Natural Hazards
and Earth System
Sciences

Open Access

Discussions

Interactive comment on "Reassessing flood frequency For the Sussex Ouse, Lewes: the Inclusion of historical Flood Information since AD 1650" *by* N. Macdonald et al.

N. Macdonald et al.

neil.macdonald@liverpool.ac.uk

NHESS manuscript nhess-2013-363, Reassessing flood frequency For the Sussex Ouse, Lewes: the Inclusion of historical Flood Information since AD 1650 by N. Macdonald, T.R. Kjeldsen, I. Prosdocimi and H. Sangster

Point to point authors' reply to the Referees' reports. Please find below responses addressing the reviewers comments, specific points have been directly addressed in the revised paper attached and are not detailed here. We would like to thank the three reviewers for their detailed comments on the paper, which have improved it considerably.

C3123

Referee #1 The broad content of this paper is an important one and the methodology and sensitivity analysis of sufficient originality that the paper is recommended for publication. However, I have serious problems with the detail. In particular I am concerned with the reliability of the data used to demonstrate the method. The answers to these questions may (or may not) be in previously published papers but for this paper to stand alone, it needs to address more thoroughly the reliability and homogeneity of the historic record. We thank the reviewer for their positive comments and recommendation. We hope that the comments below and additional materials added into the paper help address these concerns.

The paper makes comparisons between short gauged records and records extending using historical information. It makes no mention of the alternative method of extending a local record by the use of pooled information from similar catchments. Whilst it is legitimate in the context of the paper to omit a pooled estimate, I believe that the method should at least be mentioned as it is the recommended first choice of FEH. If I were an EA flood risk manager, I would definitely like to see a comparative pooled estimate. This is an interesting point, currently within the FEH-WINFAP software the option to include a sight which suffers some tidal flow is not easily done, to explore this a little further we have forced WINFAP to run with an estimated catchment and by keeping all catchment descriptors the same as the tidal point, this results in a return frequency of a 100 year flood if 177.9m3s-1, a discharge smaller than that derived from any of the methods applied within this paper.

The paper skips too easily over the problems associated with the use of historical flood information in flood risk assessment. Whilst you note that \during the largest flows, relatively minor modifications within the channel and catchment may have minimal impact on flood discharge", this needs to be addressed more comprehensively than reference to a previous paper. Within the space provided we are unable to go into detailed examination of all the different elements of the study, as this it is not the primary focus of the paper and has previously been examined in greater detail elsewhere (as commented

C3124

upon by Reviewer III) we do not feel it would be appropriate to explore this further here.

How well does the measurement of flow at the two gauging stations correspond with the flow at Lewes? You state (page 7 last line) that \no other significant flows enter the system between the town and Ouse (presumably you mean Gold) Bridge". However there is quite a large catchment area which includes the Longford Stream and the Bevern Stream. At least you might indicate the additional area contributed by this ungauged inflow. These contributions have been added to the paper, the total area contributing to the system at Lewes is approximately 100km2, though the flow contributing to the system at high flow is c. 5-9m3s-1.

What changes have occurred in the catchment and channel above the point of interest (Lewes) that could have affected the homogeneity of the historical record of discharge with the more recent gauged record? Your remarks about drainage and canalisation are relevant but you fail to make the connection. Presumably the works to permit navigation to Balcombe (which is a long way upstream) involved straightening and deepening the river such that it would increase its capacity to carry floods (and less on the floodplain) and reducing the attenuation in the reach especially from the confluence of the Ouse and Uck to Lewes. Google maps shows some continuing effects of this canalisation for example at the Pellbrook cut. As this navigation fell out of use the effects of the canalisation would have diminished as the channel silted up (possibly again increasing the attenuation and allowing more overbank flow). All this since 1790! These changes could have affected the proportion of flow generated in the headwaters that reaches Lewes and caused the historical flow information to be non-homogeneous with the gauged record. River canalisation has been found to have a major effect on flood frequency in Ireland, for example Bailey and Bree (1981). At least you might indicate whether there is significant overbank storage in this reach between Ouse and Uck confluence and Lewes and whether it is protected by flood banks. Pearce (2002) indicates the presence of floodplain flooding in the Middle Ouse in the Oct 2000 flood. This point is discussed in the paper (7169, lines 25-29) and that reviewer is correct,

it is very difficult to ascertain the impact of channelization on the river flows, particularly during large flood events, even within contemporary events the uncertainties in discharge estimates is greater than that during normal flows. This section of the paper refers to the flooding upstream of Lewes and cites the Pearce paper. This section was added after initial brief review.

