Adaptive Environmental Monitoring for Canada's Oil Sands Report

2024-12-17

Submitted to:
Alberta Environment and Protected Areas

KILGOUR & ASSOCIATES LTD. **and** **INNOTECH ALBERTA**
www.kilgourassociates.com

Project Number: AEP1265

ACKNOWLEDGEMENTS

This report was prepared by the following:

Keith Somers, PhD, Kilgour & Associates Ltd., Dorset, Ontario

Bruce Kilgour, PhD, Kilgour & Associates Ltd., Ottawa, Ontario

Rajiv Neal Tanna, (PhD), InnoTech Alberta, Calgary, Alberta

This report received constructive feedback from the following reviewers:

Kelly Munkittrick, PhD, University of Calgary, Calgary, Alberta

Rod Hazewinkel, MSc, Alberta Environment, Edmonton, Alberta

Sawyer Stoyanovich, PhD, Kilgour & Associates Ltd., Ottawa, Ontario

This work was partially funded under the Oil Sands Monitoring Program and is a contribution to the Program but does not necessarily reflect the position of the Program.

EXECUTIVE SUMMARY

Adaptive monitoring was suggested as a guiding principle for the Oil Sands Monitoring (OSM) Program based on its successful application in the federal Environmental Effects Monitoring programs within the Pulp and Paper Effluent Regulations and Metal and Diamond Mining Effluent Regulations. Implementation of adaptive monitoring principles leads to more efficient use of resources. That is, adaptive monitoring is intended to focus monitoring effort on areas of greatest concern, where scaled monitoring efforts can help identify the causes of unanticipated, undesirable change and propose candidate solutions, while maintaining a base level of monitoring in areas of lesser concern. Further, when fully adopted and implemented, adaptive monitoring has the potential to efficiently demonstrate to stakeholders the level to which the environmental resource has been successfully managed.

Adaptive monitoring can be applied to various indicators of environmental condition, including valued ecosystem components, pressures, pathways (stressors), and responses. Adaptive monitoring benefits from an understanding of the relationships between valued components and the monitored indicators. That is, monitoring of indicators should tie back quantitatively or narratively to the valued components that stakeholders have agreed are important.

Adaptive monitoring requires thresholds for triggering changes in monitoring tiers. To avoid ambiguity, we suggest the term *threshold* be used to connote the notion of a numeric or narrative value for an indicator, that when crossed, triggers change in monitoring. That is, we use *trigger* as a verb or an action that follows when a threshold is crossed. In the context of the Oil Sands Operational Framework, thresholds here are akin to "limits of change", or levels of indicators that prompt changes in monitoring focus.

Baseline thresholds are defined here as numeric or narrative values that indicate the limit of change to a baseline condition. Baselines can represent a pristine pre-European contact condition, a local present-day best-available condition, or simply a local pre-project condition. Baseline thresholds are typically numeric, based on some estimate of the average condition with some measure of variability, and may be represented by a value that is a percentile of the baseline data set. Numeric or narrative models of the relationships between anthropogenic pressures and valued components are used to predict or forecast expected conditions for valued components and indicators. *Forecast thresholds*, therefore, are for the expected condition. If underlying pressures are not expected to change the condition of an indicator, then baseline and forecast thresholds may be the same numeric or narrative value. We propose that *thresholds of unacceptable change* are numeric or narrative values that stakeholders agree represent a condition to avoid. Conditions that cannot be recovered from may be used to inform thresholds of unacceptable change.

The 2021/2022 OSM workplans were required to incorporate adaptive monitoring principles. Despite the history of adaptive monitoring in aquatic environments, the approach to adaptive monitoring in the 2021/2022 workplans was variable. The purpose of this report is to provide a generic framework for application of adaptive monitoring principles and to further harmonize adaptive monitoring methodologies among OSM disciplines.

In this proposed framework, adaptive monitoring has the following fundamental tiers. Periodic ambient (surveillance) monitoring involves regular assessment and reporting of ambient conditions relative to expected conditions. When surveillance monitoring detects variations beyond baseline or forecast thresholds (i.e., beyond those expected), those observations trigger focused monitoring, including confirmation (to verify the last result), or investigations of cause (to improve our understanding of the relationship between pressures, stressors, pathways, and responses). When investigation of cause results in a revised understanding of how pressures influence responses, those new understandings can be used to inform models that would be used in scenario analysis to update forecasts. If the updated forecast indicates unacceptable present or future conditions, investigation of solutions with subsequent mitigation may be required. If mitigations are implemented and follow-up effectiveness monitoring confirms effective mitigation, models, thresholds, and triggers may be updated, while routine surveillance monitoring may resume. Identification of mitigations is not within scope of OSM, but the scope could be expanded to include effectiveness monitoring to confirm the effectiveness of mitigations.

The following recommendations are provided:

- Recommendation 1 – The Oversight Committee to provide a formal definition of baseline condition, to be quantified by Technical Committees
- Recommendation 2 – The Oversight Committee to define acceptable and unacceptable thresholds of change for the VECs and associated indicators
- Recommendation 3 – OSM workplans to incorporate VEC indicators based on established conceptual models
- Recommendation 4 – TACs to identify predictive models, thresholds, and actions triggered by threshold exceedances
- Recommendation 5 – Oversight Committee to establish a process for engaging stakeholders regarding the interpretation of regional model predictions, thresholds, and observed results
- Recommendation 6 – Funding requests for monitoring should be supported by a description of the models and associated predictions that will be tested
- Recommendation 7 – The Oversight Committee, with the SIKIC, the ICBMAC and the TACs to determine the criteria to trigger a change from Surveillance Monitoring to the Reduced Monitoring tier
- Recommendation 8 – OSM to provide oversight and guidance to TAC when Periodic Monitoring thresholds are exceeded
- Recommendation 9 – The Oversight Committee, with the SIKIC and TACs to consider how to define thresholds for non baseline forecasts
- Recommendation 10 – TACs develop IOC plans for approval by the SIKIC and the OC
- Recommendation 11 – TACs update predictive models after IOC (and IOS), and use those model predictions to update thresholds
- Recommendation 12 – OSM to consider expanding scope to include IOS, mitigation and effectiveness monitoring tiers associated with Focused Monitoring.

TABLE OF CONTENTS

1.0 INTRODUCTION 1
1.1 OIL SANDS HISTORICAL MONITORING 1
1.2 CRITICAL MONITORING REVIEWS 2
1.3 JOINT OIL SANDS MONITORING (JOSM) 3
1.4 OIL SANDS MONITORING GOVERNANCE STRUCTURE 3
1.5 ADAPTIVE MONITORING 4

2.0 AMBIENT MONITORING.......... 5
2.1 TYPES OF MONITORING 5
2.1.1 Compliance Monitoring.......... 6
2.1.2 Effectiveness Monitoring 7
2.1.3 Baseline Monitoring.......... 7
2.1.4 Environmental Effects Monitoring 8
2.1.5 Cumulative Effects Monitoring..........10
2.1.6 Adaptive Monitoring..........11
2.2 INDICATORS USED IN MONITORING 12
2.2.1 Conceptual Site Model 12
2.2.2 Valued Components..........14
2.2.3 Stressor and Effects-Based Indicators..........16
2.2.4 Surrogate Indicators 16
2.2.5 Pathway/Diagnostic Indicators..........17
2.2.6 Indicators based on Models and Forecasts 18
2.3 INDICATOR THRESHOLDS 19
2.3.1 Using Indicator Thresholds to Trigger Changes in Monitoring..........23
2.3.2 Baseline Thresholds..........25
2.3.3 Forecast Thresholds..........25
2.3.4 Limits of Change Thresholds..........26
2.4 CASE STUDIES 26
2.4.1 Canadian EEM 26
2.4.2 Aquatic Environment Monitoring in the Elk Valley..........29
2.4.3 Alberta Biodiversity Monitoring Institute (ABMI) Framework 32

3.0 PROPOSED ADAPTIVE MONITORING FRAMEWORK FOR OIL SANDS..........33
3.1 OVERVIEW 33
3.2 COMPONENTS AND TIERS 34
3.2.1 Baseline Monitoring..........34
3.2.2 Periodic Monitoring..........35
3.2.3 Focused Monitoring..........36
3.3 APPLICATION TO OIL SANDS MONITORING 38
3.3.1 Baseline Monitoring Component..........38
3.3.2 Periodic Monitoring Component 40

4.0 SUMMARY AND CONCLUSIONS..........41

5.0 LITERATURE CITED..43

List of Figures

Figure 1 Governance structure of the Oil Sands Monitoring Program....................................... 4
Figure 2 An example adaptive monitoring flowchart 11
Figure 3 The DPSIR-based conceptual model as applied to the oil sands...............................13
Figure 4 Example of a cartoon-based conceptual model with anticipated effect pathways for oil sands project interaction with surface waters..14
Figure 5 Relationships between project vision, goals, objectives, indicators, and criteria for aquatic systems...15
Figure 6 Diagnostic response patterns for EEM surrogate benthos and sentinel fish indicators 18
Figure 7 Example dose-response curves linking stressor and effect indicators 19
Figure 8 Conceptual model illustrating monitoring thresholds and triggers. 24
Figure 9 Decision tree for metal mining EEM 28
Figure 10 Schematic of primary components of the ABMI proposed cumulative effects monitoring framework (from Burton et al., 2014) 33
Figure 11 Simplified adaptive monitoring flowchart with baseline, periodic and focused monitoring components 34

List of Acronyms and Abbreviations

ABMI – Alberta Biodiversity Monitoring Institute
AEMP - Alberta Environmental Monitoring Panel
AOSERP – Alberta Oil Sands Environmental Research Program
AWTF - Active Water Treatment Facility
BCMOE – British Columbia Ministry of the Environment
CCME – Canadian Council of Ministers of the Environment
CEMA – Cumulative Environmental Management Association
CEQG - Canadian Environmental Quality Guideline
CES – Critical Effect Size
CNSC - Canadian Nuclear Safety Commission
DFO – Department of Fisheries and Oceans Canada
DPSIR – Driver, Pressure, State, Indicator, Response
EA – Environmental Assessment
ECCC – Environment and Climate Change Canada
EEM – Environmental Effects Monitoring
EMA – British Columbia Environmental Management Act
EPEA - Environmental Protection and Enhancement Act
EMP – Enhanced Monitoring Program
EVWQMP - Elk Valley Water Quality Management Plan
FLNRO – British Columbia Ministry of Forests, Lands, Natural Resource Operations
ICBM – Indigenous Community Based Monitoring
ICBMAC - Indigenous Community Based Monitoring Advisory Committee
IOC – Investigation of Cause
IOS – Investigation of Solutions

JOSM – Join Canada-Alberta Oil Sands Monitoring Program
KAL – Kilgour & Associates Ltd.
KNC - Ktunaxa First Nations Council
LARP - Lower Athabasca Regional Plan
LTRN – Long-Term River network
MDMER – Metal and Diamond Mining Effluent Regulations
MMER - Metal Mining Effluent Regulations
NRBS – Northern River Basins Study
NREI – Northern Rivers Ecosystem Initiative
OC – Oversight Committee
OSM – Oil Sands Monitoring
PERD - Panel on Energy Research and Development
PPER – Pulp & Paper Effluent Regulations
RAMP – Regional Aquatics Monitoring Program
SBEB – Science-Based Environmental Benchmark
SIKIC – Science and Indigenous Knowledge Integration Committee
SD – Standard Deviation
TAC – Technical Advisory Committee
VEC – Valued Ecosystem Component
WBEA – Wood Buffalo Environmental Association

1.0 INTRODUCTION

1.1 Oil Sands Historical Monitoring

The development of Alberta's oil sands is a priority for the governments of Alberta and Canada. Development of the oil sands involves changes to the natural landscape including the removal of overburden and the loss and conversion of terrestrial and aquatic ecosystems, resulting in emissions to air, groundwater, and surface water (Charpentier et al., 2009; Environment Canada, 2012; Giesy et al., 2010; E. N. Kelly et al., 2009, 2010). The environmental effects of oil sands projects on land, air and water are assessed through social and scientific processes to ensure that anticipated effects are acceptable to the government before approval and licensing. Regardless of the assessment process, scientific uncertainties regarding the magnitude and nature of environmental effects always remain. Post-approval monitoring of the environment is therefore required to verify predictions made in the assessment process.

Canada's oil sands have been mined on a large scale since 1967 (Humphries, 2008). Monitoring in the oil sands region has been ongoing since 1919 when the Water Survey of Canada was established and commenced measuring flow volumes on the Lower Athabasca River (Cronmiller & Noble, 2018b.). Monitoring of other environmental components started in the 1960s and has continued through various federal and provincial initiatives. Alberta has been monitoring water quality of the Lower Athabasca River via the Long-Term River Network (LTRN) Program since 1968 (Cronmiller & Noble, 2018b.). The Alberta Oil Sands Environmental Research Program (AOSERP) was created in 1975 as an independent program managed jointly by Alberta and Canada to fund, direct, and coordinate environmental research in the oil sands region. Canada withdrew from the program in March of 1979. The program was operated by Alberta until 1985, ultimately producing over 220 reports, articles and papers describing natural environment (air, land, water) interactions with oil sands operations.

The Northern River Basins Study (NRBS) was co-funded by Canada and Alberta, with participation by the Northwest Territories. Over 4.5 years, $12 million was spent on understanding the relationships between industrial, agricultural, municipal, and other development pressures on the Peace, Athabasca, and Slave river systems. Some of the work was focused on the Athabasca River in the vicinity of Fort McMurray and oil sands operations. The NRBS culminated in a final report published in 1996, suggesting that oil sands operations were affecting wildlife, while mixed function oxidase induction was observed in fish from the Clearwater and Steepbank rivers. The report identified expanding oil sands operations as a perceived risk to aquatic organisms, while questioning the potential effects of oil sands related hydrocarbons on physiological responses in fish.

The Northern Rivers Ecosystem Initiative (NREI, 1998-2003; Arciszewski et al., 2018; Bailey et al., 2004; Bisset et al., 2018; Canadian Nuclear Safety Commission, 2016; Dowdeswell et al., 2010; Environment Canada & Alberta Environment, 2004) was a second five-year joint program between Alberta, Canada, and the Northwest Territories in partial response to the conclusion of the NRBS. For example, NREI programs included studies to: (1) quantify naturally occurring hydrocarbons in the Athabasca River; (2) quantify endocrine disruption in fish populations; (3) assess fish abnormalities; and (4) establish indicators for environmental monitoring.

In 1996 when NRBS was concluding, the province of Alberta required (via licenses) that all operating oil sands companies participate in a regional aquatics monitoring program. There was no requirement that companies work together. However, commencing in 1997, several companies (Suncor Energy Inc., Oil Sands, Syncrude Canada Ltd., and Shell Canada Limited) began a collaborative Regional Aquatics Monitoring Program (RAMP) to meet license requirements. The RAMP was intended to integrate monitoring activities and to identify potential cumulative effects that could be addressed (Golder, 1998). As more companies were licensed to mine oil sands, many (but not all) participated in the RAMP collective. Government representatives and scientists (from Alberta Environment, Environment Canada, Fisheries and Oceans Canada, for example) also participated in RAMP steering committees to ensure that the program was rigorous. Annual technical reports were developed to summarize data collected during the prior year, and to assess trends over time. Monitoring was focused on the aquatic environment, as required by the operating license, issued by the province. Generally RAMP monitoring was designed to assess the predictions that were made in environmental assessments that supported oil sands development (Hatfield Consultants et al., 2009). The RAMP committee met the license requirements through to and including 2013, after which Alberta and Canada assumed full responsibility of monitoring air, land, and water in the vicinity of oil sands operations through the Joint Oil Sands Monitoring (JOSM) Program (see below).

Environment Canada established the Panel on Energy Research and Development (PERD) in 1998. PERD operated for four years, conducting research focused on assessing the effects of naturally occurring hydrocarbons on aquatic receptors (Environment Canada & Alberta Environment, 2004).

The Cumulative Environmental Management Association (CEMA) was a multi-stakeholder advisory group that provided advice to the provincial and federal governments. CEMA's goals were to recommend management frameworks, best practices, and implementation strategies to address cumulative effects on air, land, water, and biodiversity (Cronmiller & Noble, 2018a). CEMA did not collect monitoring data, but the current water management frameworks under the Lower Athabasca Regional Plan are based largely on CEMA's work (Cronmiller & Noble, 2018b).

1.2 Critical Monitoring Reviews

Critical reviews of oil sands-related monitoring commenced around 2009. Kelly et al. (2009, 2010) documented metals and other constituents entering the Athabasca River mainstem and tributaries via snowmelt. Kelly et al. were critical of monitoring in the oil sands in general, but specifically of RAMP for failing to detect the influences of snowmelt on water quality of the mainstem Athabasca River. The Royal Society of Canada completed its review in late 2010 (Gosselin et al., 2010). That review recognized the important work by RAMP, WBEA and CEMA, but recommended that the activities of those groups be integrated. The Panel recommended: (1) that an agency be made responsible for integrating the activities of the various groups; and (2) that the collected data be used to implement "…*cumulative assessment, which requires… coordinated effort to review, analyze, and interpret regional data to set targets with a publicly accessible process that define cumulative capacity limits, as has been done with water use."*

At the time of the Royal Society Review, the RAMP regular 5-year program was being peer reviewed by an external committee of academics and government scientists with expertise in aquatic environment monitoring (Main, 2011). That review recommended (among other things): (1) expansion of the temporal and spatial scale of monitoring; (2) integration of RAMP with other monitoring programs including those assessing groundwater, air, and terrestrial habitats; and (3) integration of RAMP monitoring into a decision-making framework that clarifies when effects are deemed to exist, and what actions will be taken when an effect is discovered. Main (2011) suggested that a different governance system might be applicable if RAMP were integrated with other programs. Additionally, Main (2011) recommended that RAMP include a peer-review process, while having an external Review Committee of scientists to design the program, and an independent External Science Advisory Panel to provide continuous hands-on oversight.

A federal review was commissioned by the Minister of the Environment in 2010 which concluded that monitoring of the oil sands was a fragmented collection of activities and research that would benefit from a new governance model using a holistic and integrated approach to monitoring that was adaptive, founded on credible scientific approaches, transparent, and accessible (Dowdeswell et al., 2010). The Alberta Environmental Monitoring Panel (AEMP) was created in January 2011 as an independent entity reporting to the Minister. The AEMP concluded that monitoring in the oil sands needed a new system that (1) focused on environmental effects, (2) provided a consistent approach to monitoring air, land, water, and biodiversity, and (3) was overseen by an arm's length Environmental Monitoring Commission to provide organization and integration of activities.

1.3 Joint Oil Sands Monitoring (JOSM)

The Joint Oil Sands Monitoring (JOSM) Program was subsequently established in 2012 as a provincial and federal partnership. The JOSM program was renewed in 2017 through a Memorandum of Understanding as the Oil Sands Monitoring (OSM) Program and associated framework (M. Dubé et al., 2018). The OSM Program includes representatives from Alberta Environment and Parks (AEP), Environment and Climate Change Canada (ECCC), regional First Nations and Métis communities, environmental agencies, academic and research institutions, and industry stakeholders (see Figure 1). The Operational Framework Agreement outlines a commitment to Indigenous-designed and -led monitoring programs and the inclusion of Indigenous knowledge in monitoring programs. These organizations collectively provide comprehensive monitoring data, with the overarching goal of understanding the long-term cumulative effects of oil sands development on the natural environment. The OSM Program includes air and deposition, surface water, groundwater, terrestrial biology, wetlands, and integrated analytics, with technical advisory committees (TACs) overseeing different components. The TACs are accountable for designing monitoring programs to address the priorities determined by the Oil Sands Monitoring Program Oversight Committee (OC) and consist of representatives from government, industry, other non-government organizations, and Indigenous communities.

1.4 Oil Sands Monitoring Governance Structure

The governance structure for Canada's Oil Sands Monitoring Program is illustrated in the organization chart from the Operational Framework (Figure 1; Dubé et al., 2018). There are three levels of technical input that manage routine implementation and program design: (1) the Oversight Committee (OC); the Science and Indigenous Knowledge Integration Committee (SIKIC); and (3), the Technical Advisory Committees (TACs) and the Indigenous Community Based Monitoring Advisory Committee (ICBMAC). The OC is responsible for developing and adapting the risk-based framework that will guide monitoring, as well as defining baseline, and identifying key monitoring questions. The OC is also responsible for ensuring accurate and timely reporting of environmental conditions by the SIKIC, the ICBMAC and the TACs. The ICBMAC is responsible for developing, reviewing, and recommending the use and application of "limits of change" important to Indigenous communities in the region and within scope of the OSM Program. The TACs are responsible for overseeing monitoring, analysis of data and reporting on environmental condition relative to "limits of change", and for recommending focused studies or methods development studies to improve monitoring.

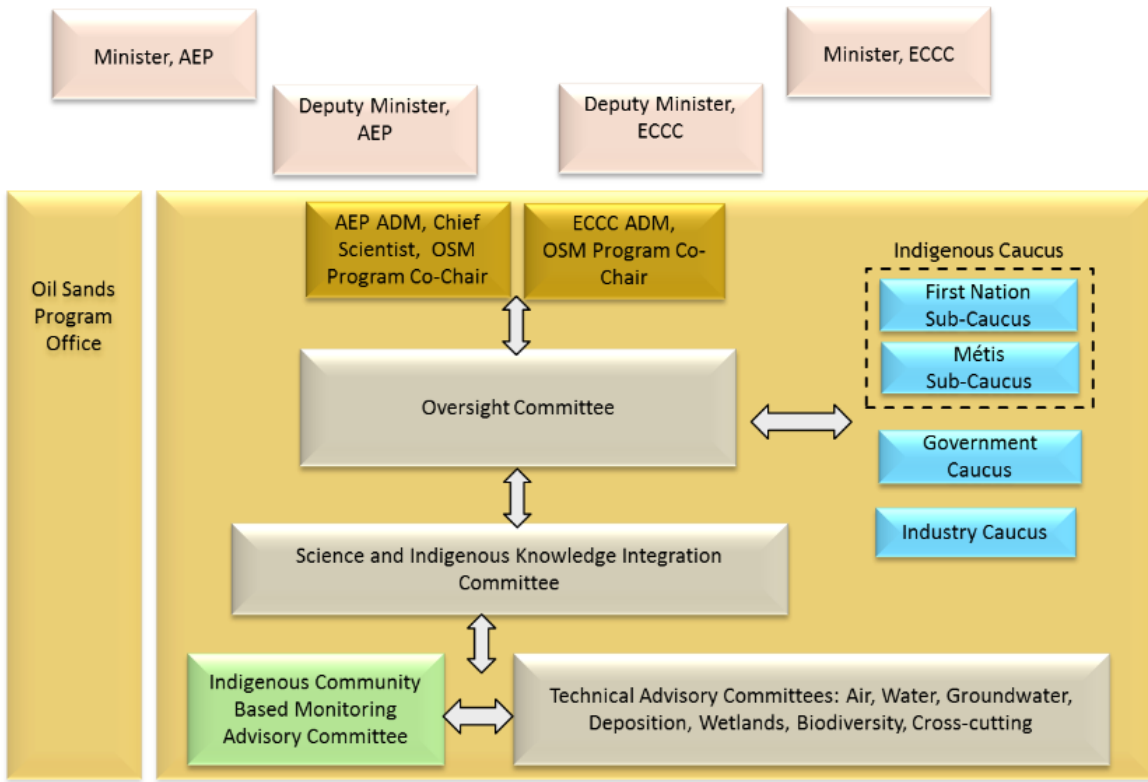

Figure 1 Governance structure of the Oil Sands Monitoring Program

Figure Notes: from Dubé et al. (2018).

