



WROCLAW UNIVERSITY
OF ENVIRONMENTAL
AND LIFE SCIENCES

Przepływy maksymalne roczne w zlewniach kontrolowanych - statystyczne metody ich oceny

Wojciech Jakubowski – Katedra Matematyki UP Wrocław

Klasyczne podejście

(analiza jednowymiarowa)

Klasyczne podejścia

Metody estymacji rozkładów prawdopodobieństwa rocznych przepływów maksymalnych:

1. FFA (Flood Frequency Analysis) – metoda intuicyjna. Z każdego roku hydrologicznego pobierana jest tylko jedna obserwacja – maksymalnego przepływu;
2. POT (Peak Over Threshold) – metoda zaproponowana przez Teodorović'a i Zelenhaisić'a w 1970 roku wymaga wyznaczenia wszystkich wezbrań przewyższających ustalony arbitralnie poziom odcięcia – przejście do przepływów maksymalnych rocznych poprzez model Poissona (CP – Compound Poisson).

Klasyczne podejścia

Metody estymacji rozkładów prawdopodobieństwa rocznych przepływów maksymalnych:

1. FFA (Flood Frequency Analysis) – metoda intuicyjna. Z każdego roku hydrologicznego pobierana jest tylko jedna obserwacja – maksymalnego przepływu;
2. POT (Peak Over Threshold) – metoda zaproponowana przez Teodorović'a i Zelenhaisić'a w 1970 roku wymaga wyznaczenia wszystkich wezbrań przewyższających ustalony arbitralnie poziom odcięcia – przejście do przepływów maksymalnych rocznych poprzez model Poissona (CP – Compound Poisson).

Zestaw około 10 różnego typu rozkładów prawdopodobieństwa

1. Pearson III
2. log-normalny
3. Gumbel
4. GEV
5. log-Pearson

POT

Badane są przepływy przekraczające zadany poziom odcięcia. Wyznacza się przepływ maksymalny wezbrania. Stosowane są dwa typy rozkładów:

