Exploratory studies into seasonal flow forecasting potential for large lakes

Kevin Sene, Wlodek Tych, Keith Beven

Author response to interactive comment by K. Engeland, July 2017

The title explains well the content of the paper: "Exploratory studies into seasonal flow forecasting potential for large lakes". I think the paper is interesting and provide some useful insights and conclusions.

Thanks for these useful comments. We have provided some responses below including items that we will change when submitting the revised version of the paper

First a general comment on the evaluation of model performance in this study. When assessing forecasting skill, the benchmark that is used as a reference for assessing skill should be given. If there is a strong seasonality in lake outflows, maybe a monthly climatology would be a better benchmark than a long term average. See e.g. Bettina Schaefli and Hoshin V. Gupta (2007) for choosing benchmark in catchments with a strong seasonality in runoff. Another performance measure that could be used is anomaly correlation coefficient.

We agree and, when looking over long time periods, as we note another issue is the non-stationarity in outflows; for example as illustrated by the step change in observed values for Lake Victoria in the early 1960s. We therefore decided to focus on estimates for the errors in annual peak values (Figures 6 and 7) since peaks tend to occur around the same time each year and are the most challenging to forecast. This would then avoid choice of a specific measure, or measures, although again as noted does not take account of timing errors. For the revised paper, we will include some more discussion of these points and possible alternative measure that could be used (e.g. as described by Schaefli and Gupta 2007).

The paper is, in general, well written, but some parts of the manuscript could benefit from more clarity in the presentation. I will give some suggestions below.

Introduction

The introduction is rather brief, and it could be useful to refer to both operational systems and research papers describing approaches that are used for seasonal forecasting of lake levels or outflow. E.g. for the great lakes in US/Canada there is an operational seasonal forecasting service: https://www.glerl.noaa.gov/data/wlevels/levels.html#modelsAndForecasts. In particular if there are other studies on forecasting the water levels in lake Victoria and Malawi could be useful. One recent example is Mulumpwa et al. (2017). I also think the introduction could better reflect the content of the paper, in particular the use of different circulation indices as a predictors for forecasting water levels. Maybe small parts of the data section could be moved to the introduction. In the end of the introduction I miss some clearly stated aims or objectives.

This is a valid point and we will extend the introduction to include more discussion of seasonal forecasting approaches for other lakes and, as suggested, move some of the data section forwards. Thanks also for the Lake Malawi citation which we had not yet seen

Case Studies

Often it is challenging to estimate outflows based on time series of lake levels since the results might be very sensitive to quality of water level observations. In particular for large lakes where one mm water level represents a large volume, using this approach for daily values, results in a lot of noise. It helps to use weekly or monthly values as in this study. I miss a more specific description of the data: What is the time resolution of water level and outflow data you used?

Sorry if this was not clear; the time resolution used was monthly. We will clarify this in the revised version

Methodology

It would be useful if you in the methods section explains more explicitly the combination of models that are used, i.e. how is the net inflow model combined with the lake response model. Further on, how are the regression and ARMA models are used, i.e. is the residuals of the lake response model the dependent variable?

Again we will clarify this in the revised paper (and yes the residuals were the dependent variable)

Results

The previous comment on telling what is used as dependent variable in the regression and the ARMA modelling is important when presenting results on lines 12-16 on page 15.

Agreed, we will include further discussion of this point

Section 3.2 "Net inflow estimates" is maybe not very precise. As I understand, you want to use this model as a simple forecasting model where forecasted precipitation is used to drive the model. Would "Net inflow forecasting" be a better sub-title?

Agreed, we will change this in the revised version

Discussion

Many of the great lakes are located in areas with a seasonal snov cover. In the introduction seasonal forecasting in snow dominated catchments is mentioned, but it could be useful to speculate on seasonal forecasting for catchments with large lakes and seasonal snow cover.

This is not something that we looked at but we are happy to include some speculation regarding this point

Figures

Figure 2: could change the scale of the y-axis to be between 0.5 and 1.5

Agreed, we will change this

Figure 4: It is difficult to see the observations (the dots). It also seems like the 95% confidence intervals are too wide since all observations are well inside this interval.

This was just a small part of the record so is perhaps not representative of the overall performance; we will revisit this and include a different figure and/or a comment in the revised version

Please comment.

Figure 5: It is difficult to see the difference between the lines.

The figure will be improved

Equations

Equation 1 and 2: I have some questions about dimensions in these equations. On the left hand side, dh/dt has the dimension length/time, so then N should also have the same dimension. Then "depth per unit area of lake surface" is confusing. I suggest to use "Volume flux per unit area of lake surface". In Equation 2 I miss a t. Either should (i) the fluxes P, E, Qc and Q0 be integrated over the time interval t in order to become water depths, or (ii) h be divided by t in order to become a flux. If the latteralternative is used, it could be useful to state that all P, E, Qc , Q0 are average fluxes over the time interval t.

We agree that the description could be improved and will do that in the revised version

Equation 7 and 8: it could be useful to avoid using a and b here since these symbols are already used in Equation 3.

Agreed, this will be changed

Equation 7: What is n and yt?

Again, this will be clarified

Equation 9: It is is difficult to understand this equation. What is A and B? previously A was used for lake surface area.

We agree that this could be confusing; however by convention A is used both for lake areas and in this general form of the transfer function equation. However in the latter case A and B are sometimes written in a bold type face so that is what we propose here

Equation 12: This equation is not necessary.

Yes, this will be removed

References:

Bettina Schaefli and Hoshin V. Gupta (2007) Do Nash values have value? Hydrol. Process. 21, 2075–2080 (2007) DOI: 10.1002/hyp.6825

M. Mulumpwa, W. W. L. Jere, M. Lazaro, A. H. N. Mtethiwa n(2017) World Academy of Science, Engineering and Technology, International Journal of Environmental and Ecological Engineering Vol:4, No:3, 2017 Forecasting Lake Malawi Water Level Fluctuations using Stochastic Models

Thanks for these two very useful references; we will include citations and a brief discussion of the findings they describe