7.0 CONCLUSIONS AND RECOMMENDATIONS

The 2008 RAMP monitoring program results have been discussed in detail, in Section 5 and Section 6. The following section provides a brief summary of results for each watershed specific to each monitoring component of RAMP. Based on results presented in Section 5, and Table 7.1-1 provides a summary of the 2008 RAMP monitoring program results, by watershed and component. In addition to the summary of results, overall conclusions as well as general comments and recommendations for each component for consideration by the RAMP Technical Program Committee and the RAMP Steering Committee are discussed in this section.

7.1 CLIMATE AND HYDROLOGY

The outlets of most major river basins in the RAMP Focus Study Area (FSA) are monitored, either by RAMP, by RAMP-member operators, or by federal or provincial governments, providing a good basis for assessing potential effects of focal projects and other oil sands developments. An increasing number of the hydrometric stations in the network monitor catchments in which development has occurred. Upstream monitoring locations are being added to compensate for this trend so that hydrometric data continues to be captured in natural watersheds for comparison.

In most of the watersheds, hydrologic changes were assessed as Negligible-Low in 2008 (Table 7.1-1). The Muskeg River, Tar River, Mills Creek, Poplar Creek, and Fort Creek were exceptions, with changes ranging from Moderate to High. Focal project activities with the greatest influence on hydrologic conditions in 2008 included:

- Reduced flows in Tar River due to the filling of the Canadian Natural Compensation Lake in spring 2008 and redirection of river flow into a tailings pond for the remainder of the year;
- Water releases into Poplar Creek via the Beaver River diversion and Poplar Creek spillway, increasing the flow in lower Poplar Creek;
- Muskeg dewatering in the Fort Creek catchment, increasing the flow in the creek; and
- Hydrologic isolation (closed circuit) of various mine sites in the Muskeg River basin, resulting in reduced flows to the river.

Measuring land use changes using satellite imagery, as was done for 2008, was a relatively simple process and provided useful results. However, there is some remaining uncertainty in defining catchment changes such as stream diversions and hydrologic isolation of developed areas.

The actual effects of the land change considered are in fact more complex than what is represented in the current assessment approach, and low flows in particular are not well estimated using this simple approach. However, although the analytical approach used for the hydrologic data includes a number of simplifications, assumptions and unknowns, it does provide a useful indication of the potential magnitudes of changes in hydrologic measurement endpoints. A more detailed and rigorous assessment could be made using the same analytical approach supported by hydrologic modeling, particularly if the model was calibrated using a watershed that was largely cleared.

Table 7.1-1 Summary assessment of RAMP 2008 monitoring results.

Watershed/Region	Differences Between *Test* and *Baseline* Conditions				Fish Populations: Health Risk from Metals and Organics in Fish Tissue						Acid-Sensitive Lakes: Variation from Long-Term Average Potential for Acidification
	Hydrology	Water Quality	Benthic Invertebrate Communities	Sediment Quality	Human Health				Fish Health	Fish Palatability	
					Sp.	Size	Sub.	Gen.			
Athabasca River	○	○	-	-	LKWH WALL	all sizes >400mm	○ ●	○ ●	LKWH ● WALL ○	LKWH ○ WALL ○	-
Athabasca Delta	-	-	○	-	-	-	-	-	-	-	-
Muskeg River	○	○	○		-	-	-	-	-	-	-
Steepbank River	○	○	●	-	-	-	-	-	-	-	-
Tar River	●	●	-	-	-	-	-	-	-	-	-
MacKay River	○	○	○	-	-	-	-	-	-	-	-
Calumet River	○	○	-	-	-	-	-	-	-	-	-
Firebag River	○	○	-	-	-	-	-	-	-	-	-
Ellls River	○	○	-	-	-	-	-	-	-	-	-
Christina River	nm	○	-	-	-	-	-	-	-	-	-
Clearwater River	nm	○	○	○	-	-	-	-	-	-	-
Fort Creek	●	○	●	○	-	-	-	-	-	-	-
Beaver River	-	●	○	○	-	-	-	-	-	-	-
McLean Creek	-	○	-	-	-	-	-	-	-	-	-
Mills Creek	●	-	-	-	-	-	-	-	-	-	-
Poplar Creek	●	○	●	○	-	-	-	-	-	-	-
Shipyard Lake	-	○	○	●	-	-	-	-	-	-	-
Isadore's Lake	nm	○	●	○	-	-	-	-	-	-	-
Big Island Lake	-	-	-	-	LKWH WALL NRPK	all sizes >600mm all sizes	○ ○ ●	○ ○ ●	All species ○	nm	-
Gardiner Lake	-	-	-	-	LKWH WALL NRPK	all sizes >500mm >700mm	○ ● ●	○ ● ●	All species ○	nm	-
Stony Mountains	-	-	-	-	-				-	-	○
West of Fort McMurray	-	-	-	-	-				-	-	○
Northeast of Fort McMurray	-	-	-	-	-				-	-	●
Birch Mountains	-	-	-	-	-				-	-	○
Canadian Shield	-	-	-	-	-				-	-	○
Caribou Mountains	-	-	-	-	-				-	-	○

