



PROPOSITION DE SUJET DE THESE

## **ANALYSE STATISTIQUE REGIONALE DES SURCOTES MARINES EXTREMES : PRISE EN COMPTE DES INCERTITUDES ET DES EVENEMENTS HISTORIQUES EXTREMES**

**Laboratoires d'accueil** : EDF R&D Laboratoire National d'Hydraulique et Environnement (LNHE, Chatou (France)) et le LHSV : Laboratoire d'Hydraulique Saint-Venant (Ecole des Ponts – EDF R&D, Chatou (France))

**Encadrants** : Damien Violeau (EDF-R&D LNHE-LHSV), Michel Benoit (EDF-R&D LNHE-LHSV), Anne Dutfoy (EDF-R&D PERICLES), Marc Andreevsky (EDF-R&D LNHE-LHSV).

**Directeur de thèse** : Damien Violeau. **Co-directeur de thèse** : Michel Benoit.

**Durée** : 3 ans (2024-2027)

### **1 Introduction et contexte général**

Cette thèse s'inscrit dans l'objectif général d'améliorer les estimations des niveaux marins hauts les plus extrêmes associés à de grandes périodes de retour en zone côtière afin de maîtriser les risques de submersion côtière et d'inondations éventuelles d'infrastructures et de zones côtières en cas d'évènement de tempête extrême. Ces niveaux marins extrêmes résultent de la conjonction d'un niveau de marée élevé, notamment au moment de la pleine mer (PM) et d'une surélévation du niveau moyen de la mer due à des effets atmosphériques, appelé « surcote météorologique ».

Parmi ces surcotes météorologiques observables, on trouve les « surcote de pleine mer » qui sont la différence, en un lieu donné, entre la hauteur de PM observée et la hauteur de PM prédite, appelée marée de pleine mer astronomique. Les instants d'occurrence de ces deux niveaux peuvent être décalés dans le temps (voir Figure 1). La surcote / décote instantanée est la différence, à un instant  $t$ , entre la hauteur d'eau observée et la hauteur d'eau prédite à cet instant. La surcote / décote a principalement une origine météorologique : elle est générée, lors du passage de dépressions ou d'anticyclones, par les variations de pression atmosphérique et par l'action du vent à la surface de la mer (le « setup » dû au déferlement des vagues peut avoir aussi une influence).

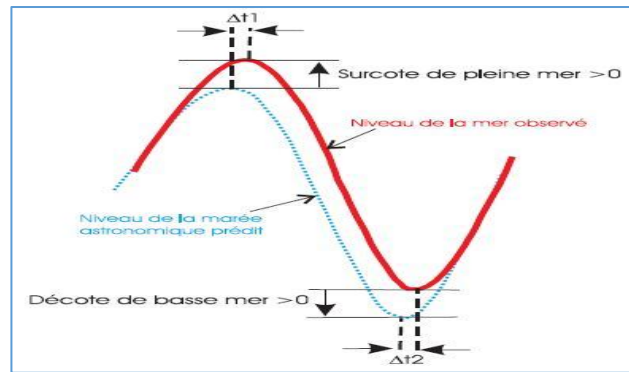


Figure 1: illustration d'une surcote de pleine mer (« skew surge » en anglais). En rouge, la marée observée en un lieu donné (par exemple à l'aide d'un marégraphe) en fonction du temps, en bleu la marée astronomique (prédictible) au même lieu.

Un niveau marin de PM  $N_{pm}$  est donc la résultante de la somme de 2 termes :

$$N_{pm} = N_a + S_{pm} \quad (1)$$

où :

- $N_a$  représente le niveau de PM de la marée astronomique, qui est prédictible et dépend de la position relative des planètes du système solaire, de la Lune et du Soleil par rapport à la Terre.
- $S_{pm}$  représente la surcote de pleine mer qui peut être positive ou négative, considérée ici comme une variable aléatoire (comme beaucoup de variables météorologiques).

Lors d'une tempête, le vent en général venant du large pousse les masses d'eau vers la côte engendrant ainsi une élévation du niveau marin par rapport à son niveau prédit et, par ailleurs, la dépression atmosphérique qui accompagne la tempête applique une force moindre de l'atmosphère sur l'océan et augmente ainsi encore le niveau marin. Lors de très grandes tempêtes (comme celle de 1953<sup>1</sup>, voir Figure 2), le risque de submersion d'une partie des côtes est donc possible.

