

Analiza 1

Marcin Borysiak

2022-04-11

Cel

Głównym celem pracy jest analiza 20 i 50 letnich poziomów zwrotu dla sezonu letniego. Wykorzystano 3 różne sposoby estymacji:

1. Biblioteka gamlss
2. Metoda maksimów blokowych
3. Metoda przekroczenia progu

Stacja

Dane do analizy pochodzą ze stacji w mieście Wisła. Kod stacji to “X249180230”. Stacja znajduje się na długości geograficznej 18.86139° , szerokości geograficznej 49.65472° i wysokości 430m nad poziomem morza. Miasto znajduje się w województwie śląskim, w powiecie cieszyńskim przy południowej granicy Polski.



Dane

Dane do analizy pochodzą ze zbioru pomiarów od 2008 do 2018 wykonywanych co 10 minut. Każdy pomiar jest oznaczony datą i godziną i wyrażony jest w stopniach celsjusza.

Biblioteka gamlss

Biblioteka gamlss implementuje funkcje dopasowujące parametry dla rozkładów prawdopodobieństwa. Umożliwia ona znalezienie właściwego rozkładu do danych i oszacowanie poziomów zwrotu.

Funkcją fitDist dopasowujemy odpowiedni model do danych.

Najlepszy rozkład znaleziony przez powyższą funkcję dla danych ze stacji Wisła to **SEP1 - The Skew Power exponential type 1-4 distribution for fitting a GAMLSS** Funkcja gęstości prawdopodobieństwa dla rozkładu ma 4 parametry:

μ, σ, ν, τ . μ i σ odpowiadają położeniu i skali rozkładu. Parametr ν określa lewy ogon rozkładu z $\tau > 1$ wskazującym na “lżejszy” ogon i $\tau < 1$ wskazującym na “cięższy” ogon. Parametr τ określa prawy ogon rozkładu w ten sam sposób co lewy. Funkcję można przedstawić za pomocą wzoru:

$$f(y|\mu, \sigma, \nu, \tau) = \frac{c}{\sqrt{2\pi}\sigma(1+z^2)^{1/2}} e^{-r^2/2}$$

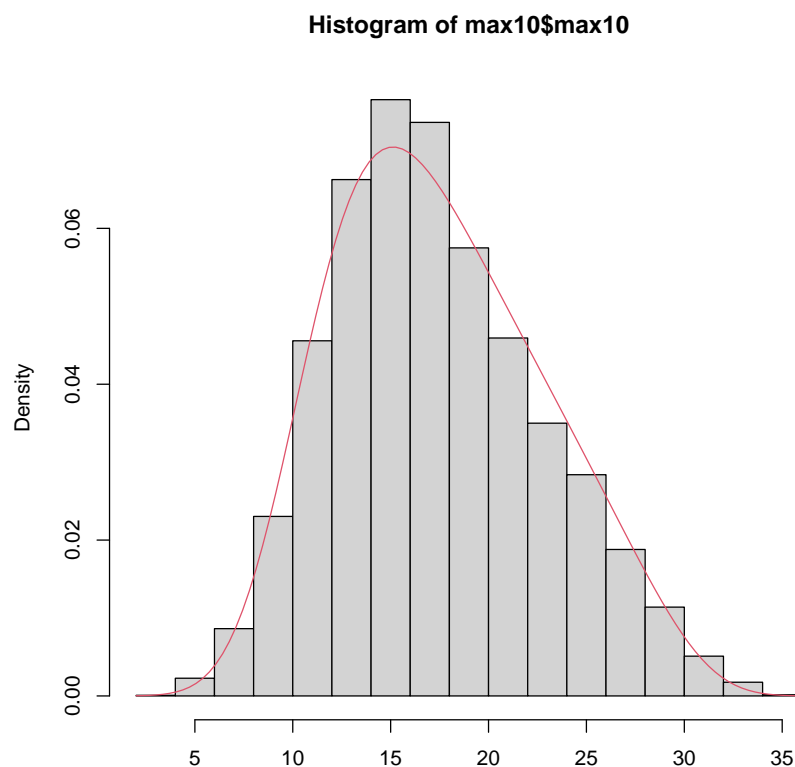


Figure 1: Histogram

Jak widać na histogramie krzywa dobrze oddaje jego kształt.