1. GEV – Generalized Extreme Value
2. GPD – Generalized Pareto Distribution

POT - wyznaczanie wezbrań

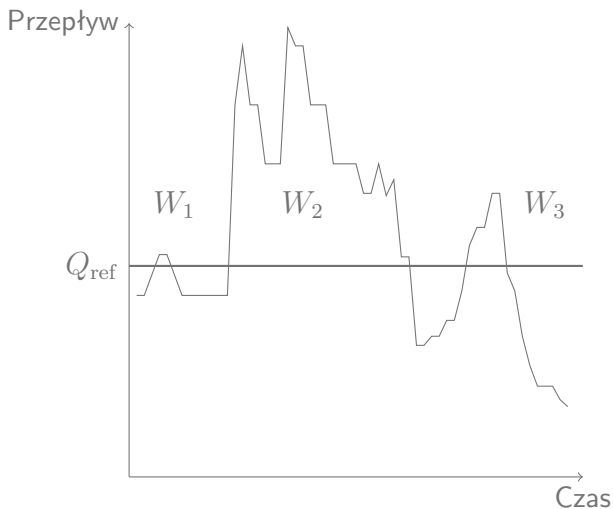


Figure: Schemat wyznaczania wezbrań

POT - wyznaczanie wezbrań

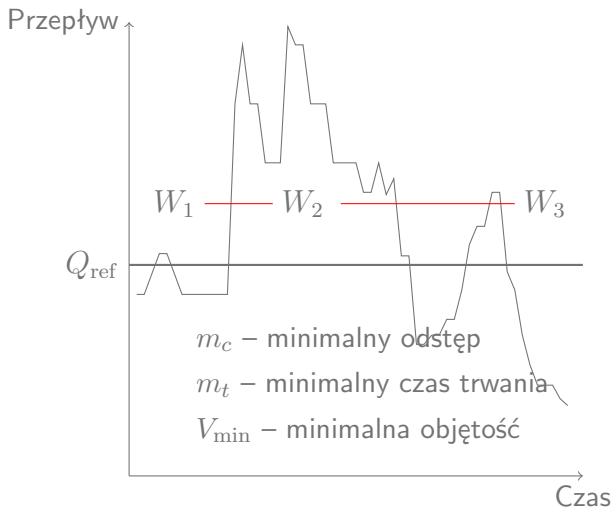


Figure: Schemat wyznaczania wezbrań

Analiza wstępna

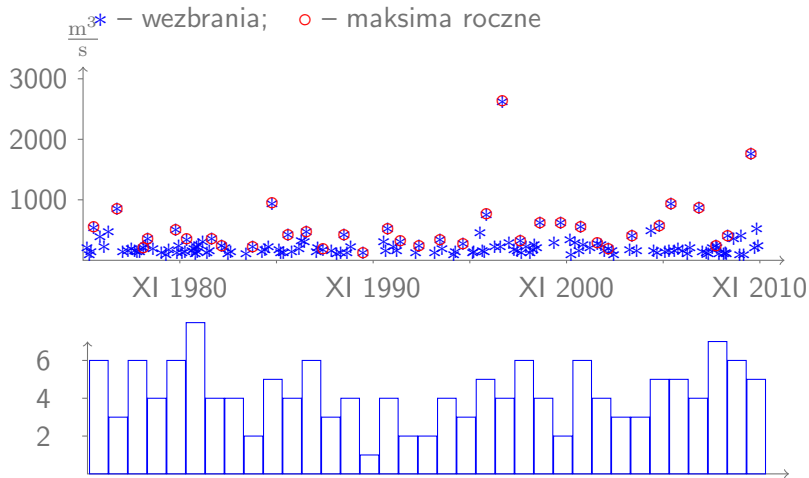


Figure: 20 Krzyżanowice

Analiza wstępna

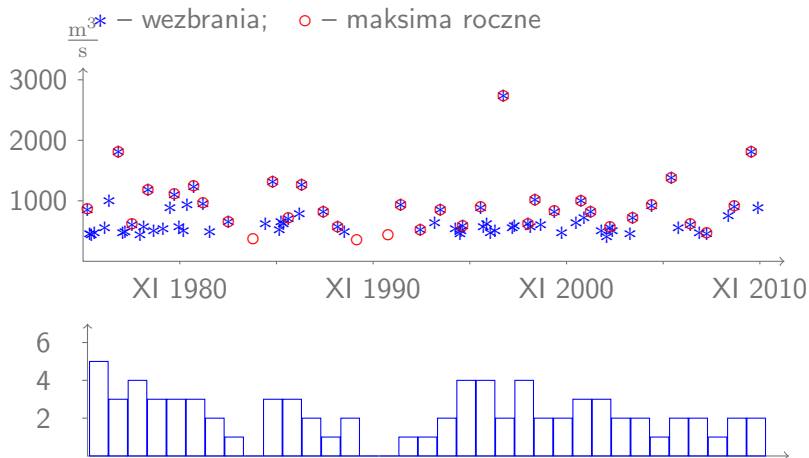


Figure: 90 Słubice

Rozkłady klasyczne, metoda FFA

1. Log-normalny z funkcją gęstości:

$$f(x, \mu, \sigma, \epsilon) = \frac{1}{\sqrt{2\pi\sigma(x-\epsilon)}} e^{-\frac{(\ln(x-\epsilon)-\mu)^2}{2\sigma^2}} \quad x > \epsilon$$

z parametrami μ – skali, σ – kształtu, ϵ – dolnego ograniczenia;

$$EX = e^{\mu+\sigma^2/2} + \epsilon \quad D^2X = (e^{\sigma^2} - 1) e^{2\mu+\sigma^2}$$

2. Pearsona III typu z funkcją gęstości:

$$\gamma(x) = \frac{1}{\Gamma(\nu)} b^\nu (x - \epsilon)^{\nu-1} e^{-b(x-\epsilon)} \quad x > \epsilon, \quad b, \nu > 0$$

z parametrami b – skali, ν – kształtu, ϵ – dolnego ograniczenia.

$$EX = \frac{\nu}{b} + \epsilon \quad D^2X = \frac{\nu}{b^2}$$

Rozkłady klasyczne, metoda FFA

3. Uogólnionych wartości ekstremalnych – GEV o dystrybucancie:

$$\text{GEV}(x, \mu, \sigma, \xi) = \begin{cases} \exp \left[- (1 + \xi(x - \mu)/\sigma)^{-\frac{1}{\xi}} \right], \\ 1 + \xi(x - \mu)/\sigma > 0, \quad \xi \neq 0 \\ \exp \left[- \exp(-(x - \mu)/\sigma) \right], \quad \xi = 0 \end{cases}$$

z parametrami μ – przesunięcia, σ – skali, ξ – kształtu.

$$EX = \begin{cases} \mu + \sigma \frac{\Gamma(1-\xi)-1}{\xi} & \xi < 1; \xi \neq 0 \\ \infty & \xi \geq 1 \end{cases}$$
$$D^2X = \begin{cases} \frac{\sigma^2}{\xi^2} (\Gamma(1 - 2\xi) - \Gamma^2(1 - \xi)) & \xi < \frac{1}{2} \\ \infty & \xi \geq \frac{1}{2} \end{cases}$$

