7.0 CONCLUSIONS AND RECOMMENDATIONS

The 2004 RAMP monitoring results have been discussed in detail, both from a watershed (Section 5) and regional (Section 6) perspective. The following section provides a brief summary of overall conclusions specific to each monitoring component of RAMP, followed by general comments and recommendations for consideration by the RAMP Technical Subcommittee and RAMP Steering Committee.

7.1 CONCLUSIONS

7.1.1 Climate and Hydrology

Precipitation in the region was well below normal in 2004. As a result, hydrological conditions in the oil sands region were generally drier than normal. Despite those conditions, which might be expected to exaggerate water supply problems, the measurement endpoints for hydrology showed negligible to low impacts of oil sands development for most of the major river basins in the region that were designated as exposed. The exception was the Muskeg River basin, which exhibited the greatest hydrologic change attributable to oil sands development, with three of the four measurement endpoints showing a 7% decrease. That level of change would have been classified as a Low impact in many oil sands EIAs. Several conclusions relating to these results were identified:

- Changes in hydrologic measurement end points were generally in a downward direction, indicating decreased stream discharges.

- Changes in measurement endpoints were generally analyzed for stations nearest the mouth of each of the monitored streams. It is possible that greater changes may have occurred at specific locations within the basins (e.g., immediately downstream of an oil sands operation) than were observed at the mouth.

- Changes in measurement endpoints were strongly dominated by changes in catchment area due to hydrologic isolation. Withdrawals and releases had relatively less influence on the measurement endpoints.

- The analysis of hydrologic impacts relied heavily on information provided by oil sands operators, which was not always available or complete. This is a source of uncertainty in the analyses.

- Although the analytical approach used for the hydrologic data includes a number of simplifications and assumptions, it does provide a useful estimate of changes in measurement endpoints. The approach is considered superior to any attempt to identify impacts based on the comparison of pre- and post-development hydrologic records, or on a comparison of similar watersheds. A more detailed and rigorous assessment could be made using the same analytical approach supported by hydrologic modeling.

7.1.2 Water and Sediment Quality

Several key measurement endpoints were used to assess the potential influence of oil sands development on watershed-specific and regional water and sediment quality. When sediment and water quality endpoints were compared to a regional reference database developed using the RAMP data, the distribution of results relative to the reference median was similar regardless of whether stations/year combinations were designated as exposed or reference. Based on these results, there was little evidence indicating effects related to ongoing oil sands development on aquatic resources measured by RAMP. Stations which in 2003 indicated potential influences of oil sands activities on water quality, namely Stanley Creek and Beaver River, did not exhibit results that exceeded regional baseline water quality characteristics in 2004. In the case of Stanley Creek, this is likely because Syncrude ceased releasing the Aurora North Clean Water Discharge stream into Stanley Creek in 2004. For the lower Beaver River, the high water level in this stream in fall 2004 relative to fall 2003 may have influenced water quality characteristics.

Generally, use of the regional reference approach, described in Section 3 and in Appendices E and F, provided an effective technique to evaluate specific exposed stations/times relative to a regional reference condition. The approach was particularly useful for water quality because it is not possible to obtain true replicate samples of water at a given station. This approach also provided an assessment tool that circumvented current limitations of time-trend analyses associated with a limited temporal dataset and consequent low statistical power.

Results obtained in 2004 highlighted a need to re-visit the specific water and sediment quality variable list and evaluate the monitoring value of each variable. Concentrations of several water quality variables have consistently been below detection limits (e.g., BOD, most PAHs, total recoverable hydrocarbons). It may be more cost-effective to identify key screening-level indicator chemicals, which if measured at concentrations of concern, would initiate more detailed follow-up analyses (i.e., a tiered design).

