

Projet : Vagues dans le golfe du Lion

Damien MARIAC

Lucien DUGOU

10 février 2026

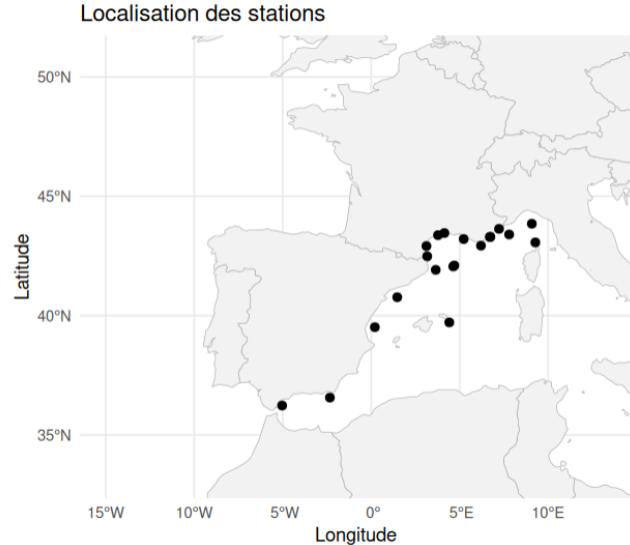
Table des matières

1	Introduction	3
2	Approche univariée	3
2.1	Maximum par bloc	3
2.1.1	Ajustement d'une GEV	3
2.1.2	Période de retour	5
2.2	Dépassement de seuil	5
3	Approche bivariée	6

1 Introduction

L'objectif de ce projet est de modéliser le comportement extrême des vagues dans le golfe du Lion à l'aide de la théorie des valeurs extrêmes étudiée en cours.

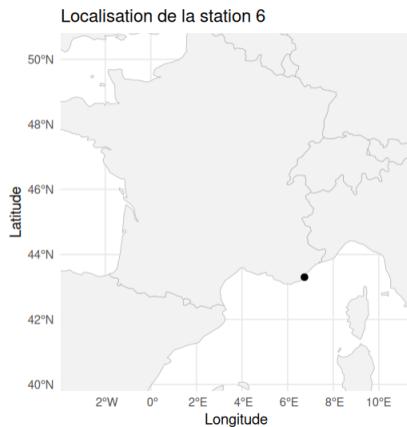
Les données utilisées pour ce projet recensent les hauteurs de vagues de 20 stations différentes de 1961 à 2012 à un rythme de 1 observation par heure.



Nous allons dans une première partie modéliser les hauteurs de vagues extrêmes de la station 6 à l'aide d'une approche univariée. Puis dans une deuxième partie nous allons modéliser la dépendance extrême entre la station 6 et 14 à l'aide d'une approche bivariée. Enfin, dans une troisième partie, nous allons modéliser la dépendance extrême entre des stations proches et éloignées à l'aide d'une approche bivariée.

2 Approche univariée

La station 6 est située à Saint Maxime et est représentée sur la carte ci-dessous.

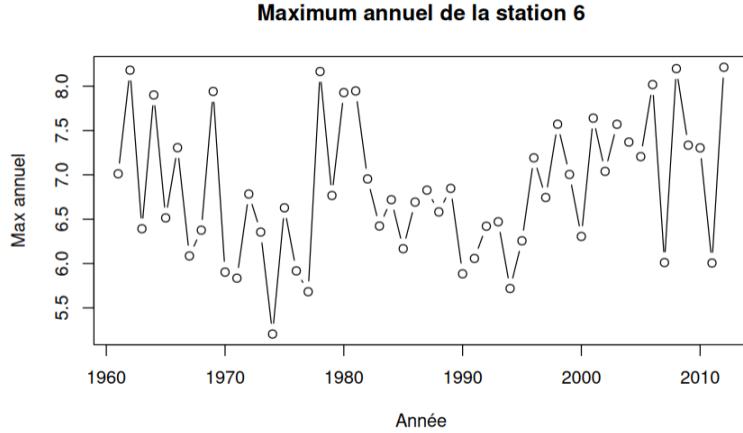


2.1 Maximum par bloc

2.1.1 Ajustement d'une GEV

Afin de modéliser les hauteurs de vagues extrêmes, nous allons utiliser la méthode du maximum par bloc. Nous allons diviser les données en blocs annuels et prendre le maximum de chaque bloc. Nous obtenons ainsi une série de 51 maxima annuels.