Rozkłady klasyczne, metoda POT

4. Uogólniony rozkład Pareto – GPD o dystrybuancie:

$\Pr(X \leq x + b) = 1 - (1 - \alpha\kappa x)^{\frac{1}{\kappa}} \quad x > 0, \alpha, b > 0, \kappa < 0$
z parametrami b – przesunięcia, α – skali, ξ – kształtu.

$$EX = \begin{cases} \frac{1}{\alpha(\kappa+1)} & \text{gdy } \kappa \in (-1; 0), \\ \infty & \text{gdy } \kappa \leq -1 \end{cases}$$
$$D^2X = \begin{cases} \frac{1}{\alpha^2(1+\kappa)^2(1+2\kappa)} & \text{gdy } \kappa \in (-\frac{1}{2}; 0) \\ \infty & \text{gdy } \kappa \leq -\frac{1}{2} \end{cases}$$

Model POT

1. Przepływ maksymalny w wezbraniu estymowany rozkładem GPD: $F(x)$;
2. Liczba wezbrań w roku estymowana rozkładem Poissona o parametrze λ ¹.

Rozkład maksymalnego przepływu rocznego:

$$\Pr(Q \leq x) = e^{-\lambda(1-F(x))}$$

¹Wang 1991

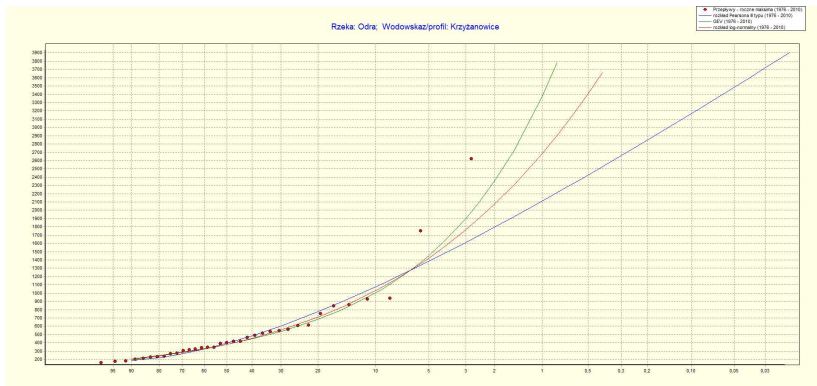
Model POT – estymacja

Estymacja dwu-krokowa, metodą największej wiarygodności
sprowadza się:

GPD: rozwiązania układu równań przyjmując parametr przesunięcia
 b jako znany

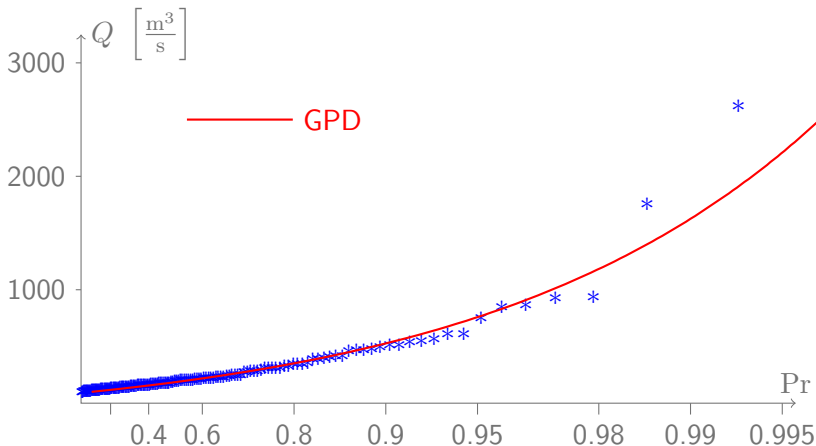
Poisson: obliczenie średniej częstości wezbrań w roku.

Krzyżanowice – roczne przepływy maksymalne

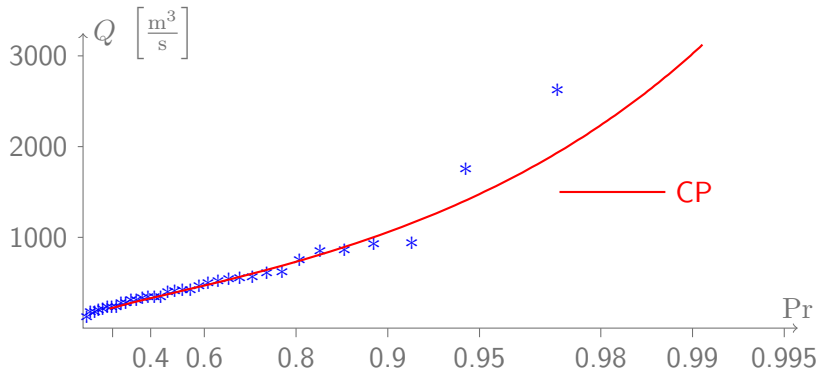


Rozkład	p -wartość	Q_{100}	Q_{200}	Q_{500}
P-III	0,56	2110	2430	2850
Log-normal	0,82	2680	3410	4570
GEV	0,85	3370	4830	7760

