

Projekt egzaminacyjny
Analiza zdarzeń ekstremalnych

Antonina Brzeska 246824

1. Krótki opis wybranych danych. Zestaw danych opisuje zanotowane wysokości fal na wybrzeżu Australijskim w miejscowości Mooloolaba. Wybrane pomiary zawierają 29108 rekordów, z czego mierzono je w okresie 20 miesięcy począwszy od stycznia 2017 roku. W każdym miesiącu pobrano podobną ilość danych (47 rekordów dziennie).

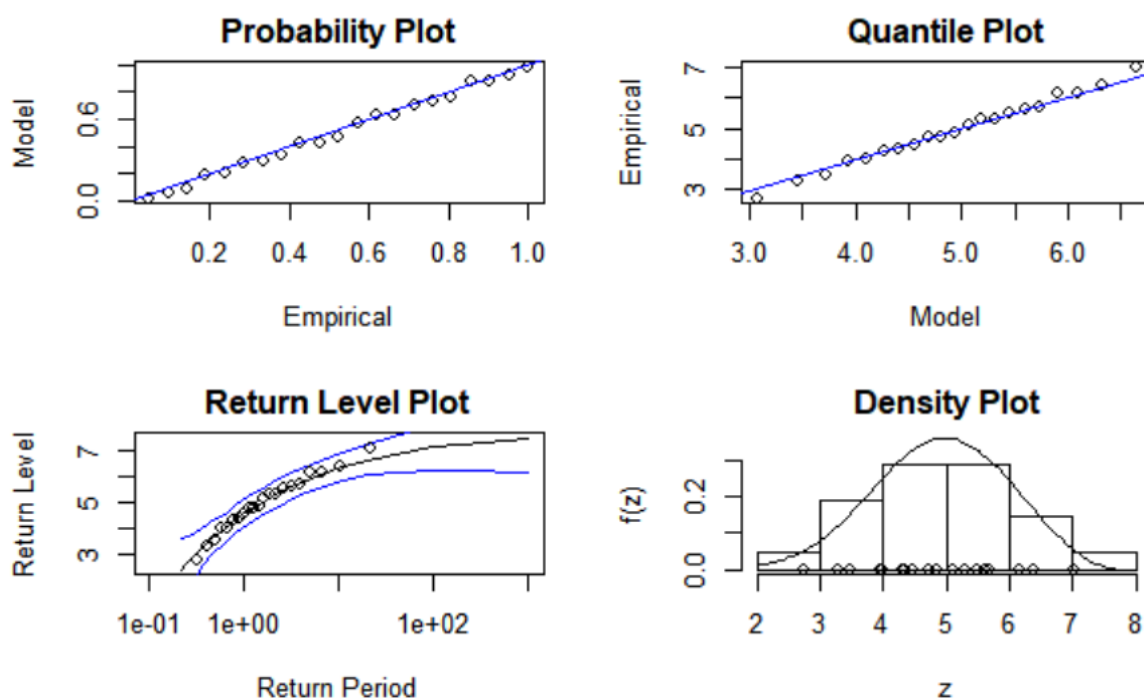
2. Analiza metodą maksimów blokowych (BMM). Wyestymowano parametry rozkładu GEV w oparciu o maksima z ustalonych równych, nienachodzących na siebie bloków o równej długości.

Dane mają dużo punktów pomiarowych (47 dziennie) więc przyjęto rozmiar bloku równy 1410, co odpowiada ilości pomiarów w bloku 30-dniowym (miesiącu).

