PROCESSUS MAX-STABLES APPLIQUÉS À L'ÉTUDE DE L'ÉVOLUTION DU TRAIT DE CÔTE NORD-MÉDITERRANÉEN

Romain Chailan[1,2,3,4] & Gwladys Toulemonde[2] & Frédéric Bouchette[3] & Anne Laurent[1]

[1] *LIRMM, UMR 5506, Université Montpellier 2 & CNRS, 34095 Montpellier.*
[2] *I3M, UMR 5149, Université Montpellier 2 & CNRS, 34095 Montpellier.*
[3] *Géosciences-M, UMR 5243, Université Montpellier 2 & CNRS, 34095 Montpellier.*
[4] *IBM FRANCE, Innovation Lab, Rue de la Vieille Poste CS81021, 34060 Montpellier.*
romain.chailan@gmail.com

Résumé. Les enjeux économiques et écologiques de l'évolution du trait de côte sont importants autant pour les acteurs privés que publics. Ce phénomène physique répond à des événements météorologiques extrêmes que nous modélisons par une approche statistique : les processus max-stables. Compte tenu des contraintes de données disponibles à long terme pour ce problème spatial, une méthodologie originale est mise en œuvre. Son but est de modéliser des forçages de houles extrêmes qui tiennent compte des structures de dépendances spatiales sous-jacentes. La première étape consiste à construire un jeu de données historiques de 51 ans à partir d'un modèle physique de vagues (WAVEWATCHIII®) et de données atmosphériques-océaniques régionales réanalysées. La seconde étape consiste à ajuster un modèle statistique max-stable sur ces simulations. Ce dernier nous permet d'extrapoler de l'information au-delà de la période de 51ans rejouée. À partir de ce modèle, des processus correspondants à des champs extrêmes de houles sont générés. Ces processus sont enfin utilisés pour alimenter un modèle physique d'évolution du trait de côte.

Mots-clés. Extrêmes Spatiaux, Processus Max-stables, Évolution du trait de côte, Modélisation de Vague historique.

Abstract. Both economic and ecologic stakes rely on the shoreline change, for both private and public actors. Shoreline change is a physical phenomenon driven by extreme meteorogical conditions. This study aims to statistically model such fluxes by using max-stable processes. Since there is a lack of observations both in time and space to apply a spatial extreme value modelling, an original approach is presented. Its goal is to seamlessly build spatial extreme wave-fluxes taking into account the underlying dependence structure. The first step consists in building a 51-year sea-states hindcast from a wave model (WAVEWATCHIII®) driven by regional atmospherical and ocean reanalyses. The second step consists in fitting a max-stable model on such simulations, allowing to extrapolate information to further timeseries than the hindcast period. From this model, processes corresponding to extremes swells are generated. These processes become inputs of a shoreline change model.

Keywords. Spatial Extremes, Max-stable Processes, Shoreline Evolution, Wave Hindcast Modelling.

1 Contexte

Le littoral et son exposition au risque d'inondation est un centre d'intérêt majeur pour les acteurs publics et privés (Rapport Ministériel (2012)). L'évolution du trait de côte – intersection terre/mer – peut avoir un impact important sur le développement économique, le tourisme, l'urbanisation et la protection de la biodiversité. Ce phénomène est largement gouverné par des événements météorologiques extrêmes comme les tempêtes, ouragans, typhons. Les vagues générées par ces événements extrêmes sont responsables d'apports et retraits de sédiments qui constituent le trait de côte. Dans le but d'analyser son évolution, l'ingénierie côtière a l'habitude d'utiliser des modèles physiques (Miller et Dean (2004)). Parmi les paramètres de ces modèles nous retrouvons comme facteur explicatif principal les houles extrêmes, observées par des houlographes. Une telle méthodologie se confronte alors à deux problèmes principaux lors de l'estimation de ces événements. Premièrement, la quantité des houlographes dans les zones d'études est faible voire inexistante ; deuxièmement les séries de temps ne sont que – au mieux – de quelques années. La présente étude centrée sur la zone nord-méditerranéenne fait face à ces problèmes en proposant une approche originale de création de champs de vagues annuels extrêmes. La méthodologie consiste dans un premier temps à se doter d'un jeu de données de hauteurs significatives de vagues (Hs) caractérisant les états de mer, dont les houles extrêmes. Cette donnée est reconstruite par modélisation numérique, ce qui nous permet d'obtenir spatialement de l'information sur une longue période historique : 51 ans. C'est à partir de ce jeu de données original que nous mettons en place l'analyse statistique. Il s'agit de mettre en application les récents travaux théoriques – (e.g. Smith (1990) ; Schlather (2002)) – et appliqués – (e.g. Blanchet (2011) ; Davison (2012)) – en théorie des valeurs extrêmes afin d'ajuster un modèle max-stable à ces données de vagues. Ce modèle est utilisé pour générer des champs de Hs extrêmes. Le produit de cette modélisation est donc directement utilisé par le modèle d'évolution du trait de côte. Ainsi il contribue à une meilleure étude de l'évolution du trait de côte en tenant compte d'une information spatiale et des structures de dépendances sous-jacentes.