Krzyżanowice – przepływy maksymalne w wezbraniach



Rozkład CP – Krzyżanowice r. Odra



Rozkład	p -wartość	Q_{100}	Q_{200}	Q_{500}
CP	0,71	3015	4060	5980

Analiza statystyczna wezbrań (POT)

Dane przepływów maksymalnych z lat 1976-2010 z 330 stacji wodowskazowych;

1. Test zgodności χ^2 na poziomie ufności 0,01 nie było podstaw do odrzucenia hipotezy o zgodności 325 przypadkach (98,5%), na poziomie ufności 0,05 – w 316 stacjach (95,8%)
2. Estymowany parametr κ rozkładu

κ	$\hat{\kappa} < -0,5$	$\hat{\kappa} < -1,0$
Liczba stacji	57 (17,3%)	8 (2,4%)

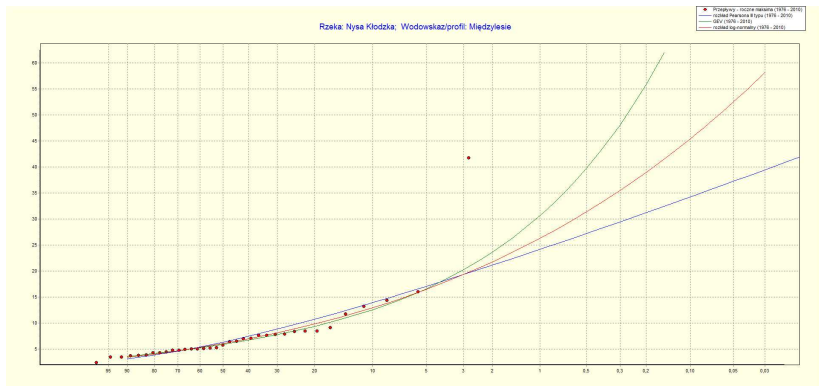
Analiza statystyczna wezbrań (POT)

Dane przepływów maksymalnych z lat 1976-2010 z 330 stacji wodowskazowych;

1. Test zgodności χ^2 na poziomie ufności 0,01 nie było podstaw do odrzucenia hipotezy o zgodności 325 przypadkach (98,5%), na poziomie ufności 0,05 – w 316 stacjach (95,8%)
2. Estymowany parametr κ rozkładu

κ	$\hat{\kappa} < -0,5$	$\hat{\kappa} < -1,0$
Liczba stacji	57 (17,3%)	8 (2,4%)

Międzylesie – roczne przepływy maksymalne



Rozkład	p -wartość	Q_{100}	Q_{200}	Q_{500}
P-III	0,05	24,2	27,2	31,2
Log-normal	0,49	26,3	31,4	39,0
GEV	0,07	30,7	39,7	55,8
CP	0,001			

$$\hat{\kappa} = -0,49978$$

Rozkład MIX'a

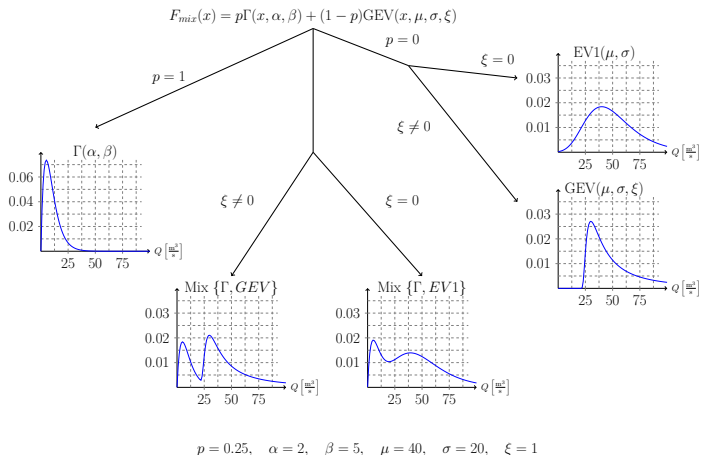
$$F_{\text{MIX}}(x) = p\Gamma(x, \alpha, \beta) + (1 - p)\text{GEV}(x, \mu, \sigma, \xi)$$

o parametrach $\theta = (p, \alpha, \beta, \mu, \sigma, \xi)$ ²
gdzie:

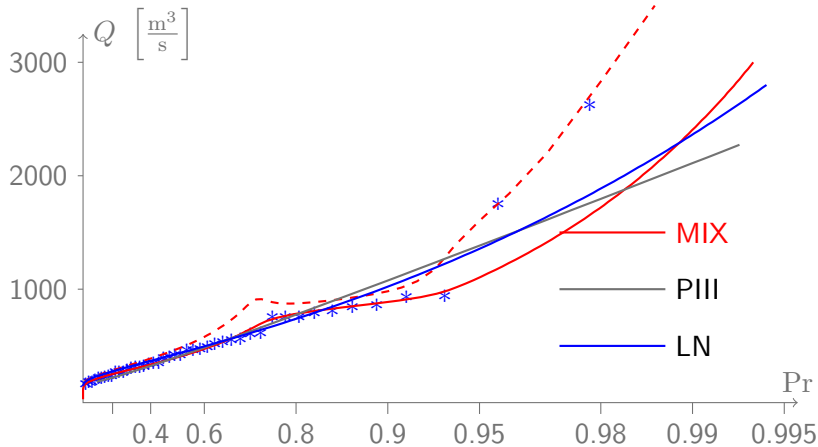
- ▶ p jest parametrem mieszania;
- ▶ $\Gamma(x, \alpha, \beta)$ jest dwuparametrowym rozkładem gamma;
- ▶ $\text{GEV}(x, \mu, \sigma, \xi)$ jest trójparametrowym uogólnionym rozkładem zdarzeń ekstremalnych.

²W. Szulczewski, W. Jakubowski,
<http://dx.doi.org/10.1007/s11269-018-2005-6>

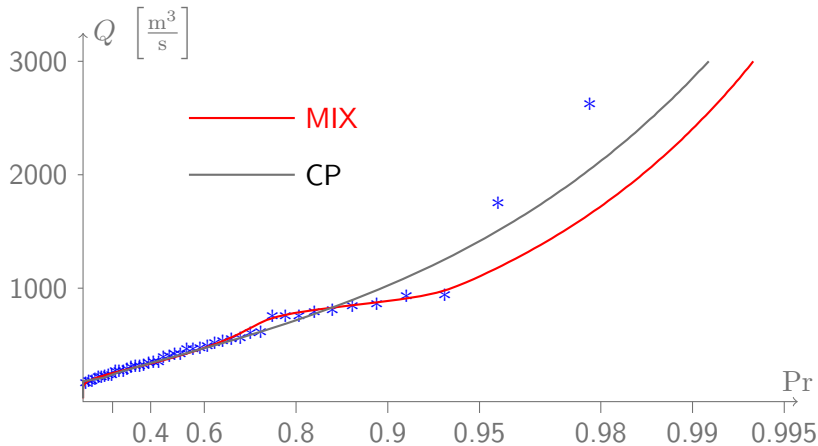
Schemat możliwych rozkładów MIX'a



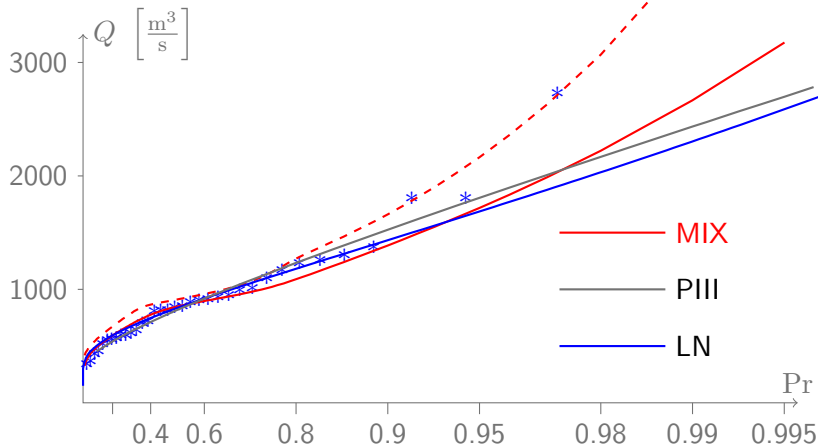
Porównanie - Krzyżanowice r. Odra lata 1966-2010



