Manuscript: Documentation of the Integrated Ecological Response Model (IERM2) for the International Upper Great Lakes Water Levels Study

Author(s): LimnoTech

Name of Reviewer: Patricia A. Chambers, Ph.D

1.	Are the objectives of the work clearly stated?	1 **2** 3 4 5
2.	Are the methods employed valid, appropriate and sufficient to address the questions, hypotheses or the problem?	1 2 3 **4** 5
3.	Are the observations, conclusions and recommendations supported by the material presented in the manuscript (e.g., data, model and analyses)?	1 **2** 3 4 5
4.	Are the assumptions used valid and are the mathematics presented correct?	1 2 **3** 4 5
5.	Is the manuscript well organized, material precise and to the point, and clearly written using correct grammar and syntax?	1 **2** 3 4 5
6.	Are all of the figures and tables useful, clear, and necessary?	1 **2** 3 4 5
7.	What is the quality of the overall work?	1 2 **3** 4 5

Recommendation (please circle your response) – **C (concept good; data in the model needs validation and information on confidence)**

A - acceptable
B - acceptable with suggestions for revision
C - acceptable if adequately revised
D - unacceptable

If you have selected **C**, do you wish to receive the revised manuscript for further review? yes **no**

Rating (Circle the rating you would like to give this manuscript. Unacceptable work should be given a score of 40 or less.)

100 90 80 70 60 50 40 30 20 10 0

concept good – 80
data in the model needs validation and information on confidence - 50

A. What is the best/most unique part of the analysis?

The authors have produced a decision model that incorporates biotic and water quantity regime information, generating outputs that show potential impacts at number of temporal and spatial scales in the Upper Great Lake System. This decision model is a valuable assessment and communications tool for both decision-makers and the concerned public.

B. What is the most critical aspect of the study/analysis? Why?

As with any model, the results are as reliable as the data that go into the model and the assumptions in the model design. In the case of IERM2, the data with the largest uncertainty are the tolerances of the nearshore aquatic life. See detailed comments below.

C. Which aspect of the analysis/modeling is weakest? Why? How can it be improved?

The coping criteria should be defined using statistical approaches to identify thresholds (e.g., regression tree analysis) or, for indicators where recommended ecological limits are available (e.g., wetland habitat required to support waterfowl populations), predicting the water quantity regime that sustains the desired ecological endpoint based on regression relationships between the indicator and water quantity. Such quantitative approaches for identifying thresholds are critical for moving forward and are the direction being used in the newer scientific literature. See detailed comments below.

D. Are there any other suggestions that are related to how this analysis may be used more effectively or the results explicated in a more understandable manner?

The IERM2 includes a large number of indicators, making it difficult for managers and informed public to absorb and weigh in decision making. I suggest omitting indicators that are derived from plant indicators, and "bundling" plant indicators on the basis of trait or attributes of species. See detailed comments below.

Confidential comments to the Co-Chair(s) of the Independent Peer Review Group

The goal of this report was to devise and implement an approach that would allow evaluation of the impacts of the current and alternative Lake Superior regulation plans on ecosystem health of the of the Upper Great Lakes system. The report describes a decision model that incorporates existing data on tolerances of nearshore aquatic life (macrophytes, benthic invertebrates and fish) to changes in the water quantity regime (magnitude, timing and duration of water level or flow). Also included in the decision model are the current and alternative Lake Superior regulation plans. By combining these two aspects (aquatic life tolerances to changes in water quantity regime and current/alternative regulation plans), the model produces graphs and tables showing where and under what regulation conditions nearshore aquatic life is likely threatened.

