adamson. 1999. The Value of Consistent Methodology in Long-term Environmental Monitoring. Environmental Monitoring and Assessment **54**:239-258.

Bennett, N. J., R. Roth, S. C. Klain, K. Chan, P. Christie, D. A. Clark, G. Cullman, D. Curran, T. J. Durbin, G. Epstein, A. Greenberg, M. P. Nelson, J. Sandlos, R. Stedman, T. L. Teel, R. Thomas, D. Veríssimo, and C. Wyborn. 2017. Conservation social science: Understanding and integrating human dimensions to improve conservation. Biological Conservation **205**:93-108.

Biggs, B. J. F. 1996. Hydraulic habitat of plants in streams. Regulated Rivers: Research & Management **12**:131-144.

Biggs, B. J. F., and C. W. Hickey. 1994. Periphyton responses to a hydraulic gradient in a regulated river in New Zealand. Freshwater Biology **32**:49-59.

Biggs, B. J. F., R. A. Smith, and M. J. Duncan. 1999. Velocity and Sediment Disturbance of Periphyton in Headwater Streams: Biomass and Metabolism. Journal of the North American Benthological Society **18**:222-241.

Biggs, B. J. F., and S. Stockseth. 1996. Hydraulic habitat suitability for periphyton in rivers. Regulated Rivers: Research & Management **12**:251-261.

Blaustein, A. R., and J. M. Kiesecker. 2002. Complexity in conservation: lessons from the global decline of amphibian populations. Ecology Letters **5**:597-608.

Boffa Miskell. 2018. Natural Character, Riverscape and Visual Amenity Assessments. Clutha/Mata-Au Water Quantity Plan Change - Stage 1. Technical Report Prepared for Otago Regional Council.

Booker, D. J. 2013. Spatial and temporal patterns in the frequency of events exceeding three times the median flow (FRE3) across New Zealand. Journal of Hydrology (New Zealand) **52**:15-39.

Booker, D. J., and A. L. Whitehead. 2022. River water temperatures are higher during lower flows after accounting for meteorological variability. River Research and Applications **38**:3-22.

Boulton, A. J. 2003. Parallels and contrasts in the effects of drought on stream macroinvertebrate assemblages. Freshwater Biology **48**:1173-1185.

Boulton, A. J., S. Findlay, P. Marmonier, E. H. Stanley, and H. M. Valett. 1998. The Functional Significance of the Hyporheic Zone in Streams and Rivers. Annual Review of Ecology and Systematics **29**:59-81.

Boulton, A. J., and P. S. Lake. 1992. Benthic organic matter and detritivorous macroinvertebrates in two intermittent streams in south-eastern Australia. Hydrobiologia **241**:107-118.

Box, G. E. P., and G. C. Tiao. 1965. A Change in Level of a Non-Stationary Time Series. Biometrika **52**:181-192.

Box, G. E. P., and G. C. Tiao. 1975. Intervention Analysis with Applications to Economic and Environmental Problems. Journal of the American Statistical Association **70**:70-79.

Bricker, O. P., and M. A. Ruggiero. 1998. Toward a national program for monitoring environmental resources. Ecological Applications **8**:326-329.

Brierley, G., M. Tadaki, D. Hikuroa, B. Blue, C. Šunde, J. Tunnicliffe, and A. Salmond. 2019. A geomorphic perspective on the rights of the river in Aotearoa New Zealand. River Research and Applications **35**:1640-1651.

Brooks, A. J., B. C. Chessman, and T. Haeusler. 2011. Macroinvertebrate traits distinguish unregulated rivers subject to water abstraction. Journal of the North American Benthological Society **30**:419-435.

Brown, T. C., J. G. Taylor, and B. Shelby. 1991. Assessing the direct effects of streamflow on recreation: A literature review. JAWRA Journal of the American Water Resources Association **27**:979-989.

Bunn, S. E., and A. H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management **30**:492-507.

Bunn, S. E., M. C. Thoms, S. K. Hamilton, and S. J. Capon. 2006. Flow variability in dryland rivers: Boom, bust and the bits in between. River Research and Applications **22**:179-186.

Burgman, M., R. Chiaravalloti, F. Fidler, Y. Huan, M. McBride, A. Marcoci, J. Norman, A. Vercammen, B. Wintle, and Y. Yu. 2022. A toolkit for open and pluralistic conservation science. Conservation Letters **n/a**:e12919.

Caissie, D. 2006. The thermal regime of rivers: a review. Freshwater Biology **51**:1389-1406.

Cash, D. W., W. C. Clark, F. Alcock, N. M. Dickson, N. Eckley, D. H. Guston, J. Jäger, and R. B. Mitchell. 2003. Knowledge systems for sustainable development. Proceedings of the National Academy of Sciences **100**:8086-8091.

Chaffin, B. C., H. Gosnell, and B. A. Cosens. 2014. A decade of adaptive governance scholarship: synthesis and future directions. Ecology and Society **19**.

Champion, P. D., and C. C. Tanner. 2000. Seasonality of macrophytes and interaction with flow in a New Zealand lowland stream. Hydrobiologia **441**:1-12.

Chess, C., B. B. Johnson, and G. Gibson. 2005. Communicating about environmental indicators. Journal of Risk Research **8**:63-75.

Chessman, B. C., K. A. Fryirs, and G. J. Brierley. 2006. Linking geomorphic character, behaviour and condition to fluvial biodiversity: implications for river management. Aquatic Conservation: Marine and Freshwater Ecosystems **16**:267-288.

Cocklin, C., I. Fraser, and M. Harte. 1994. The Recreational Value of In-Stream Flows. New Zealand Geographer **50**:20-29.

Conroy, M. J., and J. T. Peterson. 2013. Decision Making in Natural Resource Management: A Structured, Adaptive Approach. John Wiley & Sons, Ltd.

Cowx, I. G., W. O. Young, and J. M. Hellawell. 1984. The influence of drought on the fish and invertebrate populations of an upland stream in Wales. Freshwater Biology **14**:165-177.

Cundill, G., and C. Fabricius. 2009. Monitoring in adaptive co-management: Toward a learning based approach. Journal of Environmental Management **90**:3205-3211.

Danielsen, F., N. D. Burgess, A. Balmford, P. F. Donald, M. Funder, J. P. G. Jones, P. Alviola, D. S. Balete, T. Blomley, J. Brashares, B. Child, M. Enghoff, J. Fjeldså, S. Holt, H. Hübertz, A. E. Jensen, P. M. Jensen, J. Massao, M. M. Mendoza, Y. Ngaga, M. K. Poulsen, R. Rueda, M. Sam, T. Skielboe, G. Stuart-Hill, E. Topp-Jørgensen, and D. Yonten. 2009. Local Participation in Natural Resource Monitoring: a Characterization of Approaches. Conservation Biology **23**:31-42.

Davey, A. J. H., D. J. Kelly, and B. J. F. Biggs. 2006. Refuge-use strategies of stream fishes in response to extreme low flows. Journal of Fish Biology **69**:1047-1059.

David, B. O., and G. P. Closs. 2002. Behavior of a stream-dwelling fish before, during, and after high-discharge events. Transactions of the American Fisheries Society **131**:762-771.

Dewson, Z. S., A. B. W. James, and R. G. Death. 2007a. Invertebrate community responses to experimentally reduced discharge in small streams of different water quality. Journal of the North American Benthological Society **26**:754-766.

Dewson, Z. S., A. B. W. James, and R. G. Death. 2007b. A review of the consequences of decreased flow for instream habitat and macroinvertebrates. Journal of the North American Benthological Society **26**:401-415.

Diaz, R. J., and D. L. Breitburg. 2009. Chapter 1 The Hypoxic Environment. Pages 1-23 *in* J. G. Richards, A. P. Farrell, and C. J. Brauner, editors. Fish Physiology. Academic Press.

Failing, L., and R. Gregory. 2003. Ten common mistakes in designing biodiversity indicators for forest policy. Journal of Environmental Management **68**:121-132.

Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes: A Continuous View of the River is Needed to Understand How Processes Interacting among Scales Set the Context for Stream Fishes and Their Habitat. Bioscience **52**:483-498.

Ferraro, P. J., and M. M. Hanauer. 2014. Advances in Measuring the Environmental and Social Impacts of Environmental Programs. Annual Review of Environment and Resources **39**:495-517.

Ferraro, P. J., and S. K. Pattanayak. 2006. Money for Nothing? A Call for Empirical Evaluation of Biodiversity Conservation Investments. PLOS Biology **4**:e105.

Franklin, P., M. Dunbar, and P. Whitehead. 2008. Flow controls on lowland river macrophytes: A review. Science of the Total Environment **400**:369-378.

Fraser, L. H., H. A. L. Henry, C. N. Carlyle, S. R. White, C. Beierkuhnlein, J. F. Cahill, B. B. Casper, E. Cleland, S. L. Collins, J. S. Dukes, A. K. Knapp, E. Lind, R. J. Long, Y. Q. Luo, P. B. Reich, M. D. Smith, M. Sternberg, and R. Turkington. 2013. Coordinated distributed experiments: an emerging tool for testing global hypotheses in ecology and environmental science. Frontiers in Ecology and the Environment **11**:147-155.

Friberg, N., N. Bonada, D. C. Bradley, M. J. Dunbar, F. K. Edwards, J. Grey, R. B. Hayes, A. G. Hildrew, N. Lamouroux, M. Trimmer, and G. Woodward. 2011. Biomonitoring of Human Impacts in Freshwater Ecosystems: The Good, the Bad and the Ugly. Pages 1-68 *in* G. Woodward, editor. Advances in Ecological Research. Academic Press.