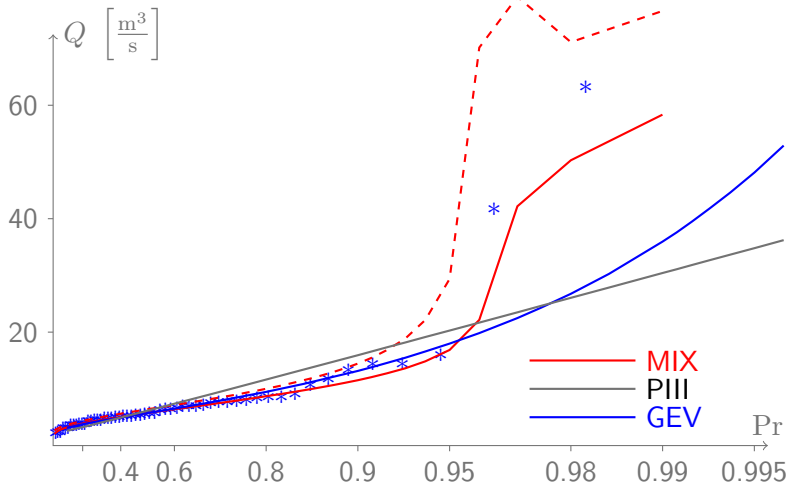
Porównanie - Krzyżanowice r. Odra lata 1966-2010



Porównanie - Słubice r. Odra lata 1966-2010

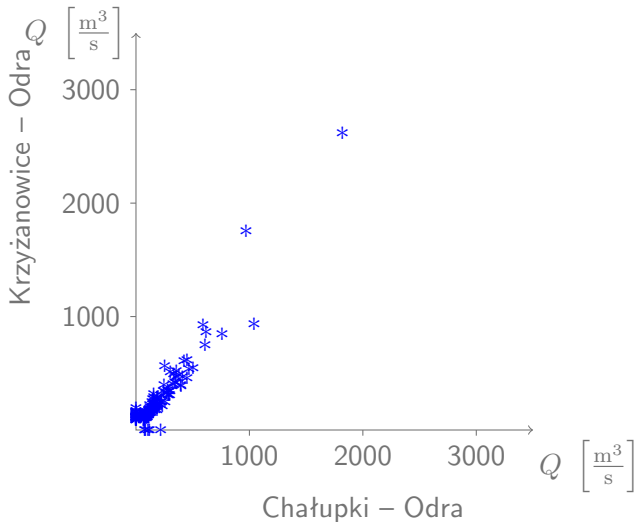


Porównanie - Międzylesie r. Nysa Kłodzka lata 1955-2009

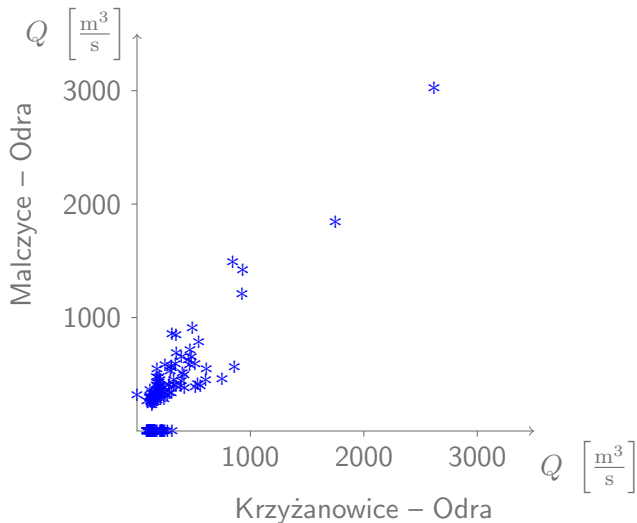


Analiza dwuwymiarowa wezbrań

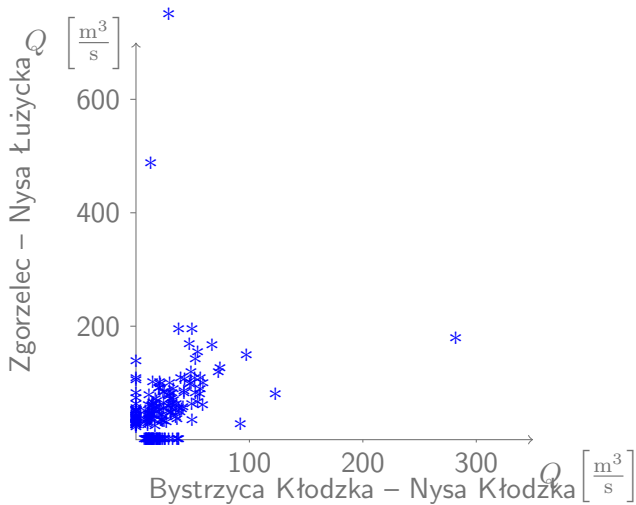
Analiza dwuwymiarowa wezbrań – rzeka Odra