1.5 Adaptive Monitoring

Adaptive monitoring of the ambient environment was suggested as a guiding principle for the OSM Program (Dubé et al., 2018) based on its successful application in the federal Environmental Effects Monitoring (EEM) programs within the Pulp and Paper Effluent Regulations and Metal and Diamond Mining Effluent Regulations (Hodson et al., 1996; Wrona & Schaefer, 2011; Arciszewski et al., 2017). In those contexts, adaptive monitoring is understood to be a process that uses quantitative threshold values to trigger changes in monitoring intensity or focus (Somers et al., 2018). That is, monitoring results may cause an increase in monitoring intensity (to confirm, or explore the extent, magnitude, or cause of unexpected effects) or a decrease in monitoring intensity if conditions are deemed stable and understood. Adaptive monitoring in Canada's oil sands is intended to focus monitoring effort on areas of greatest concern, where scaled monitoring efforts can help identify the causes of unanticipated, undesirable change and propose candidate solutions, while maintaining a base level of monitoring in areas of lesser concern (Dubé et al., 2018).

The 2021/2022 OSM workplans were required to incorporate adaptive monitoring principles (e.., Roberts, 2021). Despite the history of adaptive monitoring in aquatic environments (e.g., Hodson et al., 1996; Arciszewski et al., 2017), adaptive monitoring has not been generally or overtly applied in the oil sands region. Therefore, the 2021/2022 workplans described a variety of approaches to adaptive monitoring. The purpose of this report is to provide a generic framework based on adaptive monitoring principles, that can further harmonize adaptive monitoring methodologies used by OSM practitioners. To achieve this purpose, we first provide a review of ambient environment monitoring, and the various associated terms in present use. In this review, we provide examples generally from outside of the oil sands situation. However, some oil sands examples are included because they exemplify attributes of adaptive monitoring. Section 3.0 provides a recommended adaptive monitoring framework with a discussion as to how the existing OSM program can be adjusted to fully integrate adaptive monitoring principles.

2.0 AMBIENT MONITORING

Some of the challenges in communicating the notion of adaptive monitoring rest in how different research groups use and interpret monitoring terminology. The purpose of this section, therefore, is to clarify and define terms used in a proposed generic adaptive monitoring framework (see 3.0 below). Various terms have been developed to describe kinds of ambient environment monitoring. Section 2.1, provides a review of kinds of ambient monitoring, including the elements of adaptive monitoring. For example, indicators are attributes of the environment on which data are collected and that serve as the basis for monitoring and assessment. Section 2.2 discusses indicators, classification schemes, and approaches for selecting indicators in a monitoring program. Changes in monitoring activities can be triggered by observed values for an indicator that diverge from an expected benchmark value or exceed a threshold. In Section 2.3 we briefly review the various kinds of thresholds that are used to interpret monitoring data and that may be used to prompt or trigger changes in monitoring intensity or focus. Where relevant, we illustrate how the various terms are currently used in the Canadian monitoring context. Section 2.4 presents three case studies that use indicators and thresholds in an adaptive monitoring context.

2.1 Types of Monitoring

Adaptive monitoring implies changes in monitoring, or different kinds of monitoring to reflect uncertainties. The purpose of this section is to provide a review of the general kinds of ambient-environment monitoring approaches and to discuss how these approaches have been used within adaptive monitoring frameworks. Here, we review the following 'kinds' of monitoring:

- Compliance monitoring;
- Effectiveness monitoring;
- Baseline monitoring;
- Environmental effects monitoring, which includes:
    - Periodic monitoring:
        - Surveillance monitoring and,

        - Reduced monitoring;
    - Focused monitoring:
        - Confirmation monitoring;
        - Extent and magnitude monitoring;
        - Investigation of cause;
        - Investigation of solutions;
- Cumulative effects monitoring; and,
- Adaptive monitoring.

These terms often overlap in how they have been defined and used. Their use in the Canadian context is the focus of the next sub-sections.

2.1.1 Compliance Monitoring

In the context of the natural receiving environment, **compliance monitoring** is the process that ensures that organizations meet their environmental commitments as defined under licences or other operating certificates issued by agencies or governments (Environment Canada, 2011b). Releases to air and water are regulated and facilities that release sufficient quantities of regulated substances require a licence for release. These licences typically dictate permissible discharge limits on quantities and concentrations of constituents of concern based on well-known environmental outcomes. Monitoring at some frequency at the end of pipe or stack is normally a requirement of the licence, with exceedances of licence limits incurring consequences such as follow-up investigations or penalties and fines.

In Canada's metal mining sector, the following are regulated: metals including As, Cu, CN, Pb, Ni and Zn; suspended solids, Ra_{226}, unionized ammonia, and pH. Companies and agencies typically set up processes to try to avoid exceedances of licenced limits. In Canada's nuclear sector, monitoring of effluent quality (i.e., concentrations) is judged relative to licenced limits, but also to 'action levels'. Action Levels are concentrations of water quality constituents at different points in the treatment system that indicate when there is a potential loss of control of part of the effluent treatment and conveyance system and triggers a requirement for a specific action to be taken (Radiation Protection Regulations[1]). When Action Levels are exceeded, an investigation into cause of the exceedance is required, while actions for regaining control need to be identified and implemented as required.

In Canada's oil sands, stack emissions are licenced under the Environmental Protection and Enhancement Act (EPEA), and companies are required to monitor and report emissions for substances including SO_2, H_2S and NOx (Alberta Environment and Parks, 2021). Oil sands operators are not permitted to release process-affected water, currently. However, there are indications that process-affected waters are seeping into surface waters in some areas (Fennell & Arciszewski, 2019), which presents a challenge for monitoring

[1] https://laws-lois.justice.gc.ca/eng/regulations/SOR-2000-203/

and assessment. Because compliance monitoring applies at the scale of a facility, it is outside of scope of the OSM. In the context of this report, compliance monitoring is not further considered.

2.1.2 Effectiveness Monitoring

Effectiveness Monitoring is the process for determining if the goals of management activities are having the desired effects (Morrison & Marcot, 1995). Fisheries and Oceans Canada (DFO) uses effectiveness monitoring to determine if habitat mitigations, compensation, or offsets have had the desired effect: that is to determine if the productive capacity of the receiver has been maintained or enhanced (Fisheries and Oceans Canada, 2007). Effectiveness monitoring is also described under Environmental Effects Monitoring.

2.1.3 Baseline Monitoring

Baseline Monitoring documents ambient conditions relative to a baseline state that may be pristine or best available (most natural) condition representing what may be a desired environmental/ecological state. Baseline monitoring is normally conducted prior to projects being designed and built. As such, baseline monitoring data are often used to support environmental impact statements. The data may be used, specifically, to identify sensitive ecological areas to be avoided, and thus support situating projects. Properly designed, baseline monitoring programs and their resulting data can support subsequent assessment of project effects (Kilgour et al., 2007), and/or can support testing the accuracy of environmental outcomes that were predicted during the EIS stage (Somers et al., 2018).

Under the historical RAMP (Hatfield Consultants et al., 2009), three years of baseline water quality, sediment quality, benthic macroinvertebrate community, and fish population data were collected in tributaries prior to the development of oil sands operations. The intent of the baseline data collection was to characterize conditions immediately prior to oil sands development that would support future assessment of the effects of operations at a regional scale within watersheds for the major tributaries. Baseline monitoring is typically completed in areas that represent an undisturbed condition. However, baseline conditions in the oil sands region are not pristine because of historical mining activities, development of access roads, forestry activities, etc. (Roberts et al., 2022). Under RAMP, baseline monitoring data were typically collected from an area upstream where oil sands activities were not anticipated to occur for some time, as well as in areas anticipated to be influenced by oil sands related land cover change. However, the upstream reference areas often were overtaken by land cover change earlier than originally anticipated, essentially removing the spatial control (Environment Canada, 2011b).

A major purpose of baseline monitoring is to further our understanding of the relationships between the project, activity, associated stressor, and the response. Predictive models can provide a basis for estimating the magnitude and extent of anticipated project influences relative to baseline conditions, but often the nature of effects is expressed in simple narrative terms (Dipper, 1998). For example, the Terra Nova offshore drilling operation was predicted to impact deep-water benthos up to a radius of 1 km from the center of the operation (DeBlois et al., 2014). That statement was translated to a prediction that benthos community composition within 1 km of the drill center would differ significantly from reference areas, but beyond 1 km community composition would not differ significantly from reference areas. In effect, most narrative statements can be translated into testable hypotheses, which in the case of Terra Nova were evaluated during environmental effects monitoring (DeBlois et al., 2014).

Baseline conditions do not always need to be measured directly. In the absence of a measured baseline, environmental conditions can be estimated based on relationships between measured biophysical response and some indicator of oil sands operation. The Alberta Biodiversity Monitoring Institute (ABMI; Burton et al., 2014) and others have been modeling the relationships between bird species abundances and various measures of oil sands operational influences (Leston et al., 2020; Mahon et al., 2019; Saracco et al., 2022; Shonfield & Bayne, 2023). Similar modeling was completed with Trout-perch from the mainstem Athabasca River (Kilgour et al., 2019; McMillan et al., 2022), and benthic macroinvertebrates from tributaries to the Athabasca River (Arciszewski, 2021). Those models can be used to 'hindcast' the expected baseline condition assuming no development (i.e., no land cover change; Kilgour & Stanfield, 2006).

2.1.4 Environmental Effects Monitoring

Environmental Effects Monitoring (EEM) was established in Canada in 1992 under the Pulp and Paper Effluent Regulations (PPER) and subsequently under the Metal Mining Effluent Regulations (MMER, currently the Metal and Diamond Mining Effluent Regulations or MDMER). EEM was developed as a process to assess the effectiveness of the PPER and MMER for protecting aquatic receiving environments (Hodson et al., 1996). Monitoring is "effects" based, meaning it is focused on effects on biological responses. EEM, specifically, is focused on benthic macroinvertebrate communities and sentinel fish populations. Various study designs are used under EEM to test the significance of effects of released effluent on benthos and fish. Studies are conducted at typically 3-year intervals (Environment Canada, 2012).

The EEM program is an adaptive monitoring program with two major monitoring components, commonly described as Periodic and Focused monitoring, that are referred to as monitoring 'Tiers' (not 'types'; Arciszewski & Munkittrick, 2015; Environment Canada, 2012; Somers et al., 2018). The EEM program uses specified indicator thresholds to trigger a change in monitoring activity when a threshold is exceeded (i.e., the monitoring strategy changes according to the observed monitoring results – see below).

2.1.4.1 Periodic Monitoring

Periodic Monitoring is the general term for monitoring designed to regularly test for changes in indicators relative to baseline conditions or to test predictions from an EA (or similar) process. Predictions or forecasts based on the EA can be used to confirm that post-operational environmental conditions are consistent with model predictions (DeBlois et al., 2014; Dipper, 1998; Kilgour et al., 2007; Noble & Birk, 2011). Thresholds, such as limits of acceptable change (Cole & Stankey, 1997; Rogers et al., 2013; see below), can be based on these predictions. In effect, forecast conditions can be treated as hypotheses that are tested with monitoring data (Duinker & Greig, 2006).

The simple purpose of periodic monitoring then is to regularly test our expectations; that is, monitoring ambient conditions to see if conditions differ significantly from expectations (e.g., using models from baseline monitoring) given what we know about the presence of various pressures (e.g., oil sands mines, tailings ponds, aerial emissions, angling, roads, forestry, etc.). Often indicators are evaluated in isolation (i.e., individually; Davies & Jackson, 2006); however, Hewitt et al. (2005) have shown the added diagnostic value associated with the interpretation of multiple indicators. These studies suggest that conceptual models and forecast conditions should involve multiple indicators to embrace the potential complexity of the receiving environment.

In Canada's EEM program, periodic monitoring is termed "**Initial Monitoring**" (Ribey et al., 2002) and also "**Surveillance Monitoring"** (Arciszewski et al., 2017). The purpose of Surveillance Monitoring is to test for unexpected effects, which may be changes, or differences, from a baseline (reference) condition, or from an EIS-based (or other) model prediction (see Burton et al., 2014). Surveillance Monitoring typically involves an established sampling design and set of protocols that are repeated on a predetermined schedule (e.g., every 3 years). Facilities may transition to a reduced frequency of monitoring (e.g., every 6 years, **Reduced Monitoring**) if the observations from several cycles of Surveillance Monitoring are consistent with predictions (i.e., either no significant effects or effects ≤ specified indicator thresholds). By contrast, facilities may switch to **Focused Monitoring** if indicator thresholds are exceeded during Surveillance Monitoring (see below).

Other terms have been used to connote Surveillance Monitoring. Environment Canada's Lower Athabasca River water quality monitoring plan used the term Performance Monitoring, suggesting this monitoring approach is for verifying and/or validating whether EIS predictions are accurate (Environment Canada, 2011b). This application is consistent with surveillance as described by Burton et al. (2014). Here, and for clarity, we define Surveillance Monitoring as the stage in monitoring where the expected condition (the prediction) is tested.

In situations where surveillance monitoring data depart from model predictions, intervention may be warranted (Dubé, 2003; Fitzpatrick & Williams, 2020). It is important to note that a lack of fit between monitoring results and predictions may simply indicate an incomplete understanding of the fate and effects of the stressor(s) in the environment and a need for additional research and/or a different type of monitoring.

2.1.4.2 Focused Monitoring

Focused monitoring involves monitoring activities designed to improve our understanding and application of baseline models and predictions. If surveillance monitoring demonstrates that conditions are not as expected, this result indicates that we do not adequately understand the system. For example, if surveillance monitoring demonstrates that bird communities near an oil sands operation are not what we thought they should be, then our original/historical understanding of the effects of oil sands operations are inadequate. Focused monitoring provides the opportunity to collect additional data to refine the models and produce updated predictions that are consistent with periodic monitoring results.

Focused Monitoring, in Canada's EEM programs for pulp mills and metal mines, is initiated when Periodic Monitoring identifies significant effects, exceeding critical effect sizes, exceeding reference baselines, or exceeding predictions (i.e., whichever threshold triggers a change in monitoring). The first step in Focused Monitoring involves **Confirmation**, where Surveillance Monitoring observations are verified. If significant effects are not confirmed, Periodic Monitoring is resumed, otherwise Focused Monitoring evaluates the **extent and magnitude** of the effluent-related effects (on fish and/or benthos) followed by determining the effluent-related cause (i.e., **investigation of cause**; Arciszewski et al., 2017; Hodson et al., 1996; Somers et al., 2018).

In the case of pulp mills and metal mines, quantifying the **extent and magnitude** of effluent-related effects will involve more field studies, typically further afield of the point of effluent release. In the case of oil sands, we may be examining the extent and magnitude of effects related to land cover change, for example (Arciszewski, 2021; Burton et al., 2014). Identifying the effluent-related (or land-cover change related) cause of effects typically requires a different battery of experimental tests, such as laboratory or mesocosm studies (Environment Canada, 2012). However, **Investigation of Cause** studies can sometimes benefit from the collection of additional field data to better describe the receiving environment, including such factors as diurnal variations in dissolved oxygen, or seasonal variations in nutrient levels and water temperatures (i.e., potentially controlling factors that are not routinely documented in a typical EEM field program).

After satisfactorily determining the underlying cause of effects, monitoring may return to the Periodic tier, or it may transition to **investigation of solutions**, **mitigation**, and **effectiveness monitoring** if mitigations were implemented. In the Canadian metal mining experience, mitigations are typically not required. As such, Focused Monitoring concludes when we have a new understanding of what to expect (i.e., an updated model with updated predictions). In the absence of further changes to the underlying cause of effects (land cover or effluent loads, in our example), we should anticipate that previously observed effects will persist into subsequent Periodic Monitoring. If Periodic Monitoring results are consistent with the updated model predictions, that outcome simply confirms our understanding of the observed effects and monitoring continues periodically (Environment Canada, 2012).

In the OSM program the term "surveillance" is used to specify sampling stations that have been set up to address hydrological complexity, and to otherwise better understand contaminant dispersion, deposition and degradation (Environment Canada, 2011a). In that context, the OSM use of the term "surveillance" may be more akin to extent and magnitude monitoring described here.

2.1.5 Cumulative Effects Monitoring

Traditional EEM assesses the environmental effects of single point-source stressors. Receiving environments, however, are rarely influenced by single anthropogenic stressors, and there is often a need to evaluate the effects of multiple stressors in monitoring and assessment. In receiving environments like the Athabasca River (and its tributaries), stressors include land cover change associated with mining, built infrastructure including buildings, ponds and roads that interfere with wildlife movements, mined areas from which dust can change the quality of soils and surface waters, and expanding urban areas (among myriad other factors; Arciszewski, 2021).

Cumulative Effects Monitoring has been identified as a method for assessing the cumulative effects of multiple stressors (Duinker & Greig, 2006). Dubé (2003) proposed the pairing of stressor-based and effects-based indicators in a cumulative effects framework to monitor and assess the cumulative effects of multiple stressors. Stressor-based monitoring focuses on engineering-based estimates of chemistry-related inputs (effluents/emissions) to environmental systems. Stressor-based monitoring is predictive, but modelling the ecological consequences of predicted concentrations can be challenging. Monitoring stressors in the receiving environment confirms the modelled predictions of multiple constituents of concern, but generally lacks ecological relevance. By contrast, effects-based monitoring focuses on biological responses and is the ultimate test of the ecological significance of physico-chemical changes to the system. Effects-based monitoring is retrospective, and as such has difficulty identifying the cause when there are multiple potential causes. Dubé (2003) proposed that both stressor- and effects-based indicators are needed to effectively monitor and assess cumulative effects. As noted above, predictions (or forecasts) from multiple-stressor dose-response models can serve to establish cumulative effects thresholds.

2.1.6 Adaptive Monitoring

Adaptive monitoring implies the adjustment of monitoring approach and focus over time based on past results and current information needs. Adaptive monitoring has also been described as a process in which monitoring evolves over time to address new questions and uncertainties, and to incorporate new methods and protocols as they become available (Constable, 1991; Lindenmayer & Likens, 2009; Sutter et al., 2015). Cairns et al. (1993) highlighted the value of adaptive monitoring and showed how the detection of unacceptable results via routine surveillance monitoring would trigger follow-up confirmation, then diagnosis of cause (Figure 2). Where the cause is understood, the effects may be mitigated, and follow-up monitoring designed to confirm the effectiveness of mitigation.

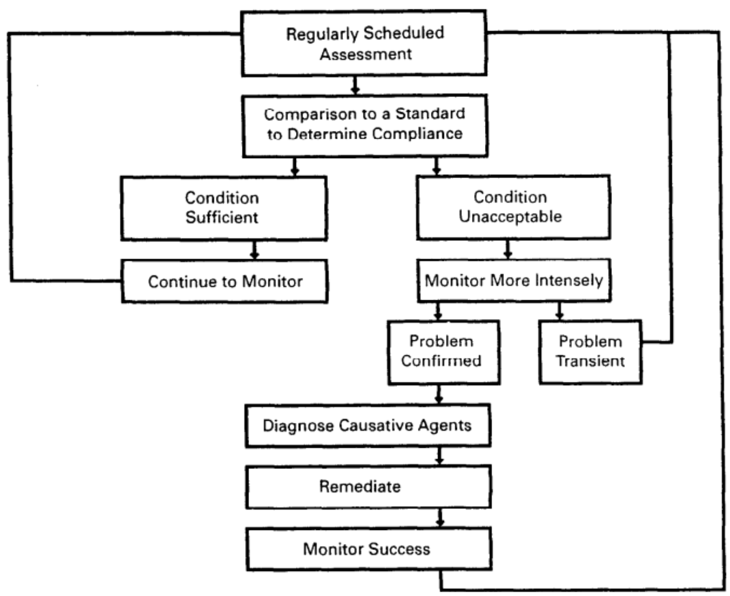

Figure 2 An example adaptive monitoring flowchart

Figure Note: From (Cairns et al., 1993)

Canadian EEM programs have followed the general flowchart described by Cairns et al. (Environment Canada, 2012; Hodson et al., 1996). That is, regular surveillance in EEM compares exposed fish and benthic communities to a reference threshold. Significant differences trigger confirmation. Confirmed differences that are large trigger investigation of cause. The adaptive nature of the Canadian EEM framework has been described several times (Arciszewski et al., 2017; Hewitt et al., 2005; Somers et al., 2018). Somers et al. (2018) described adaptive monitoring as a process of confirming predictions derived from environmental assessment (EA) or similar processes. Through the EA process baseline data are collected, the project is designed, and the project effects on the environment are predicted. The monitoring program is then designed to confirm the predictions from the EA. Where predictions are demonstrated via monitoring to be incorrect, focused monitoring ensues to (1) confirm the unexpected effect, (2) determine the cause, (3) investigate solutions, (4) implement solution, and (5) modify forecast. The cornerstone to adaptive monitoring then is monitoring to assess predictions, which requires that there be *a priori* expectations for conditions in the receiving environment.

2.2 Indicators Used in Monitoring

Although monitoring involves the collection of data, often large amounts of data, not every datum is routinely evaluated (Barbour et al., 2006; Munkittrick et al., 2019). That is, a subset of the data is rigorously summarized to draw conclusions regarding the status and trends of the resource (Johnson & Hering, 2009; Dunn & Bakker, 2011). These key parameters are often called indicators because they are representative (or indicative) of a general class of measurements (Griffith, 1998; Niemi & McDonald, 2004; Davies & Jackson, 2006). Indicators range from within-organism biochemical and physiological processes to whole ecosystem changes, although indicators based on populations, whole organisms, and specific tissues are most common (Munkittrick et al., 2019). The ecological relevance of an indicator increases with ecological complexity, as does the response time, but the likelihood of reversing change decreases. Consequently, there is a trade-off between ecological relevance and mechanistic understanding, with greater difficulty in demonstrating cause and effect when using indicators exhibiting greater ecological complexity.