Funtowicz, S. O., and J. R. Ravetz. 1993. Science for the post-normal age. Futures **25**:739-755.

Game, E. T., E. Meijaard, D. Sheil, and E. McDonald-Madden. 2014. Conservation in a Wicked Complex World; Challenges and Solutions. Conservation Letters **7**:271-277.

Gelman, A., and J. Hill. 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, America.

Graf, W. L. 2006. Downstream hydrologic and geomorphic effects of large dams on American rivers. Geomorphology **79**:336-360.

Grams, P. E., E. Head, and E. R. Mueller. 2022. Effects of flow regulation and drought on geomorphology and floodplain habitat along the Colorado River in Canyonlands National Park, Utah. River Research and Applications **38**:1266-1276.

Green, J. C. 2005. Comparison of blockage factors in modelling the resistance of channels containing submerged macrophytes. River Research and Applications **21**:671-686.

Gregg, W. W., and F. L. Rose. 1985. Influences of aquatic macrophytes on invertebrate community structure, guild structure, and microdistribution in streams. Hydrobiologia **128**:45-56.

Gregory, R., L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson. 2012. Structured Decision Making: A Practical Guide to Environmental Management Choices. John Wiley & Sons, West Sussex, UK.

Gregory, R., D. Ohlson, and J. Arvai. 2006. Deconstructing adaptive management: criteria for applications to enviromental management. Ecological Applications **16**:2411-2425.

Gregory, R. S., and R. L. Keeney. 2002. Making smarter environmental management decisions. JAWRA Journal of the American Water Resources Association **38**:1601-1612.

Griffiths, R. E., D. E. Anderson, and A. E. Springer. 2008. The morphology and hydrology of small spring-dominated channels. Geomorphology **102**:511-521.

Hakala, J. P., and K. J. Hartman. 2004. Drought effect on stream morphology and brook trout (Salvelinus fontinalis) populations in forested headwater streams. Hydrobiologia **515**:203-213.

Hamilton, S. K. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. Freshwater Biology **57**:43-57.

Hammond, J. S., R. L. Keeney, and H. Raiffa. 1998. The hidden traps in decision making. Harvard Business Review **76**:47-58.

Harmsworth, G., S. Awatere, and M. Robb. 2016. Indigenous Māori values and perspectives to inform freshwater management in Aotearoa-New Zealand. Ecology and Society **21**.

Haxton, T. J., and C. S. Findlay. 2008. Meta-analysis of the impacts of water management on aquatic communities. Canadian Journal of Fisheries and Aquatic Sciences **65**:437-447.

Hermans, L. M., J. H. Slinger, and S. W. Cunningham. 2013. The use of monitoring information in policy-oriented learning: Insights from two cases in coastal management. Environmental Science & Policy **29**:24-36.

Hoyle, J. T., C. Kilroy, D. M. Hicks, and L. Brown. 2017. The influence of sediment mobility and channel geomorphology on periphyton abundance. Freshwater Biology **62**:258-273.

Humphries, P., H. Keckeis, and B. Finlayson. 2014. The River Wave Concept: Integrating River Ecosystem Models. Bioscience **64**:870-882.

Humphries, P., A. J. King, N. McCasker, R. K. Kopf, R. Stoffels, B. P. Zampatti, and A. E. Price. 2020. Riverscape recruitment: a conceptual synthesis of drivers of fish recruitment in rivers. Canadian Journal of Fisheries and Aquatic Sciences **77**:213-225.

Jakeman, A. J., R. A. Letcher, and J. P. Norton. 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software **21**:602-614.

Joy, M. K., B. O. David, and M. Lake. 2013. New Zealand Freshwater Fish Sampling Protocols. Part 1: Wadeable Rivers and Streams. EOS Ecology.

Junk, W. J., P. B. Bayley, and R. E. Sparks. 1989. The flood pulse concept in river-floodplain systems. Pages 110-127 *in* D. P. Dodge, editor. Proceedings of the International Large River Symposium. Canadian Special Publications in Fisheries and Aquatic Sciences.

Kaenel, B. R., H. Buehrer, and U. Uehlinger. 2000. Effects of aquatic plant management on stream metabolism and oxygen balance in streams. Freshwater Biology **45**:85-95.

Kennard, M. J., B. J. Pusey, J. D. Olden, S. J. Mackay, J. L. Stein, and N. Marsh. 2010. Classification of natural flow regimes in Australia to support environmental flow management. Freshwater Biology **55**:171-193.

King, A. J., S. A. Townsend, M. M. Douglas, and M. J. Kennard. 2015. Implications of water extraction on the low-flow hydrology and ecology of tropical savannah rivers: an appraisal for northern Australia. Freshwater Science **34**:741-758.

King, A. J., K. A. Ward, P. O'Connor, D. Green, Z. Tonkin, and J. Mahoney. 2010. Adaptive management of an environmental watering event to enhance native fish spawning and recruitment. Freshwater Biology **55**:17-31.

Konrad, C. P., J. D. Olden, D. A. Lytle, T. S. Melis, J. C. Schmidt, E. N. Bray, M. C. Freeman, K. B. Gido, N. P. Hemphill, M. J. Kennard, L. E. McMullen, M. C. Mims, M. Pyron, C. T. Robinson, and J. G. Williams. 2011. Large-scale Flow Experiments for Managing River Systems. Bioscience **61**:948-959.

Kuehne, L. M., J. D. Olden, A. L. Strecker, J. J. Lawler, and D. M. Theobald. 2017. Past, present, and future of ecological integrity assessment for fresh waters. Frontiers in Ecology and the Environment **15**:197-205.

Lebel, L., and R. Daniel. 2009. The governance of ecosystem services from tropical upland watersheds. Current Opinion in Environmental Sustainability **1**:61-68.

Levin, S. A., B. Grenfell, A. Hastings, and A. S. Perelson. 1997. Mathematical and Computational Challenges in Population Biology and Ecosystems Science. Science **275**:334-343.

Lindenmayer, D. B., and G. E. Likens. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology & Evolution **24**:482-486.

Lindenmayer, D. B., and G. E. Likens. 2010. The science and application of ecological monitoring. Biological Conservation **143**:1317-1328.

Lindenmayer, D. B., M. P. Piggott, and B. A. Wintle. 2013. Counting the books while the library burns: why conservation monitoring programs need a plan for action. Frontiers in Ecology and the Environment **11**:549-555.

Lovett, G. M., D. A. Burns, C. T. Driscoll, J. C. Jenkins, M. J. Mitchell, L. Rustad, J. B. Shanley, G. E. Likens, and R. Haeuber. 2007. Who needs environmental monitoring? Frontiers in Ecology and the Environment **5**:253-260.

Lusardi, R. A., C. A. Jeffres, and P. B. Moyle. 2018. Stream macrophytes increase invertebrate production and fish habitat utilization in a California stream. River Research and Applications **34**:1003-1012.

Lusardi, R. A., A. L. Nichols, A. D. Willis, C. A. Jeffres, A. H. Kiers, E. E. Van Nieuwenhuyse, and R. A. Dahlgren. 2021. Not All Rivers Are Created Equal: The Importance of Spring-Fed Rivers under a Changing Climate. Water **13**:1652.

Lytle, D. A. 2008. Life-history and behavioural adaptations to flow regime in aquatic insects. Pages 122-138 *in* J. Lancaster and R. A. Briers, editors. Aquatic Insects: Challenges to Populations. Proceedings of the Royal Entomological Society's 24th Symposium. CAB International, Oxfordshire, UK.

Lytle, D. A., and N. L. Poff. 2004. Adaptation to natural flow regimes. Trends in Ecology & Evolution **19**:94-100.

Mace, G. M., and J. E. M. Baillie. 2007. The 2010 Biodiversity Indicators: Challenges for Science and Policy. Conservation Biology **21**:1406-1413.

Magoulick, D. D., and R. M. Kobza. 2003. The role of refugia for fishes during drought: a review and synthesis. Freshwater Biology **48**:1186-1198.

McArley, T. J., E. Sandblom, and N. A. Herbert. 2021. Fish and hyperoxia—From cardiorespiratory and biochemical adjustments to aquaculture and ecophysiology implications. Fish and Fisheries **22**:324-355.

McCann, K. S., J. B. Rasmussen, and J. Umbanhowar. 2005. The dynamics of spatially coupled food webs. Ecology Letters **8**:513-523.

McElreath, R. 2016. Statstical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press, Boca Raton, Florida.

Miller, S. W., D. Wooster, and J. Li. 2007. Resistance and resilience of macroinvertebrates to irrigation water withdrawals. Freshwater Biology **52**:2494-2510.

Ministry for the Environment. 2019. Action for healthy waterways – A discussion document on national direction for our essential freshwater. Wellington: Ministry for the Environment.

Mosley, M. P. 1983. Flow requirements for recreation and wildlife in New Zealand rivers: A review. Journal of Hydrology (New Zealand) **22**:152-174.

Naiman, R. J. 2013. Socio-ecological complexity and the restoration of river ecosystems. Inland Waters **3**:391-410.

Naiman, R. J., S. E. Bunn, C. Nilsson, G. E. Petts, G. Pinay, and L. C. Thompson. 2002. Legitimizing fluvial ecosystems as users of water: An overview. Environmental Management **30**:455-467.

Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. Trends in Ecology & Evolution **21**:668-673.

NPSFM. 2020. National Policy Statement for Freshwater Management 2020. New Zealand Ministry for the Environment.