<sup>1</sup> Voir par exemple ici : <http://tempetes.meteo.fr/Tempete-des-31-janvier-et-1er-fevrier-1953.html>, extrait : « L'arrivée de la dépression dans l'axe de la mer du Nord associée à la marée haute provoque des submersions sur de nombreux littoraux. La mer, poussée par le vent de nord à nord-ouest, s'engouffre dans l'entonnoir naturel formé par la pointe sud de la mer du Nord entre les Pays-Bas et le sud-est de l'Angleterre. À Dunkerque, le niveau des eaux s'élève de 2,40 m (atteignant à 10 cm près le niveau de déversement dans les canaux à l'intérieur des terres). Toujours à Dunkerque, deux brèches de 50 m s'ouvrent dans la digue Est. La submersion atteint la région de Calais avec des terrains et des maisons inondées. Aucune victime n'est à déplorer en France. La pluie reste faible pendant cette tempête.

Chez nos voisins la surcote (dépassement du niveau normal de la mer provoqué par la tempête) atteint 2,5 m dans le sud-est de l'Angleterre et plus de 3 m aux Pays-Bas, provoquant des inondations parfois dramatiques. Aux Pays-Bas, deux submersions se produisent à chaque marée haute pendant la tempête, dans la nuit du samedi 31 au dimanche 1er et le dimanche après-midi. Les digues s'étant déjà rompues, la submersion du dimanche après-midi provoque le plus de dégâts. »

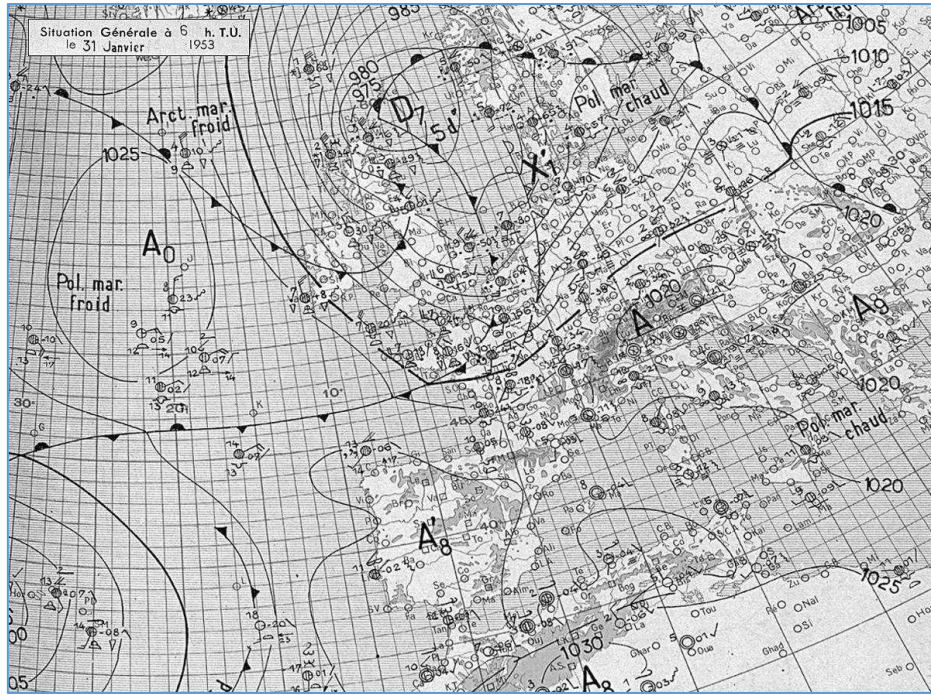


Figure 2 : carte d'analyse du 31 janvier 1953 à 07 h locales, avec le centre dépressionnaire principal (D7) situé dans la partie nord de la Mer du Nord. On observe une dépression secondaire qui se creuse au sud de l'Islande. Sources : <http://tempetes.meteo.fr/Tempete-des-31-janvier-et-1er-fevrier-1953.html>.

## 2 Contexte de la recherche et objectifs de la thèse

Les principes et règles appliqués à la protection des centrales nucléaires de bord de mer contre le risque d'inondation marine externe ont été complétés à plusieurs reprises depuis leur conception. Ils tiennent compte de l'évolution de l'état de l'art sur ce thème et du retour d'expérience à la suite d'un certain nombre d'évènements, par exemple les tempêtes Lothar et Marin de fin décembre 1999 ayant provoqué une inondation partielle de la centrale du Blayais dans l'estuaire de la Gironde, ou encore le tsunami à Fukushima au Japon (évènement sismique et non météorologique) en 2011 (voir Figure 3 ci-dessous).