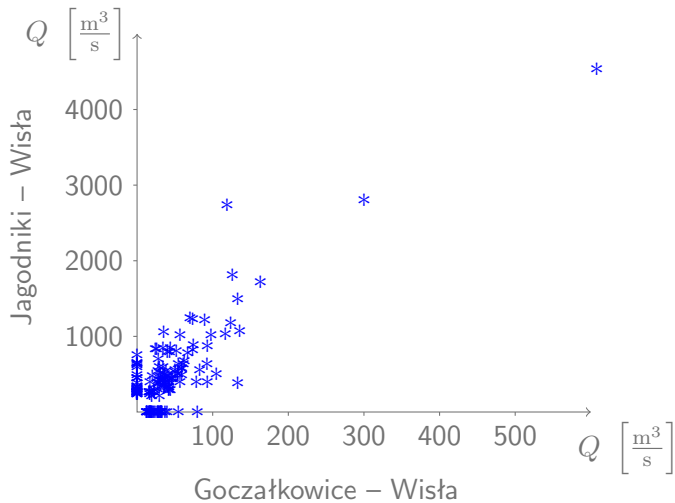
Analiza dwuwymiarowa wezbrań – rzeka Odra



Analiza dwuwymiarowa wezbrań – Sudety



Analiza dwuwymiarowa wezbrań – Wisła



Model dwuwymiarowy

Niech $\{A_i^{(1)}, A_i^{(2)}\}$, $i = 1, \dots, n$ będzie ciągiem dwuwymiarowych wzajemnie niezależnych zmiennych losowych o jednakowym rozkładzie $F(x, t)$. Tak jak w przypadku jednowymiarowym [Coles 2001] definiujemy dwuwymiarową zmienną losową wartości maksymalnych:

$$\left(M_n^{(1)}, M_n^{(2)}\right) = \left\{ \max_{i=1, \dots, n} A_i^{(1)}, \max_{i=1, \dots, n} A_i^{(2)} \right\}.$$

$A_i^{(1)}$, $A_i^{(2)}$ – określają przepływy w stacjach wodowskazowych.

$$\text{Rozkład } \Pr\left(M_n^{(1)} \leq x, M_n^{(2)} \leq t\right) = H(x, t)$$

Model dwuwymiarowy

Tajvidi [1996] pokazał

$$\bar{H}(x, t) = \left(\frac{\bar{F}_1^p(x) + k\bar{F}_1^{p/2}(x)\bar{F}_2^{p/2}(t) + \bar{F}_2^p(t)}{\bar{F}_1^p(0) + k\bar{F}_1^{p/2}(0)\bar{F}_2^{p/2}(0) + \bar{F}_2^p(0)} \right)^{\frac{1}{p}}, \quad (x, t) \geq (0, 0)$$

gdzie:

- ▶ $\bar{F}_1(x) = 1 - F_1(x) = (1 - \alpha_1\kappa_1(b_1 + x))^{1/\kappa_1}$
- ▶ $\bar{F}_2(t) = 1 - F_2(t) = (1 - \alpha_2\kappa_2(b_2 + t))^{1/\kappa_2}$
- ▶ $0 \leq k \leq 2(p - 1)$, $p \geq 2$, $\kappa_1, \kappa_2 \in (-1; 0)$, $\alpha_1, \alpha_2 > 0$
- ▶ $H(x, t) = \begin{cases} 1 - \bar{H}(x, t), & \text{gdy } (x, t) \geq (0, 0), \\ 0, & \text{w przeciwnym wypadku} \end{cases}$

F_1, F_2 są jednowymiarowymi trójparametrowymi uogólnionymi rozkładami Pareto opisującymi każdą ze zmiennych losowych $M_n^{(k)}$, $k = 1, 2$ osobno. Parametry przesunięcia b_1 oraz b_2 określają poziom odcięcia.

Model dwuwymiarowy

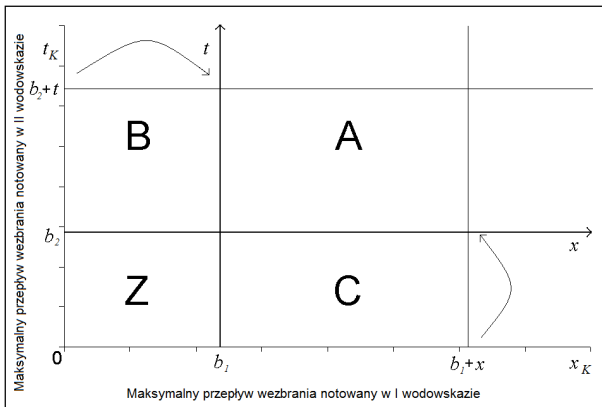


Figure: Dziedzina uogólnionego dwuwymiarowego rozkładu Pareto – Tajvidi[1996]

Model dwuwymiarowy – własności

1. Rozkład warunkowy:

$$\begin{aligned} \Pr [(M_K, N_K) \not\leq (x + b_1, t + b_2) | (M_K, N_K) \not\leq (x, t)] \\ = \frac{\bar{H}(x+b_1, t+b_2)}{\bar{H}(b_1, b_1)} \end{aligned}$$

jest również uogólnionym dwuwymiarowym rozkładem Pareto z parametrami przesunięcia $b_d + x_0$ oraz $b_t + t_0$