Olden, J. D., M. J. Kennard, and B. J. Pusey. 2012. A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. Ecohydrology **5**:503-518.

Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Yyoung. 1999. Statistical Issues for Monitoring Ecological and Natural Resources in the United States. Environmental Monitoring and Assessment **54**:1-45.

Parr, T. W., M. Ferretti, I. C. Simpson, M. Forsius, and E. Kovács-Láng. 2002. Towards A Long-Term Integrated Monitoring Programme In Europe: Network Design in Theory and Practice. Environmental Monitoring and Assessment **78**:253-290.

PCE. 2019. Focusing Aotearoa New Zealand's environmental reporting system. Report of the Parliamentary Commisioner for the Environment. Wellington, New Zealand.

Peters, D. P. C., H. W. Loescher, M. D. SanClements, and K. M. Havstad. 2014. Taking the pulse of a continent: expanding site-based research infrastructure for regional- to continental-scale ecology. Ecosphere **5**:23.

Pflüger, Y., A. Rackham, and S. Larned. 2010. The aesthetic value of river flows: An assessment of flow preferences for large and small rivers. Landscape and Urban Planning **95**:68-78.

Poff, L. N., and J. K. H. Zimmerman. 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. Freshwater Biology **55**:194-205.

Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The natural flow regime. Bioscience **47**:769-784.

Poff, N. L., J. D. Allan, M. A. Palmer, D. D. Hart, B. D. Richter, A. H. Arthington, K. H. Rogers, J. L. Meyers, and J. A. Stanford. 2003. River flows and water wars: emerging science for environmental decision making. Frontiers in Ecology and the Environment **1**:298-306.

Polasky, S., S. R. Carpenter, C. Folke, and B. Keeler. 2011. Decision-making under great uncertainty: environmental management in an era of global change. Trends in Ecology & Evolution **26**:398-404.

Power, M. E., A. Sun, G. Parker, W. E. Dietrich, and J. T. Wootton. 1995. Hydraulic Food-Chain Models: An approach to the study of food-web dynamics in large rivers. Bioscience **45**:159-167.

Puckridge, J. T., F. Sheldon, K. F. Walker, and A. J. Boulton. 1998. Flow variability and the ecology of large rivers. Marine and Freshwater Research **49**:55-72.

Quinn, J. M., and C. W. Hickey. 1990. Magnitude of effects of substrate particle size, recent flooding, and catchment development on benthic invertebrates in 88 New Zealand rivers. New Zealand Journal of Marine and Freshwater Research **24**:411-427.

RCSAG. 2020. Research for Resource Management: Regional Council Research, Science and Technology Strategy 2020. New Zealand.

Riis, T., and B. J. F. Biggs. 2003. Hydrologic and hydraulic control of macrophyte establishment and performance in streams. Limnology and Oceanography **48**:1488-1497.

Riis, T., A. M. Suren, B. Clausen, and K. Sand-Jensen. 2008. Vegetation and flow regime in lowland streams. Freshwater Biology **53**:1531-1543.

Robson, B. J. 2014. When do aquatic systems models provide useful predictions, what is changing, and what is next? Environmental Modelling & Software **61**:287-296.

Robson, B. J., D. P. Hamilton, I. T. Webster, and T. Chan. 2008. Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries. Environmental Modelling & Software **23**:369-384.

Rolls, R. J., C. Leigh, and F. Sheldon. 2012. Mechanistic effects of low-flow hydrology on riverine ecosystems: ecological principles and consequences of alteration. Freshwater Science **31**:1163-1186.

Ruan, Q., F. Wang, and W. Cao. 2021. Conflicts in Implementing Environmental Flows for Small-Scale Hydropower Projects and Their Potential Solutions—A Case from Fujian Province, China. Water **13**:2461.

Ryan, A., M. J. Colloff, and J. Pittock. 2021. Flow to nowhere: the disconnect between environmental watering and the conservation of threatened species in the Murray–Darling Basin, Australia. Marine and Freshwater Research **72**:1408-1429.

Schiller, A., C. T. Hunsaker, M. A. Kane, A. K. Wolfe, V. H. Dale, G. W. Suter, C. S. Russell, G. Pion, M. H. Jensen, and V. C. Konar. 2001. Communicating Ecological Indicators to Decision Makers and the Public. Conservation Ecology **5**.

Seeteram, N. A., V. Engel, and P. Mozumder. 2018. Implications of a valuation study for ecological and social indicators associated with Everglades restoration. Science of the Total Environment **627**:792-801.

Shearer, K. A., J. W. Hayes, I. G. Jowett, and D. A. Olsen. 2015. Habitat suitability curves for benthic macroinvertebrates from a small New Zealand river. New Zealand Journal of Marine and Freshwater Research **49**:178-191.

Shenton, W., N. R. Bond, J. D. L. Yen, and R. Mac Nally. 2012. Putting the "Ecology" into Environmental Flows: Ecological Dynamics and Demographic Modelling. Environmental Management **50**:1-10.

Smakhtin, V. U. 2001. Low flow hydrology: a review. Journal of Hydrology **240**:147-186.

Snelder, T. H., and B. J. F. Biggs. 2002. Multiscale River Environment Classification for water resources management. Journal of the American Water Resources Association **38**:1225-1239.

Snelder, T. H., and D. J. Booker. 2013. Natural flow regime classifications are sensitive to definition procedures. River Research and Applications **29**:822-838.

Sotiropoulos, J. C., K. H. Nislow, and M. R. Ross. 2006. Brook trout, Salvelinus fontinalis, microhabitat selection and diet under low summer stream flows. Fisheries Management and Ecology **13**:149-155.

Stainback, G. A., J. H. Lai, E. F. Pienaar, D. C. Adam, R. Wiederholt, and C. Vorseth. 2020. Public preferences for ecological indicators used in Everglades restoration. Plos One **15**:e0234051.

Stewardson, M. J., N. Bond, J. Brookes, S. Capon, F. Dyer, M. Grace, P. Frazier, B. Hart, A. Horne, A. King, M. Langton, R. Nathan, I. Rutherfurd, F. Sheldon, R. Thompson, R. Vertessy, G. Walker, Q. J. Wang, S. Wassens, R. J. Watts, J. A. Webb, and A. W. Western. 2021. The politicisation of science in the Murray-Darling Basin, Australia: discussion of 'Scientific integrity, public policy and water governance'. Australasian Journal of Water Resources **25**:141-158.

Stewart-Oaten, A. 2008. Chance and randomness in design versus model-based approaches to impact assessment: comments on Bulleri et al. (2007). Environmental Conservation **35**:8-10.

Stewart-Oaten, A., and J. R. Bence. 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs **71**:305-339.

Stewart-Oaten, A., W. W. Murdoch, and K. R. Parker. 1986. Environmental Impact Assessment: 'Pseudoreplication' in Time? Ecology **67**:929-940.

Stirling, A., and M. A. Burgman. 2021. Strengthening conservation science as a crisis discipline by addressing challenges of precaution, privilege, and individualism. Conservation Biology **n/a**.

Stoffels, R. J., N. R. Bond, and S. Nicol. 2018. Science to support the management of riverine flows. Freshwater Biology **63**:996-1010.

Stoffels, R. J., P. Humphries, N. R. Bond, and A. E. Price. 2022. Fragmentation of lateral connectivity and fish population dynamics in large rivers. Fish and Fisheries **23**:680-696.

Suren, A. M., B. J. F. Biggs, M. J. Duncan, L. Bergey, and P. Lambert. 2003a. Benthic community dynamics during summer low-flows in two rivers of contrasting enrichment 2. Invertebrates. New Zealand Journal of Marine and Freshwater Research **37**:71-83.

Suren, A. M., B. J. F. Biggs, C. Kilroy, and L. Bergey. 2003b. Benthic community dynamics during summer low-flows in two rivers of contrasting enrichment 1. Periphyton. New Zealand Journal of Marine and Freshwater Research **37**:53-70.

Tadaki, M., G. Brierley, and C. Cullum. 2014. River classification: theory, practice, politics. WIREs Water **1**:349-367.

Te Aho, L. 2010. Indigenous challenges to enhance freshwater governance and management in Aotearoa New Zealand - the Waikato river settlement. The Journal of Water Law **20**:285-292.

Te Aho, L. 2019. Te Mana o te Wai: An indigenous perspective on rivers and river management. River Research and Applications **35**:1615-1621.

Thompson, R. M., A. J. King, R. M. Kingsford, R. Mac Nally, and N. L. Poff. 2018. Legacies, lags and long-term trends: Effective flow restoration in a changed and changing world. Freshwater Biology **63**:986-995.

Tonkin, J. D., D. M. Merritt, J. D. Olden, L. V. Reynolds, and D. A. Lytle. 2018. Flow regime alteration degrades ecological networks in riparian ecosystems. Nature Ecology & Evolution **2**:86-93.

Turnhout, E., M. Hisschemöller, and H. Eijsackers. 2007. Ecological indicators: Between the two fires of science and policy. Ecological Indicators **7**:215-228.

Underwood, A. J. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. Journal of Experimental Marine Biology and Ecology **161**:145-178.

Verbrugge, L. N. H., W. Ganzevoort, J. M. Fliervoet, K. Panten, and R. J. G. van den Born. 2017. Implementing participatory monitoring in river management: The role of stakeholders' perspectives and incentives. Journal of Environmental Management **195**:62-69.

Vietz, G. J., and B. L. Finlayson. 2017. Chapter 5 - Geomorphological Effects of Flow Alteration on Rivers. Pages 83-100 *in* A. C. Horne, J. A. Webb, M. J. Stewardson, B. Richter, and M. Acreman, editors. Water for the Environment. Academic Press.