The selection of indicators for monitoring is largely philosophical and, in many ways, depends on what is being managed / protected, and what can be monitored. Regulations and stakeholder processes can be valuable for identifying the key environmental values that should be protected, while conceptual models can support the selection of indicators for monitoring and assessment. The factors involved in the selection of indicators are discussed below.

2.2.1 Conceptual Site Model

The suite of indicators chosen for use in a monitoring program are typically identified after formal (or informal) consideration of a conceptual site model. Conceptual models are typically stressor-based vector diagrams showing the hypothesized pathways associated with the fate and transport of specific stressors or effects-based flowcharts identifying the stressors and associated responses of selected ecosystem indicators (Ankley et al., 2010; Davidson et al., 2020). For larger projects (e.g., Atkins et al., 2011; Harwell et al., 2019), conceptual models may be constructed using formal approaches such as the DPSIR framework that focuses on Drivers, Pressures, State changes, Indicators, and Responses. In the DPSIR framework (Timmerman et al., 2011), anthropogenic processes or activities are considered drivers (e.g., oil sands operations) that release pressures (e.g., land cover change, or changes in hydrological conditions) that may cause state changes (e.g., a decline in water quantity) that alter effects indicators (e.g., a decrease in biodiversity) and elicit a management response. That is, the DPSIR framework combines stressor-based and effects-based pathways into a single construct.

Roberts et al. (2021), for example, used a modified DPSIR framework to develop a conceptual model for the Oil Sands Region that links development-related pressures and associated stressors with responses associated with valued ecosystem components (Figure 3; also see Davidson et al., 2020). The model underscores the fact that stressor pathways often involve interactions that elicit changes in multiple indicators. With these sorts of models, monitoring may focus on pressures, stressors, pathways (diagnostic variables), and effects indicators (Roberts et al. 2021). The DPSIR framework provides a structure for classifying potential variables to be monitored. Further, it forces the overt identification of valued ecosystem components which may be important when consulting stakeholders.

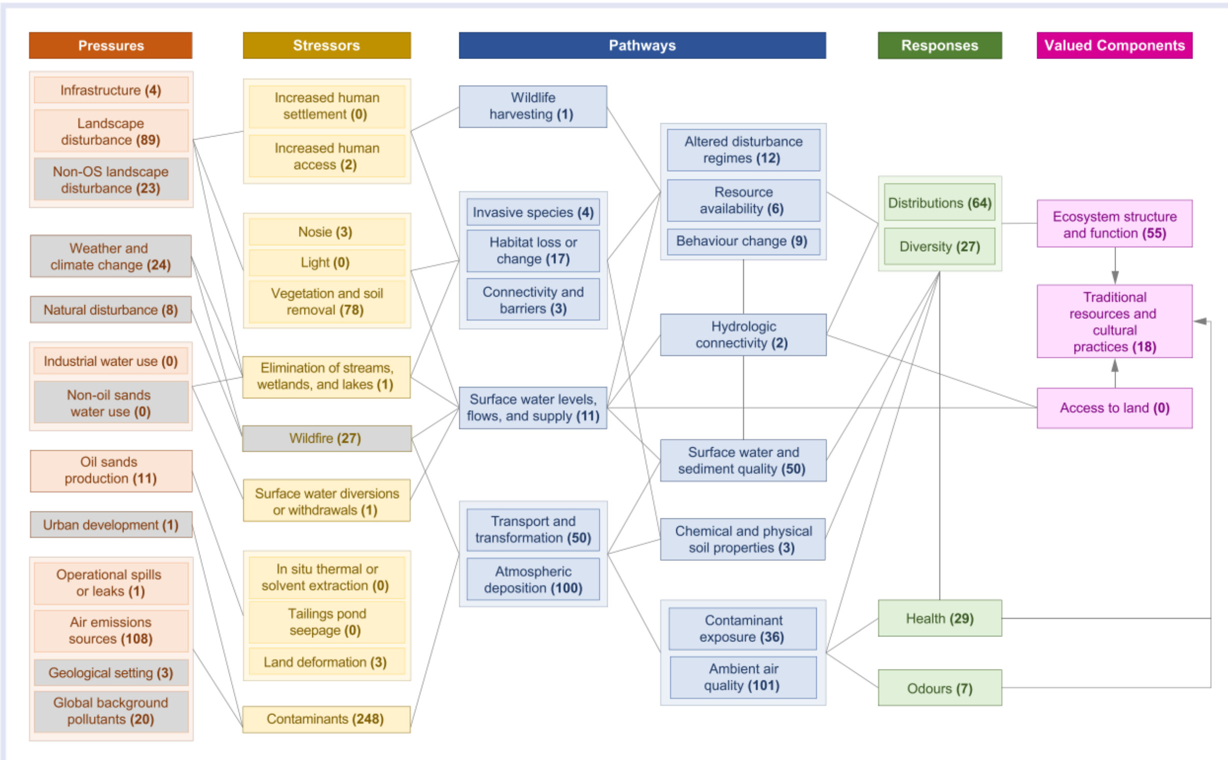

Figure 3 The DPSIR-based conceptual model as applied to the oil sands

Figure Notes: from Roberts et al. (2021)

Other types of conceptual models can include simple cartoon drawings illustrating the project within the ecosystem (Figure 4; e.g., Davidson et al., 2020). This type of model can show the project positioned on the landscape relative to major features like lakes and rivers. A conceptual model can also show potential inputs and outputs and their anticipated effects on the ecosystem, both locally and far away. These types of models are useful for demonstrating how the project will affect the environment (e.g., using adverse outcome pathways, Ankley et al., 2010) and where knowledge gaps and uncertainties of the relationship between the project and environmental response exist (Lindenmayer & Likens, 2009).

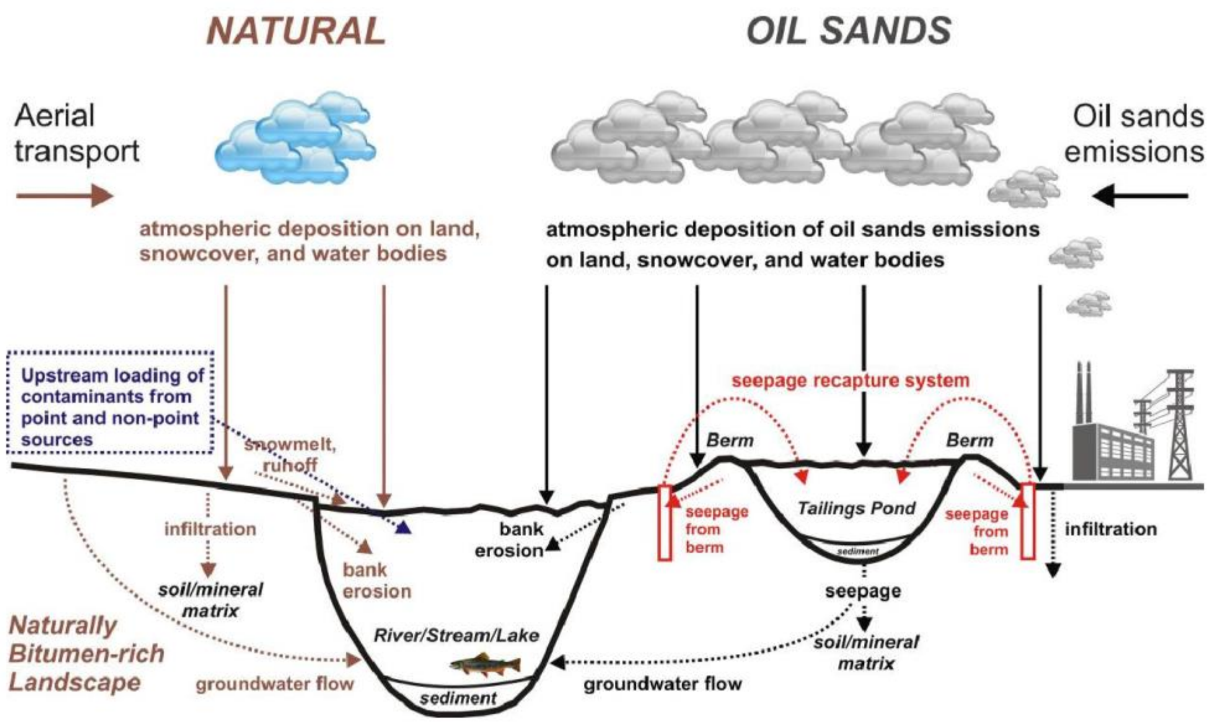

Figure 4 Example of a cartoon-based conceptual model with anticipated effect pathways for oil sands project interaction with surface waters

Figure Notes: Environment Canada (2011b)

2.2.2 Valued Components

The foundation of a monitoring program is based on the valued ecosystem components that monitoring is designed to evaluate and protect. Key environmental values (or valued ecosystem components, VECs; Figure 5**Error! Reference source not found.**) can be identified from Federal and Provincial Acts, Regulations and Policies, or from stakeholder processes. The federal Fisheries Act, for example provides protection to fish and fish habitat, while the policies of Fisheries and Oceans Canada (DFO) more specifically aim to maintain the Productive Potential (e.g., mass of fish produced per unit area of habitat; Fisheries and Oceans Canada, 2007, 2019) of fish habitat. The Fisheries Act and associated policies therefore provide guidance as to ecosystem components requiring protection. The Canadian Species at Risk Act (SARA) also protects specific species, which again provides some instruction as to key environmental values that would require protection.

Stakeholder processes such as those as part of Environmental Assessment (EA) can be valuable in identifying regionally important ecosystem components that may not have legal protection. In an EA process, stakeholders include regulators, Indigenous communities, the general public, and industry. Members of each of those groups may have knowledge about important ecosystem components for which protections may be desired, and for which loss or impairment would be considered unacceptable. The public consultation process in EA provides an opportunity for stakeholders to review project plans and associated conceptual models, and to identify valued components of the ecosystem that should not be harmed by the project (Ball et al., 2013; Duinker & Greig, 2006; e.g., see Figure 5). Valued ecosystem components may be critical species (e.g., lake trout, caribou), landscape features (e.g., spawning or calving areas), or activities (e.g., access to fishing areas; Stevenson, 1996). Generally, changes to VECs are unacceptable and anticipated effects of the project on VECs should be avoided, mitigated, or otherwise offset.

When a new facility is planned, the project may require authorization under the Canadian Impact Assessment Act (Sadler, 1995; Spellerberg, 2005). Canadian environmental assessments involve a series of steps that require ambient monitoring to document baseline environmental conditions (i.e., the collection of pre-development data) to understand regional landscape features, local geography, and ecology (Griffith, 1998; Duinker & Greig, 2006). This information is often used to develop a conceptual model that positions the project within the environment. The conceptual model usually predicts the potential effects of the project on the local environment (e.g., adverse outcome pathways, Ankley et al., 2010; Arciszewski et al., 2017). The project's conceptual model and predicted effect pathways are reviewed through a consultation process that includes regulators and other stakeholders (Constable, 1991). When unacceptable effects are anticipated, the proposal returns to an earlier point in the process to redesign the project to avoid or mitigate unacceptable effects. The stakeholder consultation component can frequently involve several iterations that result in updated project plans and revised conceptual models to accommodate stakeholder concerns (Duinker & Greig, 2006; Morgan, 2012; Baker & Westman, 2018). The conceptual model and adverse outcome pathways help the proponent and stakeholders to understand where the project will be located, how the project will interact with the landscape, anticipated environmental effects, and how those effects will be mitigated to ensure that residual effects are acceptable and balance environmental and socioeconomic needs.

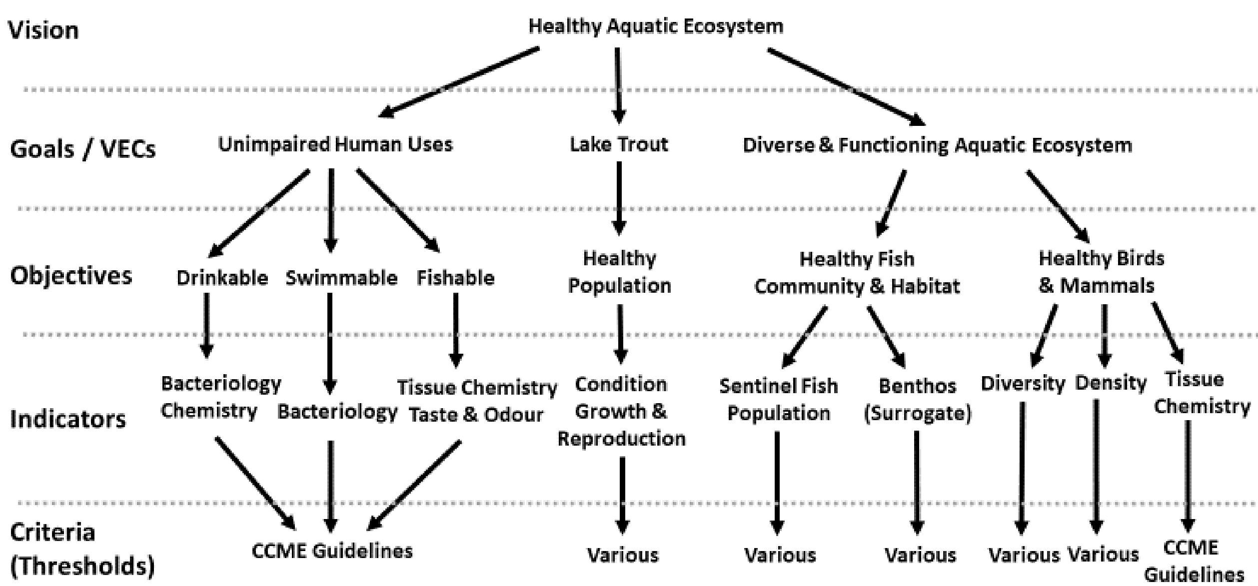

Figure 5 Relationships between project vision, goals, objectives, indicators, and criteria for aquatic systems

Figure Notes: modified from Kilgour et al. (2005, 2007)

In some situations, VECs may be defined narratively (Ball et al., 2013; Figure 5). For example, it may be important to preserve "swimmable", "fishable", and "drinkable" conditions downstream of a facility.

These types of VECs are often called ecosystem services and they can be difficult to incorporate into conceptual models and project effect pathways (Karr, 1991; Carson & Mitchell, 1993; Bell et al., 2017). Additionally, some VECs are symbolic, being relatively rare, charismatic, enigmatic, keystone, endangered, or culturally significant (Stevenson, 1996; Duinker & Greig, 2006).

Monitoring VEC indicators is logical because it is those entities that we ultimately care about. However, if only the VEC is monitored, there is the potential that monitoring fails to detect effects until after they have occurred or (worse) are irreversible. As a result, when the relationship between pressure/stressor and VEC response is well understood, monitoring programs tend to focus on stressor and pathway indicators because they provide early warning of impending effects on the VEC with the potential to mitigate effects before they are irreversible (e.g., Kilgour et al., 2005). When the relationship between pressure/stressor variables and the VEC response is uncertain, then monitoring of some surrogate for the VEC can be required.

2.2.3 Stressor and Effects-Based Indicators

Dubé & Munkittrick (2001) distinguish stressor-based and effects-based environmental monitoring. That is, stressor-based monitoring is focused on stressors (i.e., diagnostic variables, see Figure 5) that have a known relationship (numeric or narrative) with project activities. The stressor-based approach is predictive, requiring detailed knowledge of the stressor, dose-response relationships involving the stressor, as well as guidelines or criteria to identify when the stressor poses risk to VECs (Dubé, 2003). Monitoring focuses on changes in levels of the stressor in the receiver.

By contrast, effects-based monitoring focuses on biological receptors in the receiving environment. Effects-based approaches focus on monitoring changes in a biological receptor and identifying the potential cause of those changes, *a posteriori* (Dubé, 2003). Effects-based methods are regularly employed when the relationship between project activities and the VEC are imperfect. Canada's EEM programs for pulp mills and metal mines have an 'effects' focus largely because the effluents released by mills and mines were deemed complex, such that the ability to predict effects on valued components based on chemistry alone was overly challenging. Mill and mine effluents contain a number of constituents including suspended solids, and organic and inorganic substances for which our understanding of toxicological effects are well described. For those well-documented substances, making predictions of field effects can be considered reasonable. However, mill and mine effluents also contain other substances for which our toxicological knowledge is imperfect. It is those other substances that makes effects-based monitoring most valuable. With complex effluents, monitoring biological indicators that either are a VEC or are surrogates for a VEC ensures that we will be forewarned of conditions that threaten the VEC condition. Ideally, a combination of stressor- and effects-based indicators provide the greatest certainty that monitoring will detect changes before VECs are compromised, particularly when multiple stressors are involved.

2.2.4 Surrogate Indicators

Effects-based monitoring will rarely focus specifically on the VEC because harm to VECs should be avoided. That is, direct monitoring of VEC condition is not recommended since the monitoring process has the potential to impair the VEC (Kilgour et al., 2005, 2007; Harwell et al., 2019; McMaster et al., 2020). Furthermore, monitoring may not provide early warning to protect a VEC such that effects are detected only after impairments are irreversible (Constable, 1991; Kilgour & Barton, 1999). Consequently, alternate ecosystem components may be selected as **surrogate indicators** (sensu Cairns et al., 1993) that can be monitored to provide evidence of the condition of VECs (e.g., benthos, Kilgour & Barton, 1999; Niemi & McDonald, 2004; Wiens et al., 2008; Culp et al., 2020). Surrogate indicators may have enhanced sensitivity, providing an early warning of unanticipated effects on VECs (Kelly & Harwell, 1990; Cairns et al., 1993). Alternate ecosystem components may also be selected as **sentinel indicators**. "Sentinel" fish populations in Canada's EEM programs, for example, are monitored as a replacement for the broader fish community (Kilgour et al., 2007). Surrogate and sentinel indicators are typically selected in part because they will respond more quickly than the VEC to anticipated stressors and will therefore provide early warning of impending changes. As such, surrogate/sentinel indicators would also be early-warning indicators (Cairns et al., 1993; Kilgour et al., 2007; Kilgour & Barton, 1999).

2.2.5 Pathway/Diagnostic Indicators

Other classes of indicators can be 'pulled' from the conceptual site model. Pathway indicators are those variables that connect the stressor to the effect. Pathway variables may be able to support an evaluation of cause if effects are identified in a sentinel or surrogate indicator. Cairns et al. (1993) proposed that indicators that can be used to diagnose the cause of effects are "diagnostic" indicators. That is, diagnostic indicators provide insight into the cause of VECs being in a degraded condition. Cairns et al. (1993) proposed that the list of potential causative agents can be reduced by correlating effects with levels of the diagnostic variable. If the diagnostic variables are concentrations of water or sediment quality constituents (e.g., stressors; see Figure 3, Figure 5), then correlated biological responses with constituent concentrations may reduce the number of potential causes. Diagnostic indicators may include biological responses. Munkittrick & Dixon (1989) and Gibbons & Munkittrick (1994) demonstrated how variations in responses patterns of sentinel fish populations can be used to indicate the underlying cause of effects to those populations, for example. Similarly with benthos communities, it is quite common to use changes in community composition to diagnose the most likely causes of those changes. For example, an increase in total numbers of benthos and decrease in number of taxa is frequently associated with nutrient enrichment (Culp et al., 2020; Lowell et al., 2000).

In the federal EEM programs, multiple indicators are monitored to evaluate the environmental effects of pulp and paper mill and metal mine effluents (Walker et al., 2002). Hewitt et al. (2005) provided a diagnostic synthesis that identified responses of indicators, or subsets of indicators, to exposure pathways. For example, nutrient enrichment was often accompanied by an increase in condition factor, liver size and gonad size in sentinel fish species, whereas benthos abundance increased, and taxa richness was variable relative to reference conditions (Figure 6). By contrast, contaminant exposure pathways generally resulted in decreases in sentinel fish condition, liver size, and gonad size, and decreases in benthos abundance and richness. Meta-analyses of large EEM data sets suggested that these indicator response patterns were sufficiently robust to infer the predominant exposure pathways contributing to the observed effects (Hewitt et al., 2005).

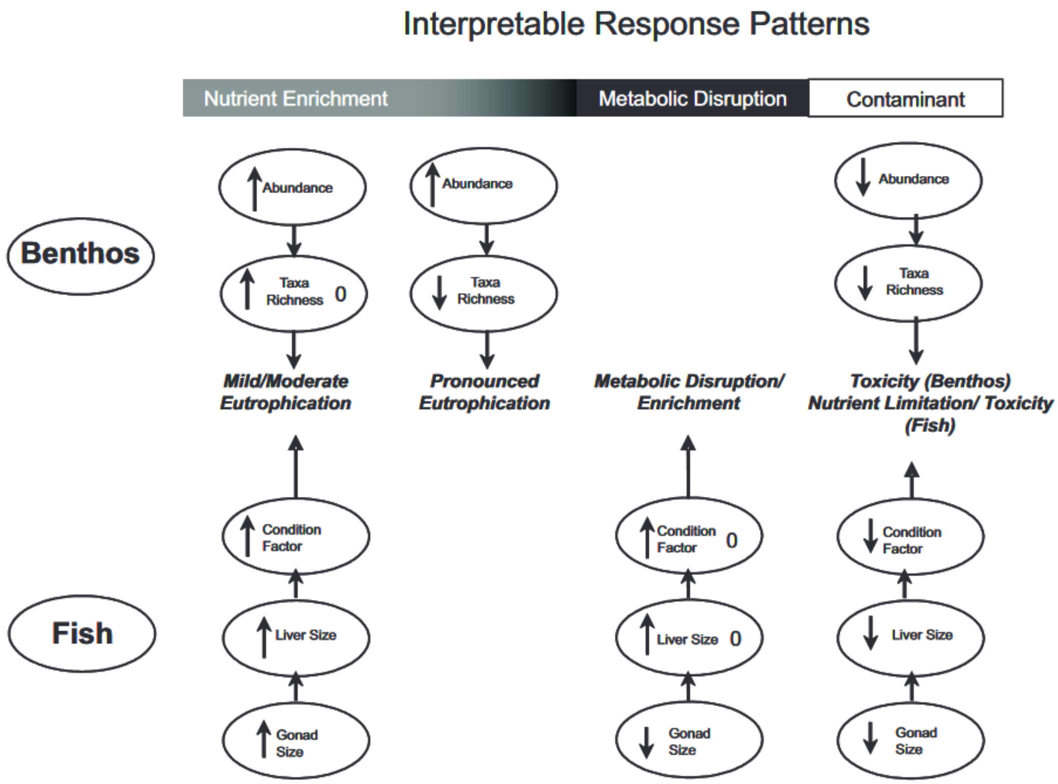

Figure 6 Diagnostic response patterns for EEM surrogate benthos and sentinel fish indicators

Figure Notes: Figure 2 from Hewitt et al. (2005)

Unfortunately, diagnostic response patterns are largely unavailable for most indicators (Gibbons & Munkittrick, 1994; Niemi & McDonald, 2004). Moreover, the responses of indicators are frequently evaluated in isolation (i.e., individually; Davies & Jackson, 2006). Hewitt et al. (2005) have shown the added diagnostic value associated with the interpretation of multiple indicators. These results suggest that conceptual models and forecast conditions should involve multiple indicators to embrace the potential complexity of the receiving environment.

