Flood Frequency Analysis in the UK: a very quick overview

Ilaria Prosdocimi (& Thomas Kjeldsen)

Department of Mathematical Sciences, University of Bath

ITT7

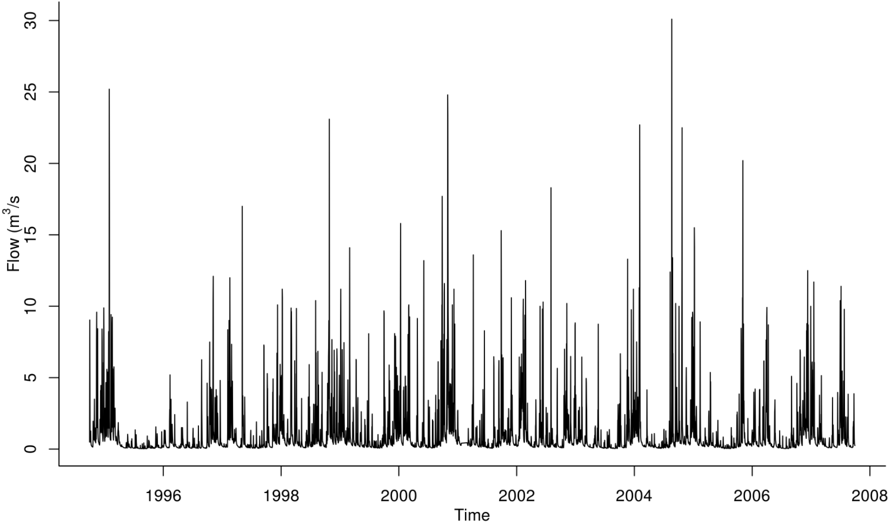
What is this about?

- Estimating the relationship between the magnitude and frequency of large events (aka the quantile function)

$$Pr[Q < q^*] = p^*$$

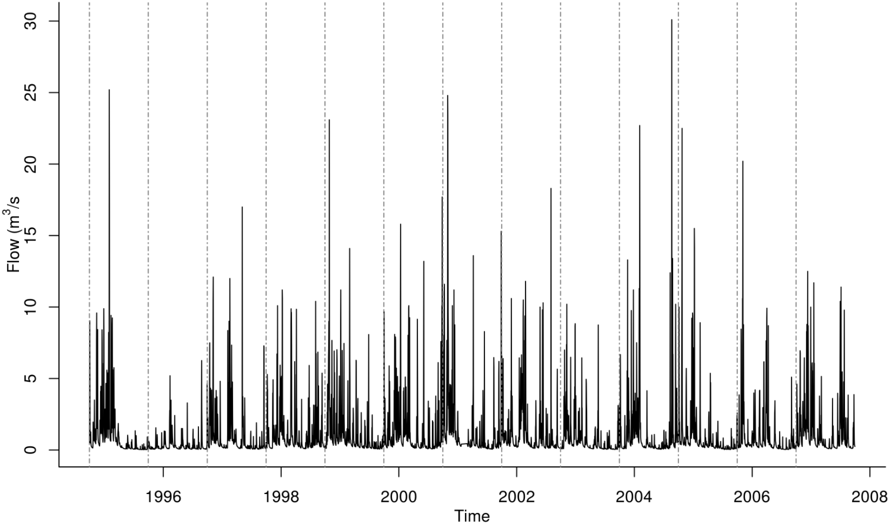
- Extreme Values Statistics: differences in hydrology and statistics
- Focus is fluvial flood risk (different practice for other risks)

What is this about?