Other upstream changes include the reservoir at Ardingley whose catchment area you fail to note and its date of construction (1978) and changes in land use (you mention forestry in the upper catchment but the catchment has much woodland – has this changed). We accept this comment. The catchment area and construction date included in paper. It is impossible to say how much change has occurred in terms of forestry cover since 1650, but a visual assessment of the first OS map from the 1870s indicates that roughly the same area is covered by forestry.

To match the gauged and historical flows at Lewes, there needs to be some specific point or points where the height (for example on a bridge) or depth (in one or more properties) can be obtained for both gauged and historical flows. You note (for gauged flows) that "estimates are derived using a single stage-discharge relationship", so presumably you used such information. You give no indication in Section 3.1 that points in Lewes have been matched up with the gauged flow assessment from the two gauging stations. With respect to historical floods, the fact that boats were sailing about" gives no real indication of the depth and hence of the discharge though I agree it gives a qualitative assessment of the magnitude of the flood. Similarly the information included in Table 1 is very general and unless there is more specific information available in Macdonald 2004, I would be skeptical of the assigned discharge figures for each of the events. The descriptive accounts used relate to events within a small spatial area, which has remained relatively stable in terms of construction over the last couple of hundred years, as such the descriptive accounts are used with reference to the extent and level of the waters, we do not expect the discharges to be exact, more a reflection of the ranking of events (see Stedinger and Cohn 1986) and the broader magnitude of

the events relative to events within the gauged series and the accompanying descriptions. We appreciate that greater confidence in estimates would be achieved through the use of epigraphic levels, as used in a number of previous studies, but in this case these were not available.

The title notes since 1650" but this figure is not mentioned again in the paper and the Abstract says \back to 1750". The title is linguistically correct but one could equally say since 1450" and is therefore confusing. We believe that the reviewer's comment concerning the title is actually a reflection of bad practice and have therefore left the paper title as initially submitted. The suggestion that the paper title should reflect the start date of the estimated flood magnitudes is not in line with current methodological thinking, as indicated by the paper submitted by Strupczewski to this Special Issue. The start of the historical period according to the first large historical flood results in a distortion of the flood series, with subsequent impacts on the prediction of flood risk and trend from the resultant dataset, as such through the use of historical records it is better to use the start of the descriptive historical record.

P 4 line 2 Not sure what the upside down question mark is meant to indicate. Spanish? Accept - This has appeared during formatting - removed P 4 line 10. What area (and what proportion of the total) does the Ardingly reservoir catchment impound? The reservoir impounds approximately 20km2 of catchment, which equates to 5% of the total catchment area. Details of catchment area added to paper P 4 line 12 High water here refers to tidal level but \high water" in line 24 refers (presumably) to groundwater. The latter should be specified to avoid confusion. Accept - the later modified to read high groundwater Page 5 Section 2.1 Channel management. This section is interesting but you fail to indicate how this is relevant to the use of historical information in flood estimation. In this section we aim to highlight the principal changes to the channel upstream of Lewes; in section 3 paragraph 2 we state that we acknowledge the limitations and the challenges of channel cross modification during the historical period and propose means of reduce this uncertainty. Sec 2 and 3 In neither section do you

give catchment areas either to the point of flooding or to the gauging stations or to the confluence. This is surely a key piece of information to put the study in context and should be included. (You do give it for the Winterborne stream and for the Lewes Levels). Accept - all details added to the paper P 7 line 6 Single stage discharge relationship - for Where? Accept - clarified for Lewes P 7 line 29 You refer here to Ouse Bridge but do not mention such a bridge name in your bridges section - it is not clear which reach you refer to. Although there may be no significant tributaries between the gauging stations and Lewes, there is a significant catchment area (Fig 1) which would be contributing to flows at Lewes. As noted above you do not indicate what proportion of the catchment this represents. Accept - correction added Section 3.1 It is not clear how you combined these records. Did you simply add the annual max from the two stations whether or not the AMAX occurred on the same day? Alternatively did you take the AMAX for the larger flow and add the daily maximum for the other station on the same day? Alternatively did you take the AMAX from the larger flow and add the flow at the same time from the other station (perhaps adjusted by travel time)? Do the annual maxima at the two stations tend to occur on the same day and at a similar time? Did you make any allowance for ungauged inflow in the reach to Lewes - or alternatively attenuation due to channel and floodplain storage? Is there any way of checking whether the Rank order of the gauged floods created from the two gauging stations corresponds with the Rank order at Lewes - either in terms of measured levels at Lewes or descriptions of flood damage? An explanation of how the flows are derived has been added to the paper – To generate a single comparable AMAX series, the IFP for the two stations were added together where they occurred within one day (time of peak is not recorded), where no IPF is available for one of the stations maximum daily flow from the NRFA was added to the IPF, this may result in some underestimation of the total discharge, but flows are unlikely to have been substantial if not recorded within the IPF series. Additional flows at Lewes are relatively small (< 9m3s-1 and are not considered - no data on attenuation within the floodplain is held. P 8 line 3 When you say complete series do you mean complete series of AMAX or complete series of daily

