Institute of Geophysics
Polish Academy of Sciences

Flood quantile estimates related to model and optimization criteria

Iwona Markiewicz
Witold G. Strupczewski
Krzysztof Kochanek

XXXII International School of Hydraulics, May 28-31, 2012 Lochow, Poland

Objective

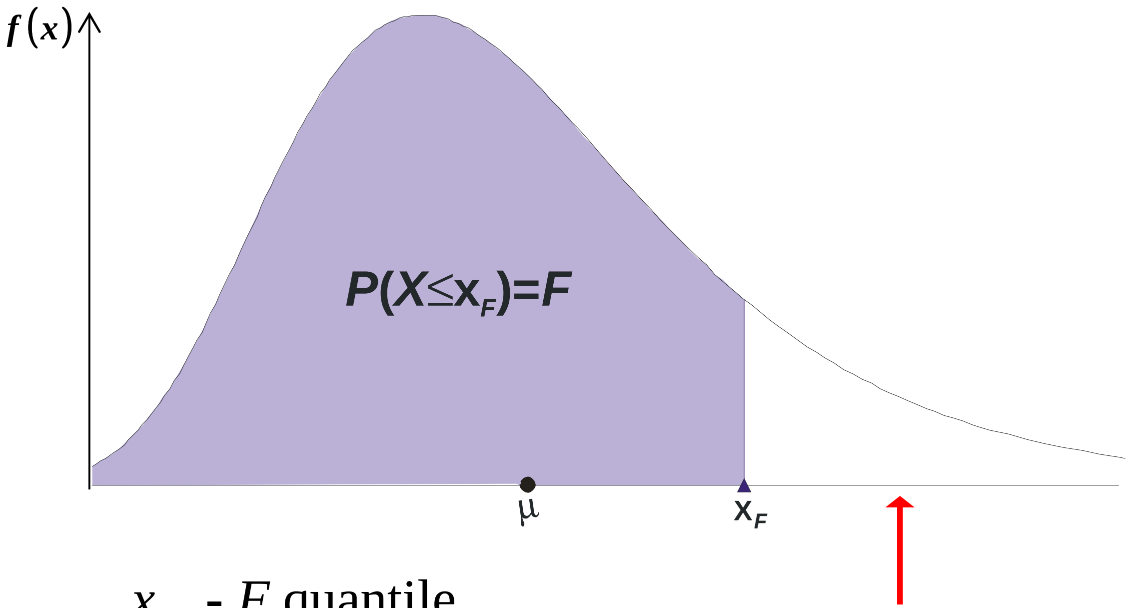
WHERE ARE WE HEADING IN FLOOD QUANTILES ESTIMATION?

Whether the theory meets the expectations of the practice?

Introduction

flood quantile = probable size of flood flows within a given return period *T* years

Flood Frequency Analysis (FFA) = estimation of upper quantiles of assumed probability density function of annual or partial duration maximum flows



Return period *T*

$$T = \frac{1}{1-F} = \frac{1}{P(X > x_F)}$$

$$T = 10 \leftrightarrow F = 0.9$$

$$T = 100 \leftrightarrow F = 0.99$$

$$T = 1000 \leftrightarrow F = 0.999$$

x_F - F quantile

$$P(X \leq x_F) = \int_{-\infty}^{x_F} f(x)\,dx = F$$

Application of flood quantile estimates

- designing hydraulic structures (spillways, dams, bridges, levees,etc.)
- determining the limits of flood zones with varying degree of flood risk
- estimating the risk of exploitation of floodplains
- valuation of some insurance premiums

Flood near the village of Swiniary, in Central Poland, May 25, 2010.
(AP Photo/Czarek Sokolowski)

Avon & Severn Rivers, Tewkesbury, UK, 2007.

Flood quantiles estimation

Issue:

- **choice of probability model**
- **choice of optimization criterion**

Discrimination procedure:

- **best-fit model to empirical data, primarily in upper quantiles**

Problems

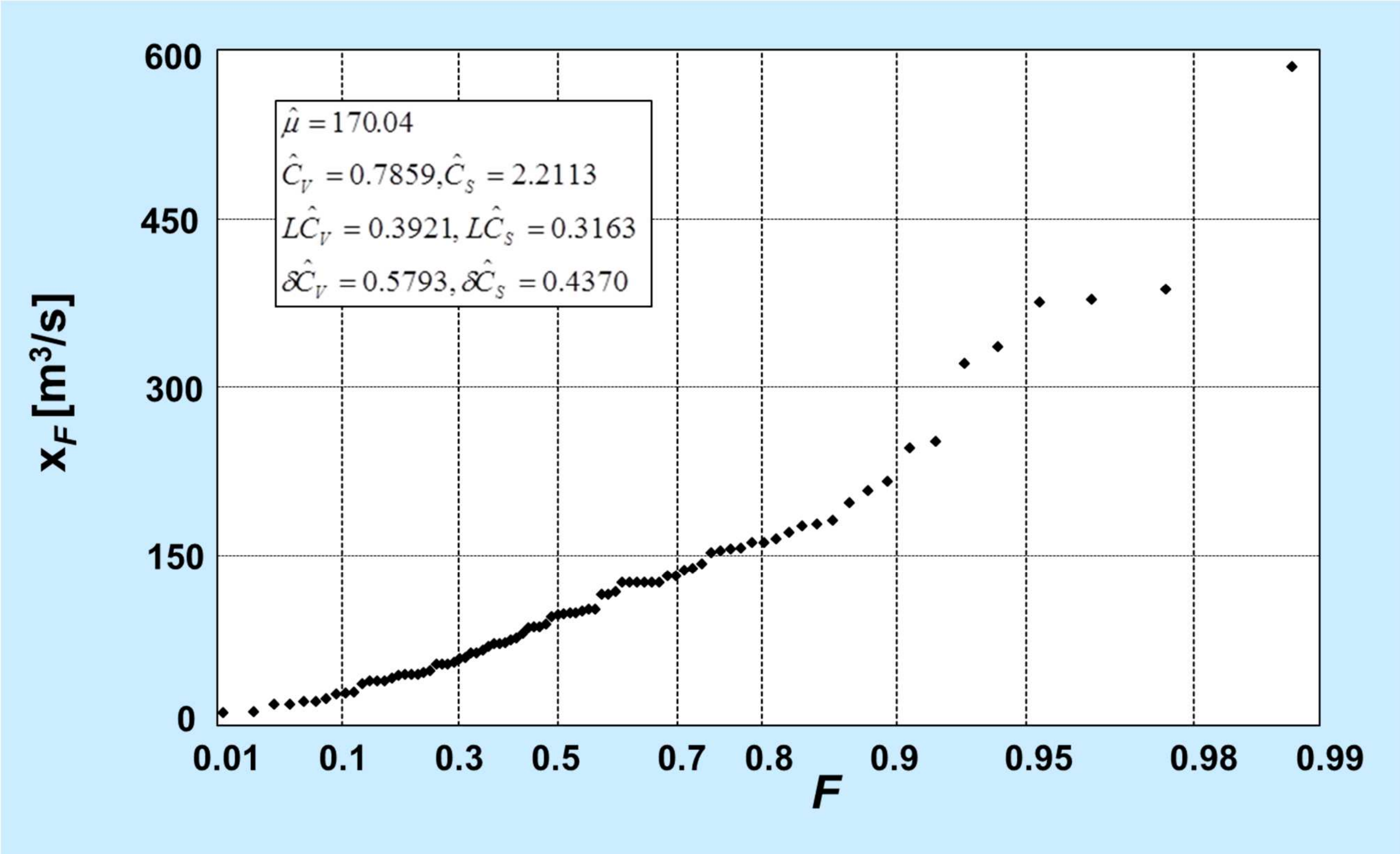
- Unknown probability distribution function of annual peaks
- Short time series ($N \ll T$)
- Error corrupted data
- Simplifying assumptions:
  - independent and identically distributed (i.i.d.) data
  - stationarity of relatively long series