○ Negligible-Low
● Moderate
● High

"-" program was not completed in 2008.

nm - not measured in 2008.

Hydrology: Calculated on differences between observed *test* and estimated *baseline* hydrographs: ± 5% - Negligible-Low; ± 15% - Moderate; > 15% - High.

Note: As not all hydrology measurement endpoints are calculated for each watershed because of differing lengths of the hydrographic record for 2008, hydrology results above are for those endpoints that were calculated.

Note: All calculated hydrology measurement endpoints in the Muskeg River watershed were assessed as Negligible-Low with the exception of Annual Maximum Daily Discharge which was assessed as Moderate.

Note: all calculated hydrology measurement endpoints in the Fort Creek watershed were assessed as High with the exception of Annual Maximum Daily Discharge which was assessed as Negligible-Low.

Water Quality: Classification based on adaptation of CCME water quality index.

Note: water quality at all stations in the Athabasca River was assessed as Negligible-Low with the exception of station ATR-SR-E which was assessed as Moderate.

Benthic Invertebrate Communities: Classification based on statistical differences in measurement endpoints between *baseline* and *test* areas as well as comparison to regional baselines.

Sediment Quality: Classification based on adaptation of CCME sediment quality index.

Fish Populations: Uses various USEPA and Health Canada criteria for risks to human health, fish health, and tainting from fish tissue concentrations of various substances.

LKWH-lake whitefish; WALL-walleye; NRPK-northern pike

Note: The classification of risk to human health for fish populations was Negligible-Low below the size class specified.

Note: For Fish Population Human Health Classification - Sub. refers to subsistence fishers; Gen. refers to general consumers as defined by Health Canada.

Acid-Sensitive Lakes: Classification based the frequency in each region with which values of seven measurement endpoints in 2008 were more than twice the standard deviation from their long-term mean in each lake.

Obtaining useful hydrometric information from small, poorly-defined and beaver-inhabited streams continues to be a challenge. In addition, minor weaknesses in the hydrologic monitoring network include:

- lack of a robust monitoring station on the Athabasca River;
- lack of winter monitoring on the Tar River and Calumet River; and
- lack of a hydrometric station at the mouth of the Christina River.

7.2 WATER QUALITY

Based on watershed-specific and regional analyses, the following waterbodies exhibited changes from historical and/or regionally typical water quality in 2008:

- Lower Tar River - increases in nutrients and various major ions, and deviation of several variables above regional baseline conditions;
- Lower Beaver River - regionally high concentrations of major anions (sulphate and chloride);
- Isadore's Lake - although water quality remains generally within regional baseline conditions, despite increasing concentrations of several major ions, including calcium, magnesium and sulphate; and
- Shipyard Lake - although water quality remains within regional baseline conditions, concentrations of several major ions, including sodium, chloride, sulphate, and boron have been and are increasing.