2.2.6 Indicators based on Models and Forecasts

Where the system is quantitatively well described, quantitative models can be used in scenario analysis to forecast effects of projects or activities (Duinker & Greig, 2006). In these situations, a knowledge of the stressor gradient and the dose-response pattern can produce detailed predictions of stress and the anticipated biological responses to the gradient.

The response of an indicator to a stressor concentration gradient is known as a dose-response curve (see **Error! Reference source not found.**A, 7B; Davies & Jackson, 2006; Sánchez-Bayo & Goka, 2012). Historically, dose-response curves have been used to develop water-quality criteria for receiving waters (Dunn & Bakker, 2011; Gaudet et al., 1995), although species sensitivity distributions (SSD) are also used for this purpose (Belanger et al., 2017; Fox et al., 2021). Dose-response curves can be used to forecast environmental effects for non-traditional indicators like benthos abundance or density (i.e., number of organisms per m^2 sampling area; Figure 7C; Gore et al., 2001; Gray, 1989; Sánchez-Bayo & Goka, 2012). When a dose-response curve is available (e.g., **Error! Reference source not found.**D; Pearson & Rosenberg, 1978), this information can be used in the conceptual model to predict values for a given indicator spatially relative to the project footprint.

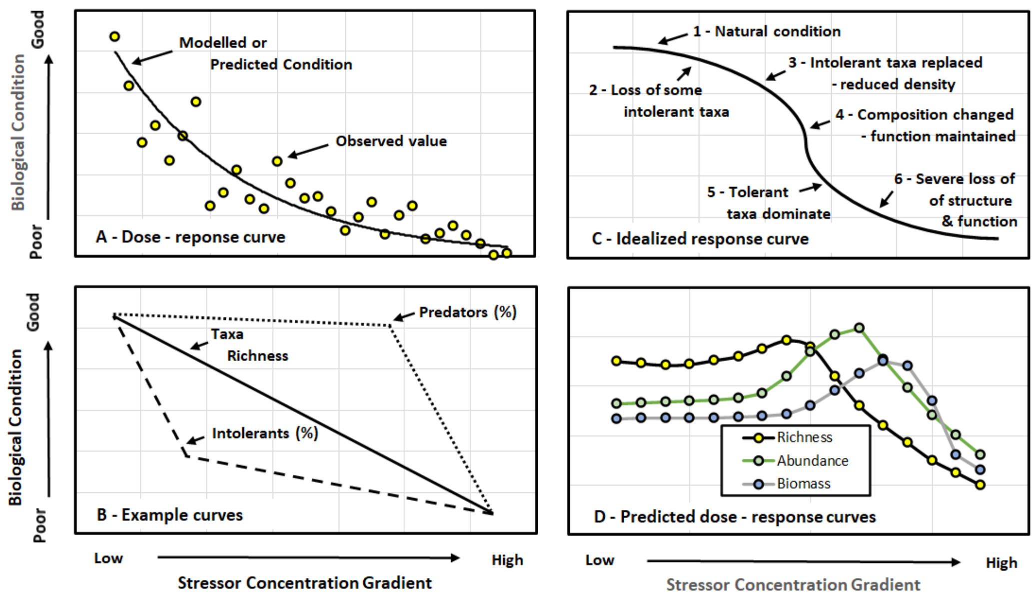

Figure 7 Example dose-response curves linking stressor and effect indicators

Figure Notes: from Karr (2006), Davies and Jackson (2006), Pearson and Rosenberg (1978)

Conceptual models may be tasked with the challenge of incorporating cumulative effects associated with nearby projects that may impinge on the environmental footprint of the new facility (Duinker & Greig, 2006; Squires & Dubé, 2013; Wong et al., 2019). Generally, these types of models include additional terms to accommodate zones of overlap among multiple stressors. Typical cumulative-effects models are accumulative models, where effects are simply added together (Cronmiller & Noble, 2018a, 2018b.). Predicting the interactive effects of multiple stressors from different sources is challenging (Brown et al., 2013; Dale & Beyeler, 2001; Gieswein et al., 2017), often because the interactions are not well known, and a thorough understanding of biological responses to combinations of stressors is largely lacking (Culp et al., 2020; Gaudet et al., 1995; Pilière et al., 2014). Recent progress with cumulative-effects models (Jones et al., 2017; Mahon et al., 2019; Spears et al., 2021) suggests that more detailed predictive models are possible. As a result, monitoring could accommodate more complicated cumulative-effects predictions based on conceptual models that incorporate the interactive effects of multiple stressors and the anticipated biological responses to those stressor interactions.

2.3 Indicator Thresholds

Various terms are used to capture the notion of a numeric or narrative value that prompts a change in course in monitoring, such as threshold, trigger, and benchmark. The term trigger can be used as both a noun or verb, leading to some confusion as to whether a trigger is a thing or an action. Here we use the term threshold, which per the Meriam-Webster dictionary is "a level, point, or value above which something is true or will take place and below which it is not or will not". Exceeding a threshold, therefore, implies that we have exceeded a value (numeric or narrative) for an indicator that we would otherwise wish to not exceed, and thus potentially 'trigger' a change in monitoring in the context of this framework.

2.3.1 Types of Thresholds

The Somers et al. (2018) adaptive monitoring framework proposed the use of indicator thresholds, numeric values that when exceeded trigger a change in monitoring focus. The OSM operational framework intends to use "limits of change" to trigger changes in monitoring focus and intensity. The phrase "limits of change" is relatively new and captures a broad set of terms capturing the notion of numeric values or narrative statements that are used in monitoring programs to guide the interpretation of data. The purpose of this section is to review the various terms that can be classified as 'limits of change', and that have been used in Canadian monitoring programs for interpreting data, making decisions, and triggering changes in monitoring (Arciszewski et al., 2017; InnoTech & Kilgour & Associates Ltd., 2019; Munkittrick et al., 2009, 2019; Munkittrick & Arciszewski, 2017; Somers et al., 2018). We will show below that there are numerous terms that have been used in a similar context and that collectively, they are all *thresholds* that can be used to inform or adjust monitoring activities. The material below is taken largely from our previous report on Limits of Change (InnoTech & Kilgour & Associates Ltd., 2019). The list of terms that may involve potential thresholds includes:

1. Goals
2. Desired Outcomes
3. Standards
4. Performance Standards
5. Limits
6. Action Levels
7. Guidelines
8. Objectives
9. Site-Specific Objectives
10. Targets
11. Ambient Limits
12. Benchmarks
13. Science-Based Environmental Benchmarks
14. Site Performance Objectives
15. Critical Effect Sizes
16. Normal Ranges
17. Natural Ranges and,
18. Threshold of Unacceptable Change.

Goals and ***Desired Outcomes*** are broad narrative statements about the intended environmental result of government Acts, Regulations and Policies (Alberta, 2012, 2012b, 2012c, 2012d, 2015a). Goals and Desired Outcomes are generally not associated with measurable indicators and so are not easily classified as thresholds that can trigger a change in monitoring, but the terms are listed here to provide perspective and links between measured entities and higher-level goals and objectives. Goals and Desired Outcomes can be considered likely to be achieved if: (1) facilities are 'performing' according to their operational standards, performance standards, and regulatory limits; and (2) if environmental conditions are as expected (currently and in the future) as stipulated in planning documents.

The term ***Standard*** is typically used to define the values of contaminants that can be emitted by various industrial or commercial processes, i.e., to define the 'performance standard' of manufactured goods or facilities (Canadian Council of Ministers of the Environment (CCME), 2012). ***Performance Standards*** in Alberta are numeric values that set performance criteria of a sector or facility. Canada for example has set a greenhouse gas performance standard for industry relative to 2005 levels (Alberta, 2015a). Performance Standards are alluded to in the Lower Athabasca Regional Plan (LARP) and the LARP Air Quality Management Plan (Alberta, 2012). *Standards* can also apply to environmental ambient conditions (i.e., CCME Canada-wide environmental standards; see Figure 5). Canadian Council of Ministers of the Environment (CCME) (2012) uses the term "Air Quality Standard" as health and environment-based numerical values of outdoor air concentrations of pollutants. In the sense of defining a minimum environmental quality, the term Standard is similar to the use of the term Guideline as is applied to other environmental media. *Standards* can also apply to concentrations of constituents in consumer goods. Health Canada is responsible for setting nutritional quality standards, which are enforced by the Canadian Food Inspection Agency.

Limits are values applied in Licenses (e.g., License Limits) and that define the allowable releases of constituents from facilities. Alberta, along with other provinces, territories and federal departments, license emissions to air and water, and the consumption of water under various Acts and Regulations (e.g., see Table 2 in Alberta, 2015a). *Limits* are developed based on protection of ambient air, water, or soil receiving the effluent. *Limits* are sometimes derived by calculating the concentrations/loads that can be released while still achieving ambient environmental quality guidelines. For air, *Limits* for aerial emissions are based on results from dispersion models, and with the intention of meeting Alberta Ambient Air Quality Guidelines or some other Standard at the edge of an industrial property. For liquid effluents, *Limits* are developed using the procedures detailed in Alberta Environment (1995). Water quality-based effluent *Limits* may also be the basis of licensed mixing zones (i.e., an area in the receiving environment that may have concentrations of constituents of concern that exceed water quality *guidelines*). Chemical specific *Limits* involve restricting release concentrations to those that will meet instream quality *guidelines*. Other approaches are possible for setting *Limits* for liquid effluent quality based on whole effluent toxicity or biological studies (e.g., benthos community surveys).

In Canada's nuclear sector, ***Action Levels*** are radiation doses or concentrations of constituents of concern that if reached indicate a potential loss of control[2]. As it relates to releases of aerial emissions or liquid effluent, exceedance of an Action Level prompts evaluation of cause, and identification and implementation of remedial measures to regain control (as necessary).

Guidelines are recommended narrative statements or numeric quantities for a substance that are intended to protect ecosystem values (e.g., human use or consumption, or ecological receptors). The Canadian Council of Ministers of the Environment (CCME) develops and publishes Canadian Environmental Quality Guidelines (CEQGs) for drinking water quality, recreational water quality, water quality for the protection of aquatic life, water quality for the protection of agricultural water uses, sediment quality for the protection of aquatic life, soil quality for the protection of environmental and human health, and tissue residue guidelines for the protection of wildlife consumers of aquatic biota (see Figure 5). Per British Columbia Ministry of Environment (BCMOE, 2016), "*Concentrations above water*

[2] https://laws-lois.justice.gc.ca/eng/regulations/SOR-2000-203/

quality guidelines do not imply that unacceptable risks are present, but that the potential for adverse effects is increased and additional investigation is needed for managing water resources".

CCME *Guidelines* are normally adopted by provinces and territories as ***Objectives.*** Per BCMOE (2016) *objectives* generally have greater site-specificity than *guidelines* and consider more localized factors that may modify the requirements of biota or receptors to substances. In British Columbia, water quality *guidelines* are applied at the provincial scale, while water quality *objectives* are applied at a watershed scale (BCMOE, 2016). In the Alberta context, the term *Objective* is nested within the development and application of site-specific numerical or narrative statements to direct action or management for specified waters (Alberta, 2012d). ***Site-Specific Objectives,*** which can be established by means of approvals or other practices, are synonymously referred to as *thresholds*, *targets* and/or *ambient limits*. ***Targets*** are levels that management aims to achieve or do better than, and ***Ambient Limits*** are a level beyond which the most sensitive use may not be protected. These *Thresholds* are applied in scenarios where water quality needs to be improved due to anthropogenic activities, maintained (i.e., existing conditions are better than relevant water quality guidelines), or protected. Protection of water uses is the most relevant scenario within the oil sands context, as *ambient limits* are set beyond existing conditions to enable some additional contaminant load, while ensuring water quality remains within the desired levels and preferably below relevant water quality *guidelines*.

The term ***Benchmark*** has been used in various ways. Water quality guidelines and objectives for example have been termed *benchmarks*, since they can be used to judge when there is an increasing or decreasing level of risk to a set of organisms. ***Science-Based Environmental Benchmarks*** (SBEBs) in British Columbia are derived for receiving-environment water quality variables to support a specific effluent discharge permit decision (BCMOE, 2016). SBEBs are site-specific, i.e., at the edge of a mixing zone, and are intended to protect aquatic life where there is an existing or proposed permitted activity. British Columbia defines SBEBs as "*a quantifiable receiving environment parameter or attribute protective of freshwater aquatic life that is developed by a qualified professional through a rigorous scientific process with the intent to inform management decisions and guide mitigative actions for a regulated mining activity at a specific location*." The Province of British Columbia also refers to these SBEBs as ***Site Performance Objectives*** which are at times written into water quality permits under the BC Environmental Management Act. The SBEBs are like Forecast Triggers as defined by Somers et al. (2018); a model-based prediction of expected environmental condition assuming a set of anthropogenic influences.

The term ***Critical Effect Size*** (or CES) is strictly a statistical term often representing the difference between two groups (treatment levels). Differences are standardized by expressing them relative to some measure of variation (e.g., within-group standard deviation, SD), or as percentages. Environment Canada's Pulp and Paper Effluent Regulations (PPER) and Metal and Diamond Mining Regulations (MDMER) specify critical effect sizes for surveys of benthic macroinvertebrates and sentinel fish populations. Differences in benthic endpoints between reference and exposure areas >2 standard deviations (i.e., using the within reference-area SD) prompt follow up study (confirmation and/or investigation of cause). Differences in fish population endpoints between reference and exposure areas that are >10% (condition) or 25% (growth, liver size, gonad size) also prompt follow up study. These critical effect sizes derive largely from the notion of *normal ranges* from a reference baseline condition defined using data from one-or-more reference areas (Kilgour et al., 1998; Munkittrick et al., 2000, 2009).

Apart from statistical considerations (Kilgour et al., 2017; Kilgour & Somers, 2017), there are at least five approaches for estimating ***normal ranges*** to characterize pre-disturbance, typical baseline conditions. That is, normal ranges may be derived from (1) temporal data, (2) spatial data, or (3) some combination of temporal and spatial data. Normal ranges derived from spatial and temporal data could be deemed representative of a "least impaired" condition if sampled in areas uninfluenced or least influenced by oil sands operations. Alternatively, data may be collected from areas influenced by other factors (e.g., municipal discharge, forest operations; Stoddard et al., 2006). For either of these options, normal ranges might or might not be considered a desired condition or state, depending on the purpose and objectives of the monitoring program. Alternatively, there remains the possibility of (4) hindcasting a historical baseline condition assuming that there is sufficient data supporting a relationship between current environmental condition and biological response (e.g., Bayne et al., 2021; Burton et al., 2014, p. 202; Kilgour et al., 2019). The hindcasting approach has been recently described in the 2021/2022 work plan for the Oil Sands Monitoring Terrestrial Biological Monitoring Program (Roberts, 2021).

The ***natural range*** of variability may be derived to reflect background variability associated with indicators exposed to natural disturbances (Doyon et al., 2008; Landres et al., 1999). The concept of natural ranges of variability has been applied principally to forest management and understanding the range of conditions that can arise from natural disturbances like forest fire. The normal and natural range of variability approaches can be adjusted based on risk sensitivity, which may be important for specific indicators due to cultural or social value (e.g., see Smyth et al., 2007). Baker (1992) argued that there are few to no examples of pristine landscapes in which anthropogenic disturbances are absent (also see Stoddard et al., 2006), making it philosophically challenging to implement the natural range of variability model; regardless, it remains as an alternative approach to describing what is an acceptable environmental condition.

The term ***threshold of unacceptable change*** is coined here to define a numeric or narrative value that stakeholders agree represents a condition to avoid, and potentially reflecting the bounds of a minimum acceptable state. Per Cairns et al. (1993), quantifying what is unacceptable provides clarity in an adaptive management framework. Here, it is proposed that the threshold of unacceptable change may be described numerically or narratively. Unacceptable changes may also be patterns such as degrading trends over time in the data that are uncomfortable even if there is uncertainty as to how degraded a system can be before action is absolutely required to return it to a desired state.

2.3.1 Using Indicator Thresholds to Trigger Changes in Monitoring

Cairns et al. (1993) provides a conceptual model using thresholds to evaluate observed and predicted trajectories for an indicator that was monitored over time (also see Abbasi & Abbasi, 2011). In this model (see Figure 8), the baseline range of variation for an unexposed (reference) site is represented by the yellow rectangle (i.e., based on the number of trout per km of stream at an exposed site during the pre-operational period). This baseline range can be characterized using statistics that summarize multiple years of data for a single location, or variation among multiple reference locations in one year using parametric or non-parametric methods (e.g., Kilgour et al., 2017; Arciszewski et al., 2018). Here, we are not distinguishing between pristine baseline, least-disturbed baseline, or a baseline associated with natural disturbance. Here, the baseline range means only the condition expected in the absence of the project. Additionally, the range of values for the forecast condition (with project influences) is also illustrated (i.e., the blue rectangle). The forecast range of values is the range of predicted values derived from scenario testing in the EA process, showing anticipated uncertainties (i.e., variability). The lower limits for the baseline and forecast ranges of values are thresholds. When monitoring data cross a threshold, this exceedance may be deemed sufficient evidence to trigger a change in the monitoring program (e.g., see 2.1.4 EEM, above). Ideally the range of values for the baseline and forecast conditions would be established before a project is initiated.

The grey rectangle illustrates the range of values where remediation efforts are expected to be ineffective or at least very difficult to achieve recovery (Figure 8). In this example, we illustrate numbers of trout per km of stream. In the case of trout, a decline in numbers below some critical threshold has the potential for inbreeding, where reduced numbers can reduce mating opportunities and subsequent recruitment, both factors leading to population collapse. The grey rectangle then is the range of values for the indicator where intervention and remediation may not successfully return an indicator to its baseline range of values.

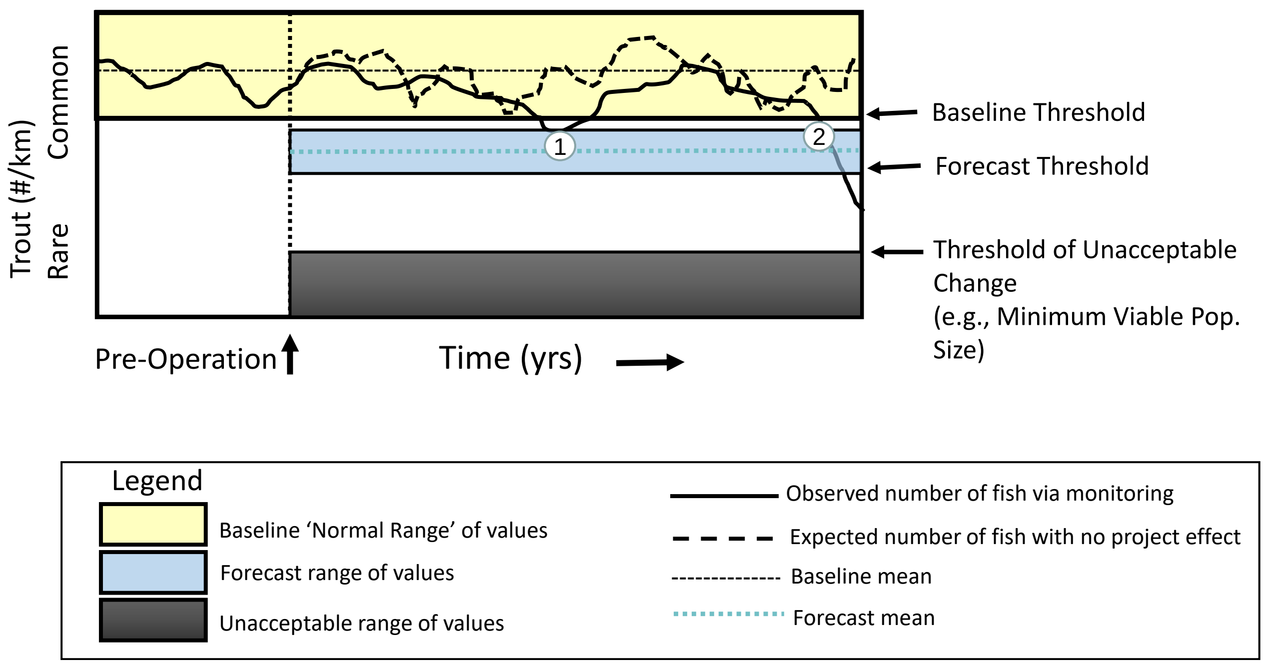

Figure 8 Conceptual model illustrating monitoring thresholds and triggers.

Figure Notes: illustrating variations in numbers of a valued ecosystem component, in this case number of trout per km in a theoretical river environment. Points (1) and (2) are discussed in the text.

To illustrate these concepts, a solid line in Figure 8 represents the time-series path of an indicator measured at an exposed site. Over time, the indicator from the exposed site gradually deviates down and away from the horizontal dashed line representing the mean for the baseline range of variation. The trajectory for the exposed indicator declines beyond the baseline threshold and then returns (see highlighted point 1). The exceedance at (1) could trigger a change to the ambient monitoring protocol (e.g., increased sampling frequency), but the indicator values reversed direction and returned to the baseline range. Subsequently at (2), the indicator crossed the baseline threshold again and then the forecast threshold (see (2) in Figure 8). Exceedance of the forecast threshold should trigger a shift from ambient monitoring to intervention/mitigation because it indicates that the system is behaving in an unexpected manner. Either the models that we used to make the forecast were insufficient, or there is some new factor influencing the system. If remediation is not initiated (or proves to be ineffective), the exposed indicator could continue to deviate from the forecast range towards the threshold of unacceptable change where recovery may no longer be effective (e.g., see Mao & Richards, 2012).

2.3.2 Baseline Thresholds

Before a new project is initiated, the area where the project will be located is usually surveyed to describe pre-project environmental conditions (Duinker & Greig, 2006). These data can be used to develop the conceptual model and associated effect pathways to support the project Environmental Assessment. The data can also be used to characterize pre-development conditions to provide baseline thresholds for the project (Squires & Dubé, 2012). For established facilities, baseline data can originate from surveillance monitoring of 'control'-type areas that are unexposed to the project (Underwood, 1992; Kilgour et al., 2017). For example, control-area data can be collected from minimally disturbed reference areas or unexposed upstream areas (e.g., McMaster et al., 2020). In some situations, baseline thresholds can be defined by pre-settlement environmental conditions instead of pre-development conditions (e.g., McCold & Saulsbury, 1996; but also see Hiers et al., 2012; Mao & Richards, 2012). Regardless of the source of the baseline thresholds, these values can be used to define the expected range of environmental conditions assuming the project has no effect (Humphrey et al., 1995; Squires & Dubé, 2012).