Figure 3 : à gauche, la centrale de Blayais après la tempête de 1999. A droite, la centrale de Fukushima au moment du début de l'impact du tsunami de 2011.

Le guide de l'Autorité de Sécurité Nucléaire n°13 (ASN, 2013) sur les protections des installations nucléaires de base contre les inondations externes, définit les caractéristiques des tempêtes dimensionnantes pour lesquelles les centrales de bord de mer doivent être résilientes. Ces caractéristiques définissent plusieurs variables dont le niveau marin  $N_{pm}$  de PM qui est atteint lors de ces tempêtes.

Ce niveau marin  $N_{pm}$  de PM est défini comme la somme des 3 termes suivants :

$$N_{pm} = PHMA + BS_{70\%}(S_{1000}) + M \quad (2)$$

Où :

- 1) PHMA est valeur de la Plus Haute Mer Astronomique au port de référence d'une centrale de bord de mer particulière. Cette valeur est donnée par le Service Hydrographique et Océanographique de la Marine (le SHOM, voir ici : <https://www.shom.fr/>).
- 2)  $BS_{70\%}(S_{1000})$  est l'estimation statistique de la borne supérieure à 70% de la surcote de pleine mer de période de retour 1000 ans au même port de référence.
- 3) M est une marge pour prendre en compte l'évolution du niveau marin (l'« eustatisme », incluant entre autre les effets potentiels du changement climatique) entre 2 réexamens réglementaires de sûreté.

Pour garantir la sûreté nucléaire vis-à-vis des inondations pour les centrales de bord de mer, le deuxième terme «  $BS_{70\%}(S_{1000})$  » doit être estimé au mieux. Une estimation précise assure, en effet, une protection réglementaire et optimale des installations tout en minimisant le coût de construction des ouvrages de protection. Cette estimation est un des éléments qui permet aux ingénieurs de dimensionner les digues et les protections nécessaires d'une plateforme nucléaire ou de vérifier que les protections existantes sont suffisantes pour garantir la sûreté. Ces vérifications ou re-vérifications se font réglementairement tous les 10 ans environ.

Néanmoins, l'estimation précise d'une surcote de PM particulièrement extrême, c'est-à-dire associée à une période de retour très grande (par exemple 1000 ans), est un exercice d'extrapolation statistique difficile, étant donnée la durée limitée des chroniques d'enregistrement, de l'ordre de quelques dizaines d'années en général. Dans ce cadre, l'analyse fréquentielle régionale des surcotes marines de PM extrêmes est une solution élégante et scientifiquement robuste, car l'estimation des surcotes extrêmes pour un site donné est réalisée à partir de statistiques sur un grand nombre d'observations issues de plusieurs marégraphes d'une région identifiée comme statistiquement homogène (c'est-à-dire une région constituée de plusieurs systèmes d'observation permettant de générer des séries de surcotes que l'on peut considérer, à la suite de différents tests, comme issues de processus aléatoires associés à des lois de probabilité similaires). Le principe de l'analyse fréquentielle régionale remonte aux années 60 et a été proposé par Dalrymple (1960) et l'approche a été développée par Hosking et Wallis (1997). Ce type d'analyse se distingue en particulier d'une analyse locale dont les calculs sont effectués à partir d'observations issues seulement d'un seul site.

A l'heure actuelle, la R&D d'EDF réalise des analyses statistiques régionales pour estimer les surcotes de pleines mers associées à de grandes périodes de retour. La méthode utilisée a été développée à la R&D d'EDF depuis 2011 (Bernardara, 2011) et se base maintenant sur les thèses de Jérôme Weiss (2014b) et de Roberto Frau (2018). Parmi les données de surcotes de PM récoltées sur lesquelles les statistiques sont effectuées, on trouve très majoritairement les surcotes enregistrées régulièrement et issues de l'analyse des enregistrements marégraphiques, ainsi que ce qu'on appelle, les « surcotes historiques », en quantité plus réduite, mais dont les valeurs en général grandes sont utiles à la sûreté : elles ont en effet tendance à « pousser les extrapolations statistiques vers le haut ». Le guide ASN 13 avec lequel EDF doit être en conformité, demande d'ailleurs de les prendre en compte (p. 29/44), comme l'indique cet extrait : « *Les valeurs des surcotes historiques sont inventoriées et prises en compte dans l'étude statistique.* ».

