



HAL
open science

Fiabilité d'une représentation " par événements " de la climatologie de vagues et de courants en Afrique de l'Ouest

Agbéko Komlan Kpogo-Nuwoklo

► **To cite this version:**

Agbéko Komlan Kpogo-Nuwoklo. Fiabilité d'une représentation " par événements " de la climatologie de vagues et de courants en Afrique de l'Ouest. Océanographie. Université de Bretagne occidentale - Brest, 2015. Français. NNT : 2015BRES0053 . tel-01975422

HAL Id: tel-01975422

<https://theses.hal.science/tel-01975422>

Submitted on 9 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



université de bretagne
occidentale



THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE

sous le sceau de l'Université européenne de Bretagne

pour obtenir le titre de

DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Mention : Océanographie Physique

École Doctorale des Sciences de la Mer (EDSM)

présentée par

Agbéko K. KPOGO-NUWOKLO

Préparée à GSM

(Comportement des Structures en Mer)

IFREMER

Fiabilité d'une représentation « par événements » de la climatologie de vagues en Afrique de l'Ouest

Thèse soutenue le 04 Novembre 2015

devant le jury composé de :

Fabrice Ardhuin

Professeur des Universités, UBO-Ifremer / président du jury

Philippe Sergent

Directeur Scientifique, CEREMA / rapporteur

Gregorio Iglesias Rodriguez

Professeur des Universités, Plymouth University / rapporteur

Valérie Quiniou-Ramus

Metoccean Specialist, TOTAL / examinateur

Pierre Ailliot

Maitre de Conférences, UBO / examinateur

Sabine Arnault

Professeur des Universités, UPMC / directrice de thèse

Christophe Maisondieu

Cadre de recherche, Ifremer / encadrant

A ma mère



Remerciements

Je tiens tout d'abord à remercier les deux rapporteurs de ma thèse, le Docteur Philippe Sergent (Directeur scientifique, CEREMA) et le Professeur Gregorio Iglesias Rodriguez (Plymouth University), qui ont pris le temps de relire et de faire de commentaires sur mon travail, malgré des délais assez courts, et en dépit de leurs nombreuses obligations. Merci pour vos précieux conseils et suggestions. Mes remerciements s'adressent également à Fabrice Ardhuin qui m'a fait l'honneur de présider le jury de thèse.

Cette thèse n'aurait pas probablement abouti sans TOTAL qui a financé ces travaux. Je tiens sincèrement à remercier particulièrement Valérie Quiniou-Ramus qui a été l'initiatrice de ce projet et n'a ménagé aucun effort pour suivre de près ces travaux.

À mes encadrants, Sabine Arnault, Michel Olagnon, Christophe Maisondieu et Zakoua Guédé, je tiens chaleureusement à vous témoigner toute ma gratitude pour votre disponibilité. Je suis ravi et fier d'avoir eu la chance de travailler avec vous. L'engagement et la rigueur dont vous avez fait preuve resteront pour moi un exemple. Je tiens à témoigner toute ma gratitude à Pierre Ailliot qui m'a constamment aidé, écouté et corrigé pendant ces trois années, son aide fut précieuse. Je remercie également tous les collègues et membres du Laboratoire Comportement des structures en Mer de l'IFREMER où j'ai passé ces trois années.

C'est le lieu de remercier ma famille qui de loin m'a apporté leur indéfectible soutien. Je pense à ma mère Akoua et à mes sœurs Épiphanie, Marie, Éssé et Aimée. Le soutien amical a également été important et je pense à Tatiana, Barthélemy, Clarisse et Casimir qui ont pris le temps de relire les chapitres de cette thèse. Je pense également à Séna, Thierry, Anani, Landy, Coco, Martin et tous les amis dont leur soutien m'a été précieux.

Last but not least, j'ai une pensée particulière pour Alphonsine, mon amie et probablement mon plus grand soutien pour ses encouragements tout au long de ces trois années.



Résumé

La connaissance des statistiques à long-terme des états de mer est primordiale dans de nombreuses applications en génie océanique et côtier : conception et dimensionnement de structures marines (plateforme pétrolière, digue, convertisseur d'énergie de vagues), gestion des zones côtières et des opérations en mer, étude de la morphodynamique côtière, récupération de l'énergie des vagues, etc. Pour ces applications, il est essentiel de pouvoir estimer avec précision la distribution (i.e. les probabilités d'occurrence) des conditions de mer sur un site d'intérêt. Un moyen qu'on peut utiliser pour estimer ces distributions est la modélisation stochastique. Elle permet de dépasser les limites, notamment en durée, imposées par les observations.

Dans ces travaux, nous proposons une nouvelle approche stochastique pour estimer une climatologie de vagues en Afrique de l'Ouest (du large de la Côte d'Ivoire au large de la Namibie), s'appuyant sur une représentation par événements des données d'états de mer. Un événement désigne un système de vagues (houle ou mer du vent) en évolution au cours du temps, observable pendant une durée significative et que l'on peut relier à un unique phénomène météorologique source (e.g. dépressions, tempêtes, etc.). En effet, les états de mer résultent fréquemment de la combinaison de plusieurs systèmes de vagues d'intensités, de fréquences et de directions différentes. En Afrique de l'Ouest, en plus de la mer du vent, deux voire trois systèmes de houles peuvent coexister. Il existe des méthodes dites de partitionnement qui permettent de décomposer un spectre caractéristique d'un état de mer en plusieurs systèmes de vagues. Par un suivi temporel de ces systèmes de vagues, il est possible d'identifier des événements de houle et de mer du vent. Nous proposons de pallier les limitations classiques dans la cohérence et l'identification des systèmes de vagues par la représentation par événements des données d'états de mers. Cette représentation permet de les structurer de manière simple et avec une base physique. Elle fait intervenir moins de paramètres et des notions de dépendance entre ceux-ci plus raisonnées. Elle permet également de reproduire la cohérence temporelle des systèmes de vagues.

Cette approche est essentiellement utilisée ici pour la modélisation stochastique de la climatologie des houles, qui contribuent à plus de 80% de l'énergie totale des états de mer en Afrique de l'Ouest. Ainsi, en utilisant comme éléments de base des événements de houle préalablement extraits à partir d'une série temporelle de spectres d'états de mer, nous avons construit un générateur stochastique d'événements de houle et reconstitué avec

succès des climatologies de houles sur des durées de longueurs arbitraires, utilisables avec confiance pour des applications en génie océanique et côtier. La démarche adoptée pour atteindre ce but peut se décomposer suivant trois étapes.

Nous avons d'abord développé une nouvelle méthode : Time-domain Expansion of Spectral Partitioning techniques (TESP)), basée sur l'algorithme de la ligne de partage des eaux dans l'espace en trois dimensions (temps, fréquence, direction), pour extraire des événements de houle à partir d'une série temporelle de spectres directionnels d'états de mer. la méthode TESP utilise la cohérence temporelle des systèmes de vagues dans le partitionnement et permet d'extraire des événements cohérents avec les considérations physiques qui les déterminent. TESP garantit l'absence de coupure des événements extraits et permet la réduction du nombre de critères empiriques, souvent nombreux, utilisés dans les méthodes existantes, réduisant de ce fait le temps de calcul. Nous avons également montré que TESP n'est pas uniquement adaptée aux données de l'Afrique de l'Ouest, mais qu'elle peut aussi s'utiliser sur des spectres issus d'une autre zone géographique ayant une dynamique différente, qu'ils proviennent des observations ou des modèles numériques de vagues.

Nous avons ensuite développé un modèle (Swell Event Model (SEM)) pour représenter chacun des événements de houle par un nombre réduit de paramètres (sept), qui sont par la suite utilisés comme variables de bases du générateur stochastique. Le modèle SEM a été validé, d'une part, sur la distribution jointe de la hauteur significative et de la période d'énergie et, d'autre part, sur un cas d'application : l'estimation de la puissance extraite par un convertisseur d'énergie de vagues. Les résultats ont montré que les événements de houle observés en Afrique de l'Ouest peuvent être modélisés par le SEM sans différence notable dans la distribution jointe de la hauteur significative et de la période d'énergie pour les applications inhérentes au génie océanique.

Enfin, nous avons construit le générateur stochastique d'événements de houles en modélisant d'abord la distribution jointe des paramètres du modèle SEM, puis le processus d'occurrence des événements de houle. Ce générateur permet la simulation d'événements de houles individuels et la reconstitution de climatologies sur des durées de longueurs arbitraires. Grâce à des tests statistiques adaptés, nous avons validé le générateur à chaque phase de sa conception. Nous avons également confronté les séries temporelles de houles simulées par le générateur avec celles de références, notamment par comparaison de la distribution des hauteurs significatives et des durées de persistance souvent utilisées dans la gestion d'opérations en mer. Les résultats ont montré un bon accord entre la climatologie reconstituée et celle de référence, permettant de conclure que notre générateur peut valablement servir à la simulation de données d'états de mer en Afrique de l'Ouest pour toute la gamme des applications en génie océanique et côtier.



Abstract

Accurate estimation and representation of long-term sea conditions is a major issue in many ocean and coastal engineering applications : design of coastal and offshore structures, preparation of marine operations, coastal zone management, wave energy harvesting, etc. The main requirement is to estimate accurately the distribution (i.e. the probability of occurrence) of sea conditions at a location of interest. Stochastic modeling is commonly used to estimate these distributions because it enables to overcome the limits in duration that observations impose.

In this study, a new stochastic approach, event-based representation of sea state data, is proposed to model the wave climate in West Africa (from Ivory Coast to Namibia). An event refers to a wave system (swell or wind sea) evolving over time, that can be observed for a finite, yet significant duration and that can be linked to a single meteorological source phenomenon (e.g. low pressure systems, storms, etc.). Sea states are mostly composed of several superimposed wave systems, each with its own contribution to the total energy and time-evolving characteristics. In West Africa, in addition to wind sea, two or three swells may be present at the same time in the sea states. Some existing methods allow partitioning of the sea state spectrum into its constituent wave systems. Conventionally, it is possible to identify events by tracking wave systems in time, following their dynamical evolutions. The event-based approach that we propose here involves fewer parameters and more rational dependencies between them. It provides a structure with physical meaning and temporal consistence for the representation of sea states data.

This approach is mainly used in this study for the modeling of swell climate, that contributes to more than 80 % of the total energy of sea states in West Africa. Using as foundation elements the swell events, previously extracted from a time series of directional spectra, we build a stochastic generator of swell events and successfully reconstruct swell climate for durations of arbitrary length, that can be used with confidence in ocean and coastal engineering applications. The procedure we have developed for that purpose is decomposed into three following steps.

First, we have developed a new method : Time-domain of Spectral Expansion partitioning techniques (TESP)), based on the watershed algorithm in three dimensions (time, frequency, direction), to extract swell events from a time series of directional spectra. TESP uses the temporal coherence of wave systems in the partitioning and allows to ex-

tract consistent events with respect to the dynamical evolutions of wave systems. TESP ensures the continuity of any extracted event and allows for reducing the number of tuning parameters, conventionally used in existing methods, thereby reducing the computation time. We have also shown that TESP is not only suited to West Africa data, but it can also be used for spectra from different locations, having different dynamics.

We have then developed a model (Swell Event Model (SEM)) to represent each swell event by a reduced number of parameters (seven), the values of which are used as the basis variables for the stochastic generator. SEM model is validated by the joint distribution of the significant wave heights and periods and also in an application to estimate the power extracted by a wave Energy converter. The results showed that swell events observed in West Africa can be modeled by SEM with no significant difference in the joint distribution of the significant heights and periods and for ocean and coastal engineering applications.

In the last step, we have constructed the stochastic swell events generator by modeling the joint distribution of the parameters of SEM and by modeling the occurrence process of the events. This generator allows for simulation of individual swell events and for reconstruction of swell climate over durations of arbitrary lengths. Using appropriate statistical tests, we have validated the generator at each step of its implementation. We have also compared the distribution of the global significant heights and weather windows estimated from the reconstructed swell climate with those from the reference dataset. The results showed good agreement between reconstructed climate and that of reference. These results allow to conclude that our stochastic swell events generator can reliably be used to simulate sea state data in West Africa for a large range of coastal and ocean engineering applications.