Les sections sont organisées comme suit. La Section 2 présente la production d'un jeu de données décrivant l'état de la mer nord-méditerranée sur une période historique de 51 ans. Cette donnée est ensuite utilisée dans des modèles statistiques max-stables dont la Section 3 met en avant la justification et la construction. Elle présente finalement leur intérêt dans notre contexte.

2 La donnée de vague

En océanographie, un état de mer est communément mesuré et représenté par une variable nommée la hauteur significative de vague (Hs). Elle correspond à la moyenne du tiers des vagues les plus hautes enregistrées. Un autre formalisme la décrit équivalente à

quatre fois l'écart type de la hauteur des vagues. Il s'agit en tout cas d'une représentation de l'énergie contenue dans les vagues à partir de laquelle nous pouvons en déduire – entre autres – les régimes de houles. En pratique, Hs provient de diverses sources de données : les bouées/houlographes, les satellites et les modèles numériques. La densité des houlographes en mer est faible. Par ailleurs les séries de temps d'observations sont souvent courtes – typiquement inférieures à 10 ans. Raillard *et al.* (2013) se sont intéressés à utiliser les données satellites mais la couverture spatiale limitée et irrégulière dans le temps nous semble à ce jour restrictive. Dans cette étude nous nous concentrons sur la troisième source de données possible : la donnée simulée par modèle numérique.
Notre cas d'étude est le trait de côte du Golfe du Lion, dans la zone nord-méditerranéenne française. Il est acquis que les champs de vagues de cette zone résultent de la dynamique atmosphérique et océanique d'une zone régionale, ici nommée MEDNORD, et délimitée en ses extrémités par les points A(W5,5°;N31,0°) et B(E17,5°;N44.5°).

Pour calculer la hauteur significative de vague nous utilisons un modèle de vague de troisième génération (WAVEWATCHIII®) (Tolman (2014)) dans sa version la plus récente (Mars 2014). Ce modèle permet notamment de manipuler des grilles non-structurées. La grille de calcul peut ainsi être raffinée en fonction de diverses contraintes physiques comme le profil bathymétrique ou la définition arbitraire de polygones de raffinement. La grille de calcul utilisée dans cette étude comporte près de 100 000 points, avec une résolution à la côte du Golfe du Lion de 200m et une résolution au large qui s'étale sur des distances de l'ordre de 10km. La résolution temporelle de calcul est arbitrairement fixée à 6 heures. Le modèle de vague est majoritairement forcé par de la donnée atmosphérique – i.e. vents – et océanique – i.e. courants de surface. Afin de bien exprimer les phénomènes régionaux, nous utilisons des réanalyses régionales plus précises que les modèles globaux de références. ARPERA (Herrmann et Somot (2008)) rend compte des forçages atmosphériques tandis que NEMOMED8 (Beuvier *et al.* (2010)) fournit la donnée océanique. Ces derniers nous permettent de rejouer les hauteurs significatives de vagues spatialement sur la période historique 1961 - 2012.

Afin de rendre robuste notre méthodologie, nous passons par une étape d'ajustement du modèle de vague sur des données réellement observées par des houlographes. Bien qu'une approximation réside dans cette approche, nous faisons l'hypothèse que la donnée simulée correspond à la donnée observée. Compte tenu du contexte de notre étude, nous nous assurerons que cette hypothèse est raisonnable, y compris concernant les événements extrêmes.

Le jeu de données ainsi créé est utilisé pour ajuster un modèle max-stable, comme expliqué en Section 3.

3 De la définition à l'application : les processus max-stables

Au regard de récentes applications géostatistiques dans un contexte extrême (Blanchet (2011) ; Davison (2012) ; etc.), un formalisme mathématique étendant les théories univariée et multivariée des valeurs extrêmes semble parfaitement adapté au cadre spatial de la problématique d'évolution du trait de côte : les processus max-stables. À travers cette section, nous rappelons l'évolution de la théorie des valeurs extrêmes et proposons d'accompagner le lecteur dans l'application d'un modèle – et donc générateur de processus – max-stables.