2. Całki

$$\begin{aligned} \int_0^{\infty} \int_0^{\infty} h(x, t) dx dt &= \frac{\bar{F}_d(0) + \bar{F}_t(0)}{V} - 1; \\ \int_0^{\infty} h(x, 0) dx &= 1 - \frac{\bar{F}_t(0)}{V}; \quad \int_0^{\infty} h(0, t) dt = 1 - \frac{\bar{F}_d(0)}{V} \end{aligned}$$

gdzie

$$V = \left(\bar{F}_d^p(0) + k \bar{F}_d^{p/2}(0) \bar{F}_t^{p/2}(0) + \bar{F}_t^p(0) \right)^{\frac{1}{p}}$$

wyznaczają prawdopodobieństwo każdego z obszarów A, B, C

Model dwuwymiarowy – własności

3. Rozkład warunkowy:

$$\Pr(M_K \leq x + b_1 | N_K = t + b_2) = \left(\bar{F}_1^p(x) + k \bar{F}_1^{p/2}(x) \bar{F}_2^{p/2}(t) + \bar{F}_2^p(t) \right)^{\frac{1}{p}-1} * \left(\frac{k}{2} \bar{F}_1^{p/2}(x) \bar{F}_2^{p/2-1}(t) + \bar{F}_2^{p-1}(t) \right)$$

opisuje rozkład przepływów maksymalnych wezbrań w pierwszym wodowskazie, pod warunkiem, że przepływ maksymalny w drugim wynosi $t + b_2$.

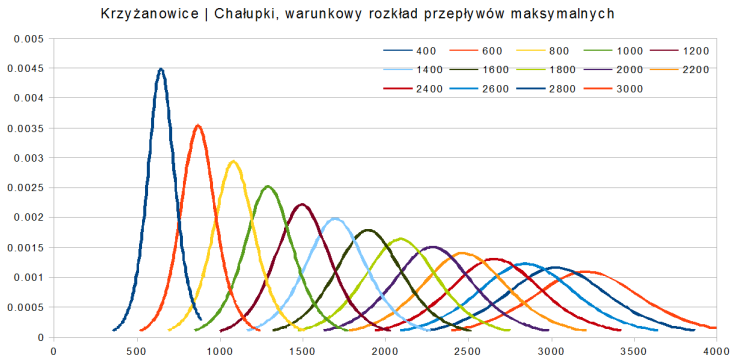
4. Współczynnik korelacji według rozkładu BGDPD wyznacza się tylko dla obszaru A i jest wyznaczalny tylko w przypadku gdy istnieją wariancje w estymowanych rozkładach brzegowych.

Model dwuwymiarowy – zastosowania

1. Współczynnik korelacji

Wodowskaz 1	Wodowskaz 2	\hat{r}
Chałupki	Krzyżanowice	0.80618
Krzyżanowice	Malczyce	0.89986
Bystrzyca Kłodzka	Zgorzelec	0.86826
Goczałkowice	Jagodniki	0.90672

2. Warunkowe rozkłady przepływów maksymalnych:



Wnioski

1. Modele jednowymiarowe:

- ▶ rozkłady standardowe (oprócz GEV) obarczone są nieznanym błędem wynikającym z estymacji parametru przesunięcia; wady tej nie ma rozkład MIX'a – wymaga on jednak przynajmniej 40-letnich ciągów obserwacji;
- ▶ jak wynika z analizy rozkładów wezbrań w 15% posterunków wodowskazowych w Polsce rozkłady przepływów maksymalnych cechują tzw. ciężkie ogony; w takim przypadku estymowany GPD (uogólniony rozkład Pareto) nie posiada skończonej wariancji; w rozkładzie MIX'a nie ma wtedy możliwości wyznaczenia obszarów ufności dla kwantyli przewyższenia;
- ▶ oszacowania średniej wody n -letniej powinno się ograniczać wody 100, maksymalnie 200 letniej.

2. Modele dwuwymiarowe:

- ▶ analiza zagrożeń powodziowych powinna odejść od metod jednowymiarowych skupiając się na badaniu związków pomiędzy wysokimi przepływami w „bliskich” sobie posterunkach wodowskazowych;
- ▶ podczas wysokich przepływów rozkład warunkowy przepływu w wodowskazie A pod warunkiem wysokiego przepływu notowanego w wodowskazie B ma dużą zmienność – potwierdza to trudności trudności dobrego oszacowania odpowiednich kwantyli.



Dziękujemy
za uwagę