↓

Flood quantile estimates are highly uncertain

Case study

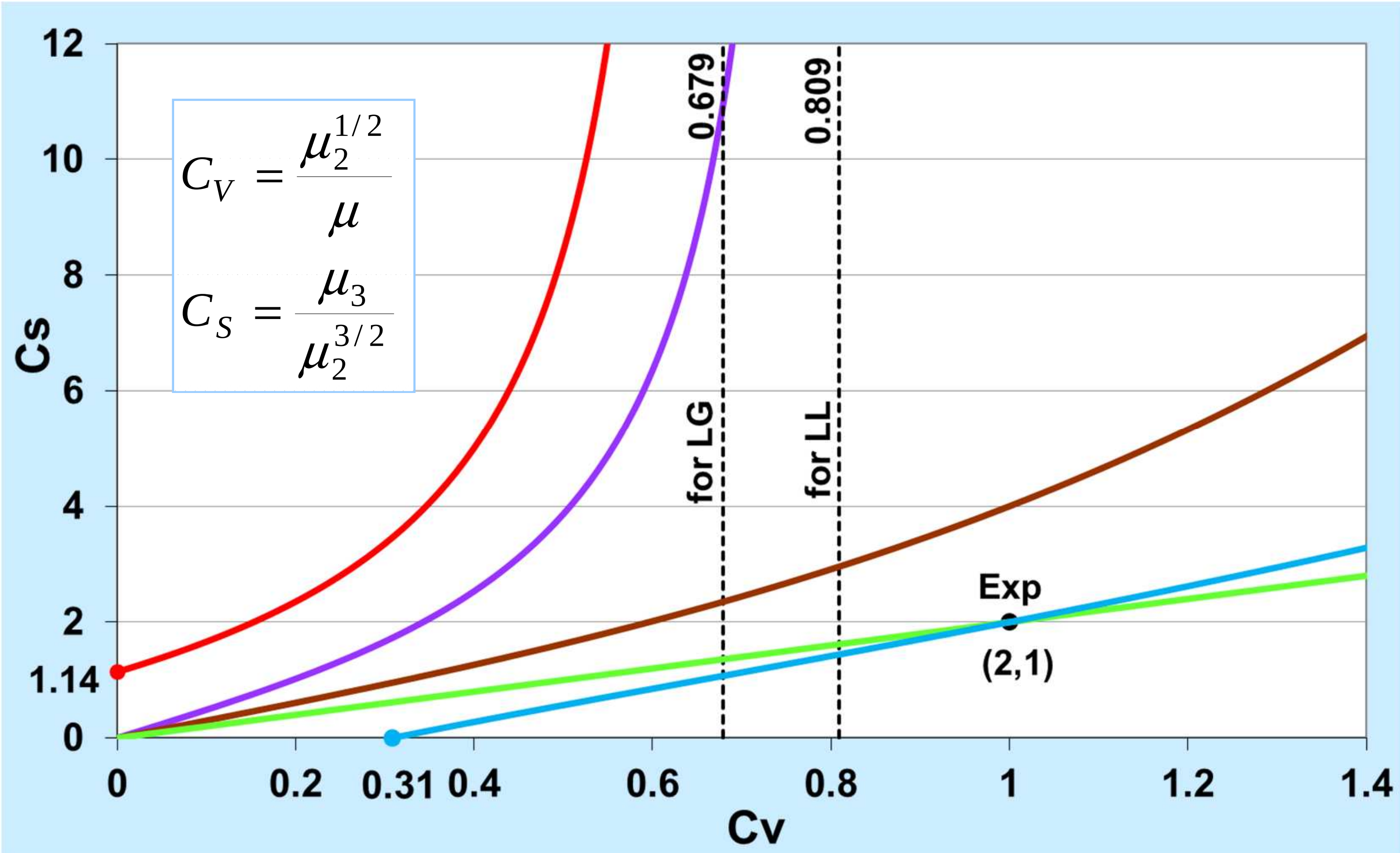
Annual peak flows for Nowy Targ on the Dunajec River, 1921-2010



Models

2 parameters
- gamma (Ga)
- Weibull (We)
- log-normal (LN)
- log-logistic (LL)
- log-Gumbel (LG)

3 parameters
- Pearson (Pe)
- Weibull (We)
- log-normal (LN)
- generalized log-logistic (GLL)
- generalized extreme value (GEV)



Optimization criteria

- Criterion of (conventional) moments - MO
- Criterion of linear moments - LM
- Maximum likelihood criterion - ML
- Criterion built on mean deviation - MD

MD	Location	Dispersion	Skewness
Measure	μ	$\delta_\mu = \int_{-\infty}^{+\infty} \lvert x-\mu \rvert \, dF(x)$	$\delta_S = \mu - med$
Dimensionless measure	-	$\delta C_V = \frac{\delta_\mu}{\mu}$	$\delta C_S = \frac{\delta_S}{\delta_\mu}$

- Markiewicz , I., Strupczewski, W.G., Kochanek, K. (2010). On accuracy of upper quantiles estimation. Hydrology and Earth System Sciences, 14: 2167-2175. DOI:10.5194/hess-14-2167-2010
- Markiewicz, I. and Strupczewski ,W.G. (2009). Dispersion measures for flood frequency analysis. *Physics and Chemistry of the Earth*, 34: 670-678. DOI 10.1016/j.pce.2009.04.003.
- Markiewicz, I., Strupczewski, W.G., Kochanek, K. and Singh, V.P. (2006). Relations between three dispersion measures used in flood frequency analysis. *Stochastic Environamental Research and Risk Assessment*, 20: 391-405. DOI 10.1007/s00477-006-0033-x.

Discrimination procedures

(1) Akaike information criterion (corrected for finite sample)

$$AICc = -2\ln\left[L\left(f\left(x/\hat{\theta}\right)\right)\right] + 2K + \frac{2K\left(K+1\right)}{N-K-1}$$

K – number of model parameters

L – likelihood function of model

The best model = this one of the lowest AICc value

Discrimination procedures

(2) Daniels characteristic

$$D_{\max} = \max_{i=1,\dots,N} \left| \frac{F(x_{i:N}) - \hat{F}_{i:N}}{1 - F(x_{i:N})} \right|$$

$F(x_{i:N})$ – theoretical probability of the i-th element of $x_{1:N} \le \dots \le x_{N:N}$

$\hat{F}_{i:N}$ – empirical probability given by the Weibull formula: $\hat{F}_{i:N} = i/(N+1)$

Sensitive within the area of probabilities approximating to one !