Fisher et Tippett (1928), Gnedenko (1943) ont introduit la théorie des valeurs extrêmes (EVT) dans le cadre univarié en s'intéressant au comportement extrême d'une variable aléatoire X – e.g. la hauteur significative de vague quotidienne. Il s'agit de décrire la distribution limite de $M_n = \max\{X_1, \ldots, X_n\}$, compte tenu que les X_i sont indépendants et identiquement distribués (i.i.d.). En faisant l'hypothèse qu'il existe $(a_n)_{n\geq 0} > 0$ et $(b_n)_{n\geq 0}$ telles que $(M_n - b_n)/a_n$ converge vers une distribution non dégénérée $G(.)$, alors la théorie des valeurs extrêmes montre que cette distribution est de type GEV (Generalised Extreme Value) de fonction de répartition :

$$GEV_{\mu,\sigma,\xi}(x) = \exp\left\{ - \left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]_+^{-\frac{1}{\xi}} \right\}, \tag{1}$$

où μ est un paramètre de localisation, σ est un paramètre d'échelle et ξ est un paramètre de forme. Ici a_+ correspond au $\max(a, 0)$, $\mu \in \mathbb{R}$, $\sigma > 0$ et $\xi \in \mathbb{R}$. Cette forme généralisée regroupe trois familles de distributions : Weibull si $\xi < 0$, Gumbel si $\xi = 0$ et Fréchet si $\xi > 0$. Notons également la propriété essentielle d'une distribution GEV : elle est invariante à une transformation affine près à son passage au max. Il s'agit de la propriété de max-stabilité.

Afin de tenir compte des structures de dépendances dans les extrêmes entre différentes variables, la théorie des valeurs extrêmes multivariée (MEV) généralise l'approche univariée. Par exemple, plaçons nous dans le cadre bivarié et considérons une même variable aléatoire en deux sites s_1 et s_2. L'analyse multivariée s'intéresse au comportement limite du max par composante $M_n = (M_{n,s_1}, M_{n,s_2})$. Par construction les distributions marginales de ce vecteur aléatoire peuvent être approximées par des GEV. Par commodité et sans perte de généralité vis-à-vis de l'étude de la structure de dépendance, nous pouvons ramener chacune de ces composantes à des distributions de type Fréchet unitaire, i.e. de fonction de répartition $F(x) = e^{-1/x}$ pour $x > 0$.

Dès lors, si $(M_{n,s_1}/n, M_{n,s_2}/n)$ converge vers une distribution non dégénérée, alors Resnick (1987) montre que cette distribution est de la forme

$$G(x, y) = \exp(-V(x, y)), x > 0, y > 0,$$

où

$$V(x,y) = 2\int_0^1 \max\left(\frac{\omega}{x}, \frac{1-\omega}{y}\right) dH(\omega),$$

et H est une distribution définie sur $[0,1]$ vérifiant $\int_0^1 \omega dH(\omega) = 1/2$.
Comme dans le cas univarié, $G(.)$ est max-stable.

Une étude multivariée nous limite à l'information contenue aux seuls sites d'observations. Or notre étude porte sur un espace continu : le trait de côte. Dans ce contexte, De Haan (1984) introduit la notion de processus max-stable : un processus aléatoire $\{Z(x), x \in \chi\}$ est max-stable s'il existe $a_n(x) > 0$ et $b_n(x) \in \mathbb{R}$ définies sur χ telles que pour les répétitions i.i.d. $Z_i(x)$,

$$\left\{\max_{i=1,\ldots,n} \frac{Z_i(x) - b_n(x)}{a_n(x)}, x \in \chi\right\} \overset{\mathcal{L}}{=} \{Z(x), x \in \chi\}.$$

De Haan (1984) montre que le processus limite

$$\max_{i=1,\ldots,n} \frac{Y_i(x) - b_n(x)}{a_n(x)} \longrightarrow Z(x),\ x \in \chi,\ n \to \infty,$$

appartient à la classe des processus max-stables.

Il existe plusieurs modèles pour construire de tels processus. Bacro et Gaetan (2012) nous rappellent les deux principales approches de construction de processus max-stables. Smith (1990), Schlather (2002), etc., utilisent une forme déterministe d'événements se déplaçant dans l'espace aléatoirement. Schlather (2002), Kabluchko (2009), etc., se reposent sur des événements à forme stochastique mais conservent une même structure de dépendance spatiale.