Il faut savoir que les données « historiques » se distinguent des données usuellement exploitées que sont les données « systématiques » enregistrées par les marégraphes : les données historiques sont, en



effet, des données reconstituées à partir de diverses sources d'information (archives, registres paroissiaux, témoignages, presse..., (Regnier, 2017)) dans le but de produire une valeur quantitative d'un évènement non directement mesuré. Ces surcotes sont en général calculées (Athimon, 2022) à partir d'une indication sur le niveau marin historique atteint lors d'une tempête d'un jour particulier et des rétro-prédictions de marée pour ce jour et au même lieu (voir l'exemple de la Figure 4). Un certain nombre d'hypothèses sont, par ailleurs, posées pour le calcul des surcotes historiques de PM (Athimon, 2022).

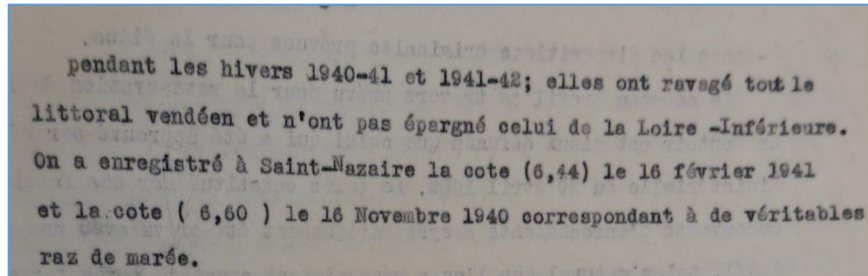


Figure 4 : archives Départementales de Loire-Atlantique (référence : 365 W 124). Photo présentant un extrait d'un rapport de remise en état de digues daté du 20/10/1943 mentionnant les niveaux d'eau observés à Saint-Nazaire lors du passage des tempêtes avec submersions du 16 novembre 1940 et 16 février 1941. N.B : Ces niveaux d'eau sont confirmés par d'autres documents exhumés des archives.

Dans la littérature, plusieurs études récentes utilisent les informations historiques pour améliorer les estimations statistiques (Bulteau, 2015 ; Frau, 2018 ; Hamdi, 2015 ; Hamdi, 2018 ; Saint-Criq, 2022). Néanmoins, des incertitudes existent concernant les surcotes de PM, qu'elles soient historiques (Athimon, 2022) ou même systématiques :

- Pour les surcotes systématiques, celles-ci sont calculées grâce, entre autres, aux rétro-prédictions (qui dépendent aussi de l'estimation de l'eustatisme<sup>2</sup>), grandeur pour laquelle des incertitudes existent selon les modes de calculs (André, 2020). On peut aussi, par ailleurs, considérer qu'une incertitude sur les mesures de niveaux marins existe, avec un impact sur la surcote.
- Pour les surcotes historiques, celles-ci sont calculées grâce à un certain nombre d'hypothèses également sources d'incertitudes (Athimon, 2021).

Ces incertitudes, qui ne sont pas prises en compte dans les estimations statistiques actuelles réalisées par EDF, peuvent l'être via une modification de la vraisemblance de l'analyse régionale, à partir de la remarque suivante : si  $x_0$  est l'observation d'une variable  $x$  que l'on considère comme aléatoire, alors si l'estimation de  $x_0$  est incertaine et comprise entre deux valeurs  $x_1$  et  $x_2$ , et si  $g(x)$  est la densité de probabilité associée à  $x$ , la probabilité  $p(x_0)$  que  $x_0$  soit comprise entre les valeurs  $x_1$  et  $x_2$  vaut :

$$p(x_0 \in [x_1, x_2]) = \int_{x_1}^{x_2} g(x) dx. \quad (3)$$

On retrouve, entre autres, un terme similaire à l'équation (3) dans l'expression de la vraisemblance d'une analyse locale réalisée par Hamdi (2015) lorsque les données historiques sont définies sur un intervalle de valeur possible. En effet, une vraisemblance étant le produit des probabilités des observations réalisées (fonction des lois de probabilité dont les paramètres sont à optimiser), les observations incertaines, c'est-à-dire comprises entre 2 valeurs, seront associées à une expression de type (3).

<sup>2</sup> Evolution du niveau marin moyen à long terme.

