Using Non-Systematic Data in Flood Frequency Analysis

Thomas Smith[1†], Ilaria Prosdocimi[1], and Thomas Kjeldsen[2]

[1] Department of Mathematical Sciences, University of Bath

[2] Department of Architecture and Civil Engineering, University of Bath

† T.G.Smith@bath.ac.uk

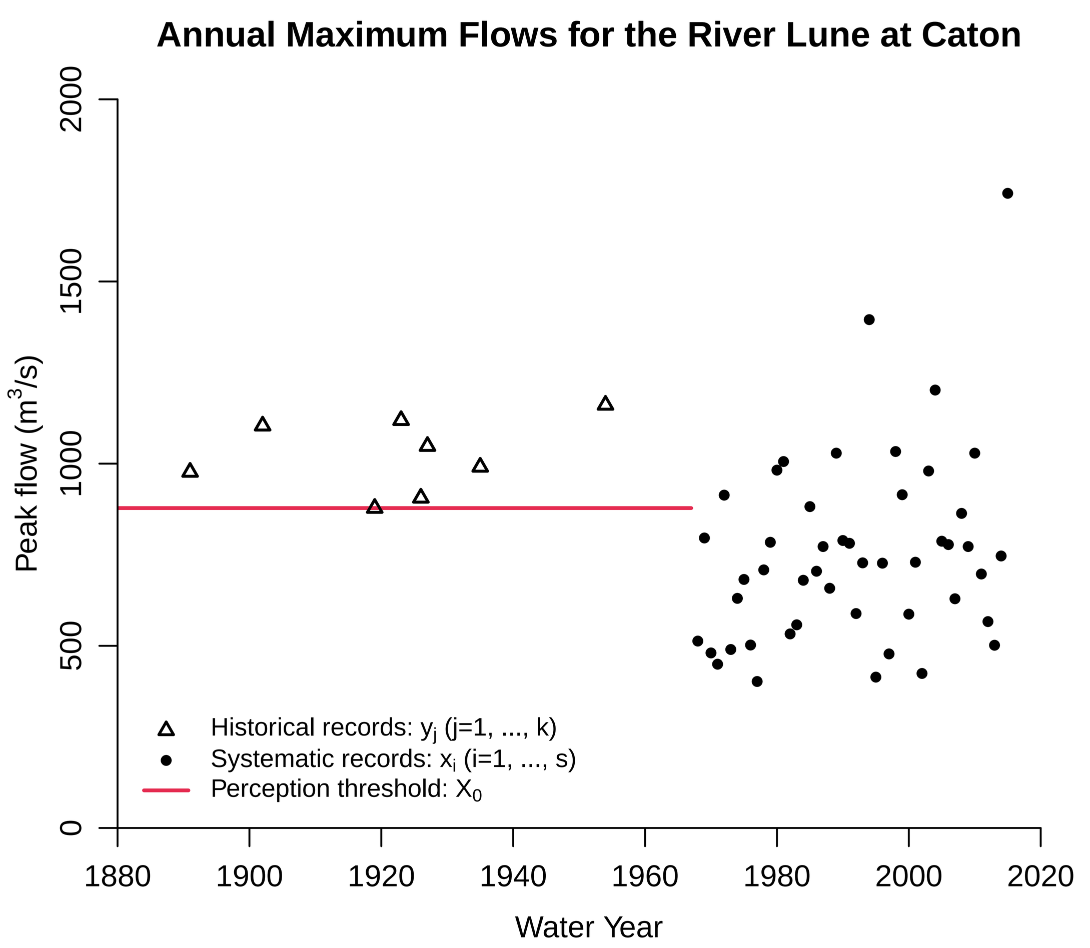
UNIVERSITY OF BATH

1. Background

Flood Frequency Analysis (FFA) aims to quantify the risk posed by floods to specific areas by, for example, estimating the size of flood that an area would expect to see on average once in 100 years. One approach towards FFA involves fitting a statistical model to the series of maximum annual river flows. In fitting such a model, hydrologists typically only use data that has been collected in the modern era through the systematic measurement of river flows. One drawback to this approach is that the duration of these records is short, being only 40-50 years long on average for the UK – far shorter than the time period for which prediction is often required.

Fortunately, there exist sources of non-systematic data which can complement the systematic record. These include historical records in the form of flood marks carved into bridges, old photographs, and newspaper reports. Inclusion of this data into the modelling procedure can improve the quality of subsequent inference.

2. Data



- Data analysed is of the river Lune at Caton, Lancashire
- The 48-year systematic record was obtained from the National River Flow Archive [1]
- Magnitude estimates for eight historical floods over a 76-year period are available from the Flood Studies Report [2]
- The December 2015 flood event exceeded every systematic and historical record
- The largest three flood events are all in the systematic record, despite it being shorter than the historical period