Additionally, one station on the Athabasca River (upstream of the Steepbank River, east bank) exhibited water quality that showed a Moderate difference from regional baseline conditions, primarily through high suspended sediments and some metals. Additional data collection in 2009 will be required to determine whether this 2008 result represents an ongoing, site-specific change in water quality or simply an atypical set of values associated with this sample. All other stations upstream and downstream of this station in the Athabasca showed consistency with historical results and with regional baseline conditions.

Aside from the above exceptions, water quality in the RAMP FSA in 2008 did not indicate significant differences with regional baseline conditions.

7.3 BENTHIC INVERTEBRATE COMMUNITIES AND SEDIMENT QUALITY

7.3.1 Benthic Invertebrate Communities

The strength of the RAMP Benthic Invertebrate Community component is the development of *baseline* data from multiple watercourses in a baseline condition. Replication within watercourses, and over time, is enabling RAMP the opportunity to extensively characterize the normal range of variability in benthic invertebrate community measurement endpoints including abundance, richness, diversity, evenness and percent of the fauna such as sensitive mayflies, stoneflies and caddisflies. Rigorous statistical techniques can be used to test for subtle variations in time trends from before to after commencement of operations, or spatially between *baseline* and operational or *test* reaches. Because statistical power is very high, subtle differences that are consistent with operations are inevitable, and were observed with these 2008 data. The regional baseline data; however, typically showed that the significant time trends or spatial variations were

subtle in comparison to natural background variability. Key findings from 2008 are provided below (Table 7.1-1):

- **Athabasca River Delta (ARD) –** Measured changes in the ARD were classified as Negligible-Low. Benthic communities in the ARD were in generally good condition in 2008 with relatively high diversity for a shifting-sand environment.

- **Muskeg River –** Measured changes in the lower, middle and upper reaches were classified as Negligible-Low on the basis that values of measurement endpoints were within normal ranges for expected baseline conditions, there was a number of sensitive taxa present and because background variation was more significant than effects-related trends. The upper reach of the Muskeg River was reclassified this year as a *test* reach on the basis of focal project activities in the upper catchment. Although there were no substantial changes measured in the upper reach, its reclassification to *test* presents significant challenges to assessments of the Muskeg River in years going forward. Future assessments of the lower, middle and upper reaches will have to rely on comparison to the regional baseline, which includes the years when the upper reach was classified as *baseline* (i.e., prior to 2008).

- **MacKay River –** Measured changes in the lower MacKay River were classified as Negligible-Low. Time trends in measurement endpoints did not indicate degrading water or habitat quality. The relatively high percent of the fauna in the lower reach as sensitive mayflies, stoneflies and caddisflies (% EPT) indicated high water and habitat quality. Measurement endpoints are well within the normal range of variability expected in *baseline* erosional reaches.

- **Fort Creek –** Measured changes in the lower Fort Creek were classified as High. The number of taxa was lower than the normal range of expected baseline conditions for a depositional reach in 2008, similar to what was observed in 2005, and total abundance and Simpson's Diversity were below the expected range of values in 2008.

- **Steepbank River –** Measured changes in the lower Steepbank River were classified as Moderate. The lower reach had lower total abundance, lower number of taxa, and lower %EPT relative to the upper reach. There has been a general decline in %EPT in the lower reach since RAMP sampling began there in 1998. Values of all benthic invertebrate community measurement endpoints with the exception of %EPT in the lower *test* reach (reach STR-E-1) were within the normal range of variation for *baseline* erosional reaches in the RAMP FSA indicating that measured changes, if real, are not yet substantial.

- **Poplar Creek** - This reach was new to RAMP in 2008, and measured changes were classified as Moderate. Diversity and evenness in this reach were below the normal range of variation for depositional reaches, and lower than the upstream *baseline* reach in the Beaver River.

- **Shipyard Lake –** Measured changes in Shipyard Lake benthos were classified as Negligible-Low. The benthic community of Shipyard Lake had as high or higher diversity, with about as many or more sensitive taxa (i.e., %EPT) compared to the *baseline* lakes. Measurement endpoints were within the range of natural variability in 2008 with the exception of Simpson's Diversity with a mean value greater than the expected range for baseline conditions.

