International Upper Great Lakes Study Subproduct Review of "Stochastic Hydrology of the Great Lakes – A Systemic Analysis"

By Thorsten Wagener, PhD, A.M. ASCE

18th May 2011

Manuscript: Stochastic Hydrology of the Great Lakes – A Systemic Analysis [A Compilation of projects completed under the Hydroclimatic Work Group for the International Upper Great Lakes Study]

Author(s): Laura Fagherazzi, Ousmane Seidou, Taesam Lee and Taha Ouarda

Name of Reviewer: Thorsten Wagener

---

1. Are the objectives of the work clearly stated? — 2

2. Are the methods employed valid, appropriate and sufficient to address the questions, hypotheses or the problem? — 2-3*

3. Are the observations, conclusions and recommendations supported by the material presented in the manuscript (e.g., data, model and analyses)? — 2-3*

4. Are the assumptions used valid and are the mathematics presented correct? — 1-2*

5. Is the manuscript well organized, material precise and to the point, and clearly written using correct grammar and syntax? — 3*

6. Are all of the figures and tables useful, clear, and necessary? — 2-3*

7. What is the quality of the overall work? — 2-3*

*see detailed comments

Recommendation (please circle your response)
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? — -

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

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

I am making some overall comments here, but the specific detailed comments can be found in the comments for transmission to authors section of this review.

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

The consequent use of time series analysis methods is very interesting and it shows the potential of such a strategy while highlighting the general limitations of any study related to long term projections of future environmental conditions. Connecting established time series tools to GCM projections is an interesting strategy.

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

The most critical aspects of this study are how the chosen (time series) approach is justified and how it is tested. I understand the interest for using time series models, though the justification should not be unreasonable. For example it is not really true that GCM output is not available as suggested ion section D (unless there are specific issues in obtaining them for this work that I am not aware of). Especially critical is how such methods are tested. The problem of calibration/validation and how much confidence can be derived from the validation is crucial of acceptance of the results. How different is the expected future from the observed past and what does this mean for the expected reliability of the models derived?

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

I miss a more detailed discussion on how confidence in the results presented has been established considering the focus on understanding the impacts of climate change on the lake system. The authors (correctly) argue that no GCM projection is necessarily the actual future. However, it would be good to discuss in more detail whether the time series models proposed can reproduce the lake system behavior given that a GCM projection of the driving variables was correct. Here it would be good to see more discussion. (See for example the uncertainty discussion in the paper by Koutsoyiannis: http://www.hydrol-earth-syst-sci.net/14/585/2010/hess-14-585-2010.html).

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 main issues with any study of this type relate to the assumptions made and the uncertainties present. I think that the main communication issue is to state these clearly in the beginning of the study and to review them at the end with respect to the conclusions drawn in the report. A bit more discussion of this would be helpful. The authors show a lot of results, which generally look convincing, but they could explain them in a bit more detail. I have made some suggestions in this regard in the comments to authors' section of this review.

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.

Signature: Date: 18th May 2011

Comments for Transmission to Authors

Objective of report: "*The objective of this section is to present a set of temporal and spatial* [Net Basin Supplies] *NBS characteristics of the different basins forming the Great Lakes system.*"

Comments on: A Statistical characteristics of the historical database

It would be helpful to the reader if the authors would include a map (in the beginning) to show exact locations of the gauges used for analysis. This would especially help to place the differences in the analysis results into context.

It would be good if specific numbers and letters were used to mark each figure so that there is no ambiguity in what the authors refer to. E.g. rather than saying the middle left plot of the 3rd figure, it would be better to say Fig. 3c.

The text provided is purely descriptive. Little attempt is made to explain the results. While this is probably on purpose, it would be very interesting to understand a bit better why there are differences between the time series.

On figures 27 to 33 it would be better to say months, rather than period, and also whether 1 equals January or not. The authors state that they show monthly hydrographs in figure 27 to 33, but the values go frequently negative. How were the time series transformed? What is the reason for the huge variability in the first 6 periods of figure 29?

Comments on: B Stochastic Modeling and Simulation of the Great Lakes System

What is the origin of the differences for the lower tail for Lake Erie (Fig. 3-12 right)?

The generated periodic series (Figs. 3-15 to 3-17) seem to be much more variable than the observed ones. Why?

Observed and generated results for Figs. 3-15 to 3-17 are mixed up in the text on paper 62.

Interesting that the water balance closes that well (p. 69). Other losses (subsurface, atmospheric) do not seem to play a big role, or are they captured in the change in storage term?

It would be helpful if the figure captions were longer and would better describe all aspects of the figures (e.g. all lines in Figs. 5-1, 5-2, 5-4, 5-6).

I am not clear whether all the observations are used for deriving the statistical models or whether some type of validation was possible?

Comments on: C Stochastic Generation of Synthetic Residual Net Basin Supply for the Great Lakes System

The authors state: "*However, the ability of GCMs in reproducing ENSO is believed to be poor.*" (p. 121) Please provide reference for this conclusion. Also, does the ability to reproduce ENSO not vary with GCM and with location?

The third bullet on page 121 refers to scenario A3. I assume this is supposed to be A2? It would also be good to briefly mention the differences between the two scenarios.

What is the general impact of ENSO on the lakes region? How much of the historical variability can be explained by ENSO?

There is some repetition of text in this chapter (e.g. p. 137).

Why would the predictive variables not change in the future (p.139-140)? This is the general question with statistical approaches (e.g. for downscaling), but it would still be good to include a couple of sentences on this point.

There are some problems with automatic referencing on page 146.

This section (C) of the report includes better explanations for figures than the previous sections.

The details (circles) on figures VII-1 and VII-2 are difficult to see. Similar comments can be made for later figures of the same type.

There seems to be a tendency for the simulations of the lake levels to be too low for the low cumulative frequency and too high for the high cumulative frequency values (e.g. IX-2). Why is that?

Comments on: D Predictability of Climate Indices with Time Series Models

The writing in this section is not as good. Should be read through another time by a native speaker.

Claiming that GCM output is not widely available is too simplistic a good starting point (p. 185). It is easy to access GCM outputs and even downscaled products are widely available. If the choice is for time series models, then that is fine, but this should not be based on such claims.

I am not fully clear why a Kalman filter is necessary here? Please explain better.

Page 200: While a calibration/validation process is used, it is not clear to me why the authors did split up the time series according to when change actually occurred (as shown in section A)? Isn't the main question whether the model can reproduce changes in the mean? Wouldn't it make more sense to split the dataset into periods with different means and show that the models can actually reproduce change?

The x and y axes are missing on the figures on page 227.

General Comments

There are some minor spelling issues that the authors will see if they read carefully through the report again. The writing also varies a bit and it would be good to improve especially part D.

What about other changes in the basin that would impact the stationarity of NBS (e.g. land cover change due to urbanization)? It would be good to discuss in the beginning why the authors chose not to include them here.