3. Assumptions

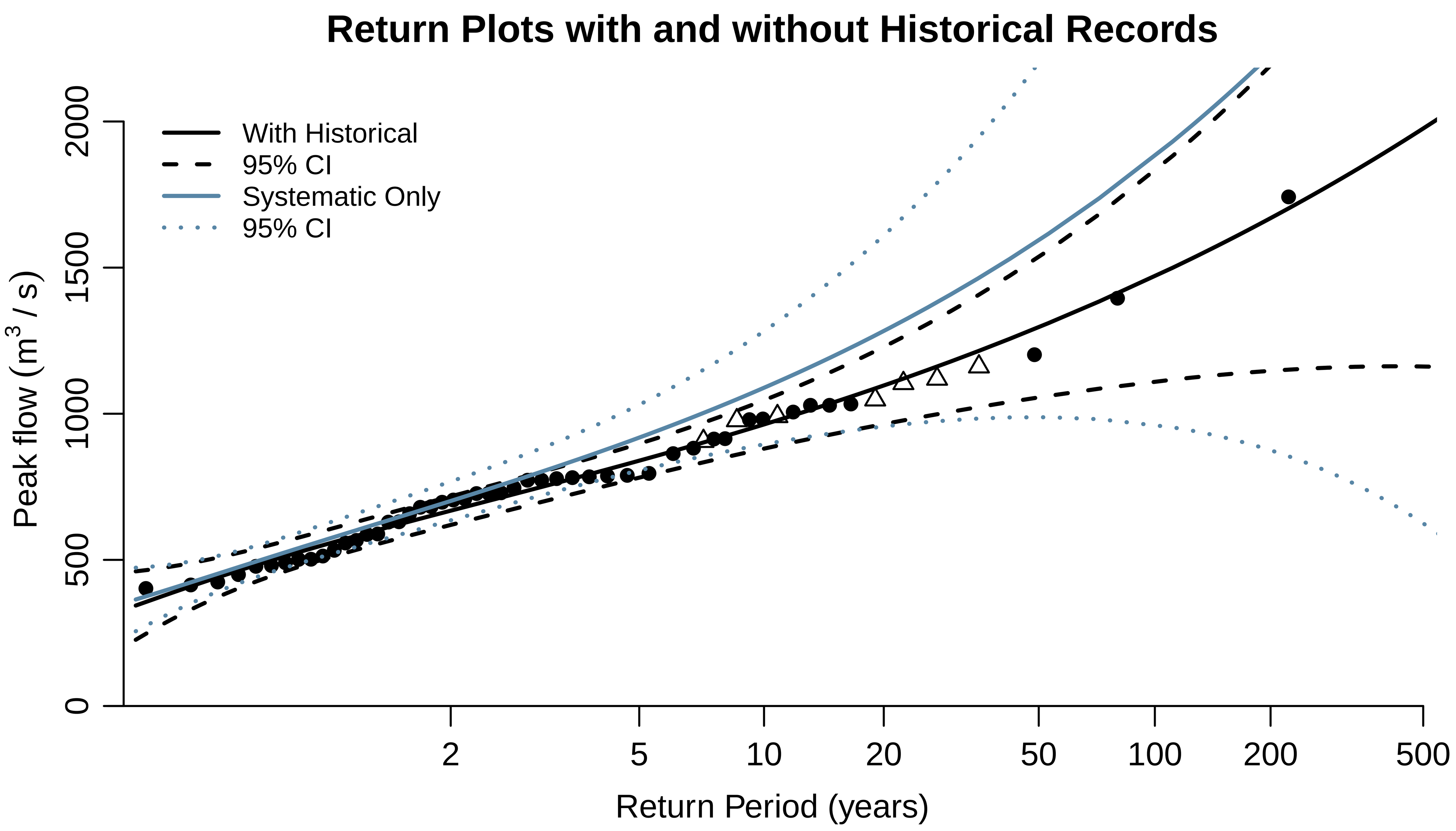
- The **perception threshold** was set to be the magnitude of the smallest historical flow, as in a previous analysis of this data [3]
- Annual maxima were assumed to originate from a **Generalised Logistic distribution** (GLO) as per UK standards [4]
- Annual maxima were also assumed to be **independent** and **identically distributed**
- It was assumed that there was **no under-recording** during the historical period

4. Methods

- Two separate GLO distributions were fitted – one with, and the other without, the historical records
- Parameters were estimated using a **maximum likelihood** method similar to the one presented in Macdonald et al. [5] using the following likelihood function:

$$\ell(D|\theta) = \prod_{i=1}^{s} f_X(x_i;\theta) \cdot \prod_{j=1}^{k} f_X(y_j;\theta) \cdot \left\{ \binom{h}{k} F_X(X_0;\theta)^{h-k} \left[1 - F_X(X_0;\theta)\right]^k \right\}$$

5. Results and Conclusions



- Including historical data has shifted the return curve to the right, resulting in a longer estimated return period for any given flow rate
  - The 100-year flood estimate decreased from 1882m^3/s to 1476m^3/s, a drop of 22%
- Confidence intervals have also narrowed substantially, especially for longer return periods
  - The width of the 95% confidence interval for the 100 year flood decreased from 1840m^3/s to 730m^3/s, a drop of 60%
- This demonstrates the potential for non-systematic records to add value to a flood frequency analysis

6. Future Work

- The systematic record is a combination of measurements from two nearby gauging stations. Whether it would be advantageous to fit a different distribution to each station has not been addressed here
- There is considerable uncertainty in the estimated magnitude of historical flows, which has not been taken into account in this analysis
- Given that the largest three flood events all occurred in the systematic record, an investigation into possible non-stationarity of the series would be worthwhile
- A Bayesian analysis would allow more the uncertainty to be captured more naturally, and would permit skew in the confidence intervals

7. References

[1] Centre for Ecology & Hydrology. National river flow archive. http://nrfa.ceh.ac.uk/data/station/peakflow/72004, 2018.

[2] NERC (Natural Environment Research Council). *Flood Studies Report: Volume IV Hydrological Data.* NERC, London, UK, 1975.

[3] I Prosdocimi. German tanks and historical records: the estimation of the time coverage of ungauged extreme events. *Stochastic Environmental Research and Risk Assessment*, 32(3):607–622, 2018.

[4] A Robson and D Reed. *Flood Estimation Handbook.* Institute of Hydrology, Wallingford, UK, 1999.

[5] N Macdonald, TR Kjeldsen, I Prosdocimi, and H Sangster. Reassessing flood frequency for the sussex ouse, lewes: the inclusion of historical flood information since ad 1650. *Natural Hazards and Earth System Sciences*, 14(10):2817–2828, 2014.