- **Isadore's Lake –** Measured changes in Isadore's Lakes benthos were classified as Moderate. Number of taxa, diversity and evenness have been lower in Isadore's Lake (on average) over the data record, than *baseline* lakes (i.e., Kearl and McClelland lakes). Multivariate analysis indicated an unusual composition in 2008 relative to the two *baseline* lakes (i.e., Kearl and McClelland lakes).

The Benthic Invertebrate Community component is developing a robust temporal database of benthic communities in a baseline condition. Some regional trends in taxa richness and % EPT have been observed in both the *test* and *baseline* data. The change in classification of the upper Muskeg River (to *test*) represents potential assessment conundrum within RAMP, that of losses of baseline locations against which to judge potential future changes. Upstream *baseline* reaches would ideally be available to assess the impacts to downstream reaches in river systems. In their absence, regional baseline conditions and results from stations prior to becoming classified as *test* will need and continue to be used to judge the significance of temporal variations.

RAMP may wish to consider the potential use of negative baselines. The RAMP benthic study design has been developed with the objective of fully characterizing natural variability, and the comparison of *test* reaches (or lakes) against the baseline condition. Many investigators in the US and Europe use negative baselines (negative controls) because they help define the degree of deviation from a baseline condition: that is, a benthic community may not be similar to the baseline condition, but might also not be in a condition considered to be impaired. The concept of the negative baseline was developed by Karr (1981) in the derivation of the Index of Biotic Integrity (IBI) for fish, and by others that followed in the derivation of like measurement endpoints (e.g., Fore *et al.* 1996, Simon 1999). RAMP may also wish to consider the addition of negative baselines for each of the major habitat classes (i.e., lake, depositional river, erosional river) to provide an additional frame of reference. The recommendation to include negative baselines into RAMP is heightened by the loss of baseline watercourses such as the upper Muskeg River.

7.3.2 Sediment Quality

Sediments in the RAMP FSA naturally contain hydrocarbons and PAHs at concentrations that may exceed environmental-quality guidelines. Spatial and temporal comparisons of sediment quality over time since monitoring by RAMP began in 1997 has not detected any consistent trends over time in concentrations of hydrocarbons or metals, any consistent differences in sediment quality between *baseline* and *test* stations, or any relationships between sediment chemistry and composition of benthic communities that would suggest a negative influence of chemicals in sediment on community structure. Sediment toxicity tests showed survival and growth rates for organisms in sediments from all evaluated *baseline* and *test* locations that were similar or greater than those observed in laboratory-control sediments.

In 2008, differences in sediment quality to the range of natural variability was assessed as Negligible-Low at all sampling locations except Shipyard Lake, which was assesses as Moderate, due to high concentrations of some PAHs and metals in the lake (Table 7.1-1). However, when data were normalized to organic carbon and percent fines, values for Shipyard Lake were similar to those for sediments from other *baseline* or *test* waterbodies examined by RAMP.

7.4 FISH POPULATIONS

The 2008 RAMP Fish Population component consisted of:

- Fish inventory – Athabasca River, Clearwater River; and
- Fish tissue analysis – Athabasca River; Big Island Lake Gardiner Lake.

Assessing potential impacts on fish populations related to focal projects and other oil sands developments is an ongoing challenge due to limitations in the ability to effectively sample all fish populations in the entire RAMP FSA, and the limited temporal data, given that not all fish programs are conducted each year. In addition to these challenges, large-bodied fish are highly migratory between and within waterbodies in this region making it difficult to determine differences between natural variability in fish populations versus differences related to focal projects and other oil sands developments. Recognizing these limitations, assessments from external reviewers have been completed to assist in refining the methodologies and design of the component.

7.4.1 Fish Inventory

In 2008, the analysis of the Athabasca River and Clearwater River fish inventories focused on spatial and seasonal trends over time of catch per unit effort, condition, and length-frequency and age-frequency distributions for large-bodied fish.