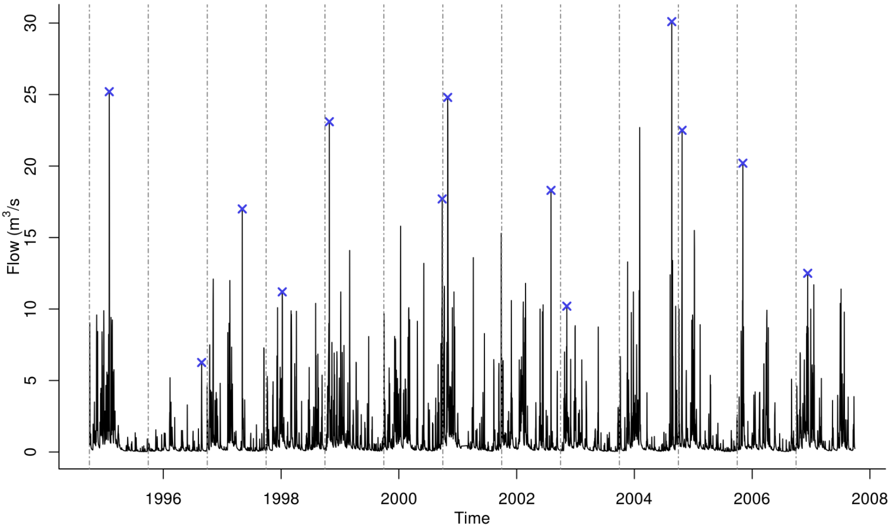


How big should the flood defences be to avoid flooding the city?

Annual Maxima



Annual Maxima



Extreme Values Approach: annual maxima

Annual Maxima

$(q_1, \dots, q_n)$, where each q_i is the maximum recorded in any given year.

$$q_i = \max\{y_j, \dots, y_{j+364}\} \text{ where } y_i \text{ is the realisations of i.i.d. } Y_i$$

The limiting distribution for q_i is the GEV distribution, which has form

$$P(Q \leq q) = F(q) = \exp\left\{-\left[1 + \xi\frac{q - \mu}{\sigma}\right]^{-1/\xi}\right\}$$

ξ is the key parameter to define the domain on which the distribution is defined (and the shape of the distribution):

$$\begin{aligned} \mu + \sigma/\xi \leq & q < \infty && \text{if } \xi > 0 \\ \infty < & q < \infty && \text{if } \xi = 0 \\ \infty < & q \leq \mu + \sigma/\xi && \text{if } \xi < 0 \end{aligned}$$

Annual Maxima

$(q_1, \ldots, q_n)$, where each q_i is the maximum recorded in any given year.

$$q_i = \max\{y_j, \ldots, y_{j+364}\} \text{ where } y_i \text{ is the realisations of i.i.d. } Y_i$$

The limiting distribution for q_i is the GEV distribution, which has form

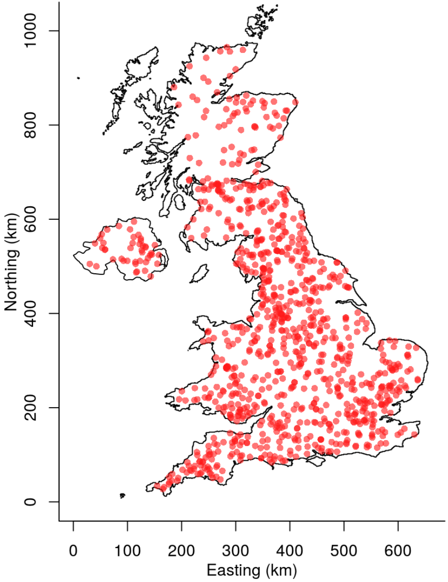
$$P(Q \leq q) = F(q) = \exp\left\{-\left[1 + \xi\frac{q-\mu}{\sigma}\right]^{-1/\xi}\right\}$$

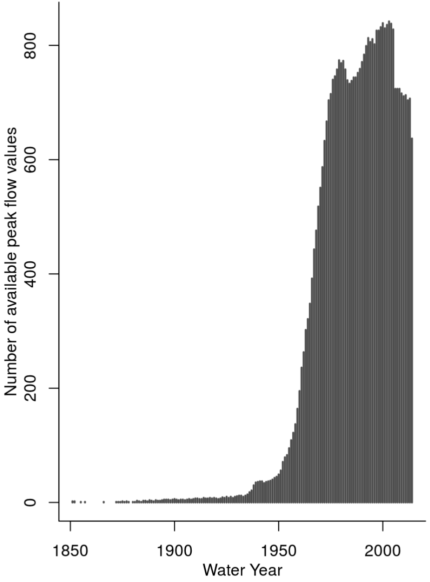
ξ is the key parameter to define the domain on which the distribution is defined (and the shape of the distribution):

$$\begin{aligned} \mu + \sigma/\xi \leq & q < \infty && \text{if } \xi > 0 \\ \infty < & q < \infty && \text{if } \xi = 0 \\ \infty < & q \leq \mu + \sigma/\xi && \text{if } \xi < 0 \end{aligned}$$