There are both parametric and non-parametric methods for quantifying baseline data ranges and for testing that indicator values are inside or outside those ranges (Kilgour et al., 2017; Arciszewski et al., 2018). Various portions of the baseline data range can be used to provide varying levels of "protection". Kilgour et al. (1998) proposed that the "normal" range of variation for baseline values might be the range of values that includes 95% of observations, or in a parametric statistical context, the baseline mean plus or minus two standard deviations. A smaller range of baseline values (e.g., mean plus/minus one standard deviation) could be used as an early warning threshold that would potentially trigger focused monitoring earlier, for example.

2.3.3 Forecast Thresholds

The Environmental Assessment process often involves the development of conceptual models and effect pathways to describe the anticipated effects of a proposed project on baseline environmental conditions (Duinker & Greig, 2006). Conceptual models can be used to design numeric risk-based models that predict the fate and transport of stressors and the responses of VECs (CNSC, 2016). Consequently, the EA process can lead to and include risk-based forecasts of environmental effects of projects (Dubé, 2003, CNSC, 2016). These risk-based forecasts can be evaluated in surveillance monitoring to determine if the predictive models are correct (Dipper et al., 1998; Fitzpatrick & Williams, 2020). Disagreement between forecast and observed conditions can suggest that the conceptual model is based on incomplete knowledge and needs to be updated (Dubé et al., 2013). If observed indicator data differ from the prediction and the associated forecast threshold, this departure can trigger a change in the monitoring program (Squires & Dubé, 2012). A small departure could be an early warning that triggers an increase in the frequency of monitoring, whereas a large departure could trigger a shift to focused studies to investigate cause. Departure from the model forecast may be a function of the original model being incorrect, or a function of the receiving environment changing or being influenced by a new factor.

Focused monitoring, then, could lead to adjustments of the conceptual and risk-based models after collecting new data to establish the extent and magnitude of the departure and the potential cause.

2.3.4 Limits of Change Thresholds

Stakeholder consultation is one of the cornerstones of the EA process (Duinker & Greig, 2006; Squires & Dubé, 2012). In addition to identifying VECs, stakeholders can specify the magnitude or thresholds of change that are acceptable (e.g., limits of acceptable change, McCool, 1996; Cole & Stankey, 1997; thresholds of unacceptable change, Rogers & Biggs, 1999; Smyth et al., 2007; Rogers et al., 2013). Monitoring data that approach thresholds of unacceptable change would likely trigger a change in monitoring to determine the underlying cause and identify / implement mitigation measures (Duarte et al., 2009; but see Mao & Richards, 2012).

2.4 Case Studies

Having reviewed different types of ambient monitoring, as well as the concepts of indicators and thresholds, below we present three case studies where adaptive monitoring principles have been applied in a comprehensive monitoring program. For each case study, we describe the process that led to indicator selection, summarize the indicators, program monitoring tiers, and how forecasts/predictions are program cornerstones.

2.4.1 Canadian EEM

Canadian EEM programs date to the early 1990s when both the PPER and MMER were amended to adjust end of pipe limits for constituents of concern. However, EEM was incorporated into the regulations because the effluents were complex, and there was residual uncertainty that the new limits would be protective of the ecological values. Environmental effects monitoring was therefore included in the regulations to confirm the prediction that receiving environments would be protected by the new regulations.

2.4.1.1 Process

The PPER and MMER (now the Metal and Diamond Mining Effluent Regulation or MDMER) fall under the federal *Fisheries Act*. The intent of the PPER and MDMER is to protect fish and fish habitat (per the *Fisheries Act*) equally across the country, in diverse receiving environments (creeks, rivers, lakes, coastal marine, offshore). Both pulp & paper and metal mining EEM programs followed extensive consultation with industry and other stakeholders. Individual monitoring programs involved site-specific considerations, given broader operational constraints (e.g., the use of monitoring tiers). Indicators and thresholds were selected, recognizing that it was fish and fish habitat that were of ultimate concern to the regulator.

2.4.1.2 Indicators

The development of the EEM programs concluded with sentinel fish populations and benthic invertebrate communities as the core components, although the EEM program regulations do not overtly state the rationale for this choice. Kilgour et al (2005) suggested that healthy fish populations are the implied VEC under the Fisheries Act and the PPER and MDMER. However, the EEM development process recognized that fish communities can be difficult to characterize, while sampling poses the additional risk of direct damage to the resource (Kilgour et al., 2005, 2007). To that end, indicators of the condition of a locally abundant 'sentinel' fish species were proposed to infer the condition of the overall fish community. The selection of sentinel species focuses monitoring on a small number of species and the sentinel serves as an indicator for healthy fish populations (Gibbons & Munkittrick, 1994; Munkittrick & Dixon, 1989; Tetreault et al., 2020). Sentinel condition can be determined by monitoring the number of individuals in the population, as well as general attributes that reflect physical (e.g., somatic tissues) and reproductive condition. Additionally, the condition of the habitat that supports the fish community is also monitored by evaluating the health of the benthos (or benthic macroinvertebrates) that inhabit the sediments (Culp et al., 2020; Karr, 2006). Here, the benthos is a surrogate indicator, where a healthy benthos community is assumed to indicate healthy fish habitat (Diaz et al., 2004; Gore et al., 2001; Tillin et al., 2008). That is, the condition of the fish habitat is not monitored directly.

Under Canada's EEM programs, sentinel fish populations and benthic invertebrate community composition are core 'effects' response variables, and EEM is *de facto* an 'effects-based' approach. Monitoring includes water quality variables as 'supporting' variables. That is, samples of water are collected from the receiving environment approximately four times per year, including when the biological sampling occurs. Surveillance monitoring occurs typically at three-year intervals. That interval is necessary to allow for reviews of study designs (submitted by proponents to the regulator), for the collection and processing of biological samples, for the interpretation of the biological data, and for the review of the interpretive reports.

2.4.1.3 Tiers

Canada's EEM framework is 'adaptive' in that the program consists of Periodic and Focused monitoring components. Under EEM, periodic monitoring consists of surveillance and reduced monitoring tiers while focused monitoring consists of confirmation monitoring, followed by various activities depending on the specific situation. Adaptive monitoring in Canada's EEM programs was first described by Hodson et al. (1996) to encompass program changes from Surveillance Monitoring to Focused Monitoring. Thresholds used to trigger changes between the program monitoring tiers have evolved over the years, as have the flowchart and associated terminology (see **Error! Reference source not found.**). For example, Arciszewski et al. (2017) introduced the notion of reduced/minimal monitoring for situations where multiple cycles of Surveillance Monitoring produced results consistent with expectations. Additionally, Somers et al. (2018) proposed to extend Surveillance Monitoring beyond testing for changes relative to baseline conditions, by including tests for changes from modelled predictions or expectations (i.e., forecast conditions).

Several pulp mills and metal mines in Canada have been required to undertake investigations of cause (Environment and Climate Change Canada, 2015; Environment Canada, 2014). Site-specific IOC conclusions are tracked by the federal government for future consideration of potential changes to the regulations. It is rare for facilities to undertake reduced monitoring despite that being a regulatory option.

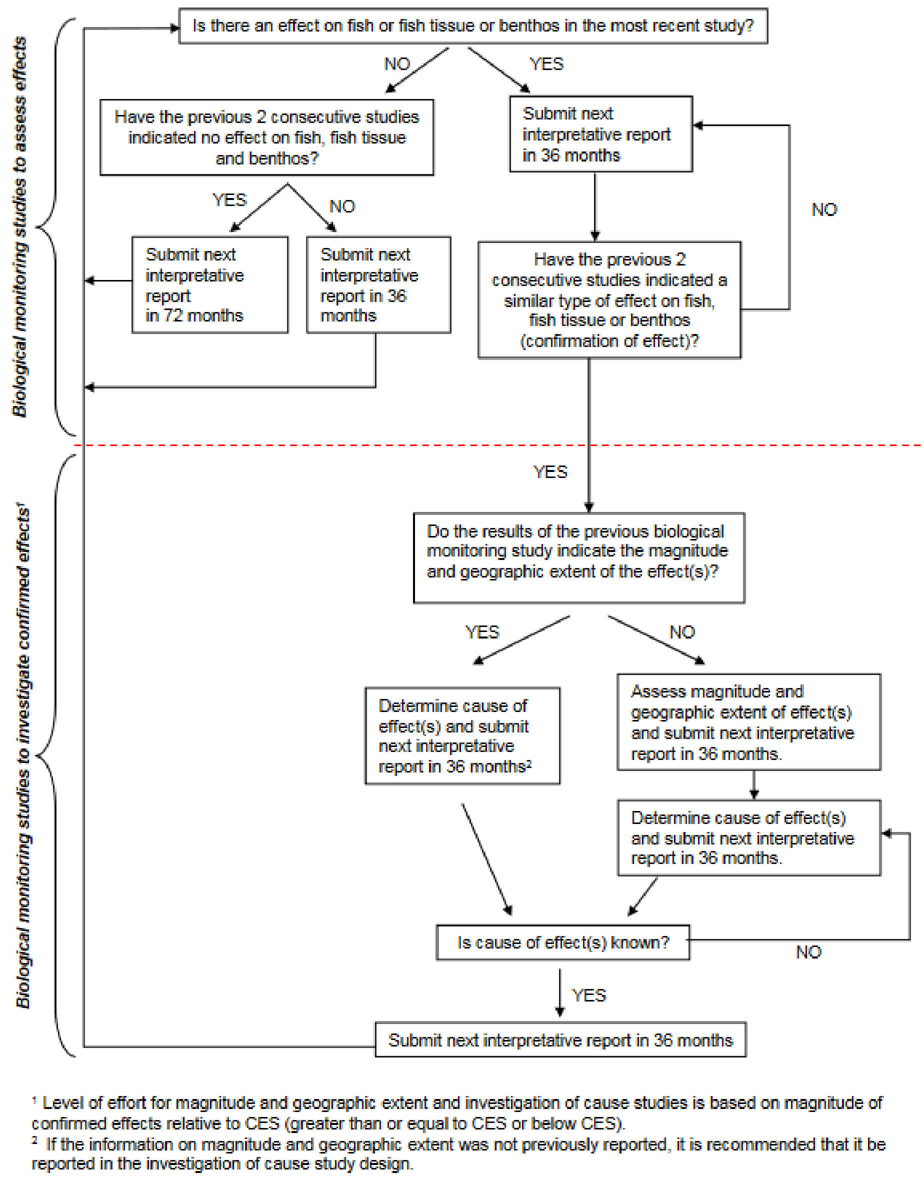

Figure 9 Decision tree for metal mining EEM

Figure Note: From (Environment Canada, 2012)

2.4.1.4 Predictions and Thresholds

Because end of pipe regulations are intended to protect receptors in the receiving environment, the expectation is that fish populations and benthic communities in exposure areas will be similar (in terms of specific indicators) to populations and communities from reference areas. Typically, reference areas are nearby (e.g., upstream) and differ from the exposure area only in the absence of exposure to treated effluent. When facilities are new, differences between reference and exposure areas are reasonably attributed to released effluent. However, across Canada many of the pulp mills and mines subject to the regulations have been releasing effluent for years (decades in some cases) prior to the EEM regulations and associated effluent limits. As such, there are numerous mills and mines where differences between reference and exposure areas are associated with historical releases, and it becomes difficult to attribute current differences to current effluent quality. Regardless, EEM monitoring programs test the general null hypothesis of no difference in indicator values between reference and exposure areas.

In the MDMER EEM process (**Error! Reference source not found.**), biological effects associated with sentinel fish populations and benthos indicators are evaluated through periodic surveillance monitoring. As above, Surveillance monitoring is typically designed to test the null hypothesis that there are no differences in indicators from reference and exposure areas. That approach involves a naïve null hypothesis since any two sites will naturally exhibit different fish populations or benthic communities, regardless of the similarity of the areas (e.g., habitat; Kilgour et al., 1998; Underwood, 1991, 1992, 1994). As such, the EEM programs adopted critical effect sizes as thresholds that trigger a change in monitoring. For fish populations, differences between reference and exposure areas need to exceed ±25% for liver size, gonad size and mean age (of adult fish) and ±10% for condition factor (Fulton's) to trigger focused monitoring. For benthic communities, differences between reference and exposure areas need to exceed ±2 SDs for density, family richness and evenness to trigger focused monitoring. For benthos, the Bray-Curtis dissimilarity index is included as an indicator of benthic community composition. There is currently no critical effect size for the Bray-Curtis index (Environment Canada, 2012). As such, any statistically significant ($p<0.1$) difference in Bray-Curtis distance between reference and exposure areas triggers Focused Monitoring.

Surveillance monitoring is repeated every three years. If no significant effects are observed after two cycles, surveillance monitoring is reduced to once every six years. If significant effects are observed, and those effects exceed critical effect sizes, focused monitoring is initiated to confirm effects. Where effects are confirmed and large (i.e., exceed the critical effect size), the cause of those effects is investigated (i.e., investigation of cause is initiated). Results from the investigation of cause can be used to set expectations for subsequent surveillance monitoring cycles.

2.4.2 Aquatic Environment Monitoring in the Elk Valley

2.4.2.1 Process

Teck Coal currently operates five coal mines in the Elk Valley of British Columbia (Teck, 2014). Coal mining has been ongoing in the Elk Valley since the late 1800s. Mining activities co-occurred with forestry and agriculture. All three activities resulted in modest expansion of residential areas in the towns of Sparwood and Fernie, with associated roadways and other infrastructure. The region is also attractive for recreational activities including downhill skiing and angling. All these activities have resulted in concerns of impacts to the natural environment. Among other processes, the province of BC is assessing/managing environmental effects in the valley through an Environmental Management Act Permit issued to Teck Coal to authorize the release of liquid effluent from Teck's five active coal mines. The Teck Coal Permit offers a variety of examples of adaptive monitoring principles.

Teck developed the Elk Valley Water Quality Management Plan (EVWQMP) to support a permit request under the BC Environmental Management Act (EMA) for release of liquid effluent to receiving surface waters. The EVWQMP was developed with advice from a technical advisory committee comprised of representatives from industry (Teck), government (BCMOE, BC Ministry of Energy and Mines, Environment Canada, USEPA, Montana State Government), the Ktunaxa First Nations Council (KNC), and an independent scientist. The Plan and its associated process, therefore, were similar to an environmental assessment in that it involved industry, government regulators and the public.

The Plan was devised in a manner that predicted the protection of human and aquatic health (Teck, 2014). The Plan determined, largely from existing literature, concentrations of selenium, sulphate, nitrate and cadmium in water and levels of calcite in substrate that would be protective of human and aquatic health in the valley. The process of developing the EVWQMP resulted in a conceptual site model that identified sources of these substances and likely receptors. The process of mining coal involves the fracturing of overburden with nitrate-based blasting agents, then removal and storage of waste overburden. Weathering of fractured waste rock results in the release of selenium, nitrate, sulphate, and cadmium, among other constituents, which variously pose risks in the aquatic receiving environments. Selenium is particularly troubling because it can bioaccumulate to levels that pose risks to fish reproductive impairment. Increasing trends in selenium in surface waters of the Elk River, combined with new scientific information about the toxicity of selenium to fish with associated fish kills, triggered the EMA permitting process for Teck's operations.

2.4.2.2 Indicators

The EVWQMP process did not result in the explicit identification of valued ecosystem components (VECs) *per se*. Rather, the conceptual site model considered typical receptors in aquatic systems (e.g., plankton, periphyton, benthic invertebrates, fish, macrophytes, amphibians, aquatic dependent birds, and aquatic dependent mammals). The choice of these receptor groups was informed by stakeholder input, with considerable input from government scientists, and industry subject matter experts. The local populations of Westslope Cutthroat Trout (*Oncorhynchus clarkia lewisi*) are a VEC because of their importance as a tourist attraction (angling). Westslope Cutthroat Trout is also the only species of fish in the Fording River upstream of a major natural barrier (Josephine Falls). Westslope Cutthroat Trout is the main point of focus for the local Elk Valley Fish and Fish Habitat Committee which has the following agency stakeholders as members: (1) BC Ministry of Forests, Lands, Natural Resource Operations (FLNRO); (2) Fisheries and Oceans Canada; (3) Ktunaxa Nation Council (KNC); and (4) Teck Coal (Bisset et al., 2018). Westslope Cutthroat Trout, further, were identified as one of five key VECs under the Cumulative Effects Management Framework process (Elk Valley Cumulative Effects Management Framework Group; EVCEMFWG, 2018); and see below).

The local Indigenous community (KNC), as a member of the Elk Valley Environmental Monitoring Committee (EMC) also has "all living things" as a guiding principle for management of resources (Ktunaxa Nation Council (KNC), 2017). The phrase "*all living things*" imbues the notion that all forms of life are VECs. The EVWQMP strives to protect all aquatic forms by setting benchmarks and associated Site Performance Objectives (SPOs) for the five constituents of concern based on the most sensitive species (Teck, 2014). The premise is that if the most sensitive species with known tolerances would be protected, then most species (and most with unknown tolerances) should also be protected.

Despite not explicitly stating what the VECs are, the EVWQMP is designed to protect the long-term sustainability of populations of Westslope Cutthroat Trout in the Elk Valley. Further, and as above, the Plan is designed to protect the most sensitive aquatic species.

2.4.2.3 Predictions and Thresholds

Mining for coal in the Elk Valley involves the blasting of overburden with nitrate-based blasting compounds. Blasted and fractured overburden is a waste material. Waste rock exposed to precipitation releases various substances but specifically selenium, sulphate, nitrate, cadmium, and carbonates. Considering selenium only, the metalloid bioaccumulates in the aquatic food chain, first by bacteria and algae. Benthic invertebrates and zooplankton obtain their tissue burdens through the consumption of contaminated bacteria and algae. Fish obtain their tissue burden through the consumption of zooplankton and benthos. Selenium tissue levels in benthos and fish are key pathway variables monitored in Teck's program (Elk Valley Environmental Monitoring Committee, 2019).

Teck (2014) developed relationships, via retrospective assessment of existing data, between concentrations of selenium in water, benthos, and fish. Concentrations in water can predict levels in benthos and fish whereas concentrations in benthos can predict levels in fish. Teck therefore routinely monitors selenium levels in water and benthos because they can be used to predict levels in fish, typically annually. Concentrations in fish are monitored less frequently, typically every three years, in part to reduce the inherent stresses caused to fish from handling and sampling of tissue plugs. Further, the benthic tissue data are considered more informative because they indicate spatial variations in bioaccumulation tendencies, which are easier to address than signals from fish which integrate selenium from considerably larger areas.

Various studies were conducted with Westslope Cutthroat Trout to identify tissue levels that cause reproductive failure (e.g., see Rudolph et al., 2008). Teck (2014) reviewed the published literature and suggested that ovary tissue levels > 25 mg Se/kg dw (dry weight) have the potential to result in 10% reduced egg viability. The 10% effect level was selected as a threshold on the basis that (1) effects at that level are generally indistinguishable from a reference/baseline condition (Suter et al., 1995), as well as (2) an expectation that populations could remain sustainable in the event of 10% reductions in growth or reproductive output (Mebane, 2010). Teck uses the term benchmark as a threshold that indicates a potential effect; with level 1 benchmarks being the 10% effect level; level 2 benchmarks a 20% effect (27 mg Se/kg dw in ovary), and level 3 benchmarks a 50% effect level (33 mg Se/kg dw in ovary). Other laboratory testing under the Permit has been conducted by Teck to confirm that the level 1 benchmark is ~25 mg Se/kg dw in ovary (Elk Valley Environmental Monitoring Committee, 2019).

2.4.2.4 Application of Tiers

The Regional Water Quality Model had been used to forecast selenium levels in water, including in Line Creek. Given the intensity of mining, the forecast was for selenium concentrations in the tributary to rise to levels that would pose risks to fish in the downstream receiving areas. Surveillance monitoring confirmed selenium levels exceeded a natural background and were consistent with model predictions. Model predictions were high enough to warrant mitigation. Engineers determined the most effective mitigation was construction of an Active Water Treatment Facility (AWTF) to remove selenium from creek water. From 2016 to 2017 monitoring indicated the facility removed 95% of the total selenium from processed creek water. An effectiveness monitoring program was also designed to test predictions related to selenium in water and biota. Engineers developed models that predicted selenium concentrations in water would be reduced, and that those reductions would result in lowered selenium levels in benthic invertebrates. Monitoring was carried out before the facility was constructed and then afterward, implementing a classic before-after-control-impact survey design. The pre-AWTF (baseline) data allowed Teck to compute not only a reference-creek regional baseline threshold, but also a baseline threshold for

Line Creek (i.e., pre AWTF) for both water and benthic tissue levels of selenium. Although the AWTF was able to reduce total selenium in the receiving environment, concentrations of selenium in benthos increased during its initial operation. Observed levels exceeded the pre-AWTF Line Creek baseline threshold. The elevated benthos tissue selenium levels exceeded the threshold and triggered confirmation monitoring. Follow up monitoring confirmed the elevated levels in benthos and that exceedance triggered investigation of the cause (i.e., focused monitoring). Follow-up study determined that the elevated levels were a result of the facility producing higher levels of selenite and other forms of organo-selenium that are considered more bioavailable. That outcome was unexpected and resulted in the company investigating solutions and adding an ozone plant to the AWTF to oxidize the final effluent and convert selenite to selenate: selenate being much less bioavailable. Subsequent effectiveness monitoring determined that selenium in benthic invertebrates reduced to approximately what it had been prior to the first installation of the AWTF. The modified AWTF therefore now reduces total selenium loads (protecting downstream surface waters) and minimizes the production of highly bioavailable selenium so that bioaccumulation rates in the nearfield receiving environment are reduced. If concentrations in benthos remain stable, there would be justification for reduced monitoring of benthic tissue selenium levels. Teck continues to monitor selenium in water and biota (i.e., benthos, fish) and benthic invertebrate communities as part of a surveillance program (Elk Valley Environmental Monitoring Committee, 2021).

2.4.3 Alberta Biodiversity Monitoring Institute (ABMI) Framework

The Alberta Biodiversity Monitoring Institute (ABMI; Burton et al., 2014) cumulative effects monitoring framework is another example of adaptive monitoring. It is described here because adaptive monitoring is applied to the terrestrial environment and it also illustrates a different approach to making predictions and estimating thresholds that will trigger changes in monitoring.