During the 2004 fall sediment sampling survey, the water level of the Athabasca River was bankfull and, as a result, sediment samples were taken from different locations than in 2003, when the wetted banks of the river often were much further within the river channel. Generally, the monitoring value of sediment samples from Athabasca River is uncertain, given annual differences in the river hydrograph, high year-to-year variability of results and the river's highly dynamic sediment bedload. It is unlikely that the stations monitored truly are depositional in nature or accumulation zones for chemicals of concern. Conversely, the Athabasca River Delta is a true depositional environment, the ultimate receiver of particulates from the oil sands region and an area of local community concern regarding environmental quality. Accordingly, it may be appropriate to increase sediment quality monitoring effort in the Athabasca River Delta, regardless of whether sediment sampling continues in the Athabasca River mainstem.

7.1.3 Benthic Invertebrate Communities

None of the five river systems potentially influenced by oil sands development (Muskeg River, Jackpine Creek, Steepbank River, Tar River, MacKay River) and one lake (Shipyard Lake) demonstrated any effects on benthic invertebrate communities that were of

concern. In general, some effects consistent with oil sands development were detected (e.g., differences in time trends in indices such as abundance, richness, and diversity), but observed differences were subtle, and reach (or lake) means for 2004 all fell within the normal range of variability for regional reference conditions.

In general, benthic communities of erosional reaches were diverse and consisted of taxa typically associated with flowing water including mayflies, stoneflies, caddisflies, chironomids, worms, and beetles. Depositional reaches were less diverse and tended to have fauna more characteristic of shifting sand environments, with greater dominance by chironomids, ceratopogonids and worms.

An overarching multivariate ordination of erosional and depostional riverine habitats demonstrated not only that the two habitat types were relatively distinct, but also that there were a number of reaches in the various watersheds that were intermediate in their faunal assemblages. The obvious gradient between erosional and depositional reaches implies that the characterization of background reference conditions would be more efficient (and potentially valid) if a landscape factor could be identified that correlates with the ecological gradient and could therefore be used as a covariable to predict the expected background condition (as per Bailey *et al.*, 1998).

The ability to detect subtle effects (in particular in the rivers) with the existing study design reflects the use of site-specific upstream reference reaches, and high replication within reaches (i.e., 15 samples). The resulting design has very high statistical power, and is thus able to detect subtle changes in time trends between reference and affected areas. Without any changes to the sampling regime, the statistical power of the program will increase over time with every sample collected. Consequently, reducing the number of stations per reach may be possible without compromising the necessary statistical power to detect potential effects.

7.1.4 Fish Populations

The 2004 RAMP Fish Populations component included several core elements in the following oil sands region watersheds:

- Fish inventory – Athabasca River, Clearwater River, Christina River, Muskeg River/Jackpine Creek;
- Fish tissue analyses – Muskeg River, Clearwater River, Winefred Lake; and
- Sentinel fish species monitoring – Muskeg River, Steepbank River.

Assessing potential impacts on fish related to oil sands development is challenging due to the limited temporal database (many program elements are not conducted on an annual basis) and alterations to the sampling design between years for some elements. These factors make it difficult to establish the level of natural variability associated with impact parameters or measurement endpoints defined for the fish program. Recognizing these limitations, the fish program is continually being refined in terms of establishing standardized protocols and methodologies. As well, non-lethal samples approaches are being incorporated to allow greater frequency of sampling while minimizing fish mortality. Concluding remarks for each of the three fish program core elements carried out in 2004 are provided below.

Fish Inventory

Historical data, as well as results from the 2004 fish inventory on the Athabasca River, indicated some level of species-specific variability in population parameters and designated measurement endpoints (e.g., relative abundance and condition factor). However, preliminary statistical analysis on the inventory dataset did not demonstrate any significant trends in this regard. Additional inventory data gathered in a standardized fashion are needed to allow for appropriate trend analysis. To date, there is little evidence to suggest that characteristics of key indicator fish populations in the Athabasca River have changed during increasing development in the oil sands region.

2004 was the second successive year a fish inventory was conducted on the Clearwater River. The fish community in the Clearwater River appears to be similar to that found in the Athabasca River mainstem. The two-year dataset is not adequate for meaningful analysis of the designated inventory measurement endpoints (e.g., relative abundance and condition factor). It is notable that no lake whitefish were captured during the fall suggesting that they do not use the Clearwater River for annual spawning migration, as is the case in the Athabasca River mainstem.