Table des matières

Remerciements	v
Table des matières	xi
Table des figures	xv
Liste des tableaux	xxi
Liste des notations	xxiii
1 Introduction	1
2 États de mer : de l'observation à la modélisation statistique	7
2.1 Observation des états de mer	9
2.1.1 Observations <i>in situ</i>	9
2.1.2 Observations satellites	10
2.1.3 Modélisation numérique des états de mer	11
2.2 Caractérisation des états de mer	11
2.2.1 Approche vague à vague	11
2.2.2 Approche spectrale	12
2.3 Statistiques à court-terme des états de mer	20
2.3.1 Distributions des crêtes et des hauteurs de vagues individuelles	20
2.3.2 Distributions des crêtes de vagues non-linéaires	21
2.4 Statistiques à long-terme (climatologie) des états de mer	23
2.4.1 Utilisation de la distribution empirique globale	24
2.4.2 Estimation des extrêmes : approche par événement ou Peak Over Threshold method (POT)	30
2.5 Conclusion	35
3 Identification des événements de houle et de mer du vent	37
3.1 Concept d'événement de houle et de mer du vent.	39
3.1.1 Dynamique des systèmes de houle et événement de houle	40

3.1.2	Dynamique des systèmes de mer du vent et événement de mer du vent	41
3.2	Méthodes existantes pour l'identification des événements de systèmes de vagues	42
3.2.1	Partitionnement	42
3.2.2	Classification	45
3.2.3	Suivi temporel des systèmes de vagues	46
3.2.4	Spectral Partitioning for Operational Parameters identification (SPOP)	47
3.2.5	Limites des méthodes existantes	48
3.3	Méthode proposée : Time-domain Expansion of Spectral Partitionning technique (TESP)	49
3.3.1	Motivations	49
3.3.2	Filtrage	51
3.3.3	Extraction	52
3.3.4	Regroupement	52
3.3.5	Estimation des paramètres caractéristiques	54
3.3.6	Classification	54
3.4	Résultats	55
3.4.1	Données <i>in situ</i> d'Afrique de l'Ouest	55
3.4.2	Données de Hindcast de la mer de Gascogne	57
3.4.3	Comparaison de TESP avec SPOP	58
3.5	Conclusion	59
4	Validation des données de hindcast à l'échelle des événements	61
4.1	Données utilisées	63
4.2	Comparaison des hauteurs significatives des états de mer	64
4.3	Comparaison, à l'échelle des événements de houle et de mer du vent, des données <i>in situ</i> et des données de hindcast	66
4.3.1	Classification des événements	66
4.3.2	Comparaison des événements de houle	69
4.3.3	Comparaison des événements de mer du vent	73
4.4	Conclusion	75
5	Modélisation d'un événement de houle	77
5.1	Choix des paramètres synthétiques utilisés pour la caractérisation d'un événement de houle	79
5.2	Modèle d'évolution de la puissance au sein d'un événement de houle	81
5.2.1	Les différentes fonctions paramétriques étudiées	82
5.2.2	Analyse des résultats de la modélisation de l'évolution de la puissance avec les 3 fonctions paramétriques	83
5.3	Modèle d'évolution de la fréquence moyenne d'énergie au sein d'un événement de houle	86
5.4	Modèle d'évolution de la direction moyenne au sein d'un événement de houle	87
5.5	Récapitulatif	89
5.6	Résultats et discussions sur la modélisation des événements de houle avec le SEM	91
5.6.1	Modélisation d'une série temporelle d'événements de houle	91

5.6.2	Comparaison de la distribution jointe des couples (H_s, T_e) issus des événements de houle observés et ceux issus du SEM.	93
5.6.3	Validation pratique : évaluation du Potentiel Houlomoteur Extractible	95
5.7	Conclusion	99
6	Générateur stochastique d'événements de houle	101
6.1	Analyses statistiques des données de paramètres du SEM	103
6.1.1	Variabilité saisonnière	103
6.1.2	Variabilité interannuelle	105
6.1.3	Étude de la dépendance entre les paramètres du SEM	106
6.2	Distributions jointes des paramètres du SEM	109
6.2.1	Présentation du schéma de la modélisation des distributions jointes des paramètres du SEM.	109
6.2.2	Modélisation des distributions jointes de $\{E, d\}$ et $\{\omega, \theta_0\}$ par utilisation des copules	111
6.2.3	Distribution conditionnelle des paramètres α, τ et T_{max} par utilisation des GLM.	128
6.3	Modélisation de l'occurrence des événements de houles	135
6.3.1	Temps d'inter-arrivée des événements de houle	135
6.3.2	Modélisation de la distribution du temps d'inter-arrivée des événements de houle	136
6.4	Reconstruction et validation des séries temporelles d'événements de houle	139
6.4.1	Reconstruction de séries temporelles d'événements de houle	139
6.4.2	Validation des séries temporelles d'événements de houle simulés	140
6.5	Conclusion	147
7	Conclusion	149
7.1	Synthèse des principales conclusions	149
7.2	Perspectives	151
A	Climatologie des mers du vent	155
A.1	Climatologie du vent sur le site Girassol	156
A.2	Corrélation entre les caractéristiques du vent et celle de la mer du vent	157
A.3	Modèles stochastiques reliant les caractéristiques du vents à celle de la mer du vent	159
A.3.1	Modèle auto-régressif de la hauteur significative conditionné par la vitesse du vent	159
A.3.2	Modèle empirique basé sur la corrélation entre l'énergie adimensionnelle et l'inverse de l'âge des vagues	160
A.4	Conclusion	161
B	Amélioration de l'estimation spectrale par utilisation de la structure temporelle des systèmes de vagues	163
C	Étude comparative de fonctions paramétriques d'évolution de la puissance au sein d'un événement de houle	175



Table des figures

1.1	Support d'exploitation pétrolière Akpo (au large du Nigéria)	2
1.2	Convertisseur d'énergie de vagues Pelamis (à gauche) ; photo illustrant l'érosion côtière à Aného (Afrique de l'Ouest) (à droite)	2
1.3	Conditions des états de mer en Afrique de l'Ouest : hauteur significative (à gauche) et période (à droite). Les états de mers sont dominés par des houles qui proviennent majoritairement du Sud-Ouest. Source : Forristall et al. [2013]	3
1.4	Schéma représentant les différentes étapes de la démarche adoptée pour la représentation par événements de la climatologie des états de mer	5
2.1	Bouée Datawell Waverider (à gauche) ; exemple de déplacements mesurés (à droite).	10
2.2	Définition d'une vague par les méthodes de "downward zero-crossing" et "upward zero-crossing". Source : Holthuijsen [2007].	12
2.3	Superposition linéaire de vagues sinusoïdales monochromatiques et monodirectionnelles indépendantes. Source : Pierson Jr et al. [1971]	13
2.4	Contribution Δvar de la portion spectrale contenue dans $(\Delta f, \Delta \theta)$ à la variance totale des vagues. Source : Holthuijsen [2007]	15
2.5	Spectre de Pierson-Moskowitz pour un état de mer de $H_s = 2.5 m$ et différentes valeurs de T_p	18
2.6	Spectre JONSWAP pour différentes valeurs de γ avec $H_s = 2.5m$ et $T_p = 11s$.	18
2.7	Modèles des fonctions de répartition directionnelle.	19
2.8	Pour des conditions de vagues de spectres à bande étroite, la hauteur H est approximativement égale au double de la hauteur de la crête η_c (gauche) ; les hauteurs des crêtes suivent une loi de Rayleigh (droite).	21
2.9	30 min de série temporelle d'élévation de la surface de la mer enregistrée au large de l'Angola ($H_s = 2.35 m$, $T_{m01} = 12.63 s$) (gauche) ; Probabilité de dépassement de seuil des crêtes de vagues avec les modèles de forristall et de Prevosto (droite)	23
2.10	21 ans de données de H_s (Hindcast) au large de l'Angola. Les données sont échantillonnées au pas de 3 h.	24

2.11	Ajustement des modèles Log-normale, Weibull à 2 paramètres et Weibull à 3 paramètres aux données de H_s mesurées au large de l'Angola	26
2.12	Contours logarithmiques de distribution jointe de H_s et de T_{02} au large de l'Angola.	30
2.13	Définition d'une tempête	31
2.14	Ajustement des maxima de H_s des tempêtes identifiées à partir des données de Hindcast au large de l'Angola. Les ajustements obtenus par le GEV et le Gumbel sont les meilleurs avec un léger avantage pour le GEV. Les p -value obtenues pour les modèles GEV, Gumbel, Weibull et GPD sont respectivement 0.86, 0.85, 0.80 et 0.16.	33
2.15	Valeurs de retour du H_s en fonction des périodes de retour pour les modèles GEV et Gumbel. Les intervalles de confiances (95%) sont représentés en trait pointillé.	33
2.16	Modèles triangulaires de tempêtes : ETS (gauche), ETDS (droite)	35
3.1	Spectre d'état de mer obtenu au large de l'Angola (Girassol) le 04 janvier 2003 à 21h00. Représentation du spectre directionnel $2D$ (à gauche) et représentation du spectre omnidirectionnel $1D$ (à droite). On note la présence de trois systèmes de vagues : une houle primaire (1), une houle secondaire (2) et une mer du vent (0).	39
3.2	Historique de spectres mesurés pendant le mois de décembre 2002 au large de l'Angola (adapté à partir de Collard et al. [2009]).	41
3.3	Illustration de l'extraction par l'algorithme de la ligne de partage des eaux en 2D : (a) spectre directionnel; (b) spectre directionnel renversé, traité comme un relief topographique et recherche des lignes de partage des eaux; (c) les bassins versants identifiés, numérotés de 1 à 4, correspondent aux partitions.	43
3.4	Exemple de partitionnement du spectre de la figure 3.1 : extraction par l'algorithme des lignes de partage des eaux (gauche), puis regroupement des systèmes (droite)	45
3.5	Illustration de la classification des systèmes de vagues : la mer du vent (système 0) est localisée dans la parabole définie par le critère de l'âge des vagues $\{H_s = 0.54m, f_p = 0.26Hz, \theta_p = 257^\circ\}$, le système 2 correspond à la houle secondaire $\{H_s = 0.46m, f_p = 0.11Hz, \theta_p = 187^\circ\}$, et le système 1 correspond à la houle primaire $\{H_s = 0.62m, f_p = 0.08Hz, \theta_p = 217^\circ\}$	46
3.6	Identification des événements avec SPOP (données <i>in situ</i> , décembre 2002). De haut en bas, les fenêtres correspondent respectivement à la série-temporelle des densités spectrales $S(f)$, la hauteur significative, la fréquence de pic et la direction de pic. Chaque événement est représenté par une couleur différente.	48
3.7	Spectre $S(f)$ instantané (gauche) et série-temporelle de spectres ($S(t, f)$) en échelle logarithmique) des états de mer observés un mois de décembre 2002 au large de l'Angola (droite).	50
3.8	Coupes dans les plans passant par 210° et 300° de la série-temporelle de spectres directionnels $S(t, f, \theta)$ du mois de janvier 2003 (gauche); spectre directionnel obtenu le 21 janvier à 3h00 (droite)	51

3.9	Extraction des événements avec l'algorithme des lignes de partage des eaux : (haut) la série-temporelle $S(t, f)$ traitée comme une topographie renversée ; (milieu) identification des bassins versants ; (bas) regroupement des séquences de partitions.	53
3.10	Identification des événements : série-temporelle de la densité spectrale (fenêtre du haut) et paramètres synthétiques des événements. Les événements de mer du vent sont représentés en couleur bleue, et les événements de houle sont chacun représentés avec une couleur différente. Sur la troisième fenêtre, la ligne pointillée en couleur noire représente la fréquence de séparation.	56
3.11	Identification des événements à partir des données de Hindcast en mer de Gascogne. L'intensité et la direction du vent sont également représentées : série-temporelle de la densité spectrale (fenêtre du haut) et paramètres synthétiques des événements. Les événements de mer du vent sont représentés en couleur bleue, et les événements de houle sont chacun représentés avec une couleur différente. L'intensité et la direction du vent sont également représentées.	57
3.12	Spectre directionnel (gauche) et omnidirectionnel (droite) du 6 mars 2010 à 3h00 :	58
3.13	Identification des événements avec SPOP (gauche) et TESP (droite)	59
4.1	Positions des sites Girassol et Akpo avec les principaux secteurs possibles de provenances des houles	63
4.2	Comparaison du H_s global sur le site Girassol : Série temporelle (en haut), nuage des points (en bas, à gauche) et diagramme quantile/quantile (en bas, à droite).	65
4.3	Comparaison du H_s global sur le site Girassol : Série temporelle (en haut), nuage des points (en bas, à gauche) et diagramme quantile/quantile (en bas, à droite).	65
4.4	Diagramme polaire de la hauteur significative et de la période de pic suivant la direction moyenne des systèmes de vagues (GIRASSOL).	68
4.5	Diagramme polaire de la hauteur significative et de la période de pic suivant la direction moyenne des systèmes de vagues (AKPO).	69
4.6	Comparaison des événements de houle sur Girassol (janvier 2003)	70
4.7	Spectres directionnels obtenus le 20 janvier 2003 à 18h00 : <i>in situ</i> (à gauche) et hindcast (droite).	71
4.8	Diagramme quantile/quantile des systèmes de houles : Sud-Ouest (à gauche), Nord-Ouest (à droite)	71
4.9	Comparaison des événements de houle sur Akpo (décembre 2003)	72
4.10	Diagramme quantile/quantile des systèmes de houles : Sud-Ouest (à gauche), Sud-Est (à droite)	72
4.11	H_s Comparaison des événements de mer du vent sur Girassol pour janvier 2003 (à gauche) et sur Akpo décembre 2003 (à droite)	73
4.12	Diagramme quantile/quantile des H_s des événements de mer du vent : sur le site Girassol (à gauche) et sur le site Akpo (à droite).	74
4.13	Comparaison des vitesses du vent sur Girassol : série temporelle (haut) ; scatter plot (en bas, à gauche) et quantile/quantile plot (en bas, à droite).	75