The best model = this one of the lowest D_{max} value

Discrimination procedures

(3) Difference between the quantile estimates

$$D_1 = \hat{x}_{1\%}^{ML} - \hat{x}_{1\%}^{MO}$$

$$D_2 = \hat{x}_{1\%}^{ML} - \hat{x}_{1\%}^{LM}$$

$$D_3 = \hat{x}_{1\%}^{ML} - \hat{x}_{1\%}^{MD}$$

$\hat{x}_{1\%}^{MO}$ – 1% quantile estimated by the criterion of moments

$\hat{x}_{1\%}^{LM}$ – 1% quantile estimated by the criterion of linear moments

$\hat{x}_{1\%}^{MD}$ – 1% quantile estimated by the criterion built on mean deviation

$\hat{x}_{1\%}^{ML}$ – 1% quantile estimated by the maximum likelihood criterion

The best model = this one of the lowest D*i* value

Flood quantile estimates for two-parameter models

1% quantile estimate value					
	Ga	We	LN	LL	LG
MO	620.64	604.11	671.1	638.76	630.68
LM	590.57	558.09	706.79	788.88	900.07
MD	603.07	570.88	726.41	827.07	914.32
ML	555.68	552.85	718.76	929.17	731.92

Discrimination procedures

Akaike information criterion					
	Ga	We	LN	LL	LG
MO	1088.7	1093.1	1081.4	1095.9	2157.2
LM	1086.7	1091.5	1080.8	1085.1	1209.4
MD	1087.4	1091.7	1080.9	1084.5	1198
ML	1085.7	1091.4	1080.8	1084.1	1142.6

Daniels characteristic					
	Ga	We	LN	LL	LG
MO	8.8241	16.12	2.1758	1.8015	1.69
LM	15.21	55.578	1.4466	0.5633	0.786
MD	12.022	37.69	1.1562	0.5436	0.7817
ML	31.978	69.362	1.2728	0.3877	1.9364

Difference between the quantile estimates					
	Ga	We	LN	LL	LG
$X_{ML}-X_{MO}$	64.96	51.265	47.661	290.4	101.24
$X_{ML}-X_{LM}$	34.886	5.2418	11.968	140.28	168.16
$X_{ML}-X_{MD}$	47.384	18.031	7.6492	102.1	182.4

Flood quantile estimates for three-parameter models

1% quantile estimate value					
	Pe	We	LN	GLL	GEV
MO	665.95	665.02	649.79	619.46	641.51
LM	650.75	635.73	678.94	719.7	699.77
MD	663.57	667.89	737.73	830.26	801.02
ML	579.71	579.1	701.3	812.19	864.17

Discrimination procedures

Akaike information criterion					
	Pe	We	LN	GLL	GEV
MO	*****	*****	1086.8	1103	1092
LM	*****	*****	1083.6	1090.9	1086.7
MD	*****	*****	1087.6	1087	1084.5
ML	1085.2	1087.7	1082.7	1086.5	1084.4

Daniels characteristic					
	Pe	We	LN	GLL	GEV
MO	3.5702	3.5053	3.2613	3.5584	3.3025
LM	4.6892	6.239	2.1049	0.9351	1.3651
MD	3.8207	3.4893	0.8734	0.5514	0.6589
ML	17.717	25.306	1.5643	0.6719	0.6594

Difference between the quantile estimates					
	Pe	We	LN	GLL	GEV
$X_{ML}-X_{MO}$	86.245	85.918	51.512	192.72	222.66
$X_{ML}-X_{LM}$	71.046	56.635	22.361	92.486	164.4
$X_{ML}-X_{MD}$	83.864	88.79	36.427	18.074	63.15

Conclusions

- ✓ **The choice of the best fitting model and, thus, hydrological design value (i.e. 1% quantile) depends on the optimization criterion and the procedure of discrimination.**
  It is characteristic for hydrological size of samples.

- ✓ **Development of flood frequency modelling is done by improvement of statistical techniques, which in turn leads to a proliferation of models, optimization criteria and discrimination procedures.**
  This causes an increase in our awareness of the uncertainty of flood quantile estimates instead of leading to clear solution.

- ✓ **The hydrologists want to have an unique value, not accepting the uncertainty !!!**

Conclusions

WHERE ARE WE HEADING IN FLOOD QUANTILES ESTIMATION?

Go back and start examining the way in which hydrological frequency analysis has been doing

THANK YOU

Acknowledgements
This work was financed by the Polish Ministry of Science and Higher Education under the Grant IP 2010 024570 titled "Analysis of the efficciency of estimation methods in flood frequency modeling".