2.4.3.1 Indicators

Bayne et al. (2021) developed a monitoring methodology that focuses on sites spanning a gradient from baseline reference conditions to highly exposed to oil sands and other anthropogenic activities. The terrestrial valued ecosystem components include vascular plants, mosses and lichens, migratory land birds, mammals and amphibians. Bayne et al. further identify measurable variables for each valued component as well as stressor variables that are used as predictors for the terrestrial indicators (e.g., land use and land cover data, contaminants including concentrations of constituents in snow, water, sediment, climate variables). Bayne et al. also identify pathway variables such as tissue concentrations of contaminants in lichen, bird eggs and amphibian larvae.

2.4.3.2 Tiers

The ABMI framework (see Figure 10) consists of an initial sampling program across a study area. Apart from anticipating that there may be influences on biodiversity elements related to known stressors (e.g., oil sands development), there is initially no understanding or expectation as to what those effects might be. As such, baseline monitoring in the AMBI framework involves a gradient design, with sampling of biodiversity elements along anticipated gradients of exposure. The data are used to identify associations between biodiversity indicators and pressure indicators (stressors) and to construct predictive models. These predictive models are then used to predict conditions for all locations in the study area. Subsequent sampling is used to confirm the veracity of the models. Models can be used to predict future development conditions (e.g., scenario testing to examine the consequences of continued land cover change). In that sense, models can be valuable in communicating the consequences of current and future resource management. The models can also be used to identify actions that could be used to mitigate observed biodiversity effects.

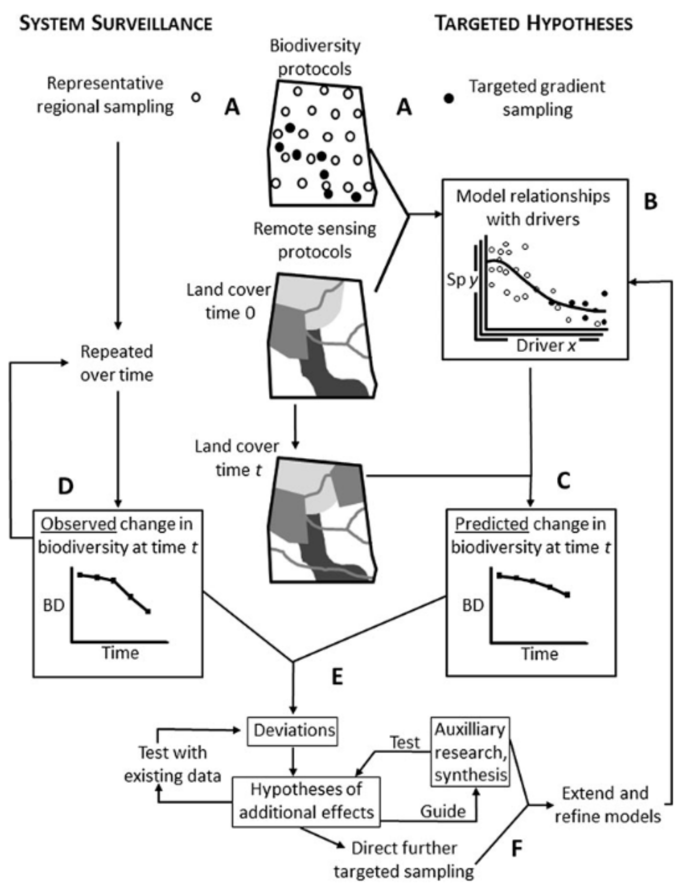

Figure 10 Schematic of primary components of the ABMI proposed cumulative effects monitoring framework (from Burton et al., 2014)

Figure notes: The approach involves a cycle of (A) sampling biodiversity elements across a study area; (B) building models relating ecological response indicators to stressor indicators; (C) use models to forecast/predict changes given current or future conditions; (D) and (E) comparing observed conditions to predicted values; and (F) carrying out research to understand deviations between model predictions and observations.

3.0 PROPOSED ADAPTIVE MONITORING FRAMEWORK FOR OIL SANDS

3.1 Overview

The purpose of this section is to propose a general framework for adaptive monitoring, and to illustrate how the framework can be used in Canada's oil sands. This section will also illustrate that some aspects of monitoring in the oil sands do not fully align with the proposed framework. For those cases, recommendations for improving alignment with the framework are provided. The proposed framework uses the EEM construct as the basis, but also incorporates fundamental aspects of the ABMI program. This framework is designed to apply equally to: (1) air and atmospheric monitoring; (2) surface water monitoring (including water quantity, quality, and biological indicators); (3) groundwater monitoring (including quantity and quality); (4) wetland monitoring; and (5) terrestrial biological monitoring (including biodiversity), per the oil sands monitoring operational framework (M. Dubé et al., 2018).

3.2 Components and Tiers

In the proposed framework, adaptive monitoring is implemented through three general components: (1) Baseline Monitoring, (2) Periodic Monitoring, and (3) Focused Monitoring (Figure 11). Each component is comprised of one-or-more monitoring "tiers" that represent different monitoring activities, often with different objectives. In the section below, the proposed framework is presented with a brief description of the OSM activities that are consistent with adaptive monitoring. A series of recommendations is also provided to further align current OSM activities with the principles of adaptive monitoring.

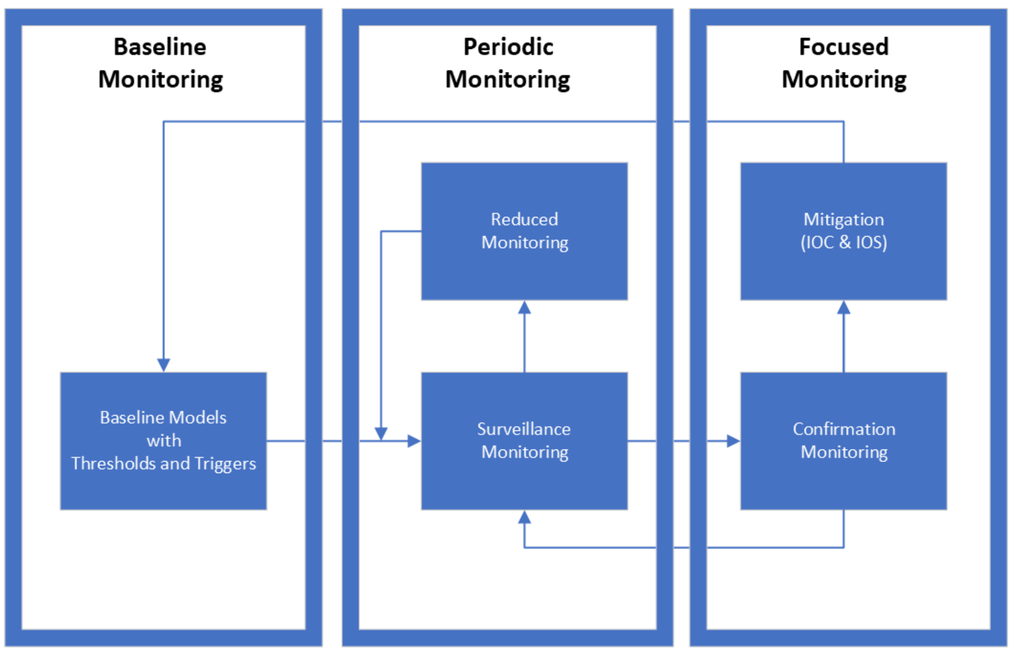

Figure 11 Simplified adaptive monitoring flowchart with baseline, periodic and focused monitoring components

3.2.1 Baseline Monitoring

Baseline Monitoring is the framework component where baseline conditions are determined, conceptual models are established, and predictive models are developed. Baseline monitoring is most likely to occur during an EA or similar exercise before the project is constructed (Figure). The EA (or similar) process will identify the VECs (see [1] in Figure), and develop conceptual models that will help identify pressures, stressor indicators, pathway indicators, and response indicators. A second baseline monitoring tier (see [2] in Figure) generates baseline data from relevant periods and/or locations to establish normal ranges for indicators and provide data to support the development of predictive models that can be used in scenario testing and developing forecasts.

Adaptive monitoring requires thresholds that trigger changes in monitoring tiers. *Baseline thresholds* are numeric or narrative values that describe the limit of acceptable change from a baseline condition. Baselines can represent a pristine pre-European contact condition, a local present-day best-available condition, or simply a local pre-project condition. Baseline thresholds are typically based on some estimate of the average condition with some measure of variability, and may be represented by a value that is a percentile of the baseline data set. Models of the relationships between anthropogenic pressures and valued component indicators are used to predict or forecast expected conditions. *Forecast thresholds*, therefore, are based on the expected condition. If underlying pressures are not expected to change the condition of an indicator, then baseline and forecast thresholds may be the same value.

In the baseline monitoring component, it is proposed that conceptual models be translated into predictive models (see [3] in Figure) using information available in peer-reviewed articles, government and consultant reports, or other initiatives. Data collected from the receiving environment may also be used to inform or develop those models (see for example Bayne et al., 2021). Predictive models are used in scenario testing to predict conditions for stressor indicators, pathway indicators and response indicators. The result will provide forecast thresholds. In nearfield environments, models may predict significant change in various indicators where forecast thresholds will represent a shift in value from the baseline condition (e.g., see Figure 8). In farfield environments, models may predict minor changes in indicators, such that the baseline condition (e.g., normal range threshold) becomes the forecast threshold.

Ideally the EA (or similar) process will also identify thresholds of unacceptable change for the VECs (**Error! Reference source not found.**). That is, models and a knowledge of what is unacceptable can be used to define thresholds of unacceptable change for stressor, pathway, and response indicators.

3.2.2 Periodic Monitoring

The second component of the proposed adaptive monitoring framework is Periodic Monitoring (Figure 12). The purpose of periodic monitoring is to regularly test predictions developed in the Baseline Monitoring component. As such, periodic monitoring would be designed to establish the burden of proof that the predictive models are robust and representative over time. Periodic monitoring consists of two tiers: (1) surveillance monitoring, and (2) reduced monitoring.

Surveillance monitoring (see [4] in Figure) is the central Periodic Monitoring tier that is designed to regularly test model predictions on a specified schedule. Where surveillance monitoring confirms predictions (i.e., no thresholds are exceeded; see Figure 8), this result provides confidence that the model predictions describe the relationships between pressures, stressors, pathway and response indicators, and that the system is understood and in control. After multiple confirmations that the observed surveillance-monitoring results fall within the predicted threshold values, there should be sufficient evidence to initiate the reduced monitoring tier (see [5] in Figure) through a reduction in the spatial extent of monitoring and/or a decrease in the temporal frequency of monitoring. By contrast, observed surveillance-monitoring indicator data that fall outside baseline and associated forecast thresholds would be unexpected, and this result would trigger the Focused Monitoring component.

3.2.3 Focused Monitoring

The Focused Monitoring component consists of a series of tiers that involve adjustments to monitoring activities, associated with unexpected monitoring outcomes during Periodic Monitoring. That is, Focused Monitoring is triggered when unexpected exceedances of baseline or forecast thresholds are observed during surveillance monitoring. The confirmation monitoring tier (see [6] in Figure) is initiated to verify that observed effects exceed thresholds. If confirmation monitoring fails to verify threshold exceedances surveillance monitoring is resumed. By contrast, if the confirmation monitoring tier affirms threshold exceedances, Focused Monitoring would proceed to the investigation of cause tier (IOC, see [7] in Figure). In the IOC tier, combinations of field-based and laboratory-based studies would be conducted to investigate the cause of the unexpected results as well as the magnitude and extent of the exceedances. Information from that investigation would be used to revise predictive models and forecasts (see [8] in Figure). Revised models could be used in scenario testing to explore the consequences of the new findings. That is, the new information is used to re-vise modelled estimates of the extent and magnitude of operations-related effects, including expected baseline conditions and associated thresholds. If model forecasts lead to the conclusion that the current (or some future) extent and magnitude of effects is unacceptable, then mitigation may be indicated. If mitigation is not necessary, then monitoring activities would return to surveillance monitoring tier in the Periodic Monitoring component. If scenario testing results indicate that mitigation is recommended, then the investigation of solutions tier (i.e., IOS, see [9] in Figure) would be initiated. Where mitigations are identified and subsequently implemented, some form of effectiveness monitoring is warranted to verify that mitigation predictions were achieved (see [10] in Figure). New information (data) derived from effectiveness monitoring could be used to revise predictive models and associated thresholds. If the model results and thresholds are deemed to be acceptable, further mitigation is not required, and the Focused Monitoring component returns to Periodic Monitoring.

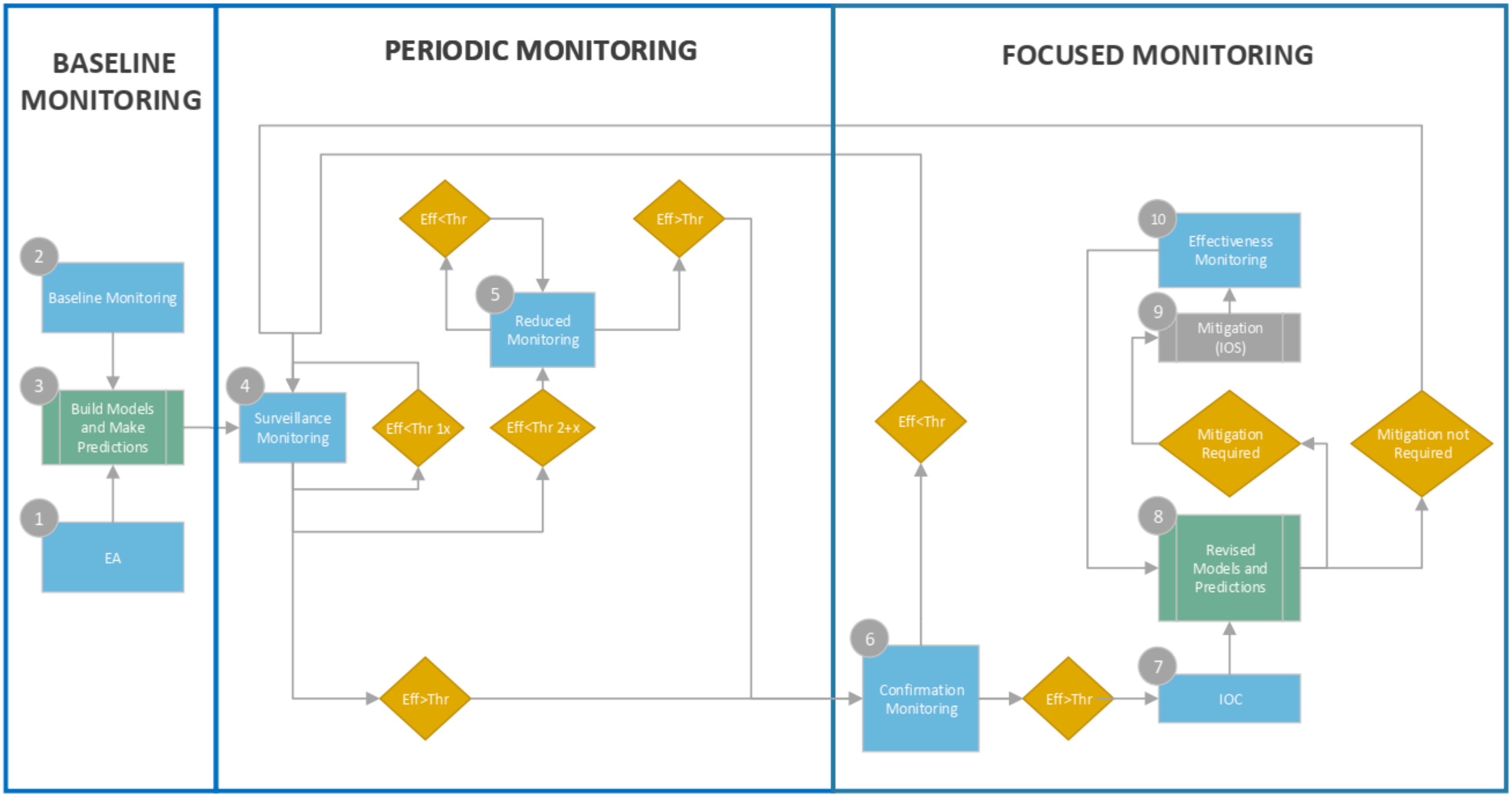

Figure 12 A proposed, generic adaptive management framework

Figure Notes: see the text for explanation of numbered tiers

3.3 Application to Oil Sands Monitoring

This section describes how current monitoring activities under the Oil Sands Monitoring Program incorporate elements of the proposed adaptive monitoring framework, and where gaps or inconsistencies exist. This description focuses on the three major components of the framework (i.e., Baseline Monitoring, Periodic Monitoring, and Focused Monitoring) and provides recommendations to align the OSM Program with the proposed adaptive monitoring framework.

3.3.1 Baseline Monitoring Component

Baseline conditions - Monitoring in Canada's oil sands region has an extensive history going back to the 1960s (e.g., Cronmiller & Noble, 2018b; Humphries, 2008; and see Chapter 1 above), and includes the period covered by RAMP (Hatfield Consultants et al., 2016), CEMA (Cronmiller & Noble, 2018b), and subsequent OSM programs.

In the proposed framework, pre-development reference data are available and could be used to quantify baseline conditions. Per the Oil Sands Monitoring Framework, the Oversight Committee (OC) is responsible for defining "baseline" (Figure 1; Dubé et al., 2018). As such, additional input from the OC may be required for the technical committees (TACs) to quantify baseline condition and establish thresholds. Consequently, it is recommended that the OC develop, or establish a mechanism to develop, thresholds of unacceptable change that can be used to support data interpretation and subsequent communication that management of the oil sands resource has been sustainable.

Recommendation 1 – The Oversight Committee to provide a formal definition of baseline condition, to be quantified by Technical Committees

Recommendation 2 – The Oversight Committee to define acceptable and unacceptable thresholds of change for the VECs and associated indicators

Conceptual models - A variety of conceptual models describing the interactions between oil sands activities and valued ecosystem components have been developed. Hatfield Consultants et al. (2009), for example, used conceptual models in the design of RAMP. The OSM Program subsequently hosted a series of workshops designed to develop conceptual models of the interactions between oil sands operations and the natural environment (Bayne et al., 2021; Swanson, 2019a, 2019b). These conceptual models considered natural drivers (e.g., weather and climate, as well as natural disturbances such as wildfire, and natural contamination such as bitumen) and oil sands related influences (e.g., land disturbance, groundwater withdrawal, surface water diversions, deep-well disposal, tailings pond creation and seepage, air emissions), pathways (including local and regional groundwater and surface water quality and quantity) and effects (e.g., traditional food quality and quantity, human health, wildlife habitat, species migration). It is recommended that the OSM workplans directly reference indicators identified in the conceptual models described by Swanson (2019a, b). Several of the 2021/2022 OSM workplans indicated that they would, as part of the workplan, go through a process of classifying the indicators that are being monitored. Bayne et al. (2021) provide a good example of identifying response indicators (i.e., VECs) that included vascular plants, mosses and lichens, migratory land birds, mammals, and amphibians, as well as the stressor indicators (e.g., landscape change, contaminants, climate change) and pathway variables (e.g., lichen tissue chemistry, waterbird egg chemistry, tadpole tissue chemistry).

Recommendation 3 – OSM workplans to incorporate VEC indicators based on established conceptual models

Predictive models - As described above, model predictions will normally originate in an EA (or similar) process. Formal Oil Sands EA processes have resulted in quantitative, qualitative and narrative predictions of how individual projects will interact with the environment (e.g., Hatfield Consultants et al., 2009). EA-like processes (e.g., government/industry workshops) have likewise produced conceptual models that describe anticipated pathways of effects, and quantify the interactions between projects and the environment (Swanson, 2019a). Groundwater models tend to be developed and applied at the scale of the 'site' (i.e., an individual oil sands operation). Water quantity models have been linked with fish habitat (and thus fish species diversity) and the potential for aboriginal navigation (Alberta, 2015). Regional air quality models have been developed and used by proponents and by CEMA to forecast acidification likelihood for small lakes in the broader region (Swanson, 2019a; Whitfield et al., 2009). Bayne et al. (2021) and Mahon et al. (2019) have developed quantitative models that describe the relationships between bird communities and various land cover and oil sands related predictors. Those quantitative models can be used to make forecasts of existing and future bird community composition, and can be used in scenario testing to determine the consequences of future oil sands development. Kilgour et al. (2019) demonstrated that Trout-perch population indicators (e.g., liver size, growth, condition factor) vary with mean air temperatures and Athabasca River flow volumes. Additional work is needed to model fish population performance indicators using oil sands pressures to predict or forecast the future Trout-perch population condition. Arciszewski (2021) recently demonstrated that benthic communities in tributaries to the Athabasca River varied in relation to climatic variables and oil sands-related land cover change. As such, these models can be used in scenario testing to assess the effects of additional oil sands development on benthic community composition. These predictive models can also be used to explore the spatial extent of anticipated effects on benthos communities, related to oil sands development. Although some modeling has been done in the oil sands region, there has been limited use of the models for scenario testing.

Within the OSM governance structure, the Technical Committees sustain the technical competencies for designing and developing predictive models. Consequently, the Technical Committees should identify relevant models to be developed and / or revised. It is anticipated that TACs will use the models to derive forecast thresholds that will be used to guide interpretation of monitoring data in Periodic Monitoring.

Recommendation 4 – TACs to identify predictive models, thresholds, and actions triggered by threshold exceedances

Stakeholder approved thresholds are an important concept in the proposed adaptive monitoring framework. That is, the various disciplines should propose thresholds for indicators based on an understanding of the relationship between oil sands operation and changes in environmental media. However, it is challenging to adequately model relationships at an appropriate scale to forecast changes across the landscape, and to have those forecasts vetted by stakeholders including the general public. This step of vetting forecasts is embedded within the EA process. However, there is currently no process for consultation with stakeholders that includes the general public within the existing oil sands monitoring process. Where monitoring results for various environmental media are within the vetted (accepted) forecast thresholds, it then becomes possible to conclude that the system is operating as expected, within an acceptable level of predicted change.

Recommendation 5 – Oversight Committee to establish a process for engaging stakeholders regarding the interpretation of regional model predictions, thresholds, and observed results

3.3.2 Periodic Monitoring Component

It is proposed here that the explicit purpose of the Periodic Monitoring component is to regularly test modelled predictions. Section 6.0 of the oil sands monitoring 2021/2022 workplan requires a description of how proposed studies will verify predictions made in environmental impact statements. Section 3.0 of the OSM workplans requires technical leads to demonstrate how monitoring will establish "*Are changes occurring…*" (leading question 1), and if there are changes "…*to what degree are changes attributable to oil sands activities, and what is the contribution in the context of cumulative effects"* (leading question 2). Leading question 1 can be rephrased to provide greater clarity and focus. For example, if this leading question is restated as "*Are unexpected changes occurring*", it would force recognition that oil sands projects are influencing the environment, potentially with measurable effects in nearfield environments, and anticipated immeasurable effects in farfield environments. Restating the leading question in this way supports a monitoring process that is focused on testing the underlying understanding of how oil sands projects interact with the receiving environment.