5.1	Fréquence de pic (f_p) et fréquence zero-crossing (f_{m02}) en fonction de la fréquence d'énergie (f_e) : Girassol (gauche) et Akpo (droite)	80
5.2	Ensemble des 201 événements de houle extraits des données Girassol (gauche); événement de houle typique (droite)	81
5.3	Modèle triangulaire pour l'évolution de la puissance au sein d'un événement de houle. La partie colorée correspond à l'énergie de l'événement (E).	83
5.4	Loi Log-normale pour $E = 700 kWh/m$ et pour différentes valeurs de μ et σ	84
5.5	Loi Gamma pour $E = 700 kWh/m$ et pour différentes valeurs de α et τ	84
5.6	Les différentes fonctions paramétriques ajustées à un événement de houle (haut); erreur relative associée (bas)	85
5.7	Boîtes à moustache de la distribution des NRMSE.	86
5.8	Suivi dans le temps d'une dépression : sur la fenêtre en dessous de chaque carte est représentée l'évolution dans le temps de l'angle sous lequel, un observateur situé sur le site Girassol voit arriver les trains de houle.	88
5.9	Histogramme des variations d'angle entre le début et la fin des événements de houle	89
5.10	Un événement et sa représentation modélisée	90
5.11	Historique d'événements de houle et leurs modélisations (données <i>in situ</i> , Girassol); les événements de houle portent chacun un numéro différent.	92
5.12	Historique d'événements de houle et leurs modélisations (données <i>in situ</i> , Akpo); les événements de houle portent chacun un numéro différent.	93
5.13	Distributions jointes des hauteurs significatives et des périodes moyennes (Données <i>in situ</i> , Girassol). Les contours sont régulièrement espacés de 0.00475 entre 0.001 et 0.02.	94
5.14	Fonction de transferts de puissance omnidirectionnel (à gauche) et uni-directionnel (à droite)	96
5.15	Modèles Jonswap et Log-normal ajustés à un spectre de houle	97
5.16	Potentiel houlomoteur extractible avec le convertisseur uni-directionnel (Girassol)	98
5.17	Potentiel houlomoteur extractible avec le convertisseur uni-directionnel (Akpo)	98
5.18	Diagramme quantile-quantile comparant la puissance extractible sur le site Akpo (à gauche) et Girassol (à droite)	99
6.1	Puissance des événements de houle identifiés à partir des données hindcast (Girassol) de l'année 2008 et SEM associés	104
6.2	Boxplot des distributions mensuelles des paramètres E , d , T_{max} et θ_0	104
6.3	Nombre annuel d'événements de houle (haut) et énergie totale annuelle (bas)	106
6.4	Nuages de points des paramètres du SEM	108
6.5	Diagramme montrant le schéma utilisé pour déterminer les distributions des paramètres du SEM	110
6.6	Illustration des différentes étapes pour la simulation d'un couple de variables aléatoires $\{x_1, x_2\}$ par utilisation d'une copule	112
6.7	Évolution mensuelle du paramètre de forme $a(k)$ et d'échelle $b(k)$ de la loi Gamma ajustée aux observations de E	114
6.8	Diagramme quantile-quantile de la comparaison entre la distribution marginale empirique de E et celle donnée par la loi Gamma	115
6.9	Diagramme quantile-quantile de la comparaison entre la distribution marginale empirique de d et celle de la loi log-normale	116

6.10	Diagramme quantile-quantile de la comparaison entre la distribution de marginale empirique de θ_0 et celle de la loi normale	117
6.11	Diagramme quantile-quantile de la comparaison entre la distribution de marginale empirique de ω et celle de la loi normale	117
6.12	Fonctions densités de probabilité des copules.	119
6.13	Distributions jointes des fonctions de répartitions empiriques de E et d (à gauche) et de ω et θ_0 (à droite) (Cas du mois d'août)	120
6.14	Comparaison des contours de probabilités de la copule empirique de $\{E, d\}$ avec ceux des 4 copules (Cas du mois d'août).	123
6.15	Distributions jointes de $\{E, d\}$: référence (à gauche) et simulée (à droite). Les histogrammes correspondent aux distributions marginales des paramètres . . .	126
6.16	Distributions jointes de $\{\omega, \theta_0\}$: référence (à gauche) et simulée (à droite). Les histogrammes correspondent aux distributions marginales des paramètres . . .	126
6.17	Nuage de points de $\{E, d\}$ simulés en comparaison avec les données de référence.	127
6.18	Nuage de points de $\{\omega, \theta_0\}$ simulés en comparaison avec les données de référence.	127
6.19	Diagnostic du GLM exprimant T_{max} en fonction de E et d	131
6.20	Diagnostic du GLM exprimant τ en fonction de d et T_{max}	132
6.21	Diagnostic du GLM exprimant α en fonction de E, τ, d et T_{max}	132
6.22	Nuage de points entre d et T_{max} simulés en comparaison avec les données de référence.	133
6.23	Nuage de points de α et τ simulés en comparaison avec les données de référence.	133
6.24	Zone de génération possible des houles du Sud-Ouest observées sur le site Girassol délimitée par les deux lignes bleues et les abords des continents	135
6.25	Définition du temps d'inter-arrivée entre deux événements de houle successifs (à gauche) ; distribution mensuelle des données du temps d'inter-arrivée (droite)	136
6.26	Histogramme de I (à gauche) ; Diagramme quantile-quantile entre distribution empirique de I et les lois Gamma (au centre) et de Weibull (à droite) (cas du mois d'août)	137
6.27	Nuage de points entre I et E (à gauche) et entre I et d (à droite)(cas du mois d'août)	137
6.28	Diagnostic du GLM exprimant I en fonction de E, τ, d, T_{max}	138
6.29	Distributions jointes de $\{d, I\}$: référence (à gauche) et simulée (à droite) (Cas du mois d'août). Les contours sont régulièrement espacés de 0.01 entre 0.01 et 0.04	139
6.30	Nuage de points entre les paramètre E et d d'un événement de houle et son suivant.	140
6.31	Exemple de simulation d'une année d'événements de houle caractérisés par leur puissance (à gauche) ; Zoom sur le mois d'août pour les paramètres P_w, f_e et θ_m (à droite)	140
6.32	Exemple d'un mois de séries temporelles d'événements de houle simulés. La Puissance, la période d'énergie et la hauteur significative déduite sont représentées avec les intervalles de fluctuation à 95% (traits en pointillés).	141
6.33	Diagramme quantile-quantile de la distribution du H_s total simulé et du H_s total de référence.	142
6.34	Exemple d'ajustement de la loi de Weibull à 3 paramètres aux données empiriques de probabilités de dépassement de seuil	143

6.35	Durée de persistance : pour l'ensemble des 21 années (en haut) ; valeur médiane (en bas)	145
6.36	Nombre de jours par mois d'accès au site	147
A.1	Approche adoptée pour la simulation des séries temporelles des caractéristiques des mers du vent	156
A.2	Histogramme (a) et la fonction de répartition empirique (b) des intensités du vent	157
A.3	Rose des vents (a) et fonction d'autocorrélation de l'intensité du vent (b) . . .	157
A.4	Série temporelle de la hauteur significative et de la direction moyenne de la mer du vent en comparaison avec le carré de la vitesse du vent et sa direction .	158
A.5	Nuage de points entre le H_s et le carré de la vitesse du vent (a) et histogramme des valeurs de la différence entre la direction du vent et celle de la mer du vent (b)	158
A.6	Diagnostic du modèle : diagramme quantile/quantile (à gauche) et série temporelle (à droite)	160
A.7	Régression entre l'énergie adimensionnelle et l'inverse de l'âge des vagues. Les intervalles de confiance sont aussi représentés	161



Liste des tableaux

2.1	Quelques familles de distributions directionnelles. Source Hauser et al. [2005]	19
2.2	Valeurs des statistiques de χ^2 , et de Kolmogorov-Smirnov (K-S)	27
3.1	Tableau résumant les valeurs des constantes utilisées pour le regroupement dans la littérature.	44
3.2	Tableau résumant les valeurs des constantes utilisées pour le suivi temporel	47
4.1	Contribution à l'énergie totale et à puissance totale des systèmes de vagues (Girassol)	67
4.2	Contribution à l'énergie totale et à puissance totale des systèmes de vagues (Akpo)	69
5.1	Tableau récapitulatif du SEM	90
5.2	NRMSE associés aux événements de houle de la figure 5.11 (Girassol)	92
5.3	NRMSE associés aux événements de houle de la figure 5.12 (Akpo)	92
5.4	Résultats du test de Kolmogorov-Smirnov appliqués aux distributions de H_s et T_e	95
6.1	Tau de Kendall entre les paramètres du SEM et p -value associée (cas du mois d'août). Les couples de paramètres pour lesquels le Tau de Kendall est significativement non nul sont indiqués en gras.	109
6.2	Valeurs mensuelles des paramètres $a(k)$ et $b(k)$ et p -value du test de Kolmogorov-Smirnov	114
6.3	Valeurs mensuelles des paramètres des lois marginales ajustées aux paramètres (d , ω et θ_0) et p -value des tests de Kolmogorov-Smirnov	116
6.4	Exemples de copules Archimédiennes bivariées	120
6.5	BIC et $AICc$ pour les différentes copules utilisées pour modéliser la dépendance entre E et d .	122
6.6	BIC et $AICc$ pour les différentes copules utilisées pour modéliser la dépendance entre ω et θ_0 .	122

6.7	Statistiques de Cramér-von Mises S_n et de Anderson-Darling A_n calculées à partir de la copule Gaussienne pour $\{E, d\}$, et pour $\{\omega, \theta_0\}$. Les p -value sont indiquées dans les parenthèses et sont obtenues à partir d'un <i>bootstrap</i> à 5000 échantillons	125
6.8	Exemples de distributions continues appartenant à la famille exponentielle et fonctions lien canoniques associées	129
6.9	Estimation des paramètre du GLM exprimant T_{max} en fonction de E et d (cas du mois d'août)	130
6.10	Estimation des paramètre du GLM exprimant τ en fonction de d et T_{max} (cas du mois d'août)	130
6.11	Estimation du GLM exprimant α en fonction de E, τ, d et T_{max} (cas du mois d'août)	130
6.12	Coefficients du GLM exprimant I en fonction de E, τ, d, T_{max}	138

Liste des notations

η	élévation de la surface de la mer
\bar{H}	hauteur moyenne des vagues
H_{rms}	hauteur quadratique moyenne
f	fréquence
k	nombre d'onde
a	amplitude des vagues
φ	phase
$S(f)$	spectre omni-directionnel d'état de mer
$S(f, \theta)$	spectre directionnel d'état de mer
$D(f, \theta)$	fonction de répartition directionnelle d'état de mer
m_p	moment d'ordre p de la densité spectrale
H_s	hauteur significative
$H_{1/3}$	valeur moyenne du tiers supérieur des hauteurs des vagues
T_{m-10}	période d'énergie
T_{m02}	période moyenne zero-crossing
T_p	période de pic
θ_p	direction principale
θ_m	direction moyenne
Q_p	paramètre de Goda
P_w	puissance des vagues
f_e	fréquence d'énergie
f_p	fréquence de pic
U_r	nombre de Ursell
Cg	vitesse de groupe des vagues
σ	étalement directionnel
ϵ	étalement spectrale
U_{10}	vitesse du vent à 10m
θ_w	direction du vent
δ	différence entre la direction du vent et celle de la mer du vent
E	énergie totale de l'événement de houle
α	paramètre de forme

τ	paramètre d'échelle
d	distance indicative parcourue par les houles
T_{max}	période maximale de houle dans l'événement
ω	vitesse angulaire moyenne, relative au point d'observation, de la tempête
θ_0	direction de provenance du premier train de houle
I	temps inter-arrivée
R	rayon de la terre
P	puissance extractible
R_p	fonction de transfert omni-directionnel du convertisseur d'énergie des vagues
R_{pd}	fonction de transfert uni-directionnel du convertisseur d'énergie des vagues
MAE	erreur relative en valeur absolue
$RMSE$	erreur quadratique moyenne
$NRMSE$	erreur quadratique moyenne normalisée
$\hat{\tau}_k$	tau de Kendall
C	copule
BIC	critère d'information bayésien
$AICc$	critère d'information d'Akaike corrigé
D^*	statistique de Kolmogorov-Smirnov
S_n	statistique de Cramèr-Von Mises
A_n	statistique de Anderson-Darling

Acronymes

SPOP	Spectral Partitioning for Operational Parameters identification
TESP	Time-domain Expansion of Spectral Partitionning technique
SEM	Swell Event Model
GLM	Generilized Linear Model

Introduction

La connaissance des états de mer est d'une grande importance lorsqu'on aborde certaines problématiques en océanographie physique et dans les disciplines connexes : conception et dimensionnement de structures marines et portuaires telles qu'une plateforme pétrolière ou une digue, gestion des zones côtières et des opérations en mer, étude de la morphodynamique côtière et de l'évolution du littoral, récupération de l'énergie des vagues, etc.