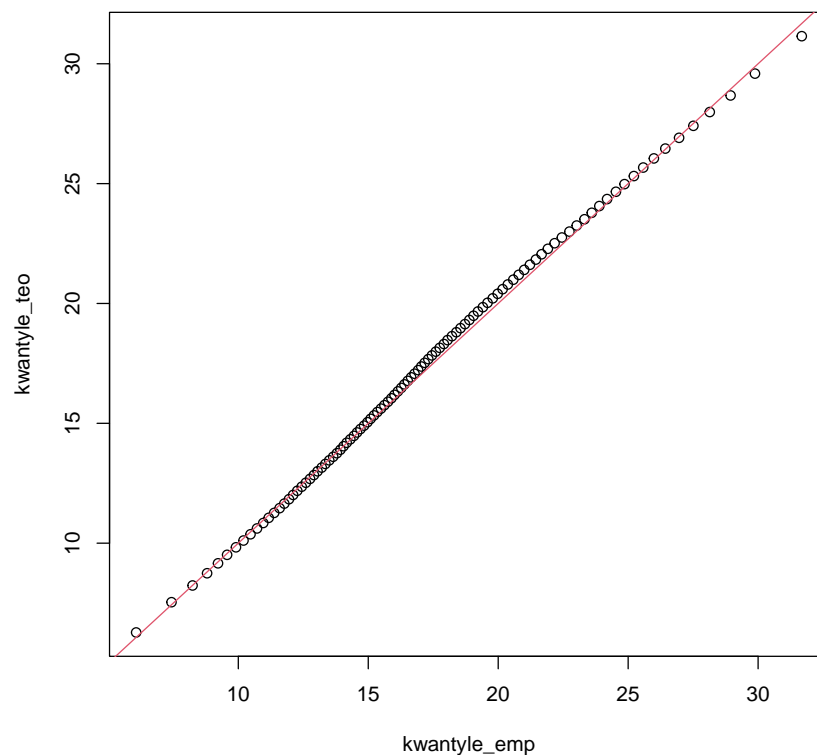


Figure 2: Porównanie kwantyli

Porównanie kwantyli empirycznych z teoretycznymi jest również dokładne.

Metoda maksimów blokowych

Metoda maksimów blokowych polega na podzieleniu danych na bloki o równym rozmiarze, z których wyliczane są maksima. Maksima są następnie wykorzystywane do obliczania parametrów do funkcji `gev()`.

Biblioteka `evir`

Biblioteka `evir` służy do obliczania zdarzeń ekstremalnych. Zawiera funkcje takie jak `GEV` i `GPD`, które zostaną użyte do metody maksimów blokowych i metody przekroczenia progu.

Rozkład `GEV`

$$G(y) = \exp(-[(y - \mu)/\sigma]_+^{-1/\xi})$$

gdzie $\sigma > 0$, $-\infty < \mu < \infty$ i $1 + \xi(y - \mu)/\sigma > 0$. Tutaj, $x_+ = \max(x, 0)$

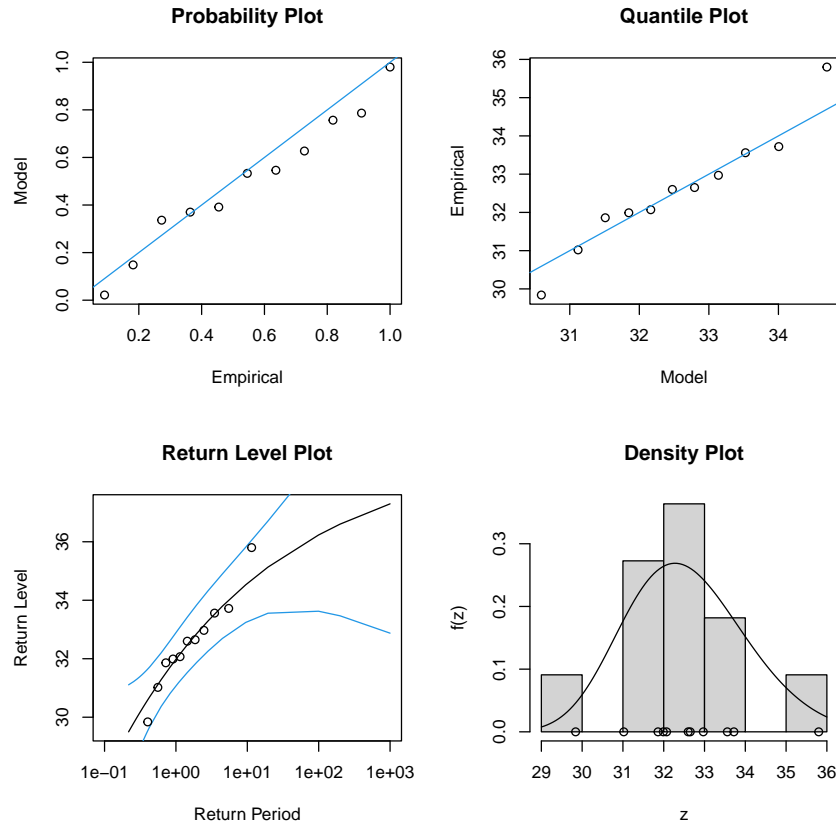


Figure 3: BMM

Odwzorowanie dla metody maksimów blokowych wydaje się bardzo niedokładne.