Vietz, G. J., M. J. Sammonds, and M. J. Stewardson. 2013. Impacts of flow regulation on slackwaters in river channels. Water Resources Research **49**:1797-1811.

Walker, K. F., F. Sheldon, and J. T. Puckridge. 1995. A perspective on dryland river ecosystems. Regulated Rivers: Research & Management **11**:85-104.

Walters, A. W., and D. M. Post. 2011. How low can you go? Impacts of a low-flow disturbance on aquatic insect communities. Ecological Applications **21**:163-174.

Walters, C. 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology **[online] 1**:1.

Waylen, K. A., and K. L. Blackstock. 2017. Monitoring for Adaptive Management or Modernity: Lessons from recent initiatives for holistic environmental management. Environmental Policy and Governance **27**:311-324.

Waylen, K. A., K. L. Blackstock, F. J. van Hulst, C. Damian, F. Horváth, R. K. Johnson, R. Kanka, M. Külvik, C. J. A. Macleod, K. Meissner, M. M. Oprina-Pavelescu, J. Pino, E. Primmer, G. Rîșnoveanu, B. Šatalová, J. Silander, J. Špulerová, M. Suškevičs, and J. Van Uytvanck. 2019. Policy-driven monitoring and evaluation: Does it support adaptive management of socio-ecological systems? Science of the Total Environment **662**:373-384.

Webb, J. A., S. C. de Little, K. A. Miller, and M. J. Stewardson. 2018. Quantifying and predicting the benefits of environmental flows: Combining large-scale monitoring data and expert knowledge within hierarchical Bayesian models. Freshwater Biology **63**:831-843.

Westgate, M. J., G. E. Likens, and D. B. Lindenmayer. 2013. Adaptive management of biological systems: A review. Biological Conservation **158**:128-139.

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems. Third edition. Academic Press, San Diego, California.

Whiting, P. J., and D. B. Moog. 2001. The geometric, sedimentologic and hydrologic attributes of spring-dominated channels in volcanic areas. Geomorphology **39**:131-149.

Whiting, P. J., and J. Stamm. 1995. The hydrology and form of spring-dominated channels. Geomorphology **12**:233-240.

Willis, A. D., A. L. Nichols, E. J. Holmes, C. A. Jeffres, A. C. Fowler, C. A. Babcock, and M. L. Deas. 2017. Seasonal aquatic macrophytes reduce water temperatures via a riverine canopy in a spring-fed stream. Freshwater Science **36**:508-522.

Winemiller, K. O. 2004. Floodplain River Food webs: Generalisations and Implications for Fisheries Management. Pages 285-309 *in* W. R. and T. Petr, editors. Proceedings of the Second International Symposium on the Management of Large Rivers for Fisheries. FAO Regional Office for Asia and the Pacific, RAP Publication 2004/16, Bangkok, Thailand.

Appendix A Participants in the design of this M&E plan

This project is a collaboration among NIWA, Cawthron and numerous regional councils and unitary authorities (councils) spanning Aotearoa New Zealand. This appendix describes the relationships among project participants. The *participants* consist of a *science team*, *core representatives* of the councils, and *additional representatives* of councils (Table A-1).

Table A-1: Membership of science team and the council core representatives.

	Science team	
1	Rick Stoffels	NIWA
2	Doug Booker	NIWA
3	Paul Franklin	NIWA
4	Robin Holmes	Cawthron Institute
5	Phil Jellyman	NIWA
	Participating councils	
1	Otago Regional Council	
2	Northland Regional Council	
3	Hawke's Bay Regional Council	
4	Nelson City Council	
5	Environment Southland	
6	Auckland Council	
7	Horizons Regional Council	
8	Environment Canterbury	
9	Gisborne District Council	
10	Taranaki Regional Council	
11	Bay of Plenty Regional Council	
12	Greater Wellington Regional Council	
13	Waikato Regional Council	

Appendix B Method of aggregating statements pertaining to the advantages and disadvantages of consistent M&E into classes

The science team and core representatives were asked to submit up to three possible advantages and up to three possible disadvantages of implementing nationally-consistent M&E in response to the NPSFM. The statements received comprise opinions of experts spanning abroad range of expertise in water resource science and management. A total of 90 statements were received (50 advantages; 40 disadvantages; Table B-1). These statements were subjectively assigned to 11 advantage *themes* and 9 disadvantage *themes*. Individual statements could be assigned to up to three themes. Certain statements contained more content than others, and so were more relevant to multiple themes. The frequencies with which statements referred to these themes has been summarised in Figure B-1 and Figure B-2.

To present the feedback on advantages/disadvantages as simply as possible, further aggregation was desirable. Inspection of Figure B-1 and Figure B-2 revealed some redundancies/similarities across themes. For example, the advantage themes *higher data quality* and *national data resource* were both presented in the context of a national-consistent program yielding *improved data quantity, quality and availability* (Table B-1), so were combined into that class. Following assignment to *classes*, we obtained seven and six advantage and disadvantage classes, respectively.

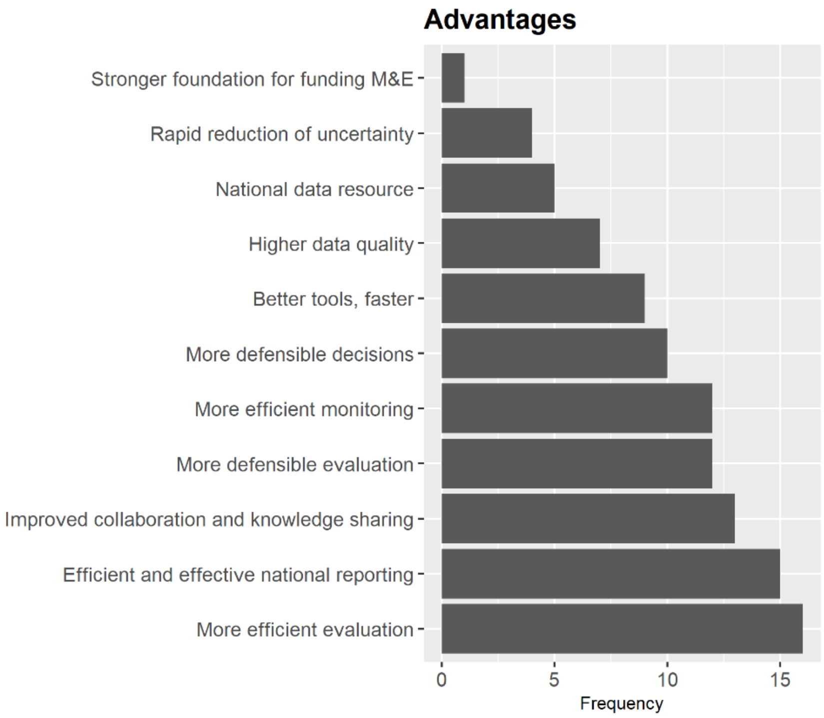

Figure B-1: Frequency of statements assigned to 11 advantage themes.

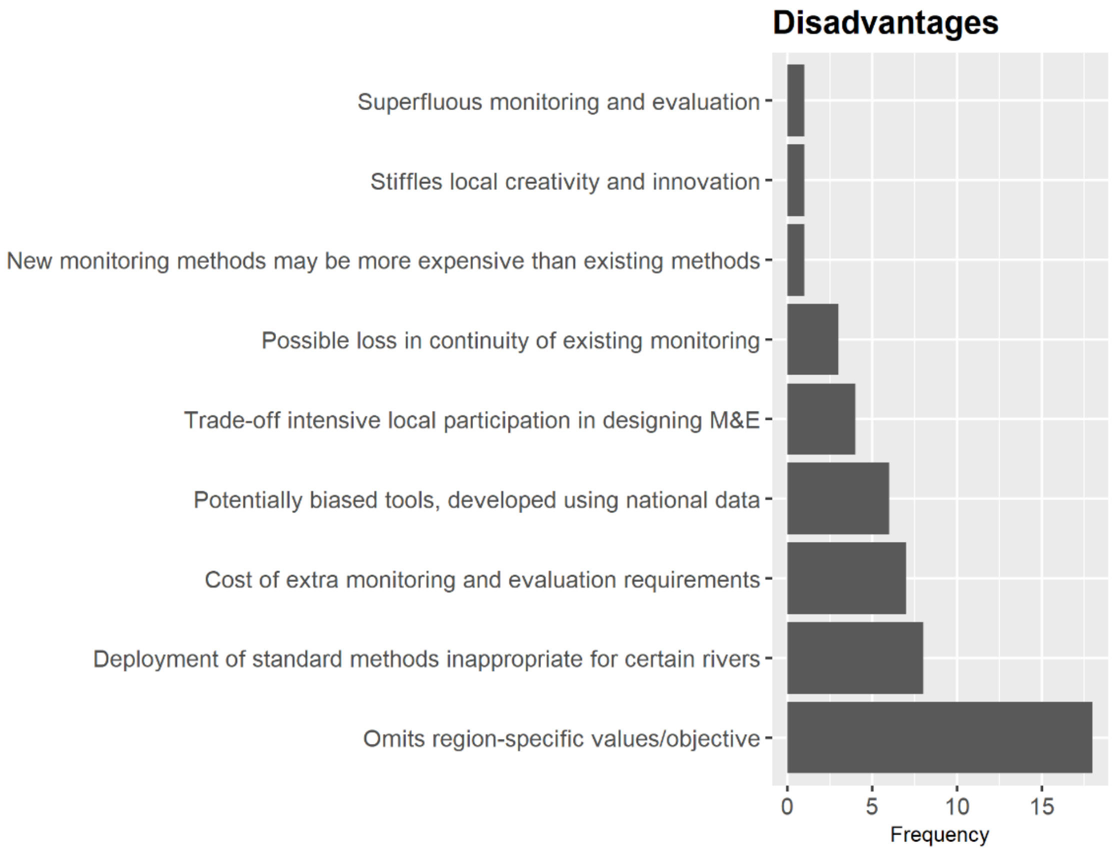

Figure B-2: Frequency of statements assigned to nine disadvantage themes.