L'état de mer est la description de la surface de la mer soumise à l'influence du vent (qui génère la mer du vent) et de la houle résiduelle. La mer du vent est le système de vagues créées à l'endroit même où souffle le vent (fetch). En quittant l'endroit où elles ont été générées, les vagues se régularisent et deviennent la houle, qui peut se propager très loin même en l'absence de vent.

Les besoins sur la connaissance des états de mer peuvent être classés en deux catégories :

- Connaissance des statistiques à court terme des états de mer : nécessaire pour estimer par exemple quantitativement le risque de rencontrer des conditions de mer indésirables lors des opérations en mer, par exemple l'installation ou la maintenance d'une plateforme pétrolière ou d'un système récupérateur de l'énergie des vagues, etc.
- Connaissance des statistiques à long-terme des états de mer (climatologie) : indispensable pour le dimensionnement des structures marines (figure 1.1), l'estimation morphodynamique du trait de côte (figure 1.2, à droite), l'estimation du potentiel récupérable de l'énergie des vagues (figure 1.2, à gauche), etc.

Pour ces différentes applications, il est essentiel de pouvoir estimer avec précision la distribution (i.e. les probabilités d'occurrence) des conditions de mer sur un site d'intérêt. Par exemple, la prévision de l'endommagement en fatigue utilise la distribution de l'ensemble des conditions de mer. De même, l'évaluation de la sécurité d'une structure requiert d'extrapoler cette distribution pour identifier les cas de charges extrêmes à considérer.

Un moyen qu'on peut utiliser pour estimer ces distributions est la modélisation stochastique. En effet, les données d'observations (e.g. données *in situ*) disponibles, ne couvrent souvent qu'une durée très limitée (généralement deux à trois ans). Même les données en sortie de modèle (hindcast) sont limitées en durée car les modèles numériques de vagues,

qui servent à les générer utilisent en forçage des données de vent, généralement indisponibles sur de très longues durées. D'où l'intérêt de la modélisation stochastique qui permet de dépasser les limites, notamment en durée, imposées par les observations ou par les données en sortie de modèle.



FIGURE 1.1 – Support d'exploitation pétrolière Akpo (au large du Nigéria)



FIGURE 1.2 – Convertisseur d'énergie de vagues Pelamis (à gauche); photo illustrant l'érosion côtière à Aného (Afrique de l'Ouest) (à droite)

Les états de mer sont généralement la combinaison de plusieurs systèmes de vagues, d'intensités, de fréquences et de directions différentes. Les spectres caractéristiques de ces états de mer présentent plusieurs pics correspondant aux systèmes de vagues qui les composent, notamment de houles et de mer du vent. En Afrique de l'Ouest (du large de la Namibie au large de la Côte d'Ivoire), en plus de la mer du vent, les états de mer sont souvent composés de plusieurs systèmes de houles (deux voire trois) qui peuvent être principalement de direction Sud-Ouest ou Nord-Ouest (voir figure 1.3). De nombreuses études ont montré l'importance d'une description fine des états de mer pour les applications inhérentes au génie océanique et côtier (Bitner-Gregersen [2005], Kerbiriou et al. [2007]). Il existe des méthodes dites de partitionnement qui permettent de décomposer un spectre caractéristique d'un état de mer présentant plusieurs pics en la combinaison de plusieurs systèmes de vagues. La modélisation paramétrique des états de mer qui en découle fait alors intervenir un nombre important de paramètres, puisqu'à chaque système de vague présent lors d'une observation est associé un modèle spectral paramétrique. Or les mesures océano-météo qui couvrent habituellement deux ou trois années, et qui peuvent permettre d'estimer la statistique des paramètres environnementaux s'il n'y a qu'un seul système

de vagues, deviennent insuffisantes pour estimer celle des combinaisons de plusieurs systèmes. La pratique actuelle pour décrire la distribution à long-terme des conditions d'état de mer consiste à dresser les statistiques respectives des systèmes de vagues observés à partir des données disponibles et à supposer l'indépendance des systèmes de vagues pour en déduire la climatologie. Cette hypothèse n'est cependant pas suffisamment conforme aux observations (Guédé et al. [2010]). De plus, les systèmes de houle ou de mer du vent affichent souvent une cohérence temporelle qui traduit leur dynamique. Cependant, cette cohérence temporelle n'est généralement pas prise en compte dans la représentation de la climatologie des états de mer.

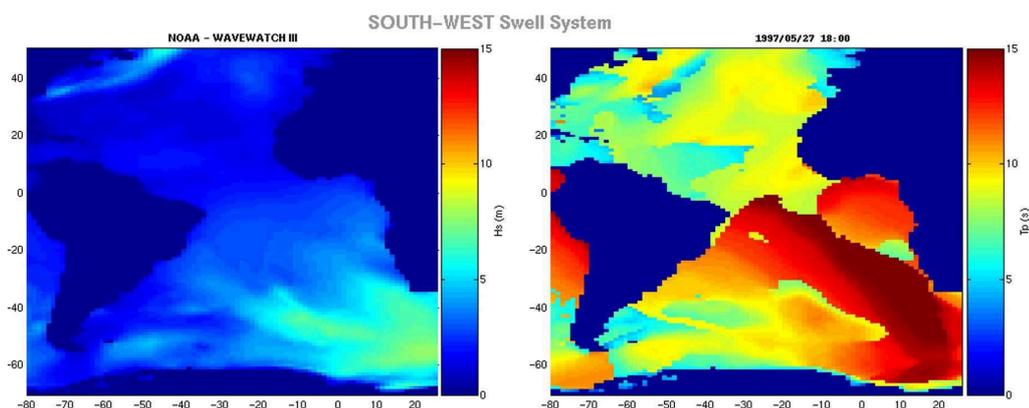


FIGURE 1.3 – Conditions des états de mer en Afrique de l'Ouest : hauteur significative (à gauche) et période (à droite). Les états de mers sont dominés par des houles qui proviennent majoritairement du Sud-Ouest. Source : Forristall et al. [2013]

Dans ces travaux, nous proposons une nouvelle approche stochastique pour estimer une climatologie de vagues s'appuyant sur une représentation par événements des données d'états de mer. Un événement désigne un système de vagues (houles ou mers du vent) en évolution au cours du temps, observable pendant une durée significative et que l'on peut relier à un unique phénomène météorologique source (e.g. dépressions, tempêtes, etc.). À partir d'une série temporelle de spectres directionnels caractéristiques des états de mers, il est possible d'identifier les événements et de reconstituer des climatologies par modélisation stochastique de ces derniers.

La représentation par événements des données d'états de mer permet la simplification de la structure des données qui deviennent une collection de séquences, chacune avec son étendue temporelle, de paramètres d'environnement. Elles font intervenir par conséquent moins de paramètres et des notions de dépendance entre eux plus raisonnées. En effet, les événements sont des entités qui ont un sens physique car ils reflètent la dynamique des systèmes de vagues. Enfin, l'identification d'événements à partir des données permet également de reproduire la cohérence temporelle des systèmes de vagues.

Dans ces travaux, l'approche par événements sera principalement utilisée pour la modélisation de la climatologie des houles qui contribuent à plus de 80% de l'énergie totale des états de mer en Afrique de l'Ouest. La démarche adoptée emploie trois étapes schématisées dans la figure 1.4 et présentées ci-après :

- La première étape consiste à identifier dans une série temporelle de spectres directionnels caractéristiques d'états de mer, les événements de houles, par utilisation

d'outils de partitionnement de spectres et de suivi temporel des systèmes de vagues.

- La deuxième étape consiste à modéliser les événements de houle identifiés. Le but est de représenter les événements, chacun avec son étendue temporelle, par un nombre réduit de paramètres tout en conservant leur cohérence temporelle.
- La dernière étape repose sur le développement d'un générateur stochastique d'événements de houle, utilisant comme variables les valeurs des paramètres du modèle introduit dans la deuxième étape et permettant la simulation d'événements de houles individuels, puis la reconstitution de climatologies sur de longues durées, pour des applications en génie océanique et côtier.

La modélisation stochastique de la climatologie des mers du vent peut s'effectuer différemment et est traitée en Annexe A. En effet, à cause d'un régime de vent très fluctuant à l'échelle de la journée en Afrique de l'Ouest, la hauteur significative des mers du vent générées est aussi fluctuante, ne favorisant pas une segmentation par événements. Dans l'annexe A, nous montrons comment reconstituer la climatologie des mers du vent à partir des modèles stochastiques reliant les caractéristiques du vent à celles des mers du vent.

Le présent mémoire s'articule autour de cinq chapitres :

Le chapitre 2 expose brièvement quelques notions de base nécessaires à la compréhension du sujet : présentation des moyens d'observations et de caractérisation des états de mers. Ce chapitre aborde également la description statistique à court-terme et à long-terme des états de mers, où il présente les modèles existants tout en proposant une étude comparative en utilisant les données d'états de mer de l'Afrique de l'Ouest.

Le chapitre 3 traite spécifiquement de l'identification des événements de houle et de mer du vent au sein d'une série temporelle d'états de mer caractérisés par leurs spectres directionnels. Les concepts d'événement de houle et d'événement de mer du vent sont d'abord expliqués, puis les méthodes existantes pour l'identification des événements sont présentées ainsi que leurs limites. On présente ensuite la méthode TESP (Time-domain Expansion of Spectral Partitionning technique) d'identification d'événements de houle et de mer du vent, développée pendant ces travaux. TESP est une extension des méthodes existantes à la dimension temporelle, permettant d'améliorer considérablement l'identification des événements par rapport à ces dernières.

Le chapitre 4 présente la validation des événements identifiés à partir des données en sortie de modèle (hindcast) par rapport aux événements identifiés à partir des données d'observations. En effet, les données de référence utilisées pour notre générateur stochastique d'événements de houles seront celles obtenues en sortie de modèles (hindcast) car par rapport aux données d'observations, elles couvrent une durée suffisamment longue pour prendre en compte l'alternance des saisons.

Dans le chapitre 5 nous présentons le modèle développé (swell event model (SEM)) pour représenter chacun des événements de houle par un nombre réduit de paramètres. La modélisation des événements de houle est une étape nécessaire à la caractérisation d'une climatologie de houles par événements. L'idée est d'identifier des formes paramétriques permettant de modéliser les évolutions des paramètres caractéristiques des états de mer dans un événement de houle. Le chapitre présente également la validation du modèle SEM, notamment sur la distribution jointe des hauteurs significatives et des périodes, puis sur

un cas d'application : l'estimation de la puissance extraite par un convertisseur d'énergie des vagues.

Le chapitre 6 présente le générateur stochastique d'événements de houle qui a été développé et qui permet la simulation d'événements de houles synthétiques, puis la reconstitution de climatologies sur des durées plus longues que celles pour lesquelles les données d'observations ou de hindcast sont disponibles. Ce générateur est construit en deux étapes. La première consiste à modéliser les distributions des paramètres du modèle SEM, permettant la simulation d'événements de houles individuels. La deuxième consiste à modéliser le processus d'occurrence des événements de houle pour reconstruire des séries temporelles réalistes d'événements de houle et reconstituer la climatologie. Ce chapitre propose également la validation du générateur, d'abord à chaque étape de sa conception, puis sur la climatologie reconstituée avec un cas d'application : la préparation d'opérations en mer.

Pour terminer, nous présentons une synthèse des différents résultats de ces travaux. Nous exposons des conclusions générales et proposons des ouvertures sur les développements à venir et perspectives de recherche à plus ou moins long terme.