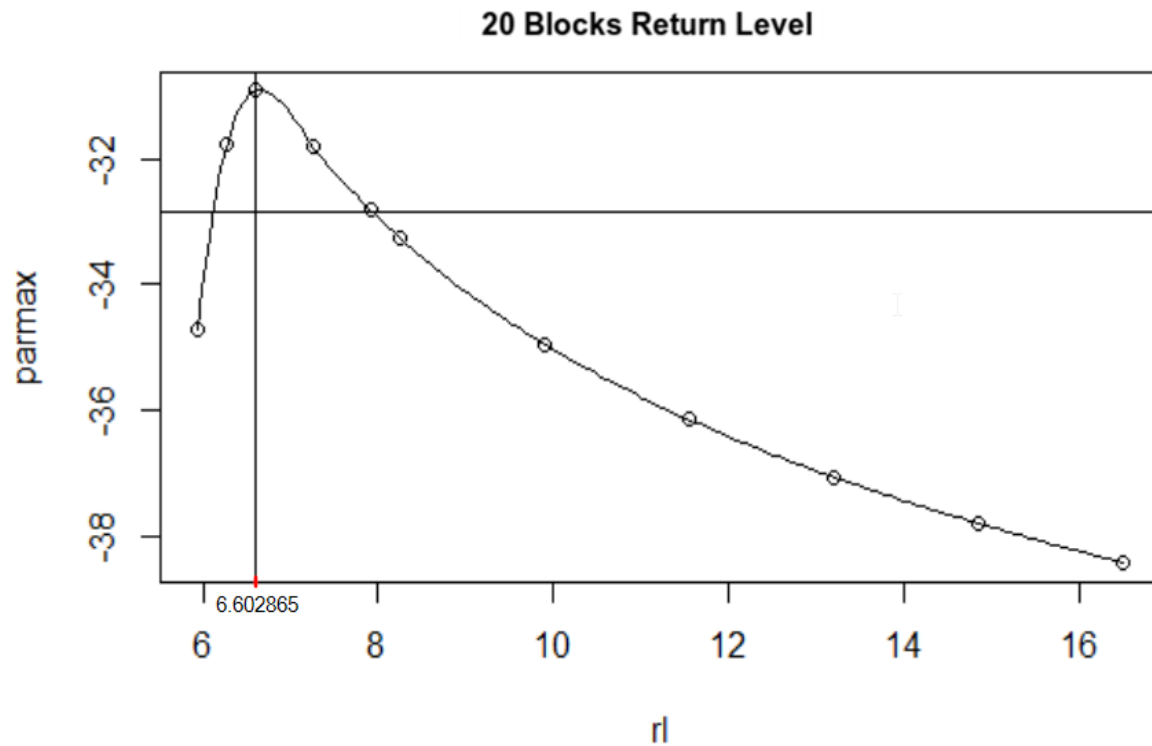
	xi	sigma	mu
Parametry rozkładu GEV	-0.3372591	1.0874319	4.5626598

$\xi < 0$, więc dystrybuenta F zbiegająca do rozkładu GEV należy do maksymalnej dziedziny przyciągania Weibulla.

Przeprowadzono analizę oceniającą dobroć dopasowania za pomocą następujących wykresów diagnostycznych:



Wyznaczono poziomy zwrotu x_{20} :



Poziomu zwrotu x_{20}
6.602865

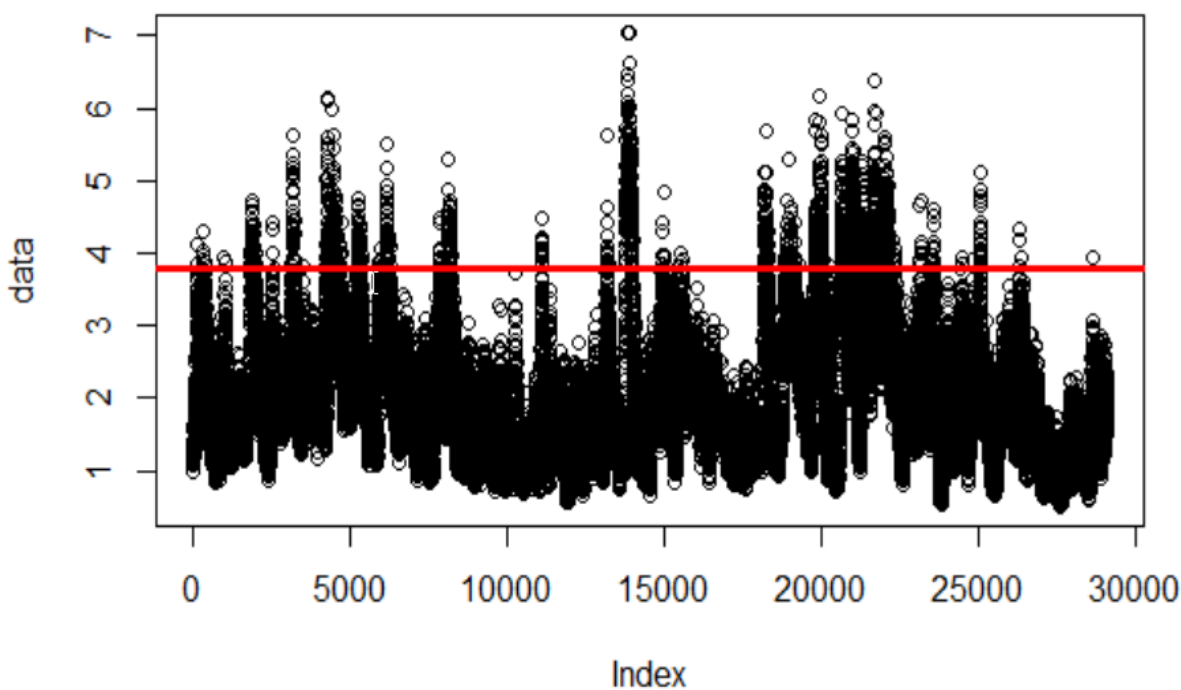
Poziom zwrotu obliczony metodą maksimów blokowych dla 20-stu bloków, z czego każdy blok zawiera średnio $30 \cdot 47$ rekordów (47 spisów dziennie przez 30 dni) wynosi 6.602865. Oznacza to, że raz na 20 miesięcy w miejscowości Mooloolaba można spodziewać się, że wysokość fal przekroczy 6,6 metra.