Table B-1: Assignment of all 90 statements concerning advantages/disadvantages of consistent M&E to themes, prior to further aggregation into classes. Statements copied verbatim from respondents.

Advantages

Theme_1	Theme_2	Theme_3	Statement
More defensible evaluation	Efficient and effective national reporting	More defensible decisions	to avoid individually argued approaches in each region and the considerable costs and stresses of that
National data resource	More defensible evaluation	Improved knowledge sharing	to have a national resource for justifying/presenting on approaches for this very controversial issue
Efficient and effective national reporting	More defensible evaluation	More defensible decisions	to generate national consistency in both M&E to hopefully generate national consistency in levels of allocation (and national reporting)
Higher data quality	Better tools, faster	More defensible evaluation	Standard methods minimise noise when combinin data from regions, increasing power to detect response
More efficient evaluation	Efficient and effective national reporting		Standard approaches to data management will facilitate more effective and efficient national reporting
Stronger partnerships			Group collaboration – A collective discussion takes advantage of expertise in some regions which benefits RCs with less experience
Higher data quality	Better tools, faster	Rapid reduction of uncertainty	larger volumes of data (both more sites and greater replication through time) provides the opportunity to apply more sophisticated models (e.g., machine learning of mixed-effects) that consider between-site as well as between-time patterns. These methods would provide more insight into patterns that might (or might not) be predictable across landscape settings

Advantages

National data resource	Efficient and effective national reporting	More efficient evaluation	A consistent approach should create a consistent data format (including recording data quality assurance) that can more easily be placed in a central repository. Archiving of data in a centralized location should create cost-savings and encourage best-practice in terms of data archiving and retrieval. Councils would therefore not have to respond to ad hoc data requests. In this respect, a consistent approach would particularly benefit national environmental reporting by MfE and StatsNZ
More defensible evaluation	More defensible decisions		A consistent approach to data collection might encourage a more consistent approach to environmental flow setting (e.g., water resource use limits in plans)
Improved knowledge sharing	Better tools, faster		Facilitates shared learning/models from comparable data
Higher data quality			Provides confidence that the best available methodologies are being used by all
More defensible decisions	Stronger partnerships	Improved knowledge sharing	A nationally-consistent approach would integrate expertise from around the country, along with healthy debate, to end up with a more robust outcome
Stronger foundation for funding M&E			The approach could help provide better structure and justification for M&E budget requirements
Efficient and effective national reporting	Improved knowledge sharing	More defensible evaluation	The approach would aid in national level reporting that allows comparison across regions
More efficient evaluation	Efficient and effective national reporting		Regional and National level reporting would be better aligned
National data resource			Data management issues may arise because of different database solutions among councils
Efficient and effective national reporting			Easy to compare between regions for national reporting purpose
More efficient monitoring	More efficient evaluation		Easy to incorporate nationally standard protocols (e.g., NEMS) starting from data collection to reporting
Efficient and effective national reporting	More defensible evaluation		Easy to integrate with NPS-FM requirements and standards for freshwater accounting purposes
Higher data quality	Better tools, faster	More defensible decisions	Statistical power of larger (and comparable!) datasets. Higher confidence in conclusions and weight of evidence to support local/regional decisions
Stronger partnerships	Improved knowledge sharing		Common ground & shared goals promotes more meaningful engagement/collaboration between regions
More efficient monitoring	More efficient evaluation		Cost sharing and efficiency in effort
More defensible evaluation	More defensible decisions	Higher data quality	The approach is based on evidence based monitoring and research and standards that is defendable in environment court
Efficient and effective national reporting	More efficient evaluation		A consistent approach is required to assess apples with apples – water resources cross regional boundaries
More efficient monitoring	More efficient evaluation		Some efficiencies in developing the approach and future amendments/development to meet resource management and legislative requirements
Better tools, faster	Efficient and effective national reporting	Rapid reduction of uncertainty	It enables national analyses that may draw out patterns/responses that cannot be detected within regions, hence allowing people to learn more from each other
More efficient monitoring	More efficient evaluation	Higher data quality	It reduces duplication of effort in developing methodologies, infrastructure, data management etc leading to lower costs and probably better quality outcomes (by pooling resources) for everyone
More efficient monitoring	More efficient evaluation		It supports transferability of staff between regions
More defensible evaluation	More defensible decisions		In theory a nationally consistent approach should limit the number of challenges on the science
More efficient monitoring	More efficient evaluation		There would be a standard approach to monitoring, management and reporting

Advantages			
Improved knowledge sharing	More efficient evaluation		Could benefit cross boundary catchments through standardistation
Higher data quality	More defensible evaluation	Better tools, faster	Ensure that data is captured using nationally established (best practice) methodologies will providing greater confidence in the data collected and in any outputs generated with the data.
Improved knowledge sharing	Efficient and effective national reporting	More defensible decisions	Improved ability to compare results across different regions with differing allocation frameworks
National data resource	More efficient evaluation	More efficient monitoring	Having a single approach means there will be stronger national knowledge resource in the methods and opportunities for technical support across Councils (and other agencies) in the methods
More efficient monitoring			Training in methods could also be delivered in a more effective and efficient manner at a national scale.
Rapid reduction of uncertainty			More rapid reduction of structural uncertainty
Better tools, faster	More defensible decisions		Faster reduction of statistical uncertainty, leading to quicker development of powerful models for eval and forecasting
More efficient evaluation	More defensible evaluation	Efficient and effective national reporting	More effective and efficient national reporting, but probably also more effective and efficient regional reporting too - borrowing strength from neighbours
Efficient and effective national reporting			It would enable national-scale reporting. National scale State of Environment reporting for water quality provides a powerful tool / mandate for the NZ public and the government to implement environmental improvements. I think that national scale reporting on the effects of abstraction would do the same
More efficient monitoring	More efficient evaluation		Nationally consistent monitoring and reporting frameworks would allow transferable data and skills between the Regions. This could save lots of money e.g., though economies-of-scale when analyzing data or training staff
Improved knowledge sharing	More defensible evaluation	More defensible decisions	Consistent ecological effects monitoring would make it easier to communicate results to the public. A consistent approach to communicating the ecological effects of water allocation will emerge from a consistent monitoring approach. Ultimately, this will enable water allocation to be managed more democratically
Efficient and effective national reporting	Improved knowledge sharing		Provides for meaningful comparisons acorss the region- i.e., comparing apples with apples
More efficient monitoring			Provides potential for efficiencies/ collaboration in data gathering as everyone is speaking the same language, development of standard templates etc
More efficient monitoring	More efficient evaluation	National data resource	Could potentially facilitate one entity taking the lead on admin, with others maybe paying a support fee= could be way chpear for NZ Inc long term
Efficient and effective national reporting			Provides spatially comparable data sets for a collated national summary of catchment state
More efficient evaluation	Better tools, faster		Great opportunity to develop river type suitable methods to measure location specific values in a standardised way
Rapid reduction of uncertainty			The ability to target 'representative' sites after initial national assessment period to fill in knowledge gaps of invertebrate (or vertebrate) with flow response
Better tools, faster			Potential to apply principles of species flow response to NREI type modelling across more areas to better estimate flow demands and allocative pressure through a range of flows
Improved knowledge sharing			Improved inter council learning on ecosystem response
More efficient monitoring			Improved inter council learning on method development

Disadvantages

Theme_1	Theme_2	Theme_3	Statement
Omits region-specific values/objective	Potentially biased tools, developed using national data	Deployment of standard methods	The very different river types we deal with (i.e., braided rivers) and avoiding inappropriately lumping methods together for very different river types

Advantages			
		inappropriate for certain rivers	
Deployment of standard methods inappropriate for certain rivers	Omits region-specific values/objective		Difficulty with large latitudinal difference the length of NZ (temperature, climate, biota etc.) as well as geological differences
Omits region-specific values/objective	Deployment of standard methods inappropriate for certain rivers	Potentially biased tools, developed using national data	In dryer regions dealing with the often unique issues of seasonal drying of rivers
Deployment of standard methods inappropriate for certain rivers			Standard methods will not work in all stream types
Omits region-specific values/objective			No single ste of indicators will detect response to low flows across all stream types
Omits region-specific values/objective			'one size fits all' might miss key differences between regions (e.g high demand for water abstraction on east coast vs not much on west coast etc), also different resourcing and budget constraints between councils can mean projects and deliverables are unrealistic for some RC's
Omits region-specific values/objective			A national approach may not recognize that in-stream values vary spatially
Deployment of standard methods inappropriate for certain rivers			For a nationally-consistent approach to be effective, there might have to be more strict control of deployment of the approach (e.g., would training be needed to ensure consistent deployment of that national approach?).
Cost of extra monitoring and evaluation requirements			Increased costs would result from councils having to carry out monitoring in addition to that which they are obliged to carry out as a consequence of existing plans or consents
Omits region-specific values/objective			may privilege widespread/dominant river flow types (or species) at the expense of spatially restricted yet important ecosystems
Trade-off intensive local participation in designing M&E			may privilege one way of assessment at the disadvantage of alternate valid approaches (e.g., cultural)
Cost of extra monitoring and evaluation requirements	Superfluous monitoring and evaluation		The approach could become too onerous in order to meet the needs of all the councils. i.e., could end up with superfluous requirements for some councils
Omits region-specific values/objective			The approach could become too sparse in order to eliminate specialty requirements for some councils, i.e., some councils would still have to come up with their own supplementary M&E requirements.
Cost of extra monitoring and evaluation requirements			The approach could result in requirements that are unable to be funded by some councils
Omits region-specific values/objective			Rivers and freshwater bodies in Northland are quite different from rest of the country (e.g., nearly half of our waterways are soft-bottomed, wide non-wadable channel, smaller catchment size with tidal influence extending further inland, sub-tropical climate). Therefore a nationally consistent approach (e.g., macroinvertebrate monitoring or deposited sediment) might not fit very well to a large number of our rivers
Potentially biased tools, developed using national data	Deployment of standard methods inappropriate for certain rivers		Developing guideline values or standards based on national dataset might skew the actual picture at regional scale, which might be the reality particularly for coastal streams with smaller catchment size and soft-sedimentary geology. This will be reflected well by some water quality parameters such as cont. DO but not so much by macroinvertebrates or fish diversity.