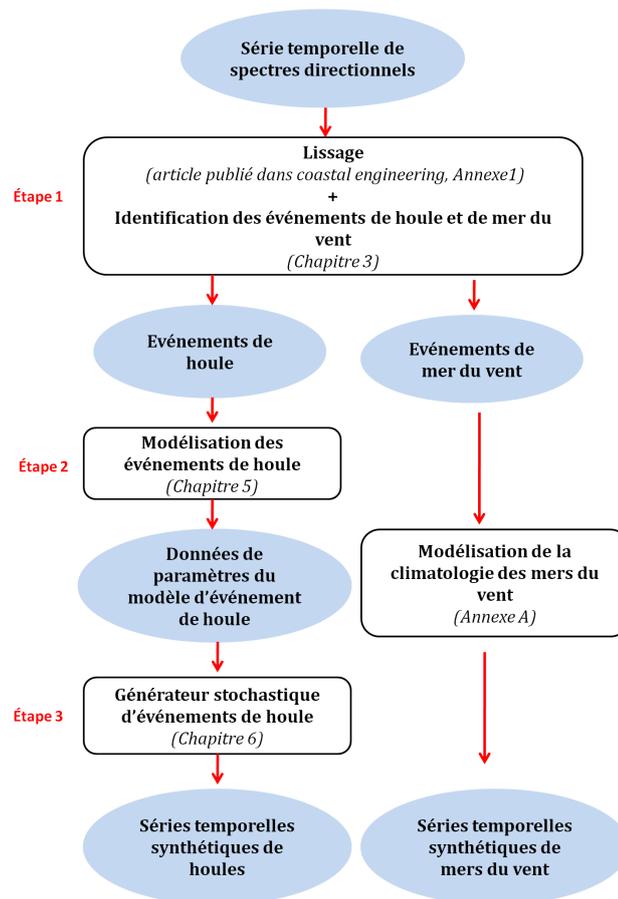


FIGURE 1.4 – Schéma représentant les différentes étapes de la démarche adoptée pour la représentation par événements de la climatologie des états de mer

États de mer : de l'observation à la modélisation statistique

Ce chapitre présente, d'une part, les notions de base nécessaires à la compréhension du sujet et, d'autre part, les méthodes généralement utilisées dans la modélisation statistique des états de mer. Le chapitre aborde dans un premier temps, les différents moyens d'observation des états de mer, puis leur caractérisation par l'analyse vague par vague ou par l'analyse spectrale. Il traite ensuite de la modélisation statistique des états de mer à court-terme (échelle de temps allant de 20 minutes à 6 heures environ) et de la modélisation statistique des états de mer à long-terme (plusieurs années). Pour cette dernière, nous proposons une étude comparative des modèles existants, en utilisant les données d'états de mer recueillies au large de l'Angola.

Le chapitre reprend quelques notions présentées en détail dans les ouvrages de référence tels que [Tucker, Pitt \[2001\]](#), [Young \[1999\]](#), [Hauser et al. \[2005\]](#), [Holthuijsen \[2007\]](#) ou le support de cours dispensés à l'ENSTA ParisTech [Ardhuin \[2012\]](#). De ce fait, il ne se substitue pas à ces ouvrages de référence vers lesquels le lecteur est invité à se reporter pour tout complément.

Sommaire

2.1	Observation des états de mer	9
2.1.1	Observations <i>in situ</i>	9
2.1.2	Observations satellites	10
2.1.3	Modélisation numérique des états de mer	11
2.2	Caractérisation des états de mer	11
2.2.1	Approche vague à vague	11
2.2.2	Approche spectrale	12
2.3	Statistiques à court-terme des états de mer	20
2.3.1	Distributions des crêtes et des hauteurs de vagues individuelles	20
2.3.2	Distributions des crêtes de vagues non-linéaires	21
2.4	Statistiques à long-terme (climatologie) des états de mer	23
2.4.1	Utilisation de la distribution empirique globale	24
2.4.2	Estimation des extrêmes : approche par événement ou Peak Over Threshold method (POT)	30
2.5	Conclusion	35

2.1 Observation des états de mer

Les dernières décennies ont vu croître la connaissance du phénomène que sont les états de mer grâce à une amélioration notable des outils d'observation. Deux moyens principaux et complémentaires sont recensés : les observations *in situ* (houlographes, profileurs ADCP) et les observations satellites (altimétrie, Synthetic Aperture Radar SAR).

2.1.1 Observations *in situ*

2.1.1.1 La perche à houle

La perche à houle mesure directement la position de la surface libre à la verticale d'un point fixe. La mesure se fait soit à partir de la résistance ou de la capacité du circuit constitué par un (ou deux) fil(s) conducteur(s) formant une boucle fermée par la surface de la mer. La mesure des vagues avec une perche à houle est cependant délicate à opérer car elle nécessite un support fixe (plateforme) et les forces exercées par les vagues sur les perches sont importantes. Les perches sont souvent associées en réseau de plusieurs perches sur une même plateforme, afin de permettre une estimation de la direction des vagues (Cavaleri et al. [1997]).

2.1.1.2 La bouée houlographe

Les accélérations verticales (pilonnement) subies par une bouée flottant en surface donnent, après une double intégration, un signal d'élévation de la surface $\eta(x, y, t)$. Suivant le type d'instrument et la présence éventuelle d'un courant, la position horizontale (x, y) n'est pas fixe mais suit, à peu près, le mouvement orbital des vagues. Cette propriété peut être assez gênante car une partie de la non-linéarité de la surface est filtrée dans le signal mesuré. En plus de cette mesure de pilonnement, qui a longtemps été la plus répandue, la direction des vagues peut être déterminée en faisant aussi une mesure des accélérations horizontales, qui donne par intégration, les déplacements horizontaux x et y (Figure 2.1), en sorte qu'il est possible de remonter à l'information spectro-directionnelle complète d'un état de mer. Précisons que les données d'observations qui seront utilisées dans ces travaux sont des mesures réalisées par des bouées houlographes.

2.1.1.3 Autres techniques d'observations *in situ*

Les vagues peuvent aussi être mesurées à l'aide de capteurs "P-U-V" ou de profileurs de courant ADCP. Le capteur "P-U-V" mesure la pression et les deux composantes de la vitesse dans le plan horizontal. L'avantage de ce moyen de mesure est qu'il est fait pour être posé au fond et constitue par conséquent le mouillage le plus simple. Cependant, il n'est pas facile de récupérer des données en temps réel.

Une variante plus récente et pratique est l'utilisation de profileurs de courant (ADCP) en visée horizontale ou verticale. La combinaison des vitesses mesurées le long des différents faisceaux acoustiques permet, en principe, une mesure intéressante du spectre directionnel. Toutefois le bruit important sur chaque mesure fait que l'ADCP en visée verticale, en pratique, ne donne pas une résolution angulaire nettement supérieure à celle d'un simple P-U-V (Ardhuin [2012]).

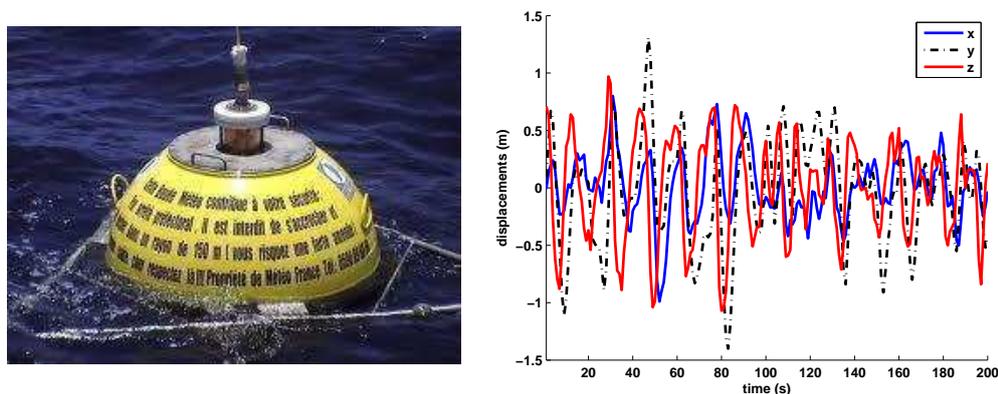


FIGURE 2.1 – Bouée Datawell Waverider (à gauche) ; exemple de déplacements mesurés (à droite).

2.1.2 Observations satellites

Les satellites sont très utilisés comme outils d'observations des états de mer et complètent les informations obtenues par les mesures in situ. Ils sont équipés de différents instruments de mesures pour l'observation des vagues tels que les altimètres et les radars à synthèse d'ouverture.

La mesure altimétrique est actuellement la seule mesure des vagues qui soit disponible de manière globale et utilisée pour la prévision opérationnelle des vagues (par assimilation de données). Ce moyen de mesure permet d'avoir la hauteur significative des états de mer de manière indirecte à partir de la forme des échos des altimètres radar. Les principales missions courantes d'altimétrie sont : Jason-2 (depuis Juin 2008), HY-2 (depuis Août 2011), Saral (depuis Février 2013), et très prochainement Sentinel-3A.

Le radar à synthèse d'ouverture ou Synthetic Aperture Radar (SAR) est un radar imageur qui effectue un traitement des données reçues afin d'améliorer la résolution en azimuth. Par inversion d'image on peut déduire des données de SAR des spectres d'état de mer et donc la direction de propagation des vagues. Il faut noter cependant que les composantes spectrales de faibles périodes qui se déplacent dans la direction azimuthale sont invisibles. Cette coupure azimuthale est la cause principale de la limitation de l'utilisation du SAR (Kerbaol et al. [1998]). Ainsi, dès que le vent est un peu fort, des vagues courtes avec une vitesse orbitale relativement élevée empêchent de voir la partie du spectre qui se propage dans la direction azimuthale sauf pour les houles les plus longues. Enfin, les données issues du SAR ont permis de mieux comprendre et de quantifier la dissipation des houles (Ardhuin et al. [2009]).

Il existe d'autres types d'instruments tels que les radars HF, les radars X qui permettent de mesurer les vagues.

Nous venons de décrire les deux principaux outils d'observations d'états de mer ayant chacun leurs spécificités et leurs disponibilités spatio-temporelles. Toutefois, le cumul de ces outils (mesures in situ et capteurs satellitaires) ne fournit pas une connaissance complète des états de mer en espace et en temps. En complément, la simulation numérique offre une solution à partir de laquelle il est possible d'étudier les états de mer avec une

forte densité spatio-temporelle pour diverses applications.

2.1.3 Modélisation numérique des états de mer

La modélisation numérique des états de mer hérite de plus de cinquante ans de développement et d'améliorations continues sur divers aspects : amélioration des modèles mathématiques et des méthodes numériques employées pour les simuler, amélioration des modèles représentant la physique des processus affectant les vagues, amélioration des méthodes de calibration et de validation de modèles à l'aide d'observations de qualité et de résolution croissantes, développement des moyens de calcul et de stockage et enfin d'amélioration des forçages atmosphériques. Généralement, les simulations numériques d'état de mer se classent en trois groupes : les simulations opérationnelles, rétrospectives et prospectives (Laugel [2013], Ardhuin, Roland [2013]). Les premières correspondent à la prévision des états de mer de quelques heures à 10 jours et répondent à la prévision des risques de submersions marines ou la gestion d'opérations en mer par exemple. Les simulations d'états de mer rétrospectives (hindcast), réalisées sur plusieurs décennies, permettent d'étudier la climatologie des vagues à l'échelle régionale ou globale, leur distribution et leur variabilité. Enfin, les simulations d'états de mer futurs permettent d'appréhender les impacts possibles du changement climatique sur la climatologie des vagues. Les modèles d'états de mer actuellement utilisés pour des applications océaniques et côtières sont dits de troisième génération (3G), car ils ne nécessitent aucune hypothèse sur la forme initiale du spectre de variance. Le précurseur des modèles 3G est le code WAM (Group [1988]) ; depuis, d'autres modèles ont été développés. Parmi les modèles les plus utilisés, citons TOMAWAC (Benoit et al. [1996]), WaveWatch III (Tolman [2002], Tolman, others [2009], Ardhuin, Jenkins [2006]) et SWAN (Booij et al. [1999]), continuellement en développement.

2.2 Caractérisation des états de mer

2.2.1 Approche vague à vague

L'approche statistique vague à vague étudie les mouvements verticaux d'une série de vagues découpées par les méthodes "downward zero-crossing" ou "upward zero-crossing" (figure 2.2). Ces méthodes de mesure découpent respectivement une vague, profil de l'élévation de la surface océanique entre deux instants successifs correspondant au niveau moyen, lors d'un mouvement descendant ou ascendant. Les deux approches sont identiques pour des vagues en eaux profondes, mais leurs résultats divergent pour les vagues en faible profondeur d'eau. Parmi les deux variantes, la méthode "downward zero-crossing" est préconisée car elle conserve le front de vague intact. Après avoir découpé une série en vagues et défini les couples hauteur-période de chaque vague isolée, différents paramètres statistiques peuvent être calculés :

Hauteur moyenne des vagues :

$$\bar{H} = \frac{1}{N} \sum_{i=1}^N H_i \quad (2.1)$$

où i est le numéro de la vague dans l'enregistrement (i.e. $i = 1$ pour la première vague, $i = 2$ pour la deuxième vague, etc.).