3. Analiza metodą przekroczeń progu (POT). Metoda polega na poddaniu analizie danych w zależności od ustalonego progu. Wybrany próg to kwantyl 95% danych. W tym przypadku kwantyl $u = 3.8$ co oznacza, że 5% mierzonych fal przekraczało wysokość 3.8 m. Wzięto następnie pod uwagę wartości nadwyżek nad ten próg i wyestymowano parametry rozkładu GPD:

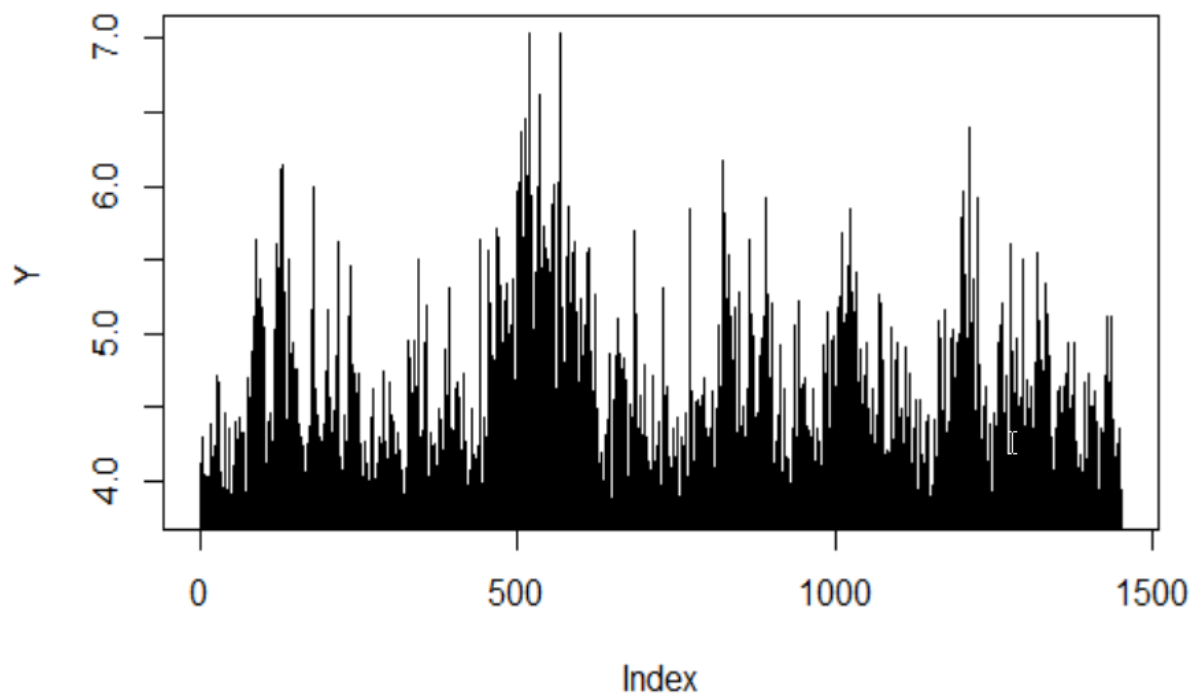
	xi	beta
Parametry rozkładu GPD	-0.1566445	0.6728065

Przeprowadzono analizę oceniającą dobroć dopasowania za pomocą następujących wykresów diagnostycznych:

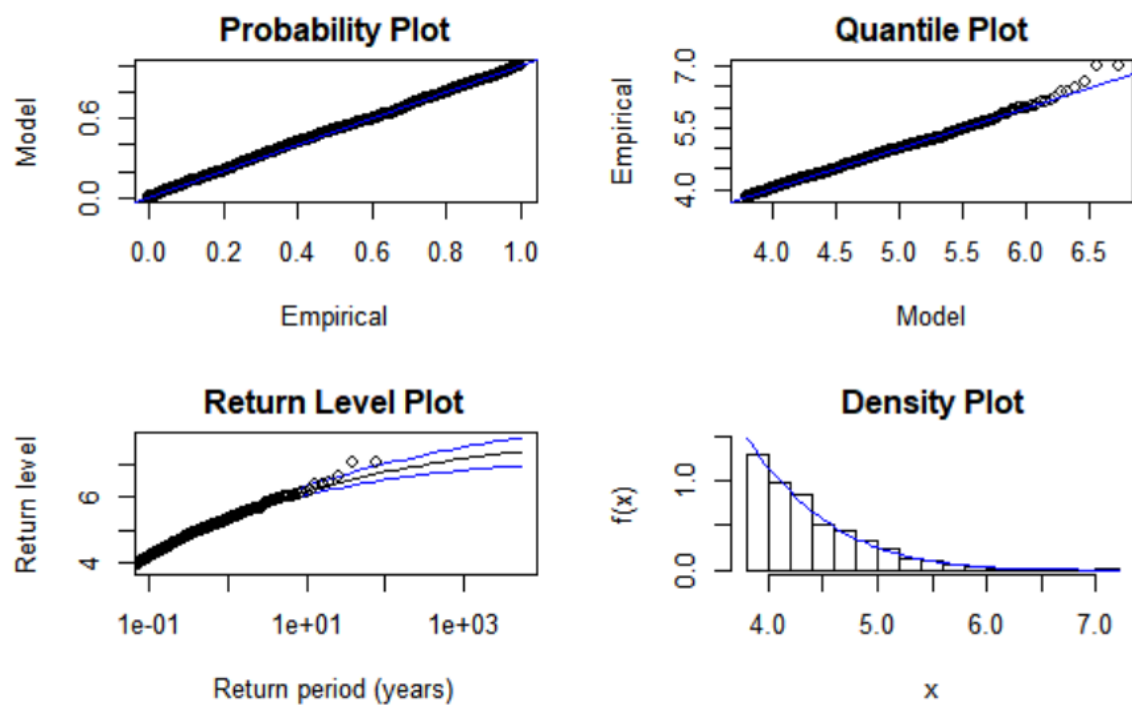
- Wykres rozrzutu z zaznaczonym progiem:



- Wykres nadwyżek nad próg:



- Dobroć dopasowania na wykresach:



Wyznaczono poziom zwrotu x_{20} :