À titre indicatif nous présentons le processus de Smith (1990). Considérons $\{(\xi_i, s_i), i \geq 1\}$ un processus de Poisson sur $(0,\infty) \times S$ et de mesure d'intensité $\xi^{-2} d\xi \times \nu(ds)$, où S est un espace arbitraire et ν une mesure définie sur S. Par ailleurs considérons $\{f(s,x), s \in S, x \in \chi\}$ une fonction non-négative qui satisfait $\int_S f(s,x)\nu(ds) = 1$, pour tout $x \in \chi$. Alors le processus max-stable communément appelé le processus tempête est défini par

$$Z(x) = \max_{i=1,\ldots,n} \xi_i f(s_i, x), x \in \chi. \tag{2}$$

Intuitivement S est vu comme l'espace des centres des tempêtes et ν leur distribution. Chaque ξ_i représente l'intensité de la tempête et $\xi_i f(s_i, x)$ la quantité totale de pluie tombée au cours de la tempête centrée sur s_i et d'intensité ξ_i. Finalement l'opérateur maximum permet d'obtenir la quantité de pluie maximale tombée après n tempêtes indépendantes.

Notre étude fait l'hypothèse que les champs de Hs extrêmes sont caractérisés par un modèle max-stable. Différents modèles sont ajustés en estimant leurs paramètres par vraisemblance composite puis comparés par des critères de sélection connus – e.g. CLIC. Le modèle sélectionné a l'intérêt de pouvoir extrapoler de l'information sur les champs Hs extrêmes au-delà de la période d'observation de 51 ans, de façon spatiale et en s'appuyant sur une structure de dépendance dans les extrêmes. Une base de scénarios types est donc simulée (Ribatet (2011)) par la génération de 200 processus issus du modèle max-stable. Nous alimentons enfin le modèle physique d'évolution du trait de côte à long terme par ces processus simulés et caractérisons ainsi les quantités d'intérêt associées.

Bibliographie

[1] Bacro, J. N., & Gaetan, C. (2012), *A review on spatial extreme modelling. In Advances and Challenges in Space-time Modelling of Natural Events.* Springer Berlin Heidelberg, 103–124.
[2] Beuvier, J., Sevault, . . . & Somot, S. (2010), *Modeling the Mediterranean Sea interannual variability during 1961–2000 : focus on the Eastern Mediterranean Transient*, Journal of Geophysical Research : Oceans (1978–2012), 115(C8).
[3] Blanchet, J., & Davison, A. C. (2011), *Spatial modeling of extreme snow depth*, The Annals of Applied Statistics, 5(3), 1699-1725.
[4] Davison, A. C., & Gholamrezaee, M. M. (2012), *Geostatistics of extremes*, Proceedings of the Royal Society A : Mathematical, Physical and Engineering Science, 468(2138), 581-608.
[5] De Haan, L. (1984), *A spectral representation for max-stable processes*, The Annals of Probability, 12(4), 1194-1204.
[6] Fisher, R. A., & Tippett, L. H. C. (1928), *Limiting forms of the frequency distribution of the largest or smallest member of a sample*, In Mathematical Proceedings of the Cambridge Philosophical Society (Vol. 24, No. 02, pp. 180-190). Cambridge University Press.
[7] Gnedenko, B. (1943), *Sur la distribution limite du terme maximum d'une serie aleatoire*, Annals of Mathematics, 423-453.
[8] Herrmann, M. J., & Somot, S. (2008), *Relevance of ERA40 dynamical downscaling for modeling deep convection in the Mediterranean Sea*, Geophysical Research Letters, 35(4).
[9] Kabluchko, Z., Schlather, M., & De Haan, L. (2009), *Stationary max-stable fields associated to negative definite functions*, The Annals of Probability, 2042-2065.
[10] Miller, J., & Dean, R. (2004), *A simple new shoreline change model*, Coastal Engineering, 51(7), 531-556.
[11] Ministère de l'Ecologie, du Developpement Durable et de l'Energie (2012), *Stratégie nationale de gestion intégrée du trait de côte : vers la relocalisation des activités et des biens*, Rapport Ministériel, repéré à `http://www.developpement-durable.gouv.fr/Strategie-nationale-de-gestion.html`.
[12] Raillard, N., Ailliot, P., & Yao, J. F. (2013), Modeling extreme values of processes observed at irregular time steps : application to significant wave height.
[13] Resnick, S. I. (1987), *Extreme Values, Regular Variation, and Point Processes.*
[14] Ribatet, M. (2011), *SpatialExtremes : modelling spatial extremes*, R package version, 1-8.
[15] Schlather, M. (2002), *Models for stationary max-stable random fields*, Extremes, 5(1), 33-44.
[16] Smith, R. L. (1990), *Max-stable processes and spatial extremes*, Unpublished manuscript, Univer.
[17] Tolman, H. L. (2014), *User manual and system documentation of WAVEWATCH III TM version 4.18*, Technical note, NOAA/NWS/NCEP/MMAB Contribution, (316).