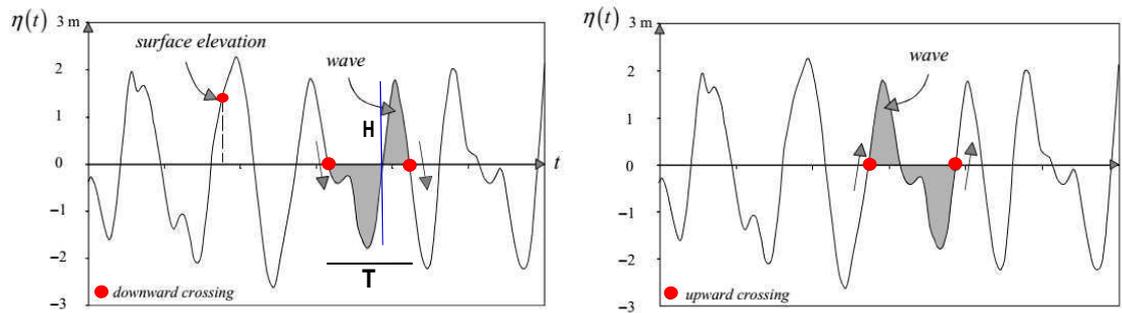


FIGURE 2.2 – Définition d'une vague par les méthodes de "downward zero-crossing" et "upward zero-crossing". Source : [Holthuijsen \[2007\]](#).

Hauteur quadratique moyenne :

$$H_{rms} = \left(\frac{1}{N} \sum_{i=1}^N H_i^2 \right)^{1/2} \quad (2.2)$$

Moyenne du tiers des vagues les plus hautes :

$$H_{1/3} = \frac{1}{N/3} \sum_{j=1}^{N/3} H_j \quad (2.3)$$

où $j = 1$ correspond à la vague la plus haute, $j = 2$ à la deuxième vague la plus haute, etc..

La hauteur $H_{1/3}$ est aussi nommée hauteur significative et est la plus utilisée car elle correspond à l'estimation visuelle des vagues obtenue par un observateur averti.

Période zero-crossing :

$$\bar{T}_0 = \frac{1}{N} \sum_{i=1}^N T_{0,i} \quad (2.4)$$

Période du tiers des vagues les plus hautes :

$$T_{1/3} = \frac{1}{N/3} \sum_{j=1}^{N/3} T_{0,j} \quad (2.5)$$

D'autres périodes caractéristiques des vagues sont utilisées, mais elles sont définies à partir du spectre d'état de mer. L'analyse spectrale, présentée dans la section 2.2.2, est l'approche la plus utilisée pour caractériser les états de mers. C'est par cette approche que les états de mers ont été caractérisés dans ces travaux.

2.2.2 Approche spectrale

Introduit par [Longuet-Higgins \[1957\]](#), le principe de l'approche spectrale correspond à une représentation de l'élévation de la surface libre en fonction du temps et de l'espace plan

de l'océan comme une superposition linéaire de vagues sinusoïdales monochromatiques et monodirectionnelles indépendantes les unes des autres (figure 2.3). L'état de mer est alors caractérisé par son spectre directionnel qui est la répartition de l'énergie suivant les fréquences et les directions.

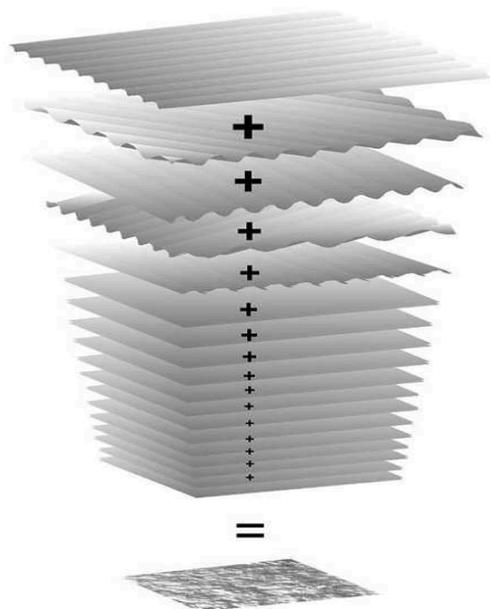


FIGURE 2.3 – Superposition linéaire de vagues sinusoïdales monochromatiques et monodirectionnelles indépendantes. Source : Pierson Jr et al. [1971]

2.2.2.1 Spectre directionnel d'un état de mer

D'après la théorie linéaire des vagues, l'élévation de la surface $z = \eta(\mathbf{x}, t)$ en un point de coordonnée \mathbf{x} dans le plan horizontal et à un instant t est donnée par

$$\eta(\mathbf{x}, t) = a \cos(\mathbf{k} \cdot \mathbf{x} - 2\pi f t + \varphi) \quad (2.6)$$

où f est la fréquence, \mathbf{k} le nombre d'onde et φ la phase. La fréquence est liée au nombre d'onde par la relation de dispersion

$$2\pi f = \sqrt{gk \tanh(kh)}, \quad (2.7)$$

avec $k = |\mathbf{k}|$, h la profondeur de l'eau. L'onde se propage dans la direction \mathbf{k}/k avec une longueur d'onde $2\pi/k$, une fréquence f et une amplitude a . La somme de toutes les vagues sinusoïdales monochromatiques et monodirectionnelles est

$$\eta(\mathbf{x}, t) = \sum_{n=1}^N a_n \cos[\mathbf{k}_n \cdot \mathbf{x} - 2\pi f_n t + \varphi_n] \quad (2.8)$$

où N est le nombre total de vagues. φ_n sont les phases supposées indépendantes et uniformément distribuées sur $[0, 2\pi]$. L'espérance et la variance d'une onde élémentaire sont

respectivement :

$$\mathbb{E} [a \cos (\mathbf{k} \cdot \mathbf{x} - \omega t + \varphi)] = \frac{1}{2\pi} \int_0^{2\pi} a \cos (\mathbf{k} \cdot \mathbf{x} - \omega t + \varphi) d\varphi = 0 \quad (2.9)$$

$$Var [a \cos (\mathbf{k} \cdot \mathbf{x} - \omega t + \varphi_n)] = \frac{1}{2\pi} \int_0^{2\pi} (a \cos (\mathbf{k} \cdot \mathbf{x} - \omega t + \varphi))^2 d\varphi = \frac{a^2}{2} \quad (2.10)$$

L'espérance et la variance étant des opérateurs linéaires, on obtient pour la somme des ondes :

$$\mathbb{E} [\eta(\mathbf{x}, t)] = 0, \quad (2.11)$$

$$Var (\eta(\mathbf{x}, t)) = \mathbb{E} [\eta^2(\mathbf{x}, t)] = \sum_{n=1}^N \frac{a_n^2}{2}. \quad (2.12)$$

En outre, l'énergie moyenne par unité de surface des vagues est

$$\bar{E} = \frac{\rho g}{8} \sum_{i=1}^N (2a_n)^2 \quad (2.13)$$

avec ρ la masse volumique de l'eau de mer et g la constante de gravité. D'où

$$\frac{\bar{E}}{2\rho g} = \sum_{i=1}^N a_n^2 = Var (\eta(\mathbf{x}, t)) \quad (2.14)$$

Les carrés des amplitudes a_n^2 sont donc reliés à l'énergie totale. De plus, il est possible de définir le spectre des amplitudes c'est à dire que la distribution des carrés des amplitudes peut être exprimée en fonction de la fréquence et de la direction. Lorsque $N \rightarrow \infty$ le spectre des amplitudes peut être transformé en un spectre continu $S(f, \theta)$,

$$S(f, \theta) = \lim_{\Delta f \rightarrow 0} \lim_{\Delta \theta \rightarrow 0} \frac{1}{\Delta f} \frac{1}{\Delta \theta} \frac{a_n^2}{2}, \quad (2.15)$$

$S(f, \theta)$ est appelé le spectre directionnel de variance d'état de mer.

Le spectre directionnel de variance peut être décomposé suivant le produit

$$S(f, \theta) = S(f)D(f, \theta) \quad (2.16)$$

où $S(f)$ est le spectre de variance omnidirectionnel (ou non-directionnel) et $D(f, \theta)$ la fonction de répartition angulaire. Elle est normalisée en sorte que

$$\int_0^{2\pi} D(f, \theta) d\theta = 1 \quad (2.17)$$

Notons au passage que le spectre directionnel de variance est relié au spectre directionnel d'énergie par la relation

$$E(f, \theta) = \rho g S(f, \theta) \quad (2.18)$$

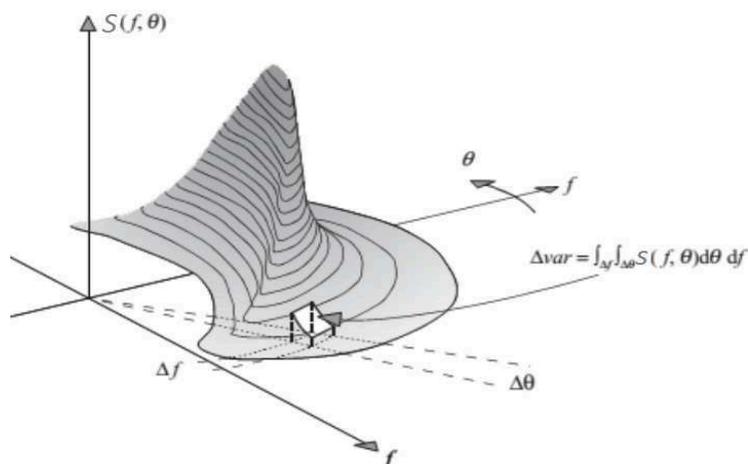


FIGURE 2.4 – Contribution Δvar de la portion spectrale contenue dans $(\Delta f, \Delta \theta)$ à la variance totale des vagues. Source : [Holthuijsen \[2007\]](#)

2.2.2.2 Paramètres spectraux et intégraux

Un certain nombre de paramètres caractéristiques d'états de mer peuvent être calculés à partir du spectre directionnel de variance. Loin d'être exhaustive, la liste de paramètres spectraux et intégraux que nous allons présenter contient principalement les paramètres qui interviennent dans le cadre de cette thèse. Pour une liste exhaustive des paramètres des états de mers, le lecteur pourra se référer au document du [IAHR, PIANC \[1986\]](#).

Hauteur significative : Cette hauteur correspond à l'impression visuelle donnée par la mer et constitue l'échelle de hauteur la plus utilisée pour les vagues aléatoires. Elle est définie par :

$$H_{m0} = 4 \left(\int_0^\infty \int_0^{2\pi} S(f, \theta) df d\theta \right)^{1/2} \quad (2.19)$$

En pratique $H_{1/3} \approx H_{m0}$, l'égalité étant démontrée dans la limite d'un spectre étroit [Longuet-Higgins \[1952\]](#). Par la suite nous prendrons H_{m0} comme la définition de H_s . L'indice "m0" signifie qu'il s'agit du moment d'ordre zéro du spectre. On définit par ailleurs le moment d'ordre p de la manière suivante,

$$m_p = \int_0^\infty \int_0^{2\pi} f^p S(f, \theta) df d\theta \quad (2.20)$$

Périodes caractéristiques : Les périodes caractéristiques les plus utilisées sont T_{-10} , T_{01} et T_{02} , elles donnent plus ou moins de poids à la partie basse fréquence du spectre.

La période T_{-10} encore appelée période énergétique (T_e) définie comme

$$T_{-10} = T_e = \frac{m_{-1}}{m_0}, \quad (2.21)$$

est souvent utilisée pour des applications liées à la récupération de l'énergie des vagues.

La période T_{02} est égale à période zero-up-crossing (T_z obtenue par l'analyse vague par vague) si on considère que l'élévation de la surface de la mer est un processus stationnaire et Gaussien :

$$T_{02} = \sqrt{\frac{m_0}{m_2}} \quad (2.22)$$

Elle est pour sa part utilisée dans le cas où on s'intéresse à un effet proportionnel à f^2 . C'est généralement le cas des forces exercées par les vagues sur les structures.

On distingue également la période moyenne T_{01} définie par

$$T_{01} = \frac{m_0}{m_1}, \quad (2.23)$$

et période de pic T_p , qui est l'inverse de la fréquence de pic f_p qui correspond à la fréquence à laquelle le spectre omnidirectionnel $S(f)$ atteint son maximum.