Sampling under Periodic Monitoring should be designed to test the modeled predictions and forecasts. In that sense, monitoring needs to be sufficiently rigorous to establish the veracity of the predictive models. The oil sands region is large, and it is not expected that model verification would rely on samples from a small subset of locations. It is anticipated that members of the TACs will devise monitoring study plans that are spatially representative (i.e., sufficient to provide stakeholders the confidence that the region is adequately covered), and that test model predictions under a variety of scenarios. If, for example, models describing bird community composition include woodland area or woodland age as a predictor (e.g., Mahon et al., 2019), then sampling locations selected for model verification should span a range of woodland sizes and woodland ages. Similarly, if the model developed for benthic community composition includes substrate texture (e.g., Arciszewski, 2021), then sampling locations should be selected to ensure that they vary in substrate texture.

Recommendation 6 – Funding requests for monitoring should be supported by a description of the models and associated predictions that will be tested

It is anticipated that TACs will assess Surveillance Monitoring data relative to model predictions, identifying where models have been supported, as well as locations where models have failed. Per the OSM Operational Framework, it is the responsibility of the ICBMAC and the TACs to assess results relative to "limits of change", or relative to baseline and forecast thresholds as expressed here. It is anticipated that the ICBMAC and TACs will assess monitoring data and make recommendations to the SIKIC and the OC regarding next steps (e.g., if surveillance monitoring results justify reduced monitoring). The OC, with the SIKIC, ICBMAC and the TACs should consider what type of evidence is required to support a change to the Reduced Monitoring tier.

Recommendation 7 – The Oversight Committee, with the SIKIC, the ICBMAC and the TACs to determine the criteria to trigger a change from Surveillance Monitoring to the Reduced Monitoring tier

When unexpected observations are encountered (i.e., thresholds are exceeded), it is anticipated that the TACs will recommend a change to the Focused Monitoring component. It is assumed that the TACs will undertake Focused Monitoring to confirm unexpected results and that the TAC will, through annual reporting, identify Confirmation Monitoring opportunities. It is further anticipated that the SIKIC and the OC will review those recommendations and provide instruction and funding to the TACs to initiate the Confirmation Monitoring tier. If for some reason the SIKIC or the OC deem that Confirmation is not required, it is anticipated they will provide that instruction to the TACs.

Recommendation 8 – OSM to provide oversight and guidance to TAC when Periodic Monitoring thresholds are exceeded

Recommendation 9 – The Oversight Committee, with the SIKIC and TACs to consider how to define thresholds for non baseline forecasts

3.2.1 Focused Monitoring

The first tier of Focused Monitoring is Confirmation Monitoring because unusual results should be verified before undertaking IOC. It is recommended here that the TACs be given the responsibility for developing investigative studies in the IOC tier. As per the OSM framework, those study plans will be reviewed and approved by the SIKIC and the OC since both the SIKIC and the OC should have the authority to approve or deny proposed IOC plans.

Recommendation 10 – TACs develop IOC plans for approval by the SIKIC and the OC

Where IOC has satisfactorily identified the cause of unanticipated effects, it is recommended that the TACs update the predictive models using this new information and use the updated models to assess scenarios including: (1) present day conditions, and (2) future conditions assuming further oil sands development. Under the present day scenario, the updated models should be used to predict conditions at locations that have not be sampled.

Recommendation 11 – TACs update predictive models after IOC (and IOS), and use those model predictions to update thresholds

The Investigation of Solutions (or Mitigation) tier involves research and consequently, it is outside of scope for OSM. In the proposed adaptive monitoring framework, IOS, mitigation and effectiveness monitoring tiers are proposed to be part of the Focused Monitoring component. Although the effectiveness monitoring tier is currently outside the scope of OSM, it is well within scope when considering a broader adaptive management context.

Recommendation 12 – OSM to consider expanding scope to include IOS, mitigation and effectiveness monitoring tiers associated with Focused Monitoring

4.0 SUMMARY AND CONCLUSIONS

Oil sands monitoring is a highly complex initiative involving multiple government agencies and expertise in air quality, water quality and quantity, aquatic ecosystems, and terrestrial ecosystems. Oil sands monitoring is meant to incorporate elements of adaptive monitoring to ensure that monitoring intensity is adjusted proportionally to the risks of environmental effects (Dubé et al., 2018). As such, there is every expectation that monitoring under OSM will escalate when there are perceived risks and uncertainties, and de-escalate when risks and uncertainties are absent.

The notion of adaptive monitoring is not new. Adaptive monitoring principles can be traced back to descriptions of environmental monitoring theory (e.g., Cairns et al., 1993) and ongoing environmental effects monitoring programs (Arciszewski & Munkittrick, 2015; Hodson et al., 1996; Somers et al., 2018).

In this proposed framework, adaptive monitoring has 3 fundamental components: Baseline Monitoring (involving the establishment of baseline conditions, appropriate indicators and associated thresholds), Periodic Monitoring (regular surveillance involving the assessment and reporting of ambient conditions relative to expected conditions), and Focused Monitoring (confirming unexpected findings, as well as investigating the extent, magnitude and cause of those effects). Implementation of this adaptive monitoring framework and associated principles has the potential for oil sands monitoring to demonstrate to stakeholders that the natural receiving environment has been protected.

The following recommendations are provided:

- Recommendation 1 – The Oversight Committee to provide a formal definition of baseline condition, to be quantified by Technical Committees
- Recommendation 2 – The Oversight Committee to define acceptable and unacceptable thresholds of change for the VECs and associated indicators
- Recommendation 3 – OSM workplans to incorporate VEC indicators based on established conceptual models
- Recommendation 4 – TACs to identify predictive models, thresholds, and actions triggered by threshold exceedances
- Recommendation 5 – Oversight Committee to establish a process for engaging stakeholders regarding the interpretation of regional model predictions, thresholds, and observed results
- Recommendation 6 – Funding requests for monitoring should be supported by a description of the models and associated predictions that will be tested
- Recommendation 7 – The Oversight Committee, with the SIKIC, the ICBMAC and the TACs to determine the criteria to trigger a change from Surveillance Monitoring to the Reduced Monitoring tier
- Recommendation 8 – OSM to provide oversight and guidance to TAC when Periodic Monitoring thresholds are exceeded
- Recommendation 9 – The Oversight Committee, with the SIKIC and TACs to consider how to define thresholds for non baseline forecasts
- Recommendation 10 – TACs develop IOC plans for approval by the SIKIC and the OC
- Recommendation 11 – TACs update predictive models after IOC (and IOS), and use those model predictions to update thresholds

Recommendation 12 – OSM to consider expanding scope to include IOS, mitigation and effectiveness monitoring tiers associated with Focused Monitoring

5.0 LITERATURE CITED

Alberta. (2012). *Lower Athabasca Region: Air quality management framework for nitrogen dioxide (NO2) and sulphur dioxide (SO2)*. Alberta Environment and Parks. https://open.alberta.ca/publications/9781460105320

Alberta. (2015). *Lower Athabasca region: Surface water quantity management framework for the Lower Athabasca River—Open Government*. Alberta Environment and Parks. https://open.alberta.ca/publications/9781460121733

Alberta. (2012b). *Capital Region air quality management framework for nitrogen dioxide (NO2), sulphur dioxide (SO2), fine particulate matter (PM2.5) and ozone (O3)*. https://open.alberta.ca/publications/9781460100653

Alberta. (2015a). *Executive Summary, Climate leadership, Report to minister*. https://open.alberta.ca/publications/climate-leadership-2015

Alberta. (2012d). *Guidance for deriving site-specific water quality objectives for Alberta rivers. Version 1.0*. https://open.alberta.ca/publications/9781460100622

Alberta. (2012c). *Lower Athabasca regional planning*. https://www.alberta.ca/lower-athabasca-regional-planning.aspx

Alberta Environment. (1995). *Water Quality Based Effluent Limits Procedures Manual*.

Alberta Environment and Parks. (2021). *Continuous Emission Monitoring System (CEMS) Code*.

Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrrano, J. A., Tietge, J. E., & Villeneuve, D. L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, *29*(3), 730–741. https://doi.org/10.1002/etc.34

Arciszewski, T. J. (2021). Exploring the Influence of Industrial and Climatic Variables on Communities of Benthic Macroinvertebrates Collected in Streams and Lakes in Canada's Oil Sands Region. *Environments*, *8*(11), Article 11. https://doi.org/10.3390/environments8110123

Arciszewski, T. J., Hazewinkel, R. R., Munkittrick, K. R., & Kilgour, B. W. (2018). Developing and applying control charts to detect changes in water chemistry parameters measured in the Athabasca River near the oil sands: A tool for surveillance monitoring. *Environmental Toxicology and Chemistry*, *37*(9), 2296–2311. https://doi.org/10.1002/etc.4168

Arciszewski, T. J., & Munkittrick, K. R. (2015). Development of an adaptive monitoring framework for long-term programs: An example using indicators of fish health: Defining Normal Framework Development. *Integrated Environmental Assessment and Management*, *11*(4), 701–718. https://doi.org/10.1002/ieam.1636

Arciszewski, T. J., Munkittrick, K. R., Scrimgeour, G. J., Dubé, M. G., Wrona, F. J., & Hazewinkel, R. R. (2017). Using adaptive processes and adverse outcome pathways to develop meaningful, robust, and actionable environmental monitoring programs. *Integrated Environmental Assessment and Management*, *13*(5), 877–891. https://doi.org/10.1002/ieam.1938

Atkins, J. P., Burdon, D., Elliott, M., & Gregory, A. J. (2011). Management of the marine environment: Integrating ecosystem services and societal benefits with the DPSIR framework in a systems approach. *Marine Pollution Bulletin*, *62*(2), 215–226. https://doi.org/10.1016/j.marpolbul.2010.12.012

Bailey, R. C., Norris, R. H., & Reynoldson, T. B. (2004). The Reference Condition Approach. In Richard H. (Ed.), *Bioassessment of Freshwater Ecosystems* (pp. 145–152). Springer US. https://doi.org/10.1007/978-1-4419-8885-0_7

Baker, J. M., & Westman, C. N. (2018). Extracting knowledge: Social science, environmental impact assessment, and Indigenous consultation in the oil sands of Alberta, Canada. *The Extractive Industries and Society*, *5*(1), 144–153. https://doi.org/10.1016/j.exis.2017.12.008

Baker, W. L. (1992). The landscape ecology of large disturbances in the design and management of nature reserves. *Landscape Ecology*, *7*(3), 181–194. https://doi.org/10.1007/BF00133309

Ball, M. A., Noble, B. F., & Dubé, M. G. (2013). Valued ecosystem components for watershed cumulative effects: An analysis of environmental impact assessments in the South Saskatchewan River watershed, Canada: VECs for Watershed Cumulative Effects Assessment. *Integrated Environmental Assessment and Management*, *9*(3), 469–479. https://doi.org/10.1002/ieam.1333

Bayne, E., Dennett, J., Dooley, J., Kohler, M., Ball, J., Bidwell, M., Braid, A., Chetelat, J., Dillegeard, E., Farr, D., Fisher, J., Freemark, M., Foster, K., Godwin, C., Hebert, C., Huggard, D., McIssac, D., Narwani, T., Nielsen, S., … Mundy, L. (2021). *A Before-After Dose-Response (BADR) Terrestrial Biological Monitoring Framework for the Oil Sands* (Oil Sand Monitoring Program Technical Report Series 7.0).

Belanger, S., Barron, M., Craig, P., Dyer, S., Galay-Burgos, M., Hamer, M., Marshall, S., Posthuma, L., Raimondo, S., & Whitehouse, P. (2017). Future needs and recommendations in the development of species sensitivity distributions: Estimating toxicity thresholds for aquatic ecological communities and assessing impacts of chemical exposures. *Integrated Environmental Assessment and Management*, *13*(4), 664–674. https://doi.org/10.1002/ieam.1841

Bell, M. D., Phelan, J., Blett, T. F., Landers, D., Nahlik, A. M., Van Houtven, G., Davis, C., Clark, C. M., & Hewitt, J. (2017). A framework to quantify the strength of ecological links between an environmental stressor and final ecosystem services. *Ecosphere*, *8*(5), e01806. https://doi.org/10.1002/ecs2.1806

Bisset, J., Tepper, H., Hussey, D., Watson, L., & Franklin, W. (2018). *The Elk Valley Fish and Fish Habitat Committee: A multi-agency consensus based approach to managing fisheries impacts in the Elk River Valley of British Columbia*. https://doi.org/10.14288/1.0374545

British Columbia Ministry of Environment (BCMOE). (2016). *A framework for the development and use of freshwater science-based environmental benchmarks for aquatic life in the Environmental Management Act Permitting for Mines. Version 1.0, March 2016.* Environmental Protection Division, Regional Operations Branch. https://www2.gov.bc.ca/assets/gov/environment/waste-management/industrial-waste/industrial-waste/mining-smelt-energy/guidance-documents/tg8_framework_for_sbebs.pdf

Brown, C. J., Saunders, M. I., Possingham, H. P., & Richardson, A. J. (2013). Managing for Interactions between Local and Global Stressors of Ecosystems. *PLOS ONE*, *8*(6), e65765. https://doi.org/10.1371/journal.pone.0065765

Burton, A. C., Huggard, D., Bayne, E., Schieck, J., Sólymos, P., Muhly, T., Farr, D., & Boutin, S. (2014). A framework for adaptive monitoring of the cumulative effects of human footprint on biodiversity. *Environmental Monitoring and Assessment*, *186*(6), 3605–3617. https://doi.org/10.1007/s10661-014-3643-7

Cairns, J., McCormick, P. V., & Niederlehner, B. R. (1993). A proposed framework for developing indicators of ecosystem health. *Hydrobiologia*, *263*(1), 1–44. https://doi.org/10.1007/BF00006084

Canadian Council of Ministers of the Environment (CCME). (2012). *Canadian Water Quality Guidelines for the Protection of Aquatic Life—Nitrate Ion*.

Canadian Nuclear Safety Commission. (2016). *REGDOC-2.9.1: Environmental Protection: Environmental Principles, Assessments and Protection Measures*. https://nuclearsafety.gc.ca/eng/acts-and-regulations/regulatory-documents/published/html/regdoc2-9-1-new/index.cfm

Carson, R. T., & Mitchell, R. C. (1993). The Value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research*, *29*(7), 2445–2454. https://doi.org/10.1029/93WR00495

Charpentier, A. D., Bergerson, J. A., & MacLean, H. L. (2009). Adaptive monitoring based on ecosystem services. *Environmental Research Letters*, *4*(1), 014005.

Cole, D. N., & Stankey, G. H. (1997). Historical development of limits of acceptable change: Conceptual clarifications and possible extensions. In S. F. McCool & D. N. Cole (Eds.), *Proceedings—Limits of Acceptable Change and related planning processes: Progress and future directions* (pp. 5–9).

Constable, A. J. (1991). The role of science in environmental protection. *Marine and Freshwater Research*, *42*(5), 527–538. https://doi.org/10.1071/mf9910527

Cronmiller, J. G., & Noble, B. F. (2018a). Integrating environmental monitoring with cumulative effects management and decision making. *Integrated Environmental Assessment and Management*, *14*(3), 407–417. https://doi.org/10.1002/ieam.4034

Cronmiller, J. G., & Noble, B. F. (2018b.). The discontinuity of environmental effects monitoring in the Lower Athabasca region of Alberta, Canada: Institutional challenges to long-term monitoring and cumulative effects management. *Environmental Reviews*, *26*(2), 169–180. https://doi.org/10.1139/er-2017-0083

Culp, J. M., Brua, R. B., Luiker, E., & Glozier, N. E. (2020). Ecological causal assessment of benthic condition in the oil sands region, Athabasca River, Canada. *Science of The Total Environment*, *749*, 141393. https://doi.org/10.1016/j.scitotenv.2020.141393

Dale, V. H., & Beyeler, S. C. (2001). Challenges in the development and use of ecological indicators. *Ecological Indicators*, *1*(1), 3–10. https://doi.org/10.1016/S1470-160X(01)00003-6

Davidson, C. J., Foster, K. R., & Tanna, R. N. (2020). Forest health effects due to atmospheric deposition: Findings from long-term forest health monitoring in the Athabasca Oil Sands Region. *Science of The Total Environment*, *699*, 134277. https://doi.org/10.1016/j.scitotenv.2019.134277

Davies, S. P., & Jackson, S. K. (2006). The Biological Condition Gradient: A Descriptive Model for Interpreting Change in Aquatic Ecosystems. *Ecological Applications*, *16*(4), 1251–1266. https://doi.org/10.1890/1051-0761(2006)016[1251:TBCGAD]2.0.CO;2

DeBlois, E. M., Tracy, E., Janes, G. G., Crowley, R. D., Wells, T. A., Williams, U. P., Paine, M. D., Mathieu, A., & Kilgour, B. W. (2014). Environmental effects monitoring at the Terra Nova offshore oil development (Newfoundland, Canada): Program design and overview. *Deep Sea Research Part II: Topical Studies in Oceanography*, *110*, 4–12. https://doi.org/10.1016/j.dsr2.2014.10.012

Diaz, R. J., Solan, M., & Valente, R. M. (2004). A review of approaches for classifying benthic habitats and evaluating habitat quality. *Journal of Environmental Management*, *73*(3), 165–181. https://doi.org/10.1016/j.jenvman.2004.06.004

Dipper, B. (1998). Monitoring and Post-auditing in Environmental Impact Assessment: A Review. *Journal of Environmental Planning and Management*, *41*(6), 731–747. https://doi.org/10.1080/09640569811399

Dowdeswell, L., Dillon, P., Ghoshal, S., Miall, A., Rasmussen, J., & Smol, J. P. (2010). *A Foundation for the Future: Building an Environmental Monitoring System for the Oil Sands* (p. 49).

Doyon, F., Yamasaki, S., & Duchesneau, R. (2008). The use of the natural range of variability for identifying biodiversity values at risk when implementing a forest management strategy. *The Forestry Chronicle*, *84*(3), 316–329. https://doi.org/10.5558/tfc84316-3

Dubé, M., Cash, K., Cronmiller, J., Abel, R., Andreeff, W., Berrade, D., Davidson, C., Dawson, J., Dersch, A., Dertien, K., Donald, G., Evans, M., Fayant, K., Glaude, B., Gosselin, J., Ilesanmi, Y., Ladouceur, B., Lawrence, L., Lee-Johnson, E., … Zhira, M. (2018). *Oil Sands Monitoring Program letter of agreement and operational framework*. Oil Sands Monitoring Program. https://open.alberta.ca/publications/9781460142363#detailed

Dubé, M. G. (2003). Cumulative effect assessment in Canada: A regional framework for aquatic ecosystems. *Environmental Impact Assessment Review*, *23*(6), 723–745. https://doi.org/10.1016/S0195-9255(03)00113-6

Dubé, M., & Munkittrick, K. (2001). Integration of Effects-Based and Stressor-Based Approaches into a Holistic Framework for Cumulative Effects Assessment in Aquatic Ecosystems. *Human and Ecological Risk Assessment: An International Journal*, *7*(2), 247–258. https://doi.org/10.1080/20018091094367

Duinker, P. N., & Greig, L. A. (2006). The Impotence of Cumulative Effects Assessment in Canada: Ailments and Ideas for Redeployment. *Environmental Management*, *37*(2), 153–161. https://doi.org/10.1007/s00267-004-0240-5

Dunn, G., & Bakker, K. (2011). Fresh Water-Related Indicators in Canada: An Inventory and Analysis. *Canadian Water Resources Journal / Revue Canadienne Des Ressources Hydriques*, *36*(2), 135–148. https://doi.org/10.4296/cwrj3602815

Elk Valley Cumulative Effects Management Framework Group (EVCEMFWG). (2018). *Elk Valley Cumulative Effects Assessment and Management Report* (p. 97). https://www2.gov.bc.ca/assets/gov/environment/natural-resource-stewardship/cumulative-effects/draft_elk_valley_ceam_12122018_draft.pdf

Elk Valley Environmental Monitoring Committee. (2019). *Permit 107517 Environmental Monitoring Committee 2019 Public Report* (p. 68). https://www.teck.com/media/2019-EMC.pdf

Elk Valley Environmental Monitoring Committee. (2021). *Permit 107517 Environmental Monitoring Committee 2021 Public Meeting Posters*. https://www.teck.com/media/2021-EMC.pdf

Environment and Climate Change Canada. (2015). *Third National Assessment of Environmental Effects Monitoring Information from Metal Mines Subject to th eMetal Mining Effluent Regulations*. Industrial Sectors, Chemicals and Waste and Environmental Protection Operations Directorates. https://publications.gc.ca/collections/collection_2016/eccc/En14-64-2016-eng.pdf

Environment Canada. (2011a). *Integrated Monitoring Plan for the Oil Sands, Expanded Geographic Extent for Water Quality and Quantity, Aquatic Biodiversity and Effects, and Acid Sensitive Lake Component* (p. 104). https://publications.gc.ca/collections/collection_2011/ec/En14-49-2011-eng.pdf

Environment Canada. (2011b). *Lower Athabasca Water Quality Monitoring Plan, Phase 1: Athabasca River Mainstem and Major Tributaries* (p. 90). https://publications.gc.ca/collections/collection_2011/ec/En14-42-2011-eng.pdf

Environment Canada. (2012). *Metal mining technical guidance for environmental effects monitoring*. https://www.ec.gc.ca/esee-eem/AEC7C481-D66F-4B9B-BA08-A5DC960CDE5E/COM-1434---Tec-Guide-for-Metal-Mining-Env-Effects-Monitoring_En_02[1].pdf

Environment Canada. (2014). *Sixth National Assessment of Environmental Effects Monitoring Data from Pulp and Paper Mills Subject to the Pulp and Paper Effluent Regulations* (p. 31).

Environment Canada, & Alberta Environment. (2004). *Northern Rivers Ecosystem Initiative, 1998 – 2003: Final Report* (p. 92). https://open.alberta.ca/dataset/2a0389a7-cbfb-4544-8b06-3f4aaa4eb3b8/resource/86900842-b4a2-4679-b88d-25040e454316/download/6375.pdf

Fennell, J., & Arciszewski, T. J. (2019). Current knowledge of seepage from oil sands tailings ponds and its environmental influence in northeastern Alberta. *Science of The Total Environment*, *686*, 968–985. https://doi.org/10.1016/j.scitotenv.2019.05.407

Fisheries and Oceans Canada. (2007). *Habitat Compliance Decision Framework*. https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/344519.pdf

Fisheries and Oceans Canada. (2019). *Fish and Fish Habitat Protection Policy Statement*. Fisheries and Oceans Canada.

