July 2020

Review of "New flood frequency estimates for the largest river in Norway based on the combination of short and long time series" by Engeland *et al*.

Manuscript ID: hess-2020-269

Review conducted by:
Dr Daniel Schillereff
Department of Geography, King's College London
daniel.schillereff@kcl.ac.uk

This manuscript reports a Holocene palaeoflood reconstruction that amalgamates sedimentary, historical and instrumental data to refine flood frequency estimates for the Glomma catchment. I was really pleased to read and review this submission because of its interest and importance. Palaeoflood research has made big strides in recent years and embedding these data into design flood estimates is the next big task – and this is precisely what the authors have achieved in this paper. The manuscript is well written, clear and most of the analysis is convincing. The amalgamation and systematic analysis of independent (at least in terms of recording method) hydrological datasets is particularly effective and commendable. There are a handful of moderately substantive areas of clarification that I suggest are necessary prior to publication. I think these will result in a clearer narrative for the broad journal audience and present a more convincing analysis. A few minor comments are also raised.

Comment #1: Structure and content of the introduction

Given the broad audience of HESS, and the likelihood that some (many?) readers may be more accustomed to studying recent datasets, I suggest the authors incorporate more detail on the palaeodata in the introduction. I specifically suggest the authors elaborate on the ways in which palaeo data can help address the "two reasons" outlined on Page 2, Lines 23-26. Whilst many readers will have a sound knowledge of return periods and design flood estimates, applying palaeo data to these processes – and the value of doing so - may be quite new.

Similarly, I think it would be useful for the authors to say more about the timescales involved: how far back in time can palaeohydrological data be obtained from lake sediments (page 2, lines 42-43) and historical records (page 3, lines 5-11)? This would give readers a platform of knowledge that will be helpful when they engage with the return period calculations later in the paper. Likewise, I think the objectives (Page 3, lines 39-43 and on to Page 4) could be a bit more specific and pay particular attention to the timescale of your analysis

I also query whether Page 3, lines 5-11 could be slimmed down. To keep the focus on your work, perhaps mention briefly that there are a range of historical archives but place particular emphasis on epigraphic sources – especially flood stones - as that is the sort of data used later in this study

Comment #2: Human modifications to the landscape and flood stationarity

I am pleased the authors mention the possibility that changes in land-use could play a role and, similarly, I was really pleased that the authors explicitly assess whether the fluvial system behaviour would have allowed bifurcation events throughout the Holocene (Page 22, Lines 18-30). I think there is scope to go further in ruling out the possibility that changes in flood frequency are driven by rather than amplified by human activity.

Page 3, line 24: it is important to state here that non-stationarity needn't only be a response to climatic variability. Human alteration of the land surface can also create a non-stationary fluvial system.

Page 5, Lines 2-5 talk about 'noteworthy land-use changes during the last 400 years, and specifically the removal of woodland cover". This is returned to briefly on Page 25, Lines 6-15 but I think this needs a more critical and in-depth evaluation. The authors acknowledge, for example, the notable rise in flood frequency rises around 500 yr BP, which happens to coincide with the assertion on Page 25 Line 8 "The mining industry that started in Norway in the 16th century required a large amount of timber which resulted in widespread deforestation also in Glomma's catchment". How confident are the authors that widespread deforestation amplified the climate driver but was not a driver of sedimentological change in its own right?

Similarly, I think the interpretation of external drivers, especially through the 2500-4000 yr BP flood-rich period, would be strengthened considerably if human interference could be ruled out. In Britain, for example, a number of fluvial and palaeolimnological studies show widespread mobilisation of sediments at that time resulting from settlement expansion. I have no idea about the mid- to late-Holocene history of human occupation in southern Norway but presenting or referring to data ruling this out possibility would really strengthen the case.

Comment #3: Evaluating the geochemical flood proxy

Overall, I find the proxy reconstruction to be convincing. I do wonder whether it could be strengthened by providing some information on the geochemical composition of the glacio-fluvial material between the two lakes. The authors state the Glomma catchment is the largest in Norway. To what extent will sediments being deposited in Flyginnsjøen during a bifurcation event be mixed with material entrained elsewhere in the Glomma catchment?