Paramètres directionnels : En définissant,

$$a_1(f) = \int_0^{2\pi} S(f, \theta) \cos \theta d\theta / \int_0^{2\pi} S(f, \theta) d\theta, \quad (2.24)$$

$$b_1(f) = \int_0^{2\pi} S(f, \theta) \sin \theta d\theta / \int_0^{2\pi} S(f, \theta) d\theta, \quad (2.25)$$

la direction moyenne des vagues de fréquence f est,

$$\theta_m(f) = \arctan \left(\frac{b_1(f)}{a_1(f)} \right) \quad (2.26)$$

et l'étalement directionnel, définit par [Kuik et al. \[1988\]](#) comme l'écart-type (en radians) de la largeur spectrale dans la limite du spectre étroit

$$\sigma_\theta(f) = \left[2\sqrt{1 - (a_1(f))^2 - (b_1(f))^2} \right]^{1/2}. \quad (2.27)$$

Les propriétés directionnelles des vagues sont généralement caractérisées par la direction moyenne $\theta_m(f_p)$ et l'étalement du pic spectral $\sigma_\theta(f_p)$. $\theta_m(f_p)$ est souvent appelé direction principale, alors que la direction moyenne sera plutôt une moyenne sur l'ensemble du spectre,

$$\theta_M = \arctan \left(\int_0^\infty b_1(f) S(f) df / \int_0^\infty a_1(f) S(f) df \right). \quad (2.28)$$

Paramètre d'acuité spectrale Q_p : Ce paramètre est aussi appelé paramètre de Goda qui l'introduisit en 1977 ([Goda \[1977\]](#)). Il définit la finesse du pic spectral et son expression est :

$$Q_p = \frac{2}{m_0^2} \int_{f_{min}}^{f_{max}} f S^2(f) df \quad (2.29)$$

Puissance des vagues P_w : La puissance des vagues est définie par

$$P_w = \rho g \int \int c_g(f, h) S(f, \theta) df d\theta \quad (2.30)$$

où $c_g(f, h)$ est la vitesse de groupe des vagues de fréquence f dans une eau de profondeur h . On peut montrer que la vitesse de groupe en eau profonde ($h \rightarrow \infty$) est donnée par $c_g(f, \infty) = g/(4\pi f)$. Ainsi, une expression approchée (en kWm^{-1}) en eau profonde de la puissance des vagues est

$$P_w = \frac{\rho g^2}{64\pi} H_s^2 T_e \approx 0.49 H_s^2 T_e \quad (2.31)$$

C'est le flux de d'énergie qui traverse $1m$ en largeur de crête de vague perpendiculairement à la direction de propagation.

2.2.2.3 Modèles de spectres en fréquence $S(f)$

Les premières mesures de spectres furent réalisées en 1944 par le Group W de l'Amirauté Britannique, après le débarquement en Normandie. Ces observations et celles qui suivirent firent ressortir que, pour les fréquences au-delà du pic, le spectre a généralement toujours la même forme. Phillips [1958] introduisit la notion de zone d'équilibre pour décrire la forme du spectre pour $f > f_p$, et proposa que seule la gravité déterminait la forme du spectre. Par analyse dimensionnelle, cela le conduisit à proposer la forme suivante

$$S(f) = \alpha (2\pi)^{-4} g^2 f^{-5}, \quad (2.32)$$

avec $\alpha = 0.008$ la constante de Phillips. L'énergie adimensionnelle des vagues est constante dans le modèle de Phillips : la surface est donc fractale et les petites vagues ont statistiquement la même forme que les grandes vagues. C'est ce type de concept qui conduisit Pierson, Moskowitz [1964] à proposer la forme suivante pour synthétiser leurs observations de mers pleinement développées, sur l'ensemble de la gamme de fréquences (voir figure 2.5),

$$S(f) = \alpha (2\pi)^{-4} g^2 f^{-5} \exp \left[-\frac{5}{4} \left(\frac{f_p}{f} \right)^{-4} \right]. \quad (2.33)$$

Il est apparu ultérieurement que pour les fetchs limités, la forme du spectre pouvait être assez différente. En particulier le pic est plus étroit par fetch court. Par ailleurs, les valeurs de $S(f)$ pour f fixé, peuvent être plus élevées que les valeurs observées pour une mer pleinement développée : il s'agit du phénomène de sur-croissance du pic, observée par Barnett, Sutherland [1968] et particulièrement mis en évidence par l'expérience de 1968-1969 du Joint North Sea Wave Project (JONSWAP, Hasselmann et al. [1973]). Un terme d'amplification du pic fut donc rajouté,

$$S(f) = S_{PM}(f) \gamma \exp \left[\frac{-(f-f_p)^2}{2\sigma_{A/B}^2 f_p^2} \right]. \quad (2.34)$$

γ est le facteur d'amplification du pic, tandis que $\sigma_{A/B}$ qui prend pour valeur $\sigma_A \approx 0.07$ si $f < f_p$ et $\sigma_B \approx 0.09$ sinon, donne la largeur relative autour de f_p qui est affectée par l'amplification du pic. La figure 2.6 montre le spectre JONSWAP pour un état de mer de $H_s = 2.5 m$ et $T_p = 11 s$ avec différentes valeurs de γ .

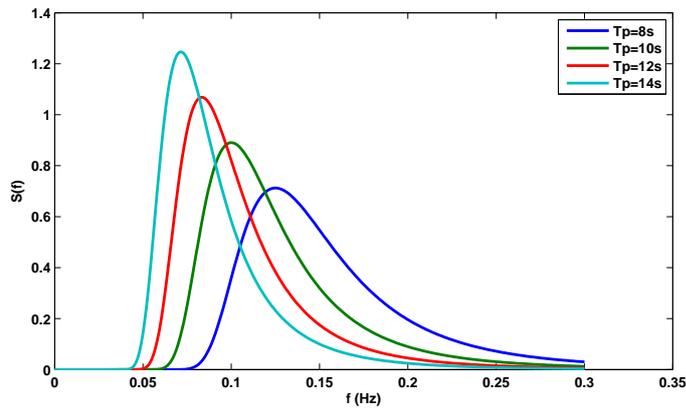


FIGURE 2.5 – Spectre de Pierson-Moskowitz pour un état de mer de $H_s = 2.5 \text{ m}$ et différentes valeurs de T_p .

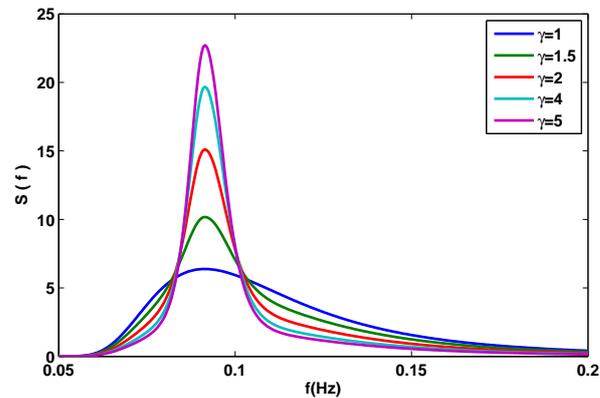


FIGURE 2.6 – Spectre JONSWAP pour différentes valeurs de γ avec $H_s = 2.5\text{m}$ et $T_p = 11\text{s}$

2.2.2.4 Modèles de fonctions de répartition directionnelle $D(f, \theta)$

Généralement on préfère ajuster une forme paramétrique à la fonction de répartition directionnelle. On trouve diverses formes paramétriques directionnelles dans la littérature, par exemple chez Benoit et al. [1997] ou Krogstad et al. [1999] : \cos^{2s} , gaussienne, wrapped-normal, Poisson, etc., chacune d'elles pouvant être déterminée à partir de la direction et de l'étalement moyens de l'état de mer. La formulation souvent utilisée est la distribution \cos^{2s} (Longuet-Higgins et al. [1963]) et sa forme qui présente le paramètre s en fonction de la fréquence (Mitsuyasu et al. [1975]). Le tableau (2.1) présente les différentes formes paramétriques de la fonction directionnelle. Une représentation graphique de ces différentes fonctions de répartition directionnelle est présentée à la figure (2.7).

Distribution	Expression de $D(f, \theta)$	Valeurs du paramètre
cos-k	$\frac{\Gamma(k+1)}{2^k \Gamma^2(\frac{k+1}{2})} \cos^k(\theta), \quad \theta \leq \frac{\pi}{2}, \quad 0$ ailleurs.	$0 < k$
cos-2s	$\frac{\Gamma(s+1)}{2\sqrt{\pi}\Gamma(s+1/2)} \cos^{2s}(\frac{\theta}{2})$	$0 < s$
Wrapped normal	$\frac{1}{2\pi} \left[1 + 2 \sum_{n=1}^{\infty} x^{n^2} \cos(n\theta) \right]$	$0 < x < 1$
Sech-2	$\frac{1}{\tanh(\beta\pi)} \frac{\beta}{2 \cosh^2(\beta\theta)}$	$0 < \beta$
Poisson	$\frac{1}{2\pi} \frac{1-x^2}{1-2x\cos(\theta)+x^2}$	$0 < x < 1$
von Mises	$\frac{1}{2\pi I_0(k)} e^{k \cos(\theta)}$	$0 < k$
Boxcar	$\frac{1}{2a}, \quad \theta \leq a, \quad 0$ ailleurs	$0 < a < \pi$
Triangulaire	$\frac{1}{2a^2} (a - \theta), \quad 0 \leq \theta \leq a, \quad 0,$ ailleurs	$0 < a < \pi/2$

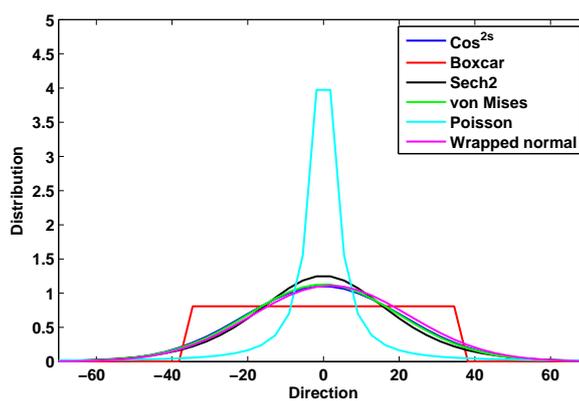
TABLE 2.1 – Quelques familles de distributions directionnelles. Source [Hauser et al. \[2005\]](#)

FIGURE 2.7 – Modèles des fonctions de répartition directionnelle.

2.3 Statistiques à court-terme des états de mer

L'hypothèse faite lorsqu'on étudie les caractéristiques des états de mer à court-terme, est que l'élévation de la surface de la mer est un processus stationnaire et Gaussien :

$$p(\eta) = \frac{1}{\sqrt{2\pi m_0}} \exp\left(-\frac{\eta^2}{2m_0}\right) \quad (2.35)$$

où $p(\eta)$ est la densité de probabilité gaussienne d'espérance $\mathbb{E}[\eta] = 0$ et d'écart type $\sigma = m_0^{1/2}$. Cette hypothèse est raisonnable pour des enregistrements d'élévation de la surface d'une durée de 15 à 30 min mais elle est parfois faite pour des durées de 3h, voire 6h (Holthuijsen [2007]).

Les statistiques à court-terme des états de mer sont utilisées pour des applications comme la résistance des structures dans les conditions extrêmes de vagues individuelles, ou la gestion d'opérations en mer. Les paramètres environnementaux qui sont souvent impliqués dans ce type de statistique sont la hauteur de vague individuelle, les hauteurs de crêtes de vagues ou encore les groupements de vagues.