Current and historical fish inventory data from the Athabasca River indicated species-specific variability in relative abundance, length-frequency distribution, and condition factor. Statistically significant differences were observed between years for condition and length-frequency distributions for the KIR fish species, with the exception of northern pike. Two significant trends over time have been observed in the Athabasca River between 1997 and 2008:

- Walleye - increase in catch per unit of effort over time; and
- Longnose sucker – decrease in catch per unit of effort over time.

With the exception of these results, there were no other significant trends that would suggest a consistent negative or positive change in the populations over time. As such, it is likely that the observed variability in measurement endpoints is a natural fluctuation in population size or growth of fish in the river during the RAMP sampling periods.

Similar to the Athabasca River, analyses of fish inventory data from the Clearwater River indicated significant differences among years for KIR fish species, but no clear negative or positive trends in length-frequency distributions, relative abundance, and condition factor. There were more instances of significant year-to-year differences in condition factor of fish captured in the fall than in the spring, likely reflecting physiological changes in fish during spawning periods.

The fish inventory results from the Athabasca and Clearwater rivers have been difficult to interpret given the variability in measurement endpoints over time and the uncertainty as to whether the variability is related to spawning and migration activities or due to stressors in the environment. Given these challenges, a summer inventory was initiated in 2008 to focus on a sampling period when fish populations were assumed to be less migratory and reflective of the local environment. A summer inventory will continue to take place each year to develop a time series of data representing more stable conditions in fish populations.

RAMP may wish to consider the following recommendations to establish more robust protocols during the fish inventories and to provide more integration with the other components of RAMP (i.e., benthos, water and sediment quality):

- Establish *baseline* reaches upstream of focal projects to assist in differentiating between natural variability versus changes in fish populations potentially related to focal projects;
- Continue to refine and increase standardization of inventory protocols for the Athabasca and Clearwater rivers. A full community assessment, rather than just focusing on large-bodied species, should be considered with emphasis on protocols developed by the US EPA Environmental Monitoring and Assessment Program;
- In response to community concerns regarding the health of fish in watercourses within the RAMP FSA, more thorough protocols for assessing fish pathology in individual fish should be considered to develop an inventory that could assist in increasing our understanding of the incidence of naturally occurring anomalies in fish populations of Northern Alberta versus anomalies related to anthropogenic influences in the environment; and
- Conduct fish community monitoring in reaches where sampling is currently conducted for benthic invertebrate communities, and water and sediment quality to provide a complete assessment of the aquatic environment within the RAMP FSA.

7.4.2 Fish Tissue

In 2008, the potential risk to human health related to fish consumption was assessed using individual and composite fish muscle samples (walleye and lake whitefish) collected from the Athabasca River and individual fish muscle samples (walleye, northern pike and lake whitefish) from Big Island Lake and Gardiner Lake. The average mercury concentration in walleye from the Athabasca River exceeded the subsistence fisher consumption guideline (0.2 mg/kg) indicating a High risk to human health of subsistence fishers and a Moderate risk to general consumers from consumption of individuals greater than 400 mm in length (Table 7.1-1). The average mercury concentration in lake whitefish from the Athabasca River was below the subsistence fisher guideline for all size classes indicating a Negligible-Low risk to human health for both subsistence fishers and general consumers (Table 7.1-1). The average mercury concentration in lake whitefish and northern pike from Big Island Lake were below the subsistence fish consumption guideline for all size classes indicating a Negligible-Low risk to human health (Table 7.1-1). The average mercury concentration in walleye in Big Island Lake exceeded the subsistence fisher guideline in fish greater than 600 mm in length indicating a High risk to human health of subsistence fishers and a Moderate risk to general consumers for consumption of fish of that size or greater. The average mercury concentration in walleye greater than 500 mm in Gardiner Lake exceeded the subsistence fisher consumption guideline indicating a High risk to human health of subsistence fishers and a Moderate risk to general consumers; average mercury concentrations in lake whitefish and northern pike were below the subsistence fisher consumption guideline across all size classes with the exception of one northern pike greater than 700 mm indicating a Negligible-Low risk to human health (Table 7.1-1). Mercury concentrations in fish muscle recorded in 2008 in the Athabasca River, Big Island Lake and Gardiner Lake were within the regional range observed in fish from other waterbodies of northern Alberta, including those beyond the influence of focal projects and other oil sands developments.