Advantages		
Omits region-specific values/objective	Potentially biased tools, developed using national data	Risk of M & E approach being captured or skewed towards issues/values/regions of highest *perceived* importance…and therefore not being of equal benefit or use around the country
Omits region-specific values/objective		Differences in biogeographic regions, catchment size-stream orders and land and water use that may be difficult to fit in a national approach
Omits region-specific values/objective		It may not necessarily account for local differences in values/desired environmental outcomes
Possible loss in continuity of existing monitoring		Possible need to forego value from/consistency with existing historical datasets if they are not the same as the new standardised methodologies
Cost of extra monitoring and evaluation requirements		Cost of retraining staff
Omits region-specific values/objective		A national approach may not take unique regional issues into account as well as a regional approach
Omits region-specific values/objective	Trade-off intensive local participation in designing M&E	Potentially not a locally collaborative approach particularly taking into account cultural values
Cost of extra monitoring and evaluation requirements		Potential cost barriers in terms of the level of monitoring requirements, data management and reporting
Possible loss in continuity of existing monitoring		A departure for existing methods may mean it is impractical/cost prohibitive to maintain any existing long-term data sets that councils have invested in
New monitoring methods may be more expenisve than existing methods		A nationally consistent method may be more expensive to deliver than existing bespoke solutions - increasing costs on Council's and ratepayers
Omits region-specific values/objective		Risk of not being applicable to region-specific values
Deployment of standard methods inappropriate for certain rivers		Difficulty of applying a standardised method to unique, region-specific circumstances
Stifles local creativity and innovation	Trade-off intensive local participation in designing M&E	Risks stiffling creative, innovative approaches employed by smaller groups
Potentially biased tools, developed using national data		A shift in focus to national-scale monitoring, evaluation (and reporting) could 'average out' important regional differences. Regions where water allocation is particularly problematic could be obscured by the national picture—which might not look so bad (e.g., because a third of waterbodies are in the National Park)
Omits region-specific values/objective	Trade-off intensive local participation in designing M&E	It will be very difficult to produce a 'one size fits all' monitoring and evaluation framework for different regions / waterbody types. Compromises in methodologies to enable a national scale approach will lead towards a 'lowest common denominator' approach when selecting methodologies. The risk is that a less bespoke approach will mean some effects, that are specific to uncommon types of waterbody (e.g., ephemeral streams), will not be recognized.
Cost of extra monitoring and evaluation requirements		Councils may end up with more work to do. They might have to continue with their regional approaches in order to maintain long-term records and regional objectives and then have to *add* another set of sites and methodologies to adhere to a national monitoring framework. Given they are already stretched for capacity this could be a major issue

Advantages	
Omits region-specific values/objective	There could be regional variations not adequately reflected by a one size fits all approach
Potentially biased tools, developed using national data	Linked to this, outputs can be used to draw assumptions/ comparisons which are not supported by the full context
Omits region-specific values/objective	regional or site-specific variation isn't able to be addressed or fell by the wayside if a nationally consistent approach was adopted
Cost of extra monitoring and evaluation requirements	Council resources/capability may not be sufficient to incorporate extra requirements of meaningful and comprehensive monitoring. May be lag in the collection of data sets until resources are sufficient and therefore gaps in the national data set
Possible loss in continuity of existing monitoring	Potentially time and budget consuming work to develop meaningful assessments of environmental variables in larger rivers
Deployment of standard methods inappropriate for certain rivers	A range of methods may need to be developed for different stream types
NA	Distinguishing temporal response ecosystems to water resource management, from climatic variability and other land use effects is difficult unless the effects are sudden and catastrophic
NA	If the consequences of bad policy only emerge sometime later in water short years, and only after the community has invested heavily in water infrastructure, how to realistic is it to expect claw back water or make other policy changes that benefit ecosystems?

Appendix C Online survey

Background

Tēnā koutou,

NIWA is working with councils to co-develop a monitoring and evaluation (M&E) plan to support the adaptive management of river flows in accordance with the NPSFM 2020. Co-development of that plan follows a process, including surveys to inform design of the M&E plan. Answers to the questions in this survey will be discussed at workshops during early December, and will inform

- what needs to be measured; and
- exactly how will measurements be used to inform decision-making, and evaluation of the ecosystem's response to previous decisions.

Many thanks for your participation.

Kā mihi

Information about you

1. Please select your workplace

☐	Northland Regional Council	☐	Nelson City Council
☐	Auckland Council	☐	Environment Canterbury
☐	Hawke's Bay Regional Council	☐	Otago Regional Council
☐	Bay of Plenty Regional Council	☐	Environment Southland
☐	Gisborne District Council	☐	Cawthron
☐	Horizons Regional Council	☐	NIWA
☐	Taranaki Regional Council	☐	
☐	Greater Wellington Regional Council	☐	

2. Approximately how many years of experience do you have working in the natural resources sector?

☐	<5 years
☐	5 - 10 years
☐	> 10 years

3. Which of the following best describes your role?

☐	Scientist
☐	Policy and planning

☐ Consenting
☐ None of the above

4. Please describe your role in one sentence (only visible to those who select "none of the above" for Q3).

Alignment of attributes/indicators with stakeholder/partner values

The questions on this page are about the values/needs of council stakeholder groups and/or partners in the broad sense. So when answering, please think beyond just the requirements of MfE / NPSFM.

Specifically, the questions on this page refer to stakeholders/partners that place high value on the ecological health of rivers. Later in the survey (subsequent pages) we ask questions relevant to all stakeholder groups, including those that benefit most from water abstraction.

The stakeholder groups that questions on this page refer to are:

A. **Consumers**: members of the general public whose values are adversely affected by deterioration of river health. E.g., anglers/fisherpeople; boaters/canoers/kayakers; those that camp/swim on/in rivers; birdwatchers; trampers;…

B. **Non-governmental organizations (NGOs)**: NGOs advocate for the conservation or wise use of rivers, but are not part of a government entity legally mandated to manage rivers and/or implement environmental policies/plans. E.g., NZ Forest & Bird

C. **Economic**: businesses whose income is negatively affected by deterioration of river health. E.g., fishing guides; ecotourism businesses.

D. **Regional resource management agencies (other than councils)**: Organisations charged with managing rivers within their legally mandated jurisdiction. E.g., some iwi/hapū; Fish and Game.

E. **Iwi/hapū:** Māori have legitimate and legal rights and interests in water management, and there is an expectation from Central government for co-management and full participation in water management decisions

5. Councils require information that directly shows how stakeholder/partner values respond to changes in river flow

☐ Strongly agree
☐ Agree
☐ Neither agree nor disagree
☐ Disagree
☐ Strongly disagree
☐ Don't know

6. Local stakeholders/partners only care whether a species is present or absent in a river; more detailed information like abundance/density or size structure of populations isn't important to them.

- [ ] Strongly agree
- [ ] Agree
- [ ] Neither agree nor disagree
- [ ] Disagree
- [ ] Strongly disagree
- [ ] Don't know

7. The amount of water/flow in a river *per se* is important to partners/stakeholders in my region

- [ ] Strongly agree
- [ ] Agree
- [ ] Neither agree nor disagree
- [ ] Disagree
- [ ] Strongly disagree
- [ ] Don't know

8. The aesthetics or "natural character" of a river reach *per se* is important to partners/stakeholders in my region

- [ ] Strongly agree
- [ ] Agree
- [ ] Neither agree nor disagree
- [ ] Disagree
- [ ] Strongly disagree
- [ ] Don't know

9. Please enter any other comments you'd like to make about the alignment of indicators with stakeholder/partner values.

Ease of reporting/communicating to/with multiple partners and stakeholders

Questions on this page are about what sort of information would be most effective when it comes to reporting and communicating environmental outcomes to *all partners and stakeholder groups*, including irrigators and agribusiness.

For clarity, **an indicator is any measurable property of the ecosystem**. It can be qualitative (e.g., subjective measurements of state like 'good' and 'bad'), or quantitative, and be a measurement of state (e.g., density of a species) or rate (e.g., rate of metabolism).

10. How important is it that indicators are easily interpreted by all partners and stakeholder groups, in terms/units that align with their values/understanding?