I commend the authors on producing a decision model that incorporates biotic and water quantity regime information, generating outputs that show potential impacts at number of temporal and spatial scales. This is a valuable assessment and communications tool for both decision-makers and the concerned public. As with any model, the results are as reliable as the data that go into the model and the assumptions in the model design. In the case of IERM2, the data with the largest uncertainty are the tolerances of the nearshore aquatic life. Considerable attention (all of Chpt 2 and most of each fact sheet) is paid to identifying relationships between ecological indicators (listed in Table 2-3) and water quantity regime (magnitude, timing and duration of water level or flow). Yet even where such relationships are quantified, they are rarely used to identify thresholds (i.e., boundaries associated with ecological impairment). For example, information on relationships between Georgian Bay wetland area connectivity and wetland SAV transformation with water level are presented in Chpt 2 (pg 35-36) and in Appendix fact sheets 08 and 09. The Zone C criteria for both these indicators are identified as ~176 m asl (in the fact sheets and Table 3-1). However, it is never explicitly stated as to how this water level was derived. Was a regression tree analysis performed on the indicator vs water level datasets to find a break point in the relationship? Or was there some pre-determined ecological condition that was judged important to protect (say, for example, 80% of wetland area is needed to sustain waterfowl populations)? In Chpt 3 (pg 53), the authors state that Coping Zone criteria were based on expert judgment, taking into consideration observed and model results; however, I could not find a statement of exactly how this was done. While professional judgment takes advantage of local knowledge on hydrology and ecology, the resulting criteria may not be objective and are not associated with any measure of reliability or variability. Ideally, criteria should be defined using statistical approaches to identify thresholds (e.g., regression tree analysis) or, for indicators where recommended ecological limits are available (e.g., wetland habitat required to support waterfowl populations), predicting the water quantity regime that sustains the desired ecological endpoint based on regression relationships between the indicator and water quantity. Such quantitative approaches for identifying thresholds are critical for moving forward and are the direction being used in the newer scientific literature (e.g., Wang et al. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: Implication to nutrient criteria development. Environ. Manage. 39:194-212.) If circumstances do not allow these quantitative approaches to be applied now, then Table 3.1 should clearly indicate how each criteria was derived (i.e., professional judgment, protection of a particular condition, regression tree analysis of X on Y, etc.) . Although Table 3-1 includes a column entitled "Goal in Avoiding Zone C", these are not quantified conditions (i.e., what constitutes a significant decline in pike abundance? Or what constitutes range compression – a reduction from long-term conditions of, say, no more than 25%?) The approach for setting thresholds needs to be clearly documented so that others can use this approach and verify it.

The IERM2 includes a large number of indicators, making it difficult for managers and informed public to absorb and weigh in decision making. I have two suggestions to reduce the number of indicators:

(1) many indicators for consumers (e.g., macroinvertebrates, fish, birds) appear to be derived from plant indicators (e.g., macroinvertebrate abundance as a function of bulrush marsh, for Saginaw Bay). My opinion is that derivation of consumer indices from plant indices is over-extrapolation of the data. I would omit these types of indicators. If there were consumer indicators that were independently derived (i.e., fish abundance in relation to sill height), then these could be retained.

(2) There are a large number of plant indicators, and some entail the same metric (e.g., areal cover) but for different plant species. There is a move in ecology to group species on the basis of attributes or traits (see Willby et al. 2000. Attribute-based classification of European hydrophytes and its relationship to habitat utilization. Freshwat. Biol. 43: 43-74). A number of the plant indicators might be combined – e.g., area of anchored emergent plants – and then criteria set for specific habitats. Doing this would also allow comparison of whether the same indicator responded differently (or not) in various geographic locations.

Signature: Date: April 14, 2011

Comments for Transmission to Authors

The goal of this report was to devise and implement an approach that would allow evaluation of the impacts of the current and alternative Lake Superior regulation plans on ecosystem health of the of the Upper Great Lakes system. The report describes a decision model that incorporates existing data on tolerances of nearshore aquatic life (macrophytes, benthic invertebrates and fish) to changes in the water quantity regime (magnitude, timing and duration of water level or flow). Also included in the decision model are the current and alternative Lake Superior regulation plans. By combining these two aspects (aquatic life tolerances to changes in water quantity regime and current/alternative regulation plans), the model produces graphs and tables showing where and under what regulation conditions nearshore aquatic life is likely threatened.

I commend the authors on producing a decision model that incorporates biotic and water quantity regime information, generating outputs that show potential impacts at number of temporal and spatial scales. This is a valuable assessment and communications tool for both decision-makers and the concerned public. As with any model, the results are as reliable as the data that go into the model and the assumptions in the model design. In the case of IERM2, the data with the largest uncertainty are the tolerances of the nearshore aquatic life. Considerable attention (all of Chpt 2 and most of each fact sheet) is paid to identifying relationships between ecological indicators (listed in Table 2-3) and water quantity regime (magnitude, timing and duration of water level or flow). Yet even where such relationships are quantified, they are rarely used to identify thresholds (i.e., boundaries associated with ecological impairment). For example, information on relationships between Georgian Bay wetland area connectivity and wetland SAV transformation with water level are presented in Chpt 2 (pg 35-36) and in Appendix fact sheets 08 and 09. The Zone C criteria for both these indicators are identified as ~176 m asl (in the fact sheets and Table 3-1). However, it is never explicitly stated as to how this water level was derived. Was a regression tree analysis performed on the indicator vs water level datasets to find a break point in the relationship? Or was there some pre-determined ecological condition that was judged important to protect (say, for example, 80% of wetland area is needed to sustain waterfowl populations)? In Chpt 3 (pg 53), the authors state that Coping Zone criteria were based on expert judgment, taking into consideration observed and model results; however, I could not find a statement of exactly how this was done. While professional judgment takes advantage of local knowledge on hydrology and ecology, the resulting criteria may not be objective and are not associated with any measure of reliability or variability. Ideally, criteria should be defined using statistical approaches to identify thresholds (e.g., regression tree analysis) or, for indicators where recommended ecological limits are available (e.g., wetland habitat required to support waterfowl populations), predicting the water quantity regime that sustains the desired ecological endpoint based on regression relationships between the indicator and water quantity. Such quantitative approaches for identifying thresholds are critical for moving forward and are the direction being used in the newer scientific literature (e.g., Wang et al. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: Implication to nutrient criteria development. Environ. Manage. 39:194-212.) If circumstances do not allow these quantitative approaches to be applied now, then Table 3.1 should clearly indicate how each criteria was derived (i.e., professional judgment, protection of a particular condition, regression tree analysis of X on Y, etc.) . Although Table 3-1 includes a column entitled "Goal in Avoiding Zone C", these are not quantified conditions (i.e., what constitutes a significant decline in pike abundance? Or what constitutes range compression – a reduction from long-term conditions of, say, no more than 25%?)   The approach for setting thresholds needs to be clearly documented so that others can use this approach and verify it.

The IERM2 includes a large number of indicators, making it difficult for managers and informed public to absorb and weigh in decision making. I have two suggestions to reduce the number of indicators:

(3) many indicators for consumers (e.g., macroinvertebrates, fish, birds) appear to be derived from plant indicators (e.g., macroinvertebrate abundance as a function of bulrush marsh, for Saginaw Bay). My opinion is that derivation of consumer indices from plant indices is over-extrapolation of the data. I would omit these types of indicators. If there were consumer indicators that were independently derived (i.e., fish abundance in relation to sill height), then these could be retained.

(4) There are a large number of plant indicators, and some entail the same metric (e.g., areal cover) but for different plant species. There is a move in ecology to group species on the basis of attributes or traits (see Willby et al. 2000. Attribute-based classification of European hydrophytes and its relationship to habitat utilization. Freshwat. Biol. 43: 43-74). A number of the plant indicators might be combined – e.g., area of anchored emergent plants – and then criteria set for specific habitats. Doing this would also allow comparison of whether the same indicator responded differently (or not) in various geographic locations.

<u>Specific Comments</u>

Table 2-2 – references are needed to support the approaches

Pg 11, paragr 2 – "primarily based on maximum or average water depth". Do you mean ***<u>minimum</u>*** water depth?

Pg 17, last parag – you discuss the need for vegetation to support consumer communities. What about vegetation preferences? Are some vegetation types or species better habitat than others?

Pg 24-25, Sect. 2.2.1.c – not clear what a P release indicator tells. What if released P is rapidly taken up by growing vegetation? How does a change in P release rate inform regulation plans?

Pg 27 – why are some indicators related to elevation and others to water depth? It would be easier to follow if only a limited number of water quantity metrics.

Pg. 30, paragr 2 – the indicator described here is call “total area of fish habitat”. However, it is more appropriately called aquatic plant cover.

Pg 31, 1st paragr – “specific flooding/dewatering rules”. What *rules*?

Pg 37 – “percentage of wetlands that would become transformed”. Transformed into what?

Pg 38, Fig 2-13. What are the lines on the graph?

Pg 45, Fig 2-17. Variance or standard error for the data?

Fig. 2-18, Fig. 3-3 – I don’t understand these target diagrams. A better explanation is needed.

Table 3.1 – For SUP-01, Zone C condition - text missing?? “water level during is less than”