I commend the authors' application of a rigorous peak detection algorithm and critical assessment of the fidelity with which the sediment record matches the gauged (post-1953) flood record. However, I found Figure 12 and the associated text a bit difficult to follow. For example, Page 18, Lines 12-15 the authors highlight that, overall, there is a good match but also segments that do not correspond, but this analysis would be strengthened had if it was easier to figure out which XRF peak linked to which flood volume bar. One idea would be to colour the vertical flood volume lines and the circle/cross shapes one of two colours when you are confident in their stratigraphic correspondence and the other colour when a match is more difficult to establish. A minor point but, personally, I think the dashed vertical lines denoting each year add substantial clutter to the graph and could be removed. Overall, Figure 12 makes a really important contribution to the paper but it's complicated and uses many different colours and shapes. Improving its aesthetics would really strengthen the paper.

Comment #4: Summer temperatures as the primary driver

I would like to see a bit more detail on the physical flood generating mechanisms. From the discussion across pages 23 and 24, it follows that warmer winter air temperatures reduce annual snow storage and, in turn, the magnitude of the spring melt. But I note July temperature is used as the primary meteorological proxy (Figures 17 and 18 and associated text) and I can't figure out why summer temperature is more important than winter or spring temperature. I would have thought winter temperature would dictate snow volumes and spring temperature would influence rate of melting. I may well have misunderstood but, given this is a fundamental aspect of the interpretation, I suggest modifying the text – and reporting a comparison with winter and/or spring temperatures, if appropriate reconstructions are available - to ensure the reader can follow the process linkages at play. Similarly, Page 23 Line 28 also mentions the importance of winter precipitation. Does a regional precipitation reconstruction exist? I'm not doubting the authors' interpretation but I found myself wondering about the role of other drivers while reading this section so there is scope to tighten up the narrative and analysis here, in my opinion.

Comments on figures and tables:

Figure 3: do the authors have a photo or aerial/satellite imagery of the bifurcation inflow zone? I am intrigued but I am struggling to visualise what this looks like on the ground. Is there a dry channel or other morphological evidence to indicate this reverse flow occurs on occasion?

Figure 4 and 5: Figure 4 suggests the 'normal-flow' inlet to Flyginnsjøen is in the same corner but not in exactly the same place as the inflow under flood conditions, whereas Figure 5 lists only one inlet. I recognise this will have minimal effect but worth clarifying this morphology.

The sequence of maps are a bit difficult to follow. For example, it's unclear whether Kongsvinger is a town. Perhaps the inset map in Figure 3 that has one red dot and the boundary of Norway could instead zoom in on a slightly larger area such that Elverum is also visible? Similarly, the locations of the lakes, gauging stations, flood stones, towns and other features mentioned in the text are spread across 3, 4, 5 and 7 and I found myself having to flick back and forth between them.

Figure 6: Given its use as a thresheld, it would be worth labelling the 1967 flood in Figure 6

Figures 14 and 15: I suggest the authors provide much more context and technical detail in the captions for Figure 14 and 15. I recognise these procedures are explored in the main text but, given their importance to the overall narrative, I think it would be really useful to present ample detail such that both figures can be interpreted on a standalone basis.

Table 3: It took me a while to figure out the source of the pairs of correlation coefficients and why some were missing. I understand why the authors have presented the data in this way and it is probably fine to do so with some additional explanation. This might be resolved by writing in the table caption which parameters were measured on which core and also being more explicit on which way round the numbers are reported. Or maybe report the coefficients for one of the cores entirely in italics?

Minor comments

Section 3.3.1: given the broad audience of the journal, I question whether all readers will be aware of the reasoning behind using and integrating two cores (and indeed two types of corer).

Page 11, line 27: the authors state they used a 9-month window in the peak detection algorithm. I like this approach as a way of considering event sequencing but what is the hydrometerological basis for the 9-month window?

Page 13, line 33: I suggest the authors report a range of layer thicknesses rather than stating "mm scale".

Page 19, Line 11: Judging by eye, there is a more prominent step in flood occurrence rate at 700 yr BP rather than 600 yr BP?

Page 19, Lines 13-14: I find the assertion that "high flood frequency in the 18th century is also recorded in the historical flood data (Fig. 6)" to be unconvincing. There are very few data points prior to the 18th century (one?). As long as the authors can be confident the 15th and 18th-century peaks are not triggered by anthropogenic landscape modification, then the sediment record speaks for itself.

Page 20, Lines 18-20: I found it difficult to follow the sequence of different approaches applied in Section 4.3. In particular, which is "case ii above" (Line 18) and which is "case iii" (Line 19)?

Technical corrections: The manuscript is, for the most part, written in clear, concise prose with ample technical detail. There are a handful of minor typographic errors - mostly inconsistent verb tenses in individual sentences. The clarity of the prose is not affected but no harm in tidying these.