- [ ] Extremely important
- [ ] Very important
- [ ] Somewhat important
- [ ] Not so important
- [ ] Not at all important
- [ ] Don't know

11. Multivariate indicies (ie. indices that combine information from several species like MCI, QMCI, fish IBI) are an effective way of reporting/communicating ecological outcomes to all partners and stakeholders

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

12. We require indicators that summarise ecosystem state and trends down to as few numbers as possible.

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

13. At least some stakeholders/partners will want to know how rivers of particular interest to them (e.g., rivers close to their property; iconic rivers) are responding to water allocation decisions.

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

14. Stakeholders/partners within your region will be most interested in indicators that show how rivers respond to water allocation decisions at what scale? (Multiple box-checks allowed)

- ☐ At the reach to segment (< 10 km) scale; just a segment of a river
- ☐ At the scale of an entire river/tributary
- ☐ At the catchment or basin scale (multiple rivers/tributaries together
- ☐ At the scale of the entire region (e.g., Canterbury)
- ☐ At the national scale
- ☐ None of the above

15. Please add any comments you'd like to make about reporting/communication objectives, and how indicator choice might affect meeting those objectives?

Sensitivity of indicators to flow management decisions

The following questions relate to what we monitor/measure tells us about the outcomes of flow management decisions.

16. Given the evaluation and reporting needs/cycles of our council, the indicators we monitor would, ideally, be sufficiently sensitive to changes in flow to exhibit significant change (improvement or deterioration) within:

- ☐ 1 year
- ☐ 3 years
- ☐ 5 years
- ☐ 10 years
- ☐ 20 years
- ☐ Don’t know

17. Managing river flows involves decisions about, for example, the magnitude (how much can be taken?), timing (when can water be taken?) and duration (for how long can river flow be less than X?) of water abstraction. The indicators used should enable us to separate/isolate how different flow management decisions affect riverine values.

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don’t know

18. Please add any comments you’d like to make about what indicators should/shouldn’t—or have/haven’t—told you about “on-the-ground” management decisions.

Defensibility or 'robustness' of decisions and evaluations

The following questions are designed to elucidate levels of interest in tools to support decision-making and evaluation of ecological response to flows. An understanding of the interest by councils in such tools will, in turn, inform monitoring design. Also included are questions aimed at deciphering your experience with decision-making and evaluation in natural resource management.

19. Water allocation decisions are likely to be contentious amongst partners/stakeholders within your region

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don’t know

20. Water allocation decisions made by my council are likely to be closely scrutinised by partners/stakeholders

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don’t know

21. Which of the below statements best describes your experience with freshwater resources disputes and/or conflict resolution:

- ☐ I have first-hand experience, e.g., through environment court, local court hearings, community/stakeholder meetings
- ☐ I don't have first-hand experience, but have been in the industry long enough to be well aware of how freshwater resource disputes arise, and how they are resolved
- ☐ I have little first-hand experience or awareness of how freshwater resource disputes arise, or how they are resolved

22. Data collected by our monitoring programs should be used to develop tools to support repeatable, transparent decision-making and reporting of outcomes

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

23. It would be useful if information derived from data we collect is transferable to sites where we do not collect any data

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

24. Evidence supported by robust, data-based statistical analyses is more convincing to partners/stakeholders

- ☐ Always
- ☐ Most of the time
- ☐ Sometimes
- ☐ Rarely
- ☐ Never
- ☐ I don't really know what forms of evidence are more/less convincing to stakeholders/partners

25. Please add any comments you'd like to make about how data collected by the monitoring program should be used to support evaluation and decision-making.

Resource and logistical constraints

26. Costs associated with monitoring numerous indicators are a concern within my council

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

27. New monitoring site requirements arising from the NPSFM should leverage off existing monitoring sites, as much as possible

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

28. New monitoring infrastructural requirements (e.g., flow gauging stations) arising from the NPSFM should leverage off existing monitoring infrastructure, as much as possible

- ☐ Strongly agree
- ☐ Agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree
- ☐ Don't know

29. Please add any comments you'd like to make about logistical/resource constraints associated with monitoring and evaluation of river flows under the NPSFM

Appendix D River type nominations of council core representatives

Table D-1: Raw river type nominations of councils. Type: council definition of types of rivers. Type_Source_Sub: labels of types assigned by the science team towards aggregating common council-defined types. Prevalence: categorical, subjective assessment by councils of the prevalence of council-defined types within their region. Magnitude: categorical, subjective assessment by councils of the relative magnitude of water abstraction within their region.

Region	Type	Type_Source_Sub	Example	Prevalence	Magnitude
Canterbury	Hill-fed rivers (hard bottomed)	HillMtnFed_CoarseSub	Ashley/Rakahuri	1_high	1_high
Canterbury	Groundwater fed streams (hard bottomed) spring	Spring_CoarseFineSub	Hart's Creek	1_high	1_high
Canterbury	drains that replace extensive wetlands	NA	NA	1_high	1_high
Otago	Hill-fed rivers w/ braid	Braided_CoarseSub	Manuherikia/Lindis/ Cardrona	1_high	1_high
Otago	Hard Bottom Streams	HillMtnFed_CoarseSub	Deep Stream	1_high	1_high
Waikato	Depositional soft bottom flashy	HillMtnFed_FineSub	Piako	1_high	1_high
Wellington	Main range gravel bed, frequent flushed	HillMtnFed_CoarseSub	Ruamahanga, Hutt	1_high	1_high
Northland	spring	Spring_FineCoarseSub	Waipao	1_high	1_high
Auckland	soft-bottom on hard-bottom streams	HillMtnFed_FineCoarseSub	Whangamaire	2_med	1_high
Auckland	springs	Spring_FineCoarseSub	Whangamaire	2_med	1_high
Auckland	hard, then soft bottom streams	HillMtnFed_FineCoarseSub	Whangamaire	2_med	1_high
Auckland	Urban Streams (soft bottomed)	Urban_FineSub	Puhinui	2_med	1_high
Auckland	intensive horticulture streams	HillMtnFed_FineCoarseSub	Whangamaire	2_med	1_high
Auckland	Streams that cross extreme geological boundaries, e.g., Basalt to silty sandstone	HillMtnFed_FineCoarseSub	Whangamaire	2_med	1_high
Bay of Plenty	Groundwater dominated, pumice systems (springs and their lower reaches)	Spring_FineCoarseSub	Pongakawa, Pukehina, Waitahanui, Wharere	2_med	1_high
Canterbury	Large Braided Rivers (hard bottomed)	Braided_CoarseSub	Waimakariri	2_med	1_high
Northland	Volcanic streams (hard/soft)	HillMtnFed_CoarseSub	Punakitere	2_med	1_high
Wellington	Foothill gravel bed	HillMtnFed_CoarseSub	Waipoua, Mangatarere	2_med	1_high
Wellington	Valley floor spring fed soft (incised - macrophyte dom)	Spring_FineSub	Papawai Stream	2_med	1_high
Bay of Plenty	Springs	Spring_FineCoarseSub	Rotorua Springs	3_low	1_high
Auckland	Dammed Rivers	LakeFed_CoarseSub	Wairoa	3_low	1_high
Canterbury	Lake fed Rivers (hard bottomed)	LakeFed_CoarseSub	Waitaki	3_low	1_high
Bay of Plenty	Lower reaches of large rivers	HillMtnFed_CoarseSub	Rangitaiki, Tarawera, Waiari, Kaituna	1_high	2_med
Hawke's Bay	Hard-bottomed streams	HillMtnFed_CoarseSub	Tukituki River, Esk River	1_high	2_med
Hawke's Bay	Soft-bottom streams (run - off)	HillMtnFed_FineSub	Awanui Stream, Taurekaitai Stream	1_high	2_med
Nelson	Hard-bottomed streams (urban and rural)	HillMtnFed_CoarseSub	Poorman Valley Stream/Stoke fan streams; Maitai,	1_high	2_med