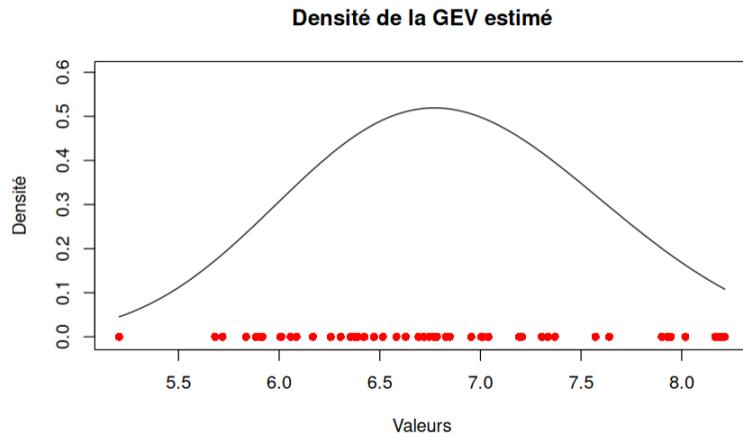
Nous avons choisi de prendre des blocs annuels car cela correspond à une période de temps suffisamment longue pour capturer les variations saisonnières et les tendances à long terme en évitant de la dépendance entre les blocs, tout en étant suffisamment courte pour éviter de perdre des informations importantes sur les événements extrêmes.



Nous pouvons donc fit une distribution de GEV sur ces maxima annuels pour estimer les paramètres de la distribution. Nous récupérons de plus les écarts types de ces paramètres pour pouvoir construire des intervalles de confiance dont nous fixons le niveau à 95%.

- $\mu = 6.560243 \pm 0.2337362$
- $\sigma = 0.7349365 \pm 0.1753661$
- $\gamma = -0.2577396 \pm 0.2722555$

Nous remarquons que le paramètre de forme γ est négatif ce qui correspond à une Weibull, cependant, l'intervalle de confiance contient zéro, ce qui n'exclue pas que la distribution pourrait être de type Gumbel (si $\gamma = 0$) ou de type Fréchet (si $\gamma > 0$).



2.1.2 Période de retour

Avec les paramètres estimés, nous pouvons calculer les périodes de retour pour 100, 500 et 1000 ans. Comme nous avons utilisé des maxima annuels, la période de retour est définie par :

Ce qui correspond respectivement à :

- $p_{100} = 8.540445 \text{ m}$
- $p_{500} = 8.83687 \text{ m}$
- $p_{1000} = 8.930977 \text{ m}$

Les périodes de retour semblent relativement proches les unes des autres, ce qui peut s'expliquer par le fait que la distribution de GEV ajustée a une queue bornée (avec un point terminal), ce qui limite la croissance des quantiles extrêmes au fur des années.

Afin de consolider ces résultats, nous pouvons utiliser une autre méthode afin d'estimer les périodes de retour, comme la méthode de dépassement de seuil. En utilisant cette méthode, nous pouvons estimer les périodes de retour pour les mêmes quantiles et comparer les résultats obtenus avec ceux de la méthode du maximum par bloc.

2.2 Dépassement de seuil

L'avantage de la méthode de dépassement de seuil est qu'elle utilise plus d'informations que la méthode du maximum par bloc, car elle prend en compte tous les événements qui dépassent un certain seuil, plutôt que de ne considérer que les maxima annuels qui pourrait mettre de côté des valeurs extrêmes. Cependant, le choix du seuil est crucial et peut influencer les résultats obtenus.

Afin de choisir un seuil approprié, nous pouvons utiliser le graphique "mean residual life plot" qui permet d'identifier une zone où les résidus moyens sont linéaires, ce qui suggère que les données suivent une distribution de type GPD au-delà de ce seuil.



En utilisant ce graphique, un seuil de 6m semble être un choix approprié car à partir de ce seuil, il y'a un virage (croissance puis décroissance) des résidus moyens.

Après implémentation de la méthode, nous obtenons paramètres avec leur intervalle de confiance à 95% suivants :

- $\sigma = 0.9649 \pm 0.08161178$
- $\gamma = -0.3613 \pm 0.06414328$

Après avoir estimé les paramètres de la GPD, nous pouvons calculer les périodes de retour pour les mêmes quantiles que précédemment. Nous obtenons ainsi :

- $p_{100} = 2.63063 \text{ m}$ donc une hauteur de vague de $6 + 2.63063 = 8.63063 \text{ m}$
- $p_{500} = 2.64841 \text{ m}$ donc une hauteur de vague de $6 + 2.64841 = 8.64841 \text{ m}$

— $p_{1000} = 2.653404$ m donc une hauteur de vague de $6 + 2.653404 = 8.653404$ m

Les périodes de retour obtenues avec la méthode de dépassement de seuil semblent ne pas "exploser" plus le temps avance du au paramètre de forme plus bas que celui de la méthode du maximum par bloc. Néanmoins, les périodes de retour obtenues avec la méthode de dépassement de seuil sont relativement proches de celles obtenues avec la méthode du maximum par bloc, ce qui nous conforte sur la hauteur maximale d'une vague dans les prochaines années.

L'analyse univariée de la station 6 nous a permis d'estimer les paramètres de la distribution de GEV et de GPD, ainsi que les périodes de retour. Nous pouvons approfondir cette analyse en modélisant la dépendance extrême entre la station 6 et une autre station, à l'aide d'une approche bivariée.

3 Approche bivariée

Dans cette section, nous allons modéliser la dépendance extrême entre la station 6 et la station 14, qui est située proche de Sète.