Although, all tainting compounds in walleye and lake whitefish muscle tissue from the Athabasca River were below guideline concentrations indicating a Negligible-Low influence on fish palatability (Table 7.1-1), the taste and odour of fish captured in the region is a concern and is the subject of ongoing research with the Canadian Oil Sands Network for Research and Development (CONRAD).

Concentrations of metals indicated a Moderate risk to lake whitefish health due to levels of copper exceeding the lethal effects threshold and selenium levels exceeding the sublethal effects threshold but a Negligible-Low risk to walleye health given all metals in composite samples were below sublethal effects and no-effects criteria (Table 7.1-1).

In collaboration with ASRD, RAMP should continue to develop a database of mercury in fish tissue from lakes and rivers within the RAMP FSA, both beyond focal project development and downstream of development given increased community concern regarding the safe consumption of fish.

7.5 ACID-SENSITIVE LAKES

Based on established criteria, over 60% of the lakes in the ASL component are considered highly sensitive or moderately sensitive to acidification. There have been only minor changes in the chemistry of the 50 ASL study lakes as a whole over the seven years of monitoring (2002-2008). Nitrate was the only ASL endpoint to show a significant change over this period. The changes in nitrate were not consistent with an acidification scenario and there is no indication that acidification is occurring to the ASL lakes from nitrogen deposition. A significant decrease in potassium in 2005 was related to the high rates of precipitation and runoff to these lakes that occurred in the fall that year. Values returned to more normal levels in the fall of 2006 and 2007.

As in 2007, critical loads of acidity were calculated using the Henriksen critical load model modified to account for the contributions of both strong and weak organic acids. Critical loads were calculated using values of runoff derived both from traditional hydrometric methods and isotopic enrichment. Using the runoff derived hydrometrically, critical loads in 2008 ranged from -0.088 keq H+/ha/y to 1.551 keq H+/ha/y with a median value of 0.308 keq H+/ha/y. Lakes located in the upland regions (the Stony Mountains, Birch Mountains, and Caribou Mountains) and in the Canadian Shield had the lowest critical load values. Using the runoff values derived from measurements of isotopic enrichment, critical loads ranged from -0.122 keq H+/ha/y to 2.137 keq H+/ha/y with a median value of 0.302 keq H+/ha/y. The mean and median critical loads were similar for the two methods. The critical loads of acidity were compared to modeled rates of acid deposition. Rates of critical load exceedance in 2008 were 36.7 % (18 of 49 ASL lakes) using hydrometrically-derived runoff estimates and 40.8 % (20 of 49 ASL lakes) using runoff estimates based on isotopic enrichment. These rates of exceedance are considerably higher than the rate of 8% reported for 399 lakes in a recent lake sensitivity report to the NOxSOx Management Working Group using the same model modifications (WRS 2006). The higher rates of critical load exceedance in the RAMP ASL lakes reflect a bias in the ASL component design that preferentially selected the most poorly buffered lakes for study. The rates of critical load exceedance in the RAMP ASL lakes were closer to rates observed in four sensitive regions in Ontario. A critical load exceedance does not necessarily mean that acidification of a lake is a certainty or imminent.

Time trend analysis (Mann-Kendall) was applied to key measurement endpoints in all 50 ASL lakes to detect changes that might indicate incipient acidification. As in previous years, most of the significant trends in measurement endpoints were inconsistent with any reasonable acidification scenario.

In 2008, a significant decrease in pH in a lake (Lake 268) was associated with a significant increase in sulphate, the primary acidifying agent. This lake is found in the physiographic subregion North-east of Fort McMurray and its chemistry will be followed in subsequent years to determine whether these trends continue.

The subregion north-east of Fort McMurray had the highest rate of measurement endpoints exceeding two standard deviations of the mean for each lake in the direction indicative of acidification. Following the criteria outlined in Section 3.5.6.3, this subregion was classified as having a Moderate indication of incipient acidification (Table 7.1-1).