Results from the 2004 Muskeg/Jackpine fish inventory indicated there was little evidence that characteristics of key indicator fish populations (i.e., designated endpoints) have changed during increasing development in the oil sands region. In general, 2004 fish population parameters (e.g., percent composition, condition factor and size distributions) were similar to those recorded in previous RAMP inventory years. Arctic grayling were not captured or observed during the 2004 inventory, which is consistent with results from 2001 and 2002. However, limited capture of grayling during the 2003 spring fish fence on the Muskeg River suggest that some individuals use the watershed for spawning.

Fish Tissue

In 2004, some individuals of target fish species from the Clearwater River (northern pike, walleye), the Muskeg River (northern pike) and Winefred Lake (northern pike, walleye, lake whitefish) exhibited elevated concentrations of mercury that exceeded consumption criteria for human health and wildlife. Consequently, there are potential human health risks from consuming the target fish species for subsistence and recreational fishers, sensitive subpopulations (i.e., children, pregnant women), and to a lesser extent, general fish consumers. Other metals and tainting compounds did not appear to pose any human health risks. However, mercury concentrations present in water and sediment in the oil sands region were generally below detection limits for all stations. Furthermore, fish tissue concentrations were similar to those observed historically during other tissue surveys in the region. These findings indicate that mercury concentrations in fish muscle tissue are naturally high in this region and not related to oil sands development.

Based on data collected on the Clearwater River in 2004, the non-lethal biopsy approach to acquiring fish tissue samples for mercury analysis can be incorporated into the RAMP fish tissue monitoring program. However, some level of lethal sampling will remain necessary for the analysis of metals and tainting compounds because sample weights acquired by the biopsy needle are too small for the testing of all required parameters. In addition further refinement to the biopsy methodology and additional comparison to corresponding fillet concentrations will be required to improve method QA/QC.

Incorporating this approach will result improved statistical power (increased sample size possible), while minimizing fish mortalities.

Sentinel Fish Species Monitoring

In 2004, the lethal sampling methodology used in previous RAMP sentinel species monitoring programs was replaced on a trial basis by a non-lethal approach. The non-lethal sentinel species monitoring program was designed to provide a comparative assessment of sculpin population health at exposure and reference sites by examining three important biological factors; population/age distribution, growth and energy storage.

Although considerable fishing effort was required at several sites, the abundance of sculpin was generally adequate to meet the needs of the non-lethal sampling approach at both exposure and reference areas. However, the issue of accurate field identification of sculpin species in the oil sands region must be resolved for future non-lethal initiatives. The most notable feature of the population data generated from the 2004 sentinel species program was the virtual absence of YOY at the Steepbank exposure site, and the poor YOY survival between sampling periods recorded for both the Steepbank and Muskeg exposure sites. Although presence/absence and survival of YOY are not currently designated as measurement endpoints for the non-lethal sentinel program, the 2004 data population data do provide some indication of potential effects at the two exposure sites. At this time, it is uncertain whether these differences represent a change from natural variation.

7.1.5 Acid Sensitive Lakes

In 2004, the Acid Sensitive Lakes component of RAMP incorporated specific analyses in response to the Scientific Peer Review comments, and discussions within the NOxSOx Management Working Group.

The 50 RAMP lakes were compared to the database of 450 regional lakes compiled by the NOxSOx Management Working Group to provide regional context for the RAMP lakes and ensure that they were representative. The vast majority of the regional lakes are of the Ca-Mg-Bicarbonate type with a smaller number having a significant contribution of sulphate/chloride to the anion charge and sodium/potassium to the cation charge. Similarly, the RAMP lakes are also dominated by Ca-Mg-Bicarbonates with fewer of the more extreme lake types identified in the large database. Chemical differences between the RAMP lakes and the population of regional lakes reflect a bias in the selection process for the RAMP program. Most of the RAMP lakes were selected for study because they were thought to represent lakes that are potentially sensitive to acid deposition. In practice, the selection of lakes that were the most poorly buffered and had the lowest values of pH appears to have provided a suitable spectrum of lake chemistry that will allow RAMP to evaluate potential changes in response to acid deposition.