Par ailleurs, des questions se posent concernant la durée équivalente d'observation que l'on peut associer aux surcotes historiques. Cette grandeur peut être définie comme une estimation crédible de la plus grande période possible qui aurait pu permettre, à partir d'un marégraphe, l'observation, au-dessus d'un seuil de sélection d'extrêmes, uniquement de ces valeurs historiques (et aucune autre). Cette durée équivalente d'observation, qui peut être plus ou moins conservative selon les hypothèses utilisées, comme celle de Frau (2018) très conservative ou encore celles de Hamdi (2015) moins conservative, a aussi un impact sur l'écriture de la vraisemblance à optimiser de l'analyse fréquentielle régionale.

Enfin, il est possible de construire une vraisemblance  $V$  à optimiser de l'analyse régionale en incluant, non seulement les surcotes de PM, mais aussi des niveaux marins historiques de PM. On utilise pour cela le fait que la probabilité associée aux processus de niveaux marins de PM peut être calculée à partir de la loi de probabilité des surcotes de PM convoluée par l'histogramme de la densité des marées de PM théoriques (voir par exemple Andreevsky (2014)). Par ailleurs, des incertitudes existent aussi sur les valeurs de ces niveaux marins historiques de PM et sont fonction des hypothèses qui permettent leur calcul.

On peut alors définir, pour l'analyse régionale, cette vraisemblance  $V$  comme le produit de 3 différents termes concernant chacun 3 types de grandeurs :

- Un terme  $V_1$  qui concerne la probabilité d'observer les surcotes systématiques et qui prend en compte les incertitudes de ces surcotes.
- Un terme  $V_2$  pour la probabilité d'observer les surcotes historiques, avec aussi la prise en compte des incertitudes.
- Et un terme  $V_3$  pour la probabilité d'observer les niveaux marins historiques qui prend en compte, comme les 2 premier termes, les incertitudes associées à ces niveaux.

La vraisemblance à optimiser de l'analyse régionale a alors une expression de la forme suivante :

$$V(\Theta) = V_1(\Theta) \times V_2(\Theta) \times V_3(\Theta) \quad (4)$$

où  $\Theta$  est le vecteur constitué des paramètres de la loi de probabilité des surcotes de PM que l'on cherche.

L'étude de Saint-Criq (2022) a déjà utilisé à la fois les surcotes systématiques et les niveaux marins historiques, mais sans prendre en compte les incertitudes associées. De plus, les hypothèses sur les durées d'observation équivalentes associées aux informations historiques ne couvrent pas tous les niveaux de conservatisme possibles.

**Dans ce contexte, l'objectif de cette thèse est de développer une nouvelle méthodologie pour la caractérisation des surcotes de PM de période de retour élevée en utilisant l'analyse régionale. Cette méthodologie doit permettre de prendre en compte les diverses incertitudes associées aux données, différentes hypothèses sur les durées d'observation associées aux surcotes historiques et inclure enfin les données de niveaux de mer historiques dans une seule vraisemblance à optimiser, et ceci, afin d'estimer les paramètres de la loi régionale.** Le but final étant une estimation toujours plus fiable de la borne supérieure à 70% de la surcote de période de retour 1000 ans pour la sûreté de chaque site nucléaire de bord de mer sur le littoral français.

### 3 Etapes indicatives de travail pour le projet de thèse

On peut envisager que le travail se composera des étapes suivantes :

- Etape 1 : étude bibliographique, avec, entre autres, les thèses de Weiss (2014), Frau (2018) et Saint-Criq (2022), permettant de prendre en main le sujet et comprendre les codes R existants : ce code permet de réaliser une analyse régionale selon Frau (2018).
- Etape 2 : développement d'un outil numérique (ex. scripts Python ou R) permettant l'extension de la méthode existante pour prendre en compte les incertitudes, avoir différentes approches pour estimer la durée d'observation des événements historiques et inclure les niveaux marins historiques aux échantillons de surcotes.
- Etape 3 : des tests et une validation seront réalisés sur les surcotes marines de PM, ainsi qu'une comparaison des résultats selon les méthodes utilisées. L'étape 3 finira par la rédaction du manuscrit.

## 4 Compétences recherchées

Statistiques des événements extrêmes naturels, océanographie physique côtière (marée, surcotes-décotes, etc.), connaissance de Linux et Windows, R et Python, probabilités et statistiques, anglais, autonomie et curiosité indispensable.

## 5 Contacts

Marc Andreevsky, Anne Dutfoy, Michel Benoit et Damien Violeau.