Fitzpatrick, P., & Williams, B. (2020). *Building the System: Follow-up, monitoring & adaptive management* [Working Paper]. University of Winnipeg. https://winnspace.uwinnipeg.ca/handle/10680/1787

Fox, D. r., van Dam, R. a., Fisher, R., Batley, G. e., Tillmanns, A. r., Thorley, J., Schwarz, C. j., Spry, D. j., & McTavish, K. (2021). Recent Developments in Species Sensitivity Distribution Modeling. *Environmental Toxicology and Chemistry*, *40*(2), 293–308. https://doi.org/10.1002/etc.4925

Gaudet, C., Lingard, S., Cureton, P., Keenleyside, K., Smith, S., & Raju, G. (1995). Canadian Environmental Quality Guidelines for mercury. *Water, Air, and Soil Pollution*, *80*(1), 1149–1159. https://doi.org/10.1007/BF01189777

Gibbons, W. N., & Munkittrick, K. R. (1994). A sentinel monitoring framework for identifying fish population responses to industrial discharges. *Journal of Aquatic Ecosystem Health*, *3*(3), 227–237. https://doi.org/10.1007/BF00043244

Gieswein, A., Hering, D., & Feld, C. K. (2017). Additive effects prevail: The response of biota to multiple stressors in an intensively monitored watershed. *Science of The Total Environment*, *593–594*, 27–35. https://doi.org/10.1016/j.scitotenv.2017.03.116

Giesy, J. P., Anderson, J. C., & Wiseman, S. B. (2010). Alberta oil sands development. *Proceedings of the National Academy of Sciences*, *107*(3), 951–952. https://doi.org/10.1073/pnas.0912880107

Golder. (1998). *Oil Sands Regional Aquatic Monitoring Program: (RAMP) 1997* [Final report for the RAMP Steering Committee].

Gore, J. A., Layzer, J. B., & Mead, J. (2001). Macroinvertebrate instream flow studies after 20 years: A role in stream management and restoration. *Regulated Rivers: Research & Management*, *17*(4–5), 527–542. https://doi.org/10.1002/rrr.650

Gosselin, P., Hrudey, S. E., Naeth, M. A., Plourde, A., Therrien, R., Van Der Kraak, G., & Xu, Z. (2010). *Environmental and health impacts of Canada's oil sands industry*. The Royal Society of Canada. https://ceaa.gc.ca/050/documents_staticpost/59540/82534/Environmental_and_Health_Impacts_Report.pdf

Gray, J. S. (1989). Effects of environmental stress on species rich assemblages. *Biological Journal of the Linnean Society*, *37*(1–2), 19–32. https://doi.org/10.1111/j.1095-8312.1989.tb02003.x

Griffith, J. A. (1998). Connecting ecological monitoring and ecological indicators: A review of the literature. *Journal of Environmental Systems*, *26*(4), 325–363.

Harwell, M. A., Gentile, J. H., McKinney, L. D., Tunnell Jr, J. W., Dennison, W. C., Kelsey, R. H., Stanzel, K. M., Stunz, G. W., Withers, K., & Tunnell, J. (2019). Conceptual Framework for Assessing Ecosystem Health. *Integrated Environmental Assessment and Management*, *15*(4), 544–564. https://doi.org/10.1002/ieam.4152

Hatfield Consultants, Kilgour & Associates Ltd., Klohn Crippen Berger Ltd., & Western Resource Solutions. (2009). *RAMP: Technical Design and Rationale* (RAMP1467.1). http://www.ramp-alberta.org/UserFiles/File/RAMP_Design_&_Rationale.pdf

Hatfield Consultants, Kilgour & Associates, & Western Resource Solutions. (2016). *Regional aquatics monitoring in support of the Joint Oil Sands Monitoring Plan, final 2015 program report*. Prepared for Alberta Environmental Monitoring, Evaluation and Reporting Agency, Edmonton, AB, Canada.

Hewitt, L. M., Dubé, M. G., Ribey, S. C., Culp, J. M., Lowell, R., Hedley, K., Kilgour, B., Portt, C., Maclatchy, D. L., & Munkittrick, K. R. (2005). Investigation of Cause in Pulp and Paper Environmental Effects Monitoring. *Water Quality Research Journal*, *40*(3), 261–274. https://doi.org/10.2166/wqrj.2005.032

Hodson, P. V., Munkittrick, K. R., Stevens, R., & Colodey, A. (1996). A Tier-Testing Strategy for Managing Programs of Environmental Effects Monitoring. *Water Quality Research Journal*, *31*(2), 215–224. https://doi.org/10.2166/wqrj.1996.013

Humphries, M. (2008). *North American Oil Sands: History of Development, Prospects for the Future*. https://apps.dtic.mil/sti/citations/ADA477532

InnoTech, & Kilgour & Associates Ltd. (2019). *Limits of change synthesis* (Report to Alberta Environment and Parks).

Jones, D. K., Mattes, B. M., Hintz, W. D., Schuler, M. S., Stoler, A. B., Lind, L. A., Cooper, R. O., & Relyea, R. A. (2017). Investigation of road salts and biotic stressors on freshwater wetland communities. *Environmental Pollution*, *221*, 159–167. https://doi.org/10.1016/j.envpol.2016.11.060

Karr, J. R. (1991). Biological Integrity: A Long-Neglected Aspect of Water Resource Management. *Ecological Applications*, *1*(1), 66–84. https://doi.org/10.2307/1941848

Karr, J. R. (2006). Seven foundations of biological monitoring and assessment. *Biologia Ambientale*, *20*(2), 7–18.

Kelly, E. N., Schindler, D. W., Hodson, P. V., Short, J. W., Radmanovich, R., & Nielsen, C. C. (2010). Oil sands development contributes elements toxic at low concentrations to the Athabasca River and its tributaries. *Proceedings of the National Academy of Sciences*, *107*(37), 16178–16183. https://doi.org/10.1073/pnas.1008754107

Kelly, E. N., Short, J. W., Schindler, D. W., Hodson, P. V., Ma, M., Kwan, A. K., & Fortin, B. L. (2009). Oil sands development contributes polycyclic aromatic compounds to the Athabasca River and its tributaries. *Proceedings of the National Academy of Sciences*, *106*(52), 22346–22351. https://doi.org/10.1073/pnas.0912050106

Kelly, J. R., & Harwell, M. A. (1990). Indicators of ecosystem recovery. *Environmental Management*, *14*(5), 527–545. https://doi.org/10.1007/BF02394708

Kilgour, B. W., & Barton, D. R. (1999). Associations between stream fish and benthos across environmental gradients in southern Ontario, Canada. *Freshwater Biology*, *41*(3), 553–566. https://doi.org/10.1046/j.1365-2427.1999.00402.x

Kilgour, B. W., Dubé, M. G., Hedley, K., Portt, C. B., & Munkittrick, K. R. (2007). Aquatic Environmental Effects Monitoring Guidance for Environmental Assessment Practitioners. *Environmental Monitoring and Assessment*, *130*(1), 423–436. https://doi.org/10.1007/s10661-006-9433-0

Kilgour, B. W., Munkittrick, K. R., Hamilton, L., Proulx, C. L., Somers, K. M., Arciszewski, T., & McMaster, M. (2019). Developing Triggers for Environmental Effects Monitoring Programs for Trout-Perch in the Lower Athabasca River (Canada). *Environmental Toxicology and Chemistry*, *38*(9), 1890–1901. https://doi.org/10.1002/etc.4469

Kilgour, B. W., Munkittrick, K. R., Portt, C. B., Hedley, K., Culp, J., Dixit, S., & Pastershank, G. (2005). Biological Criteria for Municipal Wastewater Effluent Monitoring Programs. *Water Quality Research Journal*, *40*(3), 374–387. https://doi.org/10.2166/wqrj.2005.041

Kilgour, B. W., & Somers, K. M. (2017). Challenges with the use of normal ranges in environmental monitoring. *Integrated Environmental Assessment and Management*, *13*(2), 444–446. https://doi.org/10.1002/ieam.1874

Kilgour, B. W., Somers, K. M., Barrett, T. J., Munkittrick, K. R., & Francis, A. P. (2017). Testing against "normal" with environmental data: Testing Against "Normal" with Environmental Data. *Integrated Environmental Assessment and Management*, *13*(1), 188–197. https://doi.org/10.1002/ieam.1775

Kilgour, B. W., Somers, K. M., & Matthews, D. E. (1998). Using the normal range as a criterion for ecological significance in environmental monitoring and assessment. *Écoscience*, *5*(4), 542–550. https://doi.org/10.1080/11956860.1998.11682485

Kilgour, B. W., & Stanfield, L. W. (2006). Hindcasting reference conditions in streams. In *Influences of landscape on stream habitats and biological assemblages* (pp. 623–639). American Fisheries Society, Symposium 48.

Ktunaxa Nation Council (KNC). (2017). *Ktunaxa Nation Annual Report 2017* (p. 40). https://www.ktunaxa.org/wp-content/uploads/AGA-FINAL-2017.pdf

Landres, P. B., Morgan, P., & Swanson, F. J. (1999). Overview of the Use of Natural Variability Concepts in Managing Ecological Systems. *Ecological Applications*, *9*(4), 1179–1188. https://doi.org/10.1890/1051-0761(1999)009[1179:OOTUON]2.0.CO;2

Leston, L., Bayne, E., Dzus, E., Sólymos, P., Moore, T., Andison, D., Cheyne, D., & Carlson, M. (2020). Quantifying Long-Term Bird Population Responses to Simulated Harvest Plans and Cumulative Effects of Disturbance. *Frontiers in Ecology and Evolution*, *8*. https://www.frontiersin.org/articles/10.3389/fevo.2020.00252

Lindenmayer, D. B., & Likens, G. E. (2009). Adaptive monitoring: A new paradigm for long-term research and monitoring. *Trends in Ecology & Evolution*, *24*(9), 482–486. https://doi.org/10.1016/j.tree.2009.03.005

Lowell, R. B., Culp, J. M., & Dubé, M. G. (2000). A weight-of-evidence approach for Northern river risk assessment: Integrating the effects of multiple stressors. *Environmental Toxicology and Chemistry*, *19*(4), 1182–1190. https://doi.org/10.1002/etc.5620190452

Mahon, C. L., Holloway, G. L., Bayne, E. M., & Toms, J. D. (2019). Additive and interactive cumulative effects on boreal landbirds: Winners and losers in a multi-stressor landscape. *Ecological Applications*, *29*(5), e01895. https://doi.org/10.1002/eap.1895

Main, C. (2011). *Regional Aquatics Monitoring Program (RAMP) Scientific Review* (p. 160). Alberta Innovates — Technology Futures. http://www.ramp-alberta.org/UserFiles/File/RAMP%202010%20Scientific%20Peer%20Review%20Report.pdf

McMaster, M. E., Tetreault, G. R., Clark, T., Bennett, J., Cunningham, J., Ussery, E. J., & Evans, M. (2020). Baseline white sucker health and reproductive endpoints for use in assessment of further development in the alberta oil sands. *International Journal of Environmental Impacts: Management, Mitigation and Recovery*, *3*(3), 219–237. https://doi.org/10.2495/EI-V3-N3-219-237

McMillan, P. G., Feng, Z. Z., Deeth, L. E., & Arciszewski, T. J. (2022). Improving monitoring of fish health in the oil sands region using regularization techniques and water quality variables. *Science of The Total Environment*, *811*, 152301. https://doi.org/10.1016/j.scitotenv.2021.152301

Morgan, R. K. (2012). Environmental impact assessment: The state of the art. *Impact Assessment and Project Appraisal*, *30*(1), 5–14. https://doi.org/10.1080/14615517.2012.661557

Morrison, M. L., & Marcot, B. G. (1995). An evaluation of resource inventory and monitoring program used in national forest planning. *Environmental Management*, *19*(1), 147–156. https://doi.org/10.1007/BF02472011

Munkittrick, K. R., & Arciszewski, T. J. (2017). Using normal ranges for interpreting results of monitoring and tiering to guide future work: A case study of increasing polycyclic aromatic compounds in lake sediments from the Cold Lake oil sands (Alberta, Canada) described in Korosi et al. (2016). *Environmental Pollution*, *231*, 1215–1222. https://doi.org/10.1016/j.envpol.2017.07.070

Munkittrick, K. R., Arciszewski, T. J., & Gray, M. A. (2019). Principles and Challenges for Multi-Stakeholder Development of Focused, Tiered, and Triggered, Adaptive Monitoring Programs for Aquatic Environments. *Diversity*, *11*(9), Article 9. https://doi.org/10.3390/d11090155

Munkittrick, K. R., Arens, C. J., Lowell, R. B., & Kaminski, G. P. (2009). A review of potential methods of determining critical effect size for designing environmental monitoring programs. *Environmental Toxicology and Chemistry*, *28*(7), 1361–1371.

Munkittrick, K. R., & Dixon, D. G. (1989). A holistic approach to ecosystem health assessment using fish population characteristics. *Hydrobiologia*, *188*(1), 123–135. https://doi.org/10.1007/BF00027777

Munkittrick, K. R., McMaster, M. E., Van Der Kraak, G., Portt, C. B., Gibbons, W. N., Farwell, A., & Gray, M. A. (2000). *Development of Methods for Effects Basd Cumulative Effects Assessment Using Fish Popualtions: Moose River Project.* SETAC Press.

Niemi, G. J., & McDonald, M. E. (2004). Application of Ecological Indicators. *Annual Review of Ecology, Evolution, and Systematics*, *35*(1), 89–111. https://doi.org/10.1146/annurev.ecolsys.35.112202.130132

Noble, B., & Birk, J. (2011). Comfort monitoring? Environmental assessment follow-up under community–industry negotiated environmental agreements. *Environmental Impact Assessment Review*, *31*(1), 17–24. https://doi.org/10.1016/j.eiar.2010.05.002

Pearson, T. H., & Rosenberg, R. (1978). Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology Annual Review*, *16*, 229–311.

Pilière, A., Schipper, A. M., Breure, T. M., Posthuma, L., de Zwart, D., Dyer, S. D., & Huijbregts, M. A. J. (2014). Unraveling the relationships between freshwater invertebrate assemblages and interacting environmental factors. *Freshwater Science*, *33*(4), 1148–1158. https://doi.org/10.1086/677898

Ribey, S. C., Munkittrick, K. R., McMaster, M. E., Courtenay, S., Langlois, C., Munger, S., Rosaasen, A., & Whitley, G. (2002). Development of a Monitoring Design for Examining Effects in Wild Fish Associated with Discharges from Metal Mines. *Water Quality Research Journal*, *37*(1), 229–249. https://doi.org/10.2166/wqrj.2002.015

Roberts, D. R. (2021). *2021-2022 OSM Work Plan Application*.

Roberts, D. R., Hazewinkel, R. O., Arciszewski, T. J., Beausoleil, D., Davidson, C. J., Horb, E. C., Sayanda, D., Wentworth, G. R., Wyatt, F., & Dubé, M. G. (2022). An integrated knowledge synthesis of regional ambient monitoring in Canada's oil sands. *Integrated Environmental Assessment and Management*, *18*(2), 428–441. https://doi.org/10.1002/ieam.4505

Rogers, K., Saintilan, N., Colloff, M. J., & Wen, L. (2013). Application of thresholds of potential concern and limits of acceptable change in the condition assessment of a significant wetland. *Environmental Monitoring and Assessment*, *185*(10), 8583–8600. https://doi.org/10.1007/s10661-013-3197-0

Rudolph, B.-L., Andreller, I., & Kennedy, C. J. (2008). Reproductive Success, Early Life Stage Development, and Survival of Westslope Cutthroat Trout (Oncorhynchus clarki lewisi) Exposed to Elevated Selenium in an Area of Active Coal Mining. *Environmental Science & Technology*, *42*(8), 3109–3114. https://doi.org/10.1021/es072034d

Sadler, B. (1995). Canadian Experience with Environmental Assessment: Recent Changes in Process and Practice. *Australian Journal of Environmental Management*, *2*(2), 112–130. https://doi.org/10.1080/14486563.1995.10648322

Sánchez-Bayo, F., & Goka, K. (2012). Evaluation of suitable endpoints for assessing the impacts of toxicants at the community level. *Ecotoxicology*, *21*(3), 667–680. https://doi.org/10.1007/s10646-011-0823-x

Saracco, J., Pyle, P., Kaschube, D., Kohler, M., Godwin, C., & Foster, K. (2022). Demographic declines over time and variable responses of breeding bird populations to human footprint in the Athabasca Oil Sands Region, Alberta, Canada. *The Condor*. https://doi.org/10.1093/ornithapp/duac037/6731964

Shonfield, J., & Bayne, E. M. (2023). Weak support for cumulative effects of industrial disturbance on three owl species in Alberta's boreal forest. *Avian Conservation and Ecology*, *18*(1). https://doi.org/10.5751/ACE-02409-180109

Smyth, R. L., Watzin, M. C., & Manning, R. E. (2007). Defining Acceptable Levels for Ecological Indicators: An Approach for Considering Social Values. *Environmental Management*, *39*(3), 301–315. https://doi.org/10.1007/s00267-005-0282-3

Somers, K. M., Kilgour, B. W., Munkittrick, K. R., & Arciszewski, T. J. (2018). An Adaptive Environmental Effects Monitoring Framework for Assessing the Influences of Liquid Effluents on Benthos, Water, and Sediments in Aquatic Receiving Environments: An Adaptive Environmental Effects Monitoring Framework. *Integrated Environmental Assessment and Management*, *14*(5), 552–566. https://doi.org/10.1002/ieam.4060

Spears, B. M., Chapman, D. S., Carvalho, L., Feld, C. K., Gessner, M. O., Piggott, J. J., Banin, L. F., Gutiérrez-Cánovas, C., Solheim, A. L., Richardson, J. A., Schinegger, R., Segurado, P., Thackeray, S. J., & Birk, S. (2021). Making waves. Bridging theory and practice towards multiple stressor management in freshwater ecosystems. *Water Research*, *196*, 116981. https://doi.org/10.1016/j.watres.2021.116981

Spellerberg, I. F. (2005). *Monitoring Ecological Change* (Second Edition). Cambridge University Press.

Squires, A. J., & Dubé, M. G. (2013). Development of an effects-based approach for watershed scale aquatic cumulative effects assessment: Development of an Effects-Based Approach. *Integrated Environmental Assessment and Management*, *9*(3), 380–391. https://doi.org/10.1002/ieam.1352

Stevenson, M. G. (1996). Indigenous Knowledge in Environmental Assessment. *Arctic*, *49*(3), 278–291.

Stoddard, J. L., Larsen, D. P., Hawkins, C. P., Johnson, R. K., & Norris, R. H. (2006). Setting Expectations for the Ecological Condition of Streams: The Concept of Reference Condition. *Ecological Applications*, *16*(4), 1267–1276. https://doi.org/10.1890/1051-0761(2006)016[1267:SEFTEC]2.0.CO;2

Sutter, R. D., Wainscott, S. B., Boetsch, J. R., Palmer, C. J., & Rugg, D. J. (2015). Practical guidance for integrating data management into long-term ecological monitoring projects. *Wildlife Society Bulletin*, *39*(3), 451–463. https://doi.org/10.1002/wsb.548

Swanson, S. (2019a). *Oil Sands Monitoring Program: Integration Workshop Reports (Part 1 of 2)* (7.1; OSM Technical Report). Ministry of Environment and Parks. https://open.alberta.ca/dataset/6dfe5f22-668f-45cd-9de2-1422c9ad6878/resource/1e0d7f93-aad2-40ee-964a-0b5c1c724e3e/download/rpt7.1osmintegrationreports20190619-v1.pdf

Swanson, S. (2019b). *Oil Sands Monitoring Program: Recommendation Report (Part 2 of 2)* (7.2; OSM Technical Report). Ministry of Environment and Parks. https://open.alberta.ca/dataset/4f609d3f-9ab9-4f6f-809c-41d0369425ba/resource/1717edf5-e81c-4b42-9441-ef26c6afdb99/download/rpt7.2osmintegrationrecom20190619-v1.pdf

Teck. (2014). *Elk Valley Water Quality Plan*. Teck. https://www.teck.com/media/2015-Water-elk_valley_water_quality_plan_T3.2.3.2.pdf

Tetreault, G. R., Bennett, C. J., Clark, T. W., Keith, H., Parrott, J. L., & McMaster, M. E. (2020). Fish Performance Indicators Adjacent to Oil Sands Activity: Response in Performance Indicators of Slimy Sculpin in the Steepbank River, Alberta, Adjacent to Oil Sands Mining Activity. *Environmental Toxicology and Chemistry*, *39*(2), 396–409. https://doi.org/10.1002/etc.4625

Tillin, H. M., Rogers, S. I., & Frid, C. L. J. (2008). Approaches to classifying benthic habitat quality. *Marine Policy*, *32*(3), 455–464. https://doi.org/10.1016/j.marpol.2007.06.008

Timmerman, J. G., Beinat, E., Termeer, C. J. a. M., & Cofino, W. P. (2011). Developing transboundary river basin monitoring programmes using the DPSIR indicator framework. *Journal of Environmental Monitoring*, *13*(10), 2808–2818. https://doi.org/10.1039/C1EM10092K

Underwood, A. J. (1991). Beyond BACI: Experimental designs for detecting impacts on temporal variations in natural populations. *Marine & Freshwater Research*, *42*(5), 569–587.

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*(2), 145–178. https://doi.org/10.1016/0022-0981(92)90094-Q

Underwood, A. J. (1994). On Beyond BACI: Sampling Designs that Might Reliably Detect Environmental Disturbances. *Ecological Applications*, *4*(1), 3–15. https://doi.org/10.2307/1942110

Walker, S. L., Hedley, K., & Porter, E. (2002). Pulp and Paper Environmental Effects Monitoring in Canada: An Overview. *Water Quality Research Journal*, *37*(1), 7–19. https://doi.org/10.2166/wqrj.2002.003

Whitfield, C. J., Aherne, J., & Watmough, S. A. (2009). Modeling Soil Acidification in the Athabasca Oil Sands Region, Alberta, Canada. *Environmental Science & Technology*, *43*(15), 5844–5850. https://doi.org/10.1021/es9005652

Wiens, J. A., Hayward, G. D., Holthausen, R. S., & Wisdom, M. J. (2008). Using Surrogate Species and Groups for Conservation Planning and Management. *BioScience*, *58*(3), 241–252. https://doi.org/10.1641/B580310

Wong, L., Noble, B., & Hanna, K. (2019). Water Quality Monitoring to Support Cumulative Effects Assessment and Decision Making in the Mackenzie Valley, Northwest Territories, Canada. *Integrated Environmental Assessment and Management*, *15*(6), 988–999. https://doi.org/10.1002/ieam.4179

Wrona, F., & Schaefer, K. (2011). *Integrated monitoring plan for the oil sands: Expanded geographic extent for water quality and quantity, aquatic biodiversity and effects, and acid sensitive lake component*. Environment Canada.