The Henriksen critical load model was modified to account for the contributions of both strong and weak organic acids, as determined from the derived equations relating these parameters to pH and DOC. At low pH, the reduction in ANC from strong acids was dominant over the increase in ANC from weak organic acids (buffering) and the critical load was decreased. At pH 7.8 and greater, the reverse was true and the critical load

increased. Overall, application of the modified model to the 2004 data decreased the median critical load by 0.108 keq/ha/y. An additional 9 lakes (20 in total) had values of the critical load exceeded by the potential acid input from all developments and industries. The significant decreases in the critical loads under the modified model imply that these lakes are much more sensitive to acidification than was previously anticipated.

Using the Mann-Kendall trend analysis, an attempt was made to detect trends in measurement endpoint parameters that would indicate incipient acidification. Unfortunately, additional years of data are required to detect definitive trends. However, parallel decreases in the sum of base cations and Gran alkalinity were observed at the highest rates of acidic deposition.

7.2 RECOMMENDATIONS

From a sampling design and technical point of view, the three primary objectives of the RAMP program include:

- To monitor aquatic environment in the oil sands area to detect and assess cumulative effects and regional trends;
- To collect baseline data to characterize variability in the oil sands area; and
- To collect and compare data against which predictions contained in environmental impact assessments (EIAs) can be assessed.

Detailed assessment and discussion regarding the ability of RAMP to meet the above technical objectives has been presented in the RAMP 5-Year Report (Golder 2003) and the Scientific Peer Review Report (Ayles *et al.,* 2004). Furthermore, the RAMP Design and Rationale documentation that is currently being developed will continue the process of evaluating and strengthening the program and its ability to effectively measure potential change in aquatic resources associated with oil sands development. Consequently, this level of evaluation will not be presented here; however, a number of recommendations are listed below, by RAMP component, for the 2005 monitoring program.

7.2.1 Climate and Hydrology

- The outlet of each major river basin in the Alberta oil sands region is monitored by either RAMP or WSC, providing a good basis for assessing impacts of oil sands development. However, the existing station on the Tar River (S15) is actually located some distance upstream of the mouth and does not measure all potential changes in the Tar River basin. Consideration should be given to monitoring discharges at the mouth of the Tar River, either by maintaining the existing S15 station, which is expected to be discontinued, establishing a new station nearer to the mouth, or both.
- It is recommended that winter hydrology monitoring be considered on the Tar River and Steepbank River. This information would improve understanding of the annual hydrograph for these rivers, as well as the accuracy of the assessment of hydrologic changes.

- Assessment of potential hydrologic changes in exposed watersheds requires detailed information on oil sands development activities that could influence overall drainage patterns and water balance. It is recommended that relevant hydrologic data, including daily water withdrawals, daily water releases, and catchment changes that have occurred during a given monitoring year be compiled by oil sands operators to improve the accuracy of the hydrologic assessment.

7.2.2 Water Quality

- Based on RAMP data collected from 2002 to 2004, station CHR-2 located on the Christina River near Janvier (CHR-2) does not appear to be a suitable reference station for CHR-1 located at the mouth. It is recommended that RAMP consider re-locating station CHR-2 or consider sampling at one or two locations between the current stations to determine where (and why) the transition in water quality occurs.

- A few specific water quality variables are rarely detected in water samples collected under RAMP, or are poorly correlated with other related variables. It is recommended that these variables be evaluated in terms of their monitoring value:

  - Biochemical Oxygen Demand (BOD) – since RAMP started, BOD has not been detectable in 82% of samples. Samples with measurable BOD did not exhibit any apparent trends or relationships with other variables;

  - Chlorophyll *a* – based on RAMP data, the relationship between chlorophyll *a* and nutrient variables is poor. Regardless, nutrient variables are better indicators of nutrient status;

  - Total recoverable hydrocarbons (TRH) – Concentrations in 90% of samples collected to date have been below detection limits. There has been no apparent pattern in samples with measurable concentrations; however, it is uncertain whether TRH should be dropped given it is a good summary variable for monitoring changes in hydrocarbon levels; and

  - Polycyclic aromatic hydrocarbons (PAH) – Concentrations of most PAHs in 100% of samples were below detection limits; exceptions include naphthalene, C1/C2-napthalene (87%-95% below detection, measurable levels<2-fold detection limit) and phenanthrene (96% below detection, measurable levels<2-fold detection limit).