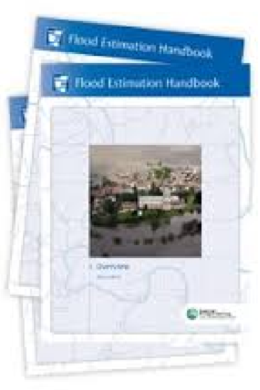
(The original Y process actually only needs to be weakly stationary)

All we need is data





The Flood Estimation Handbook

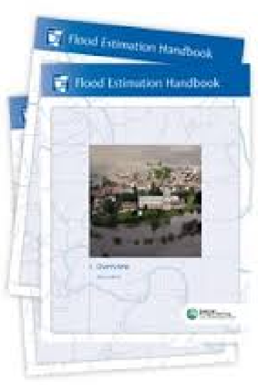


Two main approaches: ReFH and Statistical method

Statistical method:

- Estimation method: L-moments
- Region-of-influence approach
- Assume that distribution is the same for all the pooled stations after some normalisation
- Default distribution is GLO (and many others are tested)
- Can be used for ungauged catchments
- based on peak flow records

The Flood Estimation Handbook



Two main approaches: ReFH and Statistical method

Statistical method:

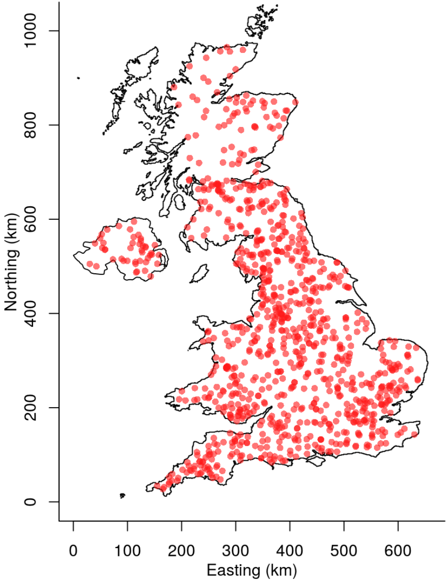
- Estimation method: L-moments
- Region-of-influence approach
- Assume that distribution is the same for all the pooled stations after some normalisation
- Default distribution is GLO (and many others are tested)
- Can be used for ungauged catchments
- based on peak flow records

Updated in early 2000s by Kjeldsen, Jones and Bayliss (et al.).

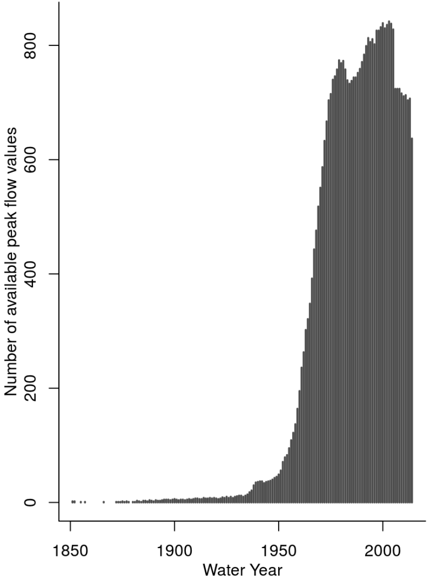
Where is the data?