Metoda przekroczeń progu

Metoda przekroczeń progu polega na ustaleniu progu, poniżej którego dane nie są uwzględniane w analizie. Do wyliczenia parametrów wykorzystany zostanie unormalizowany rozkład Pareto - GPD.

$$H(y) = 1 - \left[1 + \frac{\xi(y-\mu)}{\sigma} \right]^{-1/\xi}, \text{ gdzie}$$

$$y : y > 0$$

$$\text{i } (1 + \xi(y - \mu)/\sigma) > 0$$

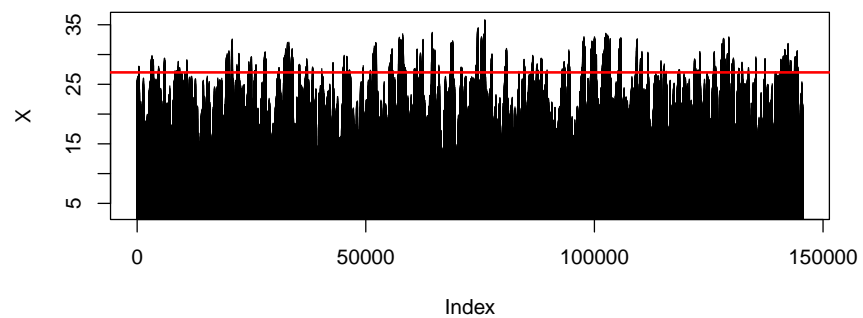


Figure 4: POT

Wykres rozrzutu przy progu 27 stopni zawiŹa próg od 25% do 35%.

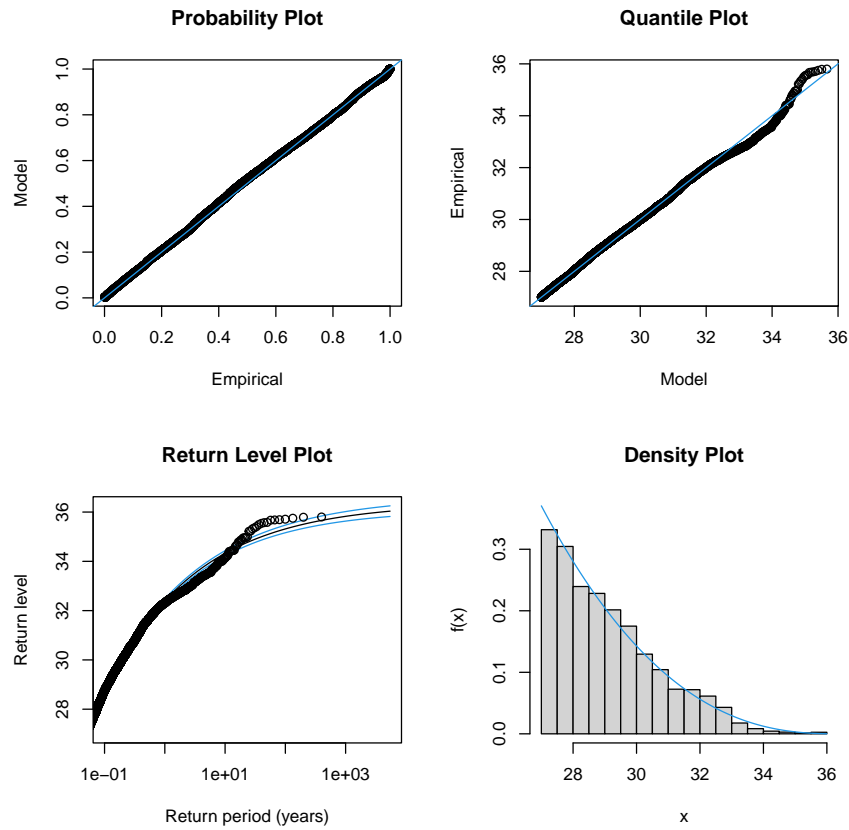


Figure 5: Wartości progowe

Wartości na ogonach są wyraźnie zniekształcone, głównie zawyżone. Dla temperatur występujących w zbiorze odwzorowanie jest dokładne.

Porównanie poziomów zwrotu dla wszystkich pór roku:

	SEP1	GEV	GPD
Lato 20	36,31206	35,13362	35,7647
Lato 50	36,71698	35,80286	35,90235
Jesień 20	37.47909	32.17075	41.03753
Jesień 50	38.4606	32.63413	45.59888
Zima 20	41.58443	20.50975	22.25799
Zima 50	42.88735	21.22244	22.64157
Wiosna 20	15.77225	29.27850	29.99343
Wiosna 50	15.78718	29.47713	30.27762