Region	Type	Type_Source_Sub	Example	Prevalence	Magnitude
			Whangamoa, Wakapuaka & Roding rivers		
Otago	hill-fed rivers sans braid	HillMtnFed_CoarseSub	Shag/kakanui	1_high	2_med
Waikato	Spring dominated pumice bed	Spring_FineCoarseSub	Waikato River	1_high	2_med
Waikato	depositional old wetlands	NA	the "drains"	1_high	2_med
Bay of Plenty	Single thread incised rivers and streams	HillMtnFed_CoarseSub	Tarawera, Ohineangaanga, Waiari, Kaituna, Aongatete, Uretara, Te Mania	2_med	2_med
Gisborne	Hard-bottomed to soft bottomed streams	HillMtnFed_FineCoarseSub	Te Arai, Taruheru, Waimata, Uawa	2_med	2_med
Gisborne	Meandering Rivers	HillMtnFed_FineSub	Waipaoa, Motu	2_med	2_med
Hawke's Bay	Spring fed hard bottom	Spring_CoarseSub	Waitio Stream	2_med	2_med
Hawke's Bay	Spring fed soft bottom	Spring_FineSub	Kahakuri Stream, Raupare Stream	2_med	2_med
Northland	Hard-bottomed streams	HillMtnFed_CoarseSub	Waiarohia	2_med	2_med
Northland	Soft-bottomed streams	HillMtnFed_FineSub	Wairua, Mangakahia	2_med	2_med
Taranaki	Soft-bottom streams	HillMtnFed_FineSub	Mangaroa	2_med	2_med
Auckland	Fully native forest cover streams	HillMtnFed_FineCoarseSub	NA	3_low	2_med
Hawke's Bay	Braided rivers	Braided_CoarseSub	Ngaruroro River	3_low	2_med
Nelson	Soft-bottom streams (peri-urban and rural)	HillMtnFed_FineSub	Todd Valley stream	3_low	2_med
Northland	Non-wadeable large rivers (soft+hard)	HillMtnFed_FineCoarseSub	Lower reaches of Northern Wairoa, Awanui	3_low	2_med
Taranaki	Spring fed	Spring_FineCoarseSub	Waiokura	3_low	2_med
Wellington	Valley floor spring fed - hard	Spring_CoarseSub	Waipipi Strea,	3_low	2_med
Auckland	tiny coastal catchments with mouth closure issues	NA	Poutawa	1_high	3_low
Taranaki	Hard-bottomed streams	HillMtnFed_CoarseSub	Punehu	1_high	3_low
Waikato	Hard-bottomed erosional streams	HillMtnFed_CoarseSub	Coromandel	1_high	3_low
Auckland	low-baseflow	NA	Rangitopuni	1_high	3_low
Bay of Plenty	Braided Rivers	Braided_CoarseSub	Motu, Kereru, Whakatane	2_med	3_low
Gisborne	Braided Rivers	Braided_CoarseSub	Waiapu, Mata. Tapuaeroa	2_med	3_low
Gisborne	Springs	Spring_FineCoarseSub	Ruatoria, Poverty Bay Flats	2_med	3_low
Otago	Springs	Spring_FineCoarseSub	Bullock Creek	2_med	3_low
Waikato	Erosional flashy	HillMtnFed_FineCoarseSub	Waipa	2_med	3_low
Wellington	Depositional soft bed/bank river (incised) - v low base flow	HillMtnFed_FineSub	Taueru, Whareama	2_med	3_low
Auckland	Urban Streams (hard bottomed)	Urban_CoarseSub	Meola	2_med	3_low
Canterbury	Urban Streams (hard bottomed)	Urban_CoarseSub	Avon/	3_low	3_low
Northland	Lake/wetland fed Rivers (soft bottom)	LakeFed_FineSub	Pouto / Aupouri	3_low	3_low
Otago	Braided/Glacial Rivers	Braided_CoarseSub	Dart/Rees/makarora	3_low	3_low

Region	Type	Type_Source_Sub	Example	Prevalence	Magnitude
Otago	Lake fed	LakeFed_CoarseSub	Clutha	3_low	3_low
Otago	Soft bottom streams	HillMtnFed_FineSub	NA	3_low	3_low
Waikato	Depositional hard bottom	HillMtnFed_CoarseSub	braided rivers	3_low	3_low
Canterbury	Soft bottom streams	HillMtnFed_FineSub	Heathcote	3_low	3_low
Nelson	Spring fed	Spring_FineCoarseSub	No comprehensive survey data, mainly wetted seeps	3_low	3_low

Appendix E Metrics nominated by councils under each NPSFM component of ecosystem health

To ensure metrics were nominated within the context of the NPSFM value ecosystem health (Section 7.1), we framed the brainstorming within the context of a basic conceptual model outlining how river flow may drive interactions between the five NPSFM components of ecosystem health (Figure 7-3). This conceptual model was presented to participants and briefly discussed prior to brainstorming. An online spreadsheet workbook was set up, within which five spreadsheets corresponded to each component of ecosystem health. Within the context of Figure 7-3 participants were asked to nominate attributes for each component (first column), and also provide some notes on how the attribute might be measured (second column) and a justification for the attribute (third column). The science team then collated all nominations, tidied them, and added an additional column denoting level of support for the attribute based on the frequency with which the attribute was nominated by participants (low; medium; high).

Table E-1: Attributes nominated within each of five components of the NPSFM compulsory value: ecosystem health. Attributes with high levels of support highlighted in grey.

Component	Metrics	Support	Measurement Notes	Why
Water Quality	DO	high	Continuous logging during low flows. As per NPSFM	Low flows and elevated water temperatures may lower DO. Critical to physiological processes.
Water Quality	Water temperature	high	Continuous logging	Low flow may elevate water teperature. Climate warming may interact with low flow to elevate water temperature. Critical to physiological processes
Water Quality	pH	low	Spot measurements (continuous logging expensive)	Stressor that may interact with flow
Water Quality	Dissolved nutrients	low	As per NPSFM	Affected by flow due to changes in volume. Stressor that may interact with flow.
Water Quality	Clarity	low	Spot measurements. As per NPSFM	Stressor that may interact with flow
Habitat Hydraulics	Depth distribution	high	Frequency distribution of depths within reach	Processes at base of food web are depth-dependent (ie biofilm growth, hence macroinvertebrate production)
Habitat Hydraulics	Discharge	high	Continuous logging of depth at known cross-sectional area	Essential attribute. Discharge is the variable most directly affected by water takes.
Habitat Hydraulics	Thalweg	med	Profile of thalweg depth over specified reach	Influences movement of organisms (particularly fish) as well as important to human use of river, including mana whenua
Habitat Hydraulics	Thalweg minimum	med	Minimum thalweg depth over specified reach	Influences movement of organisms (particularly fish) as well as important to human use of river, including mana whenua
Habitat Hydraulics	Velocity distribution	med	Frequency distribution of velocities within reach	Processes at base of food web are velocity-dependent (ie biofilm growth, hence macroinvertebrate production)
Habitat Hydraulics	Wetted area	med	Width cross-sections at known distances within reach	Determines area of aquatic habitat. Critical determinant of reach capacity to support life. Can be coupled with depth and velocity distributions to obtain absolute abundance of different depth/velocity habitats as function of flow.
Habitat Hydraulics	Habitat composition	low	Relative abundance of pool/riffle/run	These habitat categories broadly define relative abundance of habitats with different depth/velocity/sediment compositions

Component	Metrics	Support	Measurement Notes	Why
Habitat Hydraulics	Velocity maximum	low	Maximum velocity within reach (visual + measure)	A statistical property of the velocity frequency distribution - tells us upper limit of distribution (at a point in time)
Habitat Geomorphology	Streambed elevation	high	Digitial elevation model	Determines numerous ecological processes, such as longitudinal and lateral movement of matter, hydraulic habitat diversity. Determinant of natural character. Incision.
Habitat Geomorphology	Streambed elevation	high	Cross-sectional elevation (transects)	Determines numerous ecological processes, such as longitudinal and lateral movement of matter, hydraulic habitat diversity. Determinant of natural character. Incision.
Habitat Geomorphology	Particle size distribution	high	Frequency distribution of particle sizes	Critical determinant of numerous
Habitat Geomorphology	Aerial photographs	high	Aerial photographs of reach (e.g., from drone)	Indicator of numerous things, inlcuding degree of braiding, wetted width, hydrological connectivity, broad substrate characteristics, etc
Habitat Geomorphology	Degree of braiding	med	?	Natural character. Hydraulic diversity.
Aquatic life	Macroinvertebrate benthic density	high	Species-specific densities (ind m^2)	Critical determinant of fish standing crop biomas - linked to highest ecosystem values
Aquatic life	Macroinvertebrate drift density	high	Species-specific densities (ind m^3)	Critical determinant of fish standing crop biomas - linked to highest ecosystem values
Aquatic life	Periphyton composition	high	Weighted composite cover = (%Filamenteaous algae + % thick mat cover)/2	Composition determinant of processes at higher trophic levels in river food web.
Aquatic life	Macrophyte cover (x-section)	high	Cross-sectional area of river comprised of macrophyte cover.	Highly relevant to spring-fed rivers. Flow sensitive but also relevant to stakeholders (chokes streams, for example)
Aquatic life	Fish abundance	high	Netting, eFishing, or even drift-diving (trout)	High value attribute - strongest relevance to stakeholders
Aquatic life	Fish size composition	high	As above, but measuring size composition at certain times of year	Indicator of how flow regime affects age- or stage-structure of fish populations. E.g., influence of flows on recruitment.
Aquatic life	Periphyton biomass	med	Chl a.	NPSFM attribute. Determinant of processes in higher trophic levels.
Aquatic life	Macrophyte cover (lateral/reach)	med	Surface area of river reach covered by macrophytes (from above)	Highly relevant to spring-fed rivers. Flow sensitive but also relevant to stakeholders (chokes streams, for example)
Aquatic life	Fish spawning haitat cover	med	Surface area of spawning habitat for particular species	Can be negatvely (and strongly) influenced by flow.
Aquatic life	Fish IBI	low	NPSFM attribute	NPSFM directs us to measure it.
Aquatic life	Macroinvertebrate QMCI	low	NPSFM attribute	NPSFM directs us to measure it.
Aquatic life	Fish condition	low	Individual length:mass ratios	Indicator of habitat quality for fish. Adds value to abundance.
Aquatic life	eDNA	low	Species composition (presence/absence)	Broad, cost-effective indicator of species presence-absence.

Component	Metrics	Support	Measurement Notes	Why
Ecological processes	Ecosystem metabolism	med	NPSFM attribute	Will indicate the ecosytem metabolic activities influenced by algal and aquatic plant growth (P & R calculated from DO time series)
Ecological processes	Cotton strip loss	low	Rate of breakdown of cotton strips	Decomposition links to flow - microbial processes. Flow increases temperature - hence microbial activity?
Ecological processes	Food web connectivity	low	?	?
Ecological processes	Food chain length	low	Stable isotope delta-N	healthy ecosystems free of passage barriers etc. support higher trophic levels e.g., fish feeding on fish instead of inverts