E-mail : [marc.andreevsky@edf.fr](mailto:marc.andreevsky@edf.fr); [anne.dutfoy@edf.fr](mailto:anne.dutfoy@edf.fr); [michel.benoit@edf.fr](mailto:michel.benoit@edf.fr); [damien.violeau@edf.fr](mailto:damien.violeau@edf.fr)

La thèse se déroulera sur le site EDF Lab Paris-Chatou situé sur l'île des Impressionnistes à Chatou (78) : EDF-R&D, 6, quai Watier, 78400 CHATOU. Accès possible, entre autres, en RER A.

## 6 Références

André, G., Ferret, Y., Pouvreau, N., 2020 : « Prédiction de marée dans le passé pour l'estimation des surcotes historiques ». *Séminaire du Groupe de Travail Tempêtes et Submersions Historiques, BRGM*.

Andreevsky, M., Hamdi, Y., Griplet, S., Bernardara, P., Frau, R., 2020 : « Regional frequency analysis of extreme storm surges using the extremogram approach », *NHESS*, 20, p. 1705-1717.

Andreevsky, M., Kergadallan, X., Bernardara, P., Benoit, M., Gaufres, P., Trmal, C., 2014 : « Comparison of different methods for estimating extreme levels for high and low tidal range site », *La Houille Blanche*, 100:4, 26-36, DOI: 10.1051/lhb/2014035.

ASN, 2013 : « Protection des installations nucléaires de base contre les inondations externes », *Guide de l'ASN n°13*, Version du 08/01/2013.

Athimon E., Giloy N., Sauzeau T., Andreevsky M., Frau R., 2022 : « Quantification of Historical Skew Surges: Challenges and Methods », *Advances in Hydroinformatics Models for Complex and Global Water Issues – Practices and Expectations*, SimHydro 2021, special issue, p. 159-174.

Bernardara, P., Andreevsky, M., Benoit, M., 2011 : « Application of the Regional Frequency Analysis to the estimation of extreme storm surges », *J. Geophys. Res.*, 116, C02008, 1–11.

Bulteau, T., Idier, D., Lambert, J., Garcin, M., 2015 : « How historical information can improve estimation and prediction of extreme coastal water level: application to the Xynthia event at La Rochelle (France) », *NHESS*, 15, p. 1135-1147.

Darlymple, T., 1960 : « Flood Frequency Analysis, US Geological Survey », *Water Supply Paper*, 1543-A.

Hosking, J. R. M., Wallis, J. R., 1997 : « Regional Frequency Analysis. An approach based on Lmoments », *Cambridge University Press*.

Frau, R., 2018 : « Utilisation des données historiques dans l'analyse régionale des aléas marins extrêmes : la méthode FAB », *Université Paris-Est, thèse de doctorat en génie côtier, Paris*.

Frau, R., Andreevsky, M., Bernardara, P., 2018 : « The use of historical information for regional frequency analysis of extreme skew surge », *NHESS*, 18, p. 949-962.

Giloy, N., Hamdi, Y., Bardet, L., Garnier, E., Duluc, C-M., 2018 : « Quantifying historic skew surges: an example for the Dunkirk Area, France », *Natural Hazards*, 1-25.

Hamdi, Y., Garnier, E., Giloy, N., Duluc, C.-M., Rebour, V., 2018 : « Analysis of the risk associated with coastal flooding hazards: a new historical extreme storm surges dataset for Dunkirk », *France, Nat. Hazards Earth Syst. Sci.*, 18, 3383–3402, <https://doi.org/10.5194/nhess-18-3383-2018>.

Hamdi, Y., Bardet, L., Duluc C-M., Rebour, V., 2015 : « Use of historical information in extreme-surge frequency estimation: the case of marine flooding on the La Rochelle site in France », *NHESS*, 15, p. 1515-1531.

Le Roy Ladurie, E., 2004-2009 : « Histoire humaine et comparée du climat », *3 tomes, Fayard, Paris*.

Regnier, F., Frau, R., Andreevsky, M., 2017 : « Increasing skew surge database with the collection of archaeological data », *EVAN 2017, Southampton*.

Saint Cricq, L., 2022 : « Estimation des niveaux marins extrêmes en utilisant de l'information régionale et historique », *thèse*.

Weiss, J., Bernardara, P., Benoit, M., 2014a : « Modelling intersite dependence for regional frequency analysis of extreme marine events », *Water Resource Research*, 50(7), 5926–5940.

Weiss, J., 2014b : « Analyse régionale des aléas maritimes extrêmes », *Mécanique des fluides [physics.class-ph], Université Paris-Est*, <https://pastel.archives-ouvertes.fr/tel-01127291>.