2.3.1 Distributions des crêtes et des hauteurs de vagues individuelles

La distribution des maxima d'un processus Gaussien suit la loi de Rice [1944] de densité de probabilité :

$$p(\sigma, \varepsilon, \eta) = \frac{\varepsilon}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\eta^2}{2\sigma^2\varepsilon^2}\right) + \frac{\alpha\eta}{\sigma^2} \exp\left(-\frac{\eta^2}{2\sigma^2}\right) \Phi\left(\frac{\alpha\eta}{\varepsilon\sigma}\right) \quad (2.36)$$

où le facteur d'irrégularité α est définie par :

$$\alpha = \sqrt{1 - \varepsilon^2}. \quad (2.37)$$

Les paramètres ε et σ , respectivement largeur spectrale et écart type peuvent être obtenus à partir des moments du processus :

$$\varepsilon = \sqrt{1 - \frac{m_2^2}{m_0 m_4}} \quad (2.38)$$

et

$$\sigma = m_0. \quad (2.39)$$

Sous l'hypothèse de spectre à bande étroite : $\varepsilon \rightarrow 0$, les hauteurs des crêtes de vagues η_c suivent alors une loi de Rayleigh (figure 2.8),

$$p_{\eta_c}(\eta) = \frac{\eta}{m_0} \exp\left(-\frac{\eta^2}{2m_0}\right), \eta \geq 0. \quad (2.40)$$

La hauteur de vague individuelle H est approximativement égale au double de la hauteur de la crête η_c sous l'hypothèse de processus Gaussien : $H = 2\eta_c$, et ainsi

$$p_H(h) = p_{\eta_c}(\eta) \frac{d\eta}{dh} \quad (2.41)$$

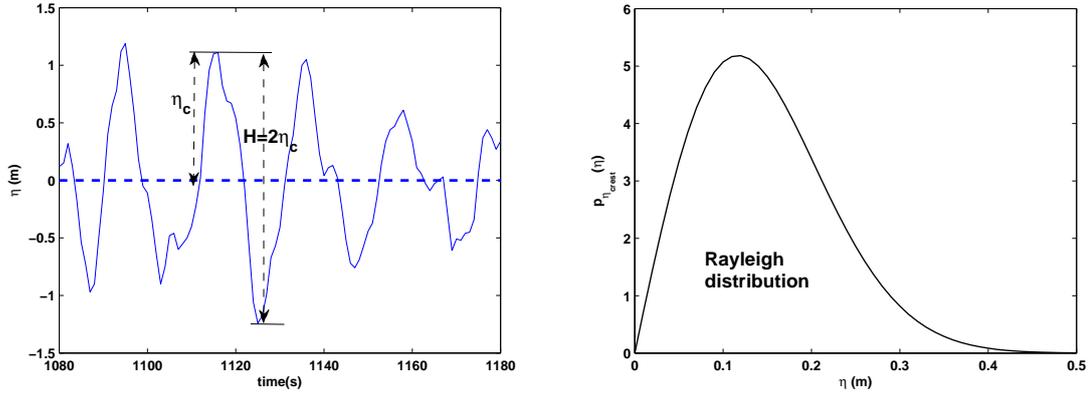


FIGURE 2.8 – Pour des conditions de vagues de spectres à bande étroite, la hauteur H est approximativement égale au double de la hauteur de la crête η_c (gauche) ; les hauteurs des crêtes suivent une loi de Rayleigh (droite).

$$p_H(h) = \frac{h}{4m_0} \exp\left(-\frac{h^2}{8m_0}\right), h \geq 0. \quad (2.42)$$

Si la distribution de Rayleigh est généralement satisfaisante en eau profonde, elle l'est moins en eau peu profonde à cause des non-linéarités. En effet, les enregistrements d'élévations de la surface libre montrent que les creux et les crêtes des vagues ne sont pas symétriques ; les crêtes sont plus pointues, les creux plus aplatis. La loi de Rayleigh sous-estime donc la distribution des crêtes de vagues [Forristall \[2000\]](#). Pour tenir compte des non-linéarités, différents auteurs ont proposé des modèles de distributions des crêtes de vagues qui sont pour la plupart basés sur la perturbation des lois dans le cas linéaire. Dans la section suivante nous présenterons deux modèles de distributions de crêtes de vagues non-linéaires.

2.3.2 Distributions des crêtes de vagues non-linéaires

Une liste quasi exhaustive des modèles existants est rapportée dans [Prevosto \[2001\]](#). Nous présenterons ici le modèle de [Forristall \[2000\]](#) et le modèle de [Prevosto \[2001\]](#).

2.3.2.1 Modèle de Forristall

Forristall utilise la théorie de second-ordre pour modéliser les non-linéarités des vagues. Dans son modèle, la probabilité qu'une crête de vague soit plus grande que η dans l'état de mer de hauteur significative H_s est

$$P(\eta_c > \eta) = \exp\left[-\left(\frac{\eta}{\alpha H_s}\right)^\beta\right]. \quad (2.43)$$

Les paramètres α et β sont données respectivement par :

$$\alpha = \frac{1}{\sqrt{8}} + 0.2568S_1 + 0.0800U_r \quad (2.44)$$

$$\beta = 2 - 1.7912S_1 - 0.5302U_r + 0.2824U_r^2 \quad (2.45)$$

où,

$$S_1 = \frac{2\pi}{g} \frac{H_s}{T_{m01}^2} \quad (2.46)$$

est la cambrure des vagues, et

$$U_r = \frac{H_s}{k_1^2 d^3} \quad (2.47)$$

le nombre de Ursell pour une profondeur d , avec k_1 le nombre d'onde correspondant à T_{m01}

2.3.2.2 Modèle de Prevosto

Le modèle de Prevosto est basé sur la transformation non-linéaire de la loi de Rayleigh. Le modèle s'écrit :

$$P(\eta_c > \eta | H_s; f_m; d) = \exp \left(-8 \frac{\left(\frac{-1 + \sqrt{1 + 4\alpha(\eta - \beta)}}{2\alpha} \right)^2}{H_s^2} \right) \quad (2.48)$$

Une relation quadratique est utilisée pour lier les crêtes linéaires et non-linéaires :

$$\eta = \eta_{lin} + (T_{nb}^D(f_m) + T_{nb}^S(f_m))\eta_{lin}^2 - T_{nb}^D(f_m) \frac{H_s^2}{8} \quad (2.49)$$

Pour tenir compte de la largeur spectrale et de l'étalement directionnel, Prevosto a considéré :

$$\widetilde{H}_s = \alpha_{H_s} H_s \quad (2.50)$$

$$\widetilde{f}_m = \alpha_{f_m} f_m \quad (2.51)$$

Les expressions de α_{H_s} et de α_{f_m} , obtenues à partir des simulations et des considérations théoriques sont respectivement données par :

$$\alpha_{h_s} = 1 - \frac{1}{2} (\tanh(k_m d) - 0.9) \sqrt{\frac{2}{1+s}} \quad (2.52)$$

$$\alpha_{f_m} = \frac{1}{1.23} \quad (2.53)$$

avec $f_m = \frac{1}{T_{02}}$ et s le paramètre de la fonction directionnelle \cos^{2s} . Les expressions des coefficients de transfert T_{nb}^S et T_{nb}^D sont complexes et détaillées dans Prevosto [2001].

La figure 2.9 montre une série-temporelle de l'élévation de la surface libre enregistrée au large de l'Angola. La hauteur significative de l'état de mer correspondant à cet enregistrement est $H_s = 2.35$ m avec une période moyenne $T_{m01} = 12.63$ s. Pour cet enregistrement, la probabilité d'avoir la hauteur de crête de vague supérieure à un seuil

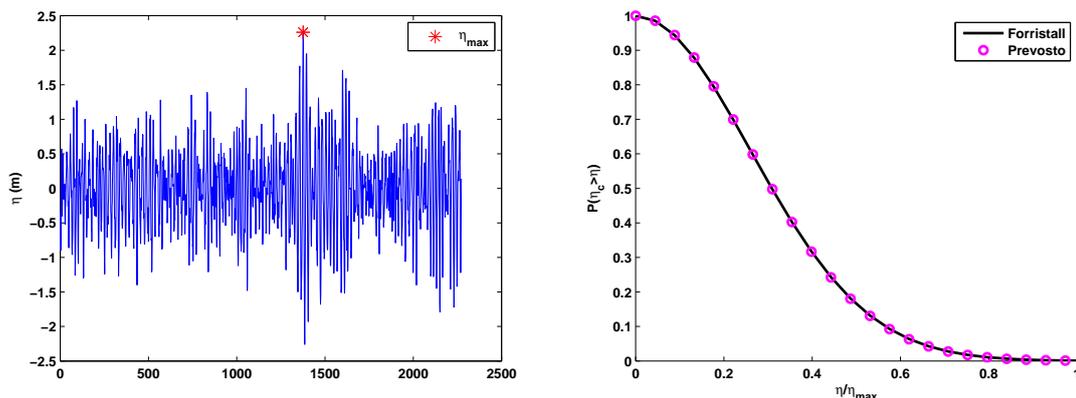


FIGURE 2.9 – 30 min de série temporelle d’élévation de la surface de la mer enregistrée au large de l’Angola ($H_s = 2.35$ m, $T_{m01} = 12.63$ s) (gauche) ; Probabilité de dépassement de seuil des crêtes de vagues avec les modèles de forristall et de Prevosto (droite)

$p(\eta_c > \eta)$ est calculée en utilisant les modèles de Forristall et de Prevosto. Les deux modèles donnent approximativement les mêmes distributions. Ce résultat a été également obtenu par Prevosto, Forristall [2002]. La différence principale entre le modèle de Forristall et celui de Prevosto est la prise en compte de l’étalement directionnel dans le modèle de Prevosto. Le modèle de Forristall est plus simple à implémenter mais le modèle de Prevosto est plus précis dans le cas des crêtes très courtes ou dans le cas des houles (Prevosto, Forristall [2002]).

2.4 Statistiques à long-terme (climatologie) des états de mer

La climatologie des états de mer est la modélisation statistique à long-terme (plusieurs années) des paramètres d’états de mer pour une région donnée. La climatologie des états de mer se distingue donc de la description statistique à court-terme où on suppose les conditions des états de mer stationnaires, généralement pour une durée de 20 min à quelques heures. Pour les périodes de temps concernées dans la climatologie, il n’est pas possible de présenter les vagues comme une série-temporelle de leur élévations. En revanche, chaque série stationnaire d’état de mer est représentée par ses paramètres caractéristiques (hauteur significative, période, direction, étalement fréquentiel, étalement directionnel). Cela procure une séquence des paramètres caractéristiques avec un pas de temps d’une à trois heures, qui peut alors être analysée pour estimer les caractéristiques statistiques à long-terme des vagues (figure 2.10).

La climatologie des états de mer est primordiale dans de nombreuses applications en génie océanique et en particulier :

- Estimation de la probabilité d’occurrences de l’ensemble des états de mer pour les calculs de dimensionnement à la fatigue des structures offshore.
- Estimation de la probabilité des états de mer extrêmes pour le dimensionnement des structures marines contre les chargements extrêmes.
- Estimation du potentiel d’énergie récupérable des vagues.

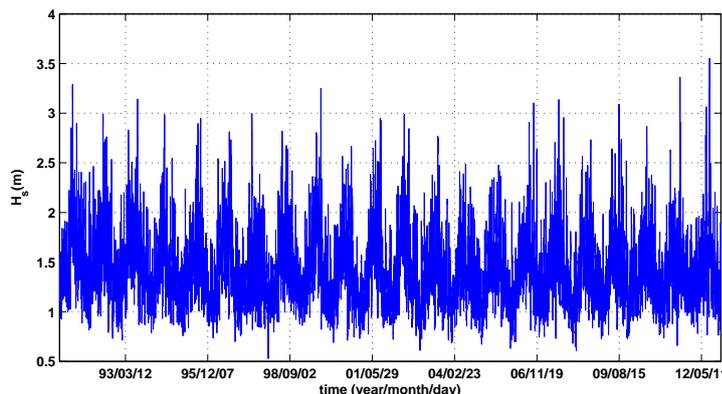


FIGURE 2.10 – 21 ans de données de H_s (Hindcast) au large de l'Angola. Les données sont échantillonnées au pas de 3 h.

- Estimation de la probabilité des fenêtres météorologiques (weather windows) à diverses périodes de l'année pour les opérations marines (installation des plates-formes offshore, des convertisseurs d'énergie de vagues, maintenance des systèmes en mer, etc.).
- Estimation de l'évolution morphodynamique du trait de côte, etc.

Deux types de méthodes sont généralement utilisées, dans la modélisation statistique à long-terme des états de mer, en fonction des applications : la méthode par distribution empirique globale (initial distribution method) et la méthode par événement ou par analyse de tempête (event method, storm method ou Peak Over Threshold method). La première utilise l'ensemble des données observées et est, de ce fait, adaptée aux applications comme le dimensionnement à la fatigue, alors que la deuxième utilise seulement les observations de hauteurs significatives de vagues au dessus d'un certain seuil et est, par conséquent, plus adaptée à l'analyse des extrêmes.

2.4.1 Utilisation de la distribution empirique globale

Dans cette méthode, il est question de déterminer la fonction de densité de probabilité jointe de la hauteur significative H_s , de la période zero-crossing T_{02} et éventuellement de la direction moyenne θ_0 . Les résultats sont présentés dans un diagramme (H_s, T_{02}) appelé scatter diagram, par secteur de direction (généralement $\Delta\theta = 45^\circ$). Ce diagramme peut être représenté par saison ou par mois (en moyennant sur un grand nombre d'années).

L'analyse peut être parfois réduite à la hauteur significative qui est souvent considérée comme le paramètre le plus important. Dans ce cas, l'analyse consiste à chercher une loi marginale de la distribution des hauteurs significatives. Cependant, il est préférable, à cause de la dépendance entre les paramètres H_s et T_{02} , de déterminer une loi de distribution jointe de ces deux paramètres. Il existe dans les deux cas des modèles que l'on ajuste aux données réelles. Dans ce qui va suivre, nous présenterons les modèles généralement utilisés de la distribution marginale de la hauteur significative $p(H_s)$