- National River Flow Archive
- Data needed to implement methods used in the Flood Estimation Handbook (FEH)
- Data also used for Water Resources Assessment/Low Flow/Water Abstraction....
- Different uses require different data
- Easy to access data: daily average flow; annual maxima (flow and stage); (peaks over threshold);
- Existing data: 15-minute flow measurements; rainfall records
- Expensive data: catchment descriptors (ungauged)

The network



Northing (km)
0
200
400
600
800
1000
0
100
200
300
400
500
600
Easting (km)



Number of available peak flow values
0
200
400
600
800
1850
1900
1950
2000
Water Year

Some ideas behind the FEH methods

The Index Flood Method: model

$$X = \frac{Q}{median(Q)}$$

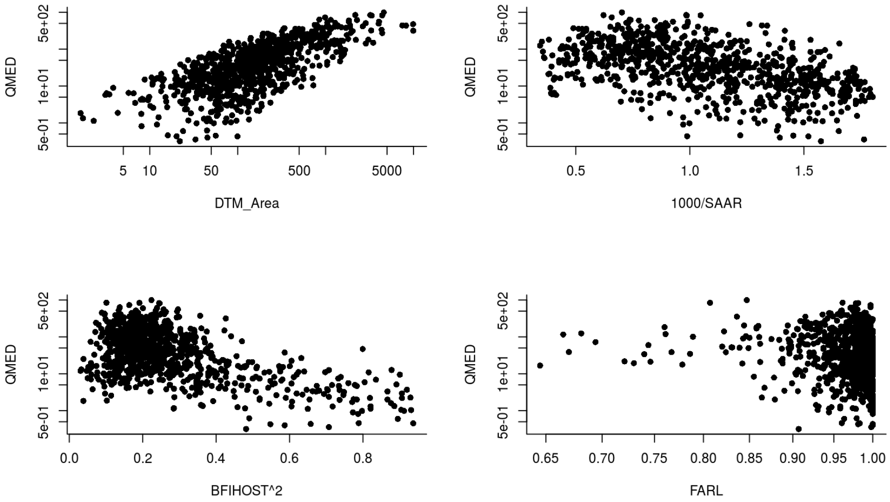
T-year Design event (quantile of interest 1-1/T):

$$x_T = QMED * z_T$$

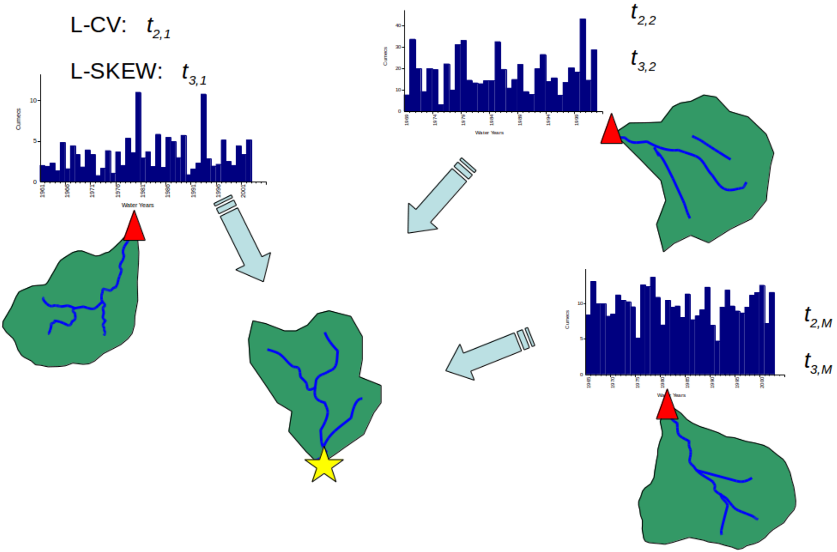
Estimates of QMED obtained from Catchment Descriptors (for ungauged locations)

Estimates of z_T obtained from a pool of catchments

QMED (index flood) estimation - ungauged



Pooling groups - region of influence



FEH assumptions and practice

Assume that region is homogeneous and all standardised peak flows come from the same distribution (i.e. same Skewness and CV).

Pooling group created using similarity distance measure:

$$SDM_{ij} = \sqrt{3.2\left(\frac{\ln AREA_i - \ln AREA_j}{1.28}\right)^2 + 0.5\left(\frac{\ln SAAR_i - \ln SAAR_j}{0.37}\right)^2 + 0.1\left(\frac{FARL_i - FARL_j}{0.05}\right)^2 + 0.2\left(\frac{FPEXT_i - FPEXT_j}{0.04}\right)^2}. \qquad (6.9)$$

Estimation of scale and skewness by averaging L-moments of pooling group.

The default distribution in the UK is the Generalised Logistic.