7.2.3 Sediment Quality

- Due to the high variability in samples/results, and the lack of true deposition zones, it is recommended that sampling of bottom sediments in the Athabasca River be re-considered.

- It is recommended that RAMP increase sediment quality monitoring effort in the Athabasca River Delta. Monitoring should be conducted at least as frequently, and at the same locations, as is done for the benthic invertebrate component (recently, conducted annually).

- The value of monitoring sediments in select lakes should be evaluated. Organic matter has to date dominated sediment samples from these shallow lakes; results do not reflect sediment accumulation of chemicals. There is a need to better define the potential linkage between lake sediments and oil sands development activities (e.g., groundwater infiltration) and whether sediment quality is the best measurement endpoint.

- There is a need to define/clarify the linkage between metals in sediments and oil sands development activities. Oil sands EIAs do not predict that metals in sediments will be affected by oil sands development; few sampling stations exhibit metal concentrations that exceed CCME guidelines.

- Given 2004 results demonstrated good correlation between CCME four-fraction total hydrocarbon analyses (instituted in 2004) and the AEP Tier 1 total hydrocarbon analyses (i.e., TVH, TEH, TRH) traditionally assessed by RAMP, the CCME four-fraction test should be used by RAMP as an indicator of total hydrocarbon concentrations in sediments. The CCME analyses provide more detailed information specific to the molecular size of measured hydrocarbons; they also have associated environmental quality guidelines and similar analytical cost.

7.2.4 Benthic Invertebrate Communities

- During future benthic surveys, it is recommended that RAMP quantify landscape features that naturally influence the condition of benthic invertebrate communities. This information will allow RAMP to develop regional reference models based on the underlying landscape features and will further improve the assessment of benthic communities in exposed watersheds.

- The statistical power of the current sampling design is high, and will get higher with time (more years of data). Given that the reaches tend to be relatively homogenous in terms of physical habitat and benthic community composition, it is recommended that RAMP consider reducing the number of stations (perhaps from 15 to 10) per reach (but the same length of reach as in the past). Sampling from seven stations per reach would exceed minimum sample sizes recommended by Environment Canada (2004) for pulp and paper and mining benthic EEM programs, and would have a negligible effect on statistical power for examining changes in time trends.

7.2.5 Fish Populations

- Prior to the 2005 field season, it is recommended that protocols for accurate field identification of sculpin species found in the oil sands region (i.e., slimy and spoonhead sculpin) be developed.

- At the earliest possible time, a second non-lethal sentinel species monitoring program should be undertaken to provide further data for comparison with the lethal approach. Currently, condition factor is the only measurement endpoint that has an assigned impact criterion. Implementation of further non-lethal sentinel species programs under RAMP will help to provide the basis for formulation of impact criteria for other proposed measurement endpoints associated with fish growth, survival and reproduction.

- A second non-lethal biopsy sampling pilot study should be conducted in conjunction with the next RAMP fish tissue sampling program.

7.2.6 Acid Sensitive Lakes

- Analysis of the measurement endpoints, including the critical load of acidity, was based on annual measurements that, over the years of the ASL program, varied between August and October. As a result, there is little information on seasonal changes and only a poor estimate of natural background variability of these parameters. Seasonal sampling on 10 of the 50 RAMP lakes has been initiated by CEMA, but the data are currently not available for analysis. It is recommended that RAMP request the results of the seasonal sampling program so this information can be incorporated into the RAMP program.

- Given the results of the trend analyses for Acid Sensitive Lakes, increased monitoring effort should be applied to those lakes receiving the highest rates of acidifying substances.

- With incorporation of the organic component, and further modifications to be applied to the runoff term of the Henriksen model, the model has now probably been pushed to its limits. It is recommended that RAMP consider adopting a more realistic dynamic model.