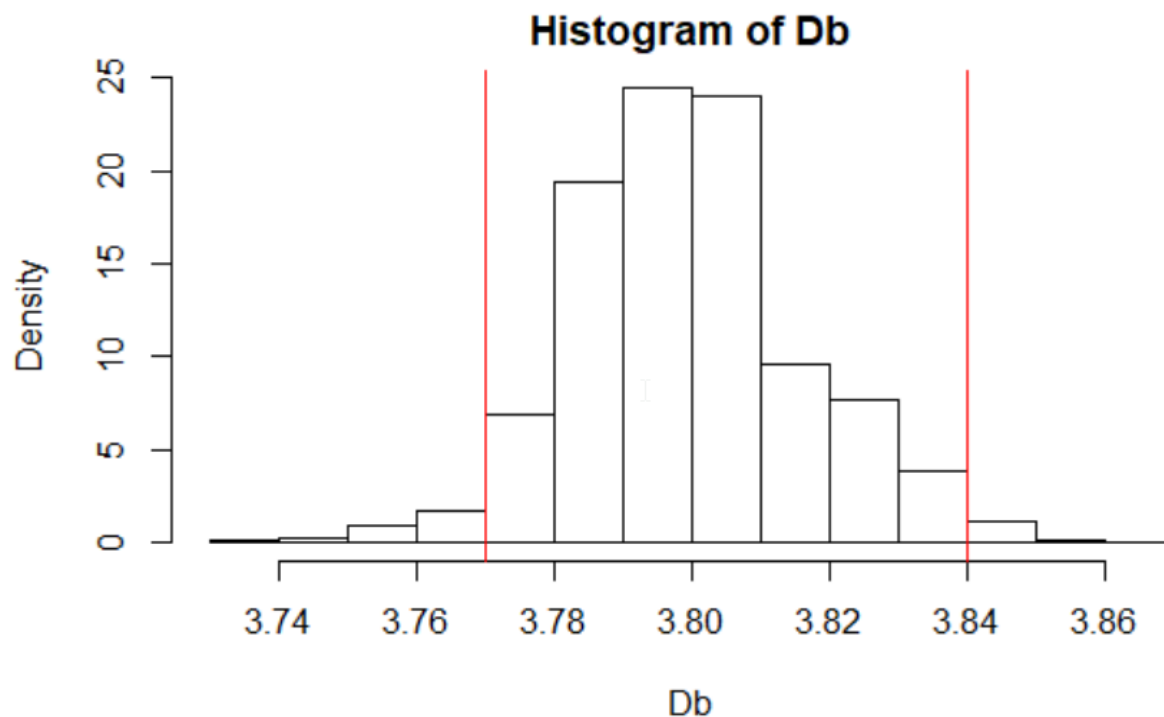
Poziomu zwrotu x_{20}
6.715564

Poziom zwrotu obliczony metodą przekroczeń progu wynosi 6.715564. Oznacza to, że średnio raz na 20 miesięcy wystąpi fala o takiej wysokości lub wyższej. Analizując powyższe wykresy można zauważyć, że model jest dobrze dopasowany.

4. Analiza metodą Bootstrap. Wygenerowano 1000 prób losowych na podstawie danych empirycznych, następnie aproksymowano dane faktyczne za pomocą danych wygenerowanych.

Kwantyl poziomy 0.95	3.803075[m]
Obciążenie	0.00307471
Kwantyl poziomy 0.95 z uwzględnieniem obciążenia	3.796925[m]

- Rozkład estymatora kwantyli na próbkach bootstrapowych:

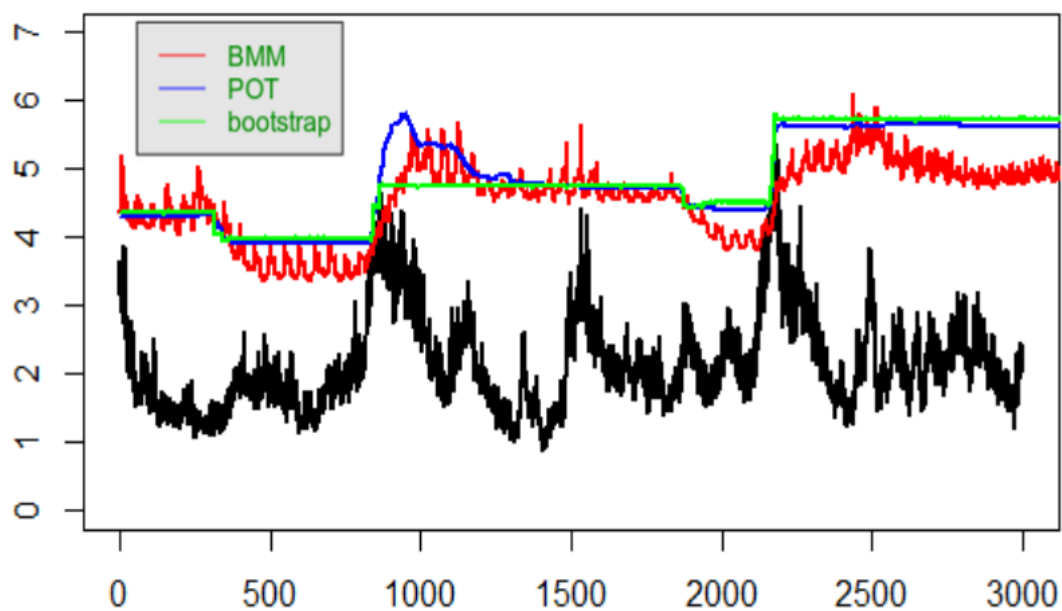


Przedział ufności na poziomie 0.95:

97.5%	2.5%
3.84	3.77

5. Backtesting. Jest to proces testowania metod przewidujących występowanie zdarzeń ekstremalnych opierając się na danych historycznych. Celem jest sprawdzenie jak dana strategia przewidzi faktyczne, aktualne dane.

Backtesting przeprowadzono w pętli 3000 iteracji testując kolejne przedziały 1000 danych metodami BMM maksimów blokowych, POT przekroczeń progu oraz bootstrapową.



Źródła danych:

- <https://www.kaggle.com/jolasa/waves-measuring-buoys-data-mooloolaba>