maxima from which AMAX are selected? A series of AMAX for each year. P12 lines 16 et seq. You acknowledge that the uncertainty estimates you quote are only those associated with the sampling uncertainty associated with the chosen distribution and use this as the basis for comparison with gauged uncertainty. I think you should add further comment in your discussion or conclusion concerning the fact that uncertainty in the magnitude of the historical floods could add significantly to the uncertainty of your 100 year RP estimate (and would be difficult to assess). This applies even to the qualitative estimates of extremes over a threshold as the evidence you have presented in the paper for the magnitudes selected would make it difficult even to say whether a particular event was above a given threshold or not. We have added the following at the end of Section 5.2 The strong differences that can be found when different perception thresholds are employed in the estimation procedure raise the question of the reliability of the estimated sizes of the historical events. In this study every effort has been made to assure a reliable set of historical data, but more modelling efforts can be done to include the uncertainties in the point estimates of the ungauged measures. See for example Neppel et al.(2010)." Section 4 p 9 to 12. I think it should be mentioned that an adaptation of the Stedinger and Cohn (1986) procedure using maximum likelihood has previously been applied by Archer (2010) (with both GLO and GEV distributions) for catchments in northeast England. This also uses alternative procedures based on (a) using the full gauged data and quantitative historical discharge data, where available and (b) using only the number of exceedences above a threshold, where quantitative estimates of historic flood discharges cannot be made. A reference to the study by Archer (2010) has been added to Section 4 in the discussion of the likelihood function. The following text has been added immediately after the Stedinger & Cohen (1986) reference (was page 7625, line 16): and previously applied for flood frequency analysis in selected British catchments by Archer (2010) using the GEV and GLO distributions." P 14 line 24-25 Fig 5 shows symbols for gauged and historical events in the key but are not included in the figure. Presumably in the third case where the historical events are only known to exceed the chosen threshold, these values cannot be plotted. Revised

References Archer, D.R. (2010) Applying historical information to flood risk assessment in north-east England, BHS Third International Symposium, Managing Consequences of a Changing Global Environment, Newcastle 2010. Bailey, A.D. and Bree, T (1981) effect of improved land drainage on river flood flows, Flood Studies Report - Five Years on, Thomas Telford Ltd, London, 1981. Pearce, J. (2000) Hydrologist experiences in times of flood: Sussex 2000, BHS 8th National Hydrology Symposium, Birmingham, 2002. The reference to the proposed papers by Pearce (2000) and Payrastre et al. (2011) are included and provides valuable support to the findings of the paper; the reviewers are thanked for pointing these papers out. We do not feel it is appropriate to include the references to Bailey and Bree (1981), as it is somewhat dated.

Referee #2 It is based on a single case study, which by the way not the most simple one; the selected locations are under tidal influence and is also exposed to ash floods of a small tributary, the river bed has been significantly modified during the historic period with a possible influence on the local stage-discharge relation, and moreover the considered location is not equipped with a stream gauge and the analysed discharge series is build from the sum of discharges measured in to upstream sections. A simpler, less questionable case study or more than one example could have been selected as an illustration. We were specifically interested in a challenging site, one that reflects the challenges presented in the application of historical data in real studies, where the hydraulic properties are often not simple and where there are a number of points requiring consideration.

The inference methods used are standard (MLE) are not up to date (evaluation of the confidence intervals based on linear approximations of the likelihood function and on the assumption of a Gaussian distribution). More rigorous inference approaches based on Bayesian MCMC methods that are now frequently used in hydrology (refs) would have been a better suited choice and helped to go further into the analysis. The choice of using the standard maximum likelihood method rather than Bayesian inference was partially taken to keep the presentation of the material in the paper simple. The results

and their interpretation would only change marginally if a Bayesian inference was used rather than standard maximum likelihood. We have added the following comment in Section 4 (at what was page 7628): "Rather than the standard numerical optimization a Bayesian MCMC approach as the one presented in Gaume et. al. (2010) could be employed to maximise the equations in (5) or (6). These methods are readily available in the nsRFA library in R (Viglione, 2013). The advantage of using a Bayesian approach is that a full posterior distribution for the parameter estimates and any relevant quantity can be obtained, and no approximations like the ones in (9) are needed. The authors experience though is that little difference can be found in the final estimates, and that the BayesMCMC function can, in some cases, give computational issues or can be fairly slow to reach convergence. The interpretation of the results discussed in Section 5 would not change if Bayesian estimates would be used" Indeed, we have fitted the model for the historical data with the Bayesian framework available in the nsRFA:BayesMCMC function in R and the 100 year return level estimates with a 95% confidence/credibility bands for increasing perception thresholds are shown in the picture below. Almost no difference can be seen for the point estimate, and the maximum likelihood estimator consistently has a lower upper bound, mostly due to the asymmetricity of the Bayesian credibility intervals. We also found that the outputs for the Bayesian inference procedure do not always give a very good convergence and can be quite different from one run to the other. In Figure R.1 100 replication of the estimated 100-year return event with 95% confidence bands for the case in which only the highest 2 historical event are kept in the model is shown. Each replication was started at similar starting points (corresponding to the L-moment estimates obtained from the systemic records plus some small noise) and each replication is based on a 3 chains of 17000 MCMC samples with 2000 observations of burn-in. Although the differences are not really large, some inconsistencies can be seen. Of course these inconsistencies in the Bayesian results might be the result of a computational issue and poor convergence: Figure R.2 show results obtained when using 3 chains of 100000 MCMC samples with 2500observations of burn in. The results are indeed more con-

C3131

sistent but this is obtained at the cost of needing 15 minutes to _t the model, instead of the 2 minutes needed for the models shown in Figure R.1.

The conclusions concerning the added value of historic data in flood frequency analysis depending on their nature (threshold level, accurate discharges, or censored data) are also not new and have been presented in more general papers recently published by Stedinger et al., or Payratre et al for instance. We accept that the papers stated show similar conclusions, though many studies often use generated data rather than site specific data, in doing so this paper adds to a relatively small number of papers with case study based analysis and \real world challenges", as Reviewer III states the[is] work was performed in the frame of 'applied hydrology paradigm' "

Figure R.1: Different estimated 100-year return levels with 95% confidence bands: quick estimation. Figure R.2: Different estimated 100-year return levels with 95% confidence bands: slow estimation.

Referee #3 The work was performed in the frame of 'applied hydrology paradigm' by which I mean the conviction that all carefully prepared and checked historic hydrological or paleo hydrological information can significantly increase our understanding of temporal and spatial patterns of river flow and, in particular, extreme events and improves the estimation of design values generated by statistical models, in this case FFA models. This paradigm is extremely important in relation to the general knowledge but ought to be supplemented by detailed statistical research to confirm or reject assumed gain in quantiles accuracy also taking into account the system of design procedures of water depending structures. Agree - we have provided some additional materials to address these points, see a detailed comment on this point above Reviewer II, pt 2

List of the sources of documentary data on historical floods available in the UK is enviable and so are the methods of data calibration and harmonisation applied by Authors. We grateful that the reviewer appreciates the challenges within this work and calibration/harmonisation challenges within this type of study

C3132

I totally agree with the statement that during the largest flows single stage-discharge relationship can be used, however it has to be noticed that uncertainty of upper limb of the rating curve imposes significant uncertainty of discharge assessment in this case. The Authors express the same opinion on the page 7628 writing that \we will only consider the sampling uncertainty, but acknowledge that especially the data uncertainty and the difference between gauged and historical events could be a significant factor." Moreover, only the part of upper quantiles' MSE - the variance - is estimated. It is obvious that the assessment of the quantiles' bias cannot be done but by MC simulations and this is out of scope of this article. It would be advisable to refer to in the text of this article and in the conclusions. Additionally it is worthy to notice, that strong sensitivity of the result to the perception level can be the signal of the bias importance in MSE of estimated upper quantiles. For an in-depth discussion of the use of MSE in this study, please refer to our reply to the 2nd comment by Referee #2.

Other discussion issue is the problem of choosing the distribution of annual maxima in the context true - false distribution. Model misspecification results in bias of parameters, consequently translated to the bias of quantiles estimates. Good asymptotic properties of estimation methods vary significantly when the model is untrue, what is the case in FFA, where we do not know the parent distribution of maxima. Authors' statement that \GLO might not be the true distribution" (p. 7627 and 7628) seems to represent excessive cognitive expectations in relation to models and the role they play. However accepting the saying of Rene Thom (I quote from memory) \The truth is not limited by the false, but by the lack of significance", it will be better perhaps to think that the model might be not good enough to describe the parent population with sufficient accuracy.

We agree with the referee's take on this, but we also would like to point out that a choice on the distribution of the annual maxima must be made in order to fit the model. For the gauged only the GLO distribution, which is the recommended distribution to be used for British AMAX data, is not rejected by formal testing and seem to be reasonable when

looking at the Lmoments diagram in Figure R.1. We have rephrased the statement concerning the GLO distribution to: "the GLO distribution might not provide a good description of the true underlying distribution. " We have also added the following at the end of the paragraph which was at the top of page 7628. "The uncertainty related to the assumed distribution cannot be avoided and it is not expected that the effect of a model misspecification would be larger when using historical data combined with gauged data."

My last remark concerns the question: Is this huge and time-consuming work to prepare all historical information worth the reduction in standard deviation of about 6 percentage points? And what about the design quantiles? For the Sussex Ouse at Lewes the 100-year flood is much lower when estimated using historical data than only the AMAX events in gauged record. Which value ought to be taken for the design?

We would encourage for design purposes that the historical augmented results should be taken for design purposes, whilst the different is c.6%, this can represent a significant increase/decrease in cost in funding flood mitigation measures. Detailed comments - addressed within the manuscript. Thanks for pointing these inaccuracies in the exposition. Regarding Comment 12: the plotting positions have been calculated using the Gringorten plotting positions.

Figure R.3: Lmoments diagram for the gauged data at Lewes.

Please also note the supplement to this comment:
http://www.nat-hazards-earth-syst-sci-discuss.net/1/C3123/2014/nhessd-1-C3123-2014-supplement.pdf
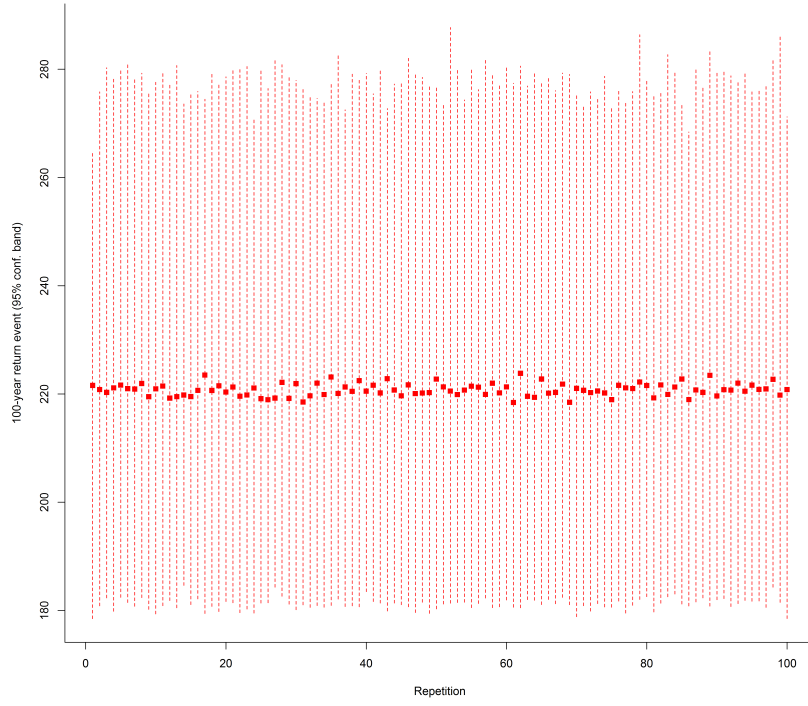
Fig. 1.

C3135

Fig. 2.

C3136

Fig. 3.

C3137