International Upper Great Lakes Study – Sub-Product Reviews, Synthesis Product Reviews, and Draft Final Study Report Reviews Template

Peer Review of Manuscripts

This manuscript has been submitted for independent peer review to the Co-Chairs of the Independent Peer Review Group (IRG) as identified in the Independent Review Plan (IRP) of the International Upper Great Lakes Study (IUGLS).

The evaluation and acceptance of the technical report (documentation) will include, as part of the review criteria, how effectively the goals of the work have been accomplished within the limits as described in the "background and context statement in Article 9.3.1.2.

Manuscripts shall be evaluated on the extent to which the authors' efforts have been covered/documented and the extent to which the reviewers can answer the review questions:

- Are the methods employed by the authors sufficient to answer the questions;
- are they being used correctly;
- are the analyses and tests appropriate for the problem at hand; and
- are the derived conclusions supportable by the model and analyses?
- Are there any other comparable methods or approaches that may/ought to be considered, which would provide more insight for the specific task under review?

Checklist for the Reviewer

Your review is:

- To provide the authors with directions as to how they could improve their analysis and technical report. Please provide clear instructions and comment objectively, remembering the efforts that they have made to prepare the manuscripts. On a separate sheet, you may provide comments for the editor that you feel are necessary. These separate comments will not be provided to the authors.

Some additional points are:

- Please document statements adequately so that authors may fully understand your concerns. You may do this using additional sheets and cross-referencing your additional comments to the specific questions below.
- Some of the questions follow a scale of 1 through 5, with 1 be the highest rank (yes -- always or excellent) and 5 being the lowest (no -- never or very poor). Please encircle your responses.

Manuscript: Formulation and Evaluation of New Control Structures in the Great Lakes System

Author(s): Tolson, Razavi, Asadzadeh

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

Recommendation (please circle your response)

C - acceptable if adequately revised

If you have selected **C**, do you wish to receive the revised manuscript for further review? yes

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

70

Comments (limit responses to one paragraph for each question; reference pages, charts, and data. Please distinguish if responses are of major or minor concerns.)

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

The best or most unique aspect of the work is the coordinated simulation of the broader systems for a multi-lake regulation plan. The authors have made a significant effort to strongly honor historical work while expanding to this broader focus.

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

The problem formulations shown in equations (20) and (21) that seek to minimize relativistic changes in lake levels from a simulated baseline are the primary drivers of the results and recommendations of the report. Additionally the "50,000 year sequence" used to test the robustness of optimized results is the primary tool for trying to evaluate variation beyond what has been historically observed.

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

Overall the weakest points of the analysis lie in the problem formulation and 50,000 year sequence. Overall the report states that an investment of $8.5 Billion will potentially reduce extreme deviations from the baseline scenario. The challenge in these findings is it is unclear what the objective in equation (20) means? What impacts would the variation have? Are there thresholds for damage? Failures? The multiobjective approach is interesting but the problem formulation seems highly limited and distant from stakeholder metrics of concern (risk, vulnerability, reliability, resilience, impact). To be clear, I am not stating that everything must be in monetary terms but it would have been easier to evaluate the results had the performance objective been formulated in the context of more traditional vulnerability or reliability measures. There is little physical meaning the in the baseline at present. It is also unclear when or where the min-max/max-min robustness formulation in equation (21) was implemented. The work has an overall view of trying to minimize time of analysis and computational demands which seems inappropriate. Given the $8.5 Billion recommendation, I would strongly submit that a more careful and computationally explorative analysis be completed before any of these recommendations are considered.

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 authors failed to specify how the 50,000 year stochastic sample is generated and consequently it is unclear how it is representative of changes in extreme conditions. Does it provide variability beyond the historical observed ranges? This is a strong weakness in the report.

Please indicate any confidential comments to the Co-Chair(s) of the Independent Peer Review Group in the space below. Comments for transmission to the author(s) should be on a separate sheet attached.

Confidential Comments:

The authors likely due to time constraints have chosen to complete the analysis using their own in-house technologies (i.e., the DDS optimization tool). They repeatedly stress the computational demands of the simulation (which is ~3 min) and use this to justify a limited exploration of the problem both from a statistical and an optimization perspective. In reality their tools are not well known and may be difficult to justify relative to other tools. The "50,000 year sequence" is wholly undefined and does not in any way clarify how it relates to expected changes in the system. Given the immense scale and cost of the proposed recommendations, the analysis of this report is very limited.

Comments for Transmission to Authors

It would be useful to have both general comments and specific comments for major and minor revision. Please use additional sheets should they be required.

Major Comments:

Page 9, Several of the simulation components have very sparse textual description. As an example the simulation of Lake Ontario simply states Plan 58DD is used based on a compiled program. This seems insufficient for a reader to understand the simulation.

Page 9, The justification and implications of the simplified St. Lawrence system simulation is lacking in the report. It is unclear how well the empirical equations perform, their biases, and how systematic errors could influence the adaptive management recommendations via optimization.

Page 10, There is a disparity of time scales in the simulation models (monthly, weekly, daily). The present report does a poor job of describing the necessary control time scale and the consequences of the selected time steps in its component simulations.

Page 11, The numerical approximation in the two point simulation of moving from hourly to daily time steps should be better supported in the authors contention that it has little impact.

Section 2.3 highlights concerns on the ability of the simulation to handle extreme outflows particularly for Lake MH, Lake St. Clair, and Lake Erie. A proposed "patch" solution is described where the outflows causing the issue are zeroed in the corresponding time step. This issue was not clearly addressed or described. It has consequences for extreme hydrologic sequences of interest throughout the adaptive management framework. Does this artificial simulation constraint strongly bias management of extremes?

Overall text in Section 2.5, preceding 2.5.1 is insufficient to provide a clear summary of the decision variables in the problem. The decision variables should be defined and summarized with their ranges in a summative table. Perhaps a detailed table could be used in the appendix and a less detailed summary table could be included in the text that explicitly list the typical 10 variables used (where the text highlights the 9 variable instance). This is strongly required to improve the clarity and reproducibility of the study.

The text support for the objective function in equation (20) is unclear. The 4 priorities claimed in this section do not make sense. There are two priorities that are mixed. Both of the formulations in equation 20 and 21 differentiate instances of violation of extremes by minimizing squared deviations when performance is worse than baseline and maximizing improvements relative to the baseline using the binary Z variable. In (21) the min-max and max-min behavior is used to proxy robustness across tested conditions. It is also unclear how the n case mixtures influence optimization. It would be informative to provide cumulative distributions for the typical objective components (i.e., a separate CDF for each term of equations 20 and 21). A Latin Hypercube sampling of the simulation would provide this and help to support the degree of variability in the components. Also is there a homogeneous impact from deviations? Across locations are deviations from the baseline equally important and with similar economic impacts? How sensitive is the system overall relative to the baseline? It unclear how much the decision variables impact baseline deviations? Should there be physical constraints beyond the simulation?

In Section 3.1 and overall in the report, it is unclear what purpose or advantage the formulation in equation 20 provides.

The first paragraph of Section 3.1 provides important methodological assumptions that would be more appropriate in the Methods. The validity of the underlying assumptions for the Lower St. Lawrence need to be better supported.

Page 24, The authors give some specifics on the solution algorithm (DDS) and computational demands in Section 3.1 but fail to place their work in the broader context of tools and computational frameworks. It is stated that a 2.7 minute simulation time requires the use of the DDS algorithm without any discussion of other tools that exist. Also it is unclear if the authors have made any effort to free their analysis to broader analysis using high performance computing which is used commonly in major water resources planning efforts by the US Army Corps and other groups (e.g., coastal hurricane analysis). Given the severe costs of the recommendations, why is a minimal analysis justified?

Page 24, The authors make reference to "pre-emption" without sufficiently describing the approach for attaining savings.

In Section 3.2 it is unclear how the 50,000 year stochastic sample is generated and consequently it is unclear how it is representative of changes in extreme conditions. Does it provide variability beyond the historical observed ranges? This is a strong weakness in the report.

Overall the report states that an investment of $8.5 Billion will potentially reduce extreme deviations from the baseline scenario. The challenge in these findings is it is unclear what the objective means? What impacts would the variation have? Are there thresholds for damage? Failures? The multiobjective approach is interesting but the problem formulation seems highly limited and distant from stakeholder metrics of concern. To be clear, I am not stating that everything must be in monetary terms but it would have been easier to evaluate the results had the performance objective been formulated in the context of more traditional vulnerability or reliability measures. There is little physical meaning the in the baseline at present.

Minor Comments:

Page 7, 50,000 year sequence is awkward phrasing

Page 7, "...the robustness of the UW plan for water supply sequences the was not optimized for." This phrasing is awkward and confusing.

Page 11, An overview illustration of components, their time scales, and the overall assumptions in the simulations would be beneficial.

Page 13, The term quarter monthly seems awkward. Why not use weekly? Weekly is less ambiguous as a time step measure given the non-standard duration of months.

Page 13, How important is the effect of the monthly time step for Lake Superior following Plan 77A on adaptive options?

Page 14, Figure 1 needs improved labeling on y-axis representation of releases. It may also help to re-label "e" and "f" given that they are not functional slopes. You could improve your notation to have the thresholds more uniquely and explicitly labeled.

Page 18, 1st text after (19), change "that level" to "the level"
The authors in Section 3.1 state six trial optimization runs where used to avoid poor results. It would be beneficial to report the variance in the results attained. Also the quality of the optimization is justified by the failure of the pattern search by Torczon, 1997 to improve results. This seems like a weak baseline justification. The authors could have demonstrated search dynamics for at least one case for a larger number of evaluations and potentially baselined DDS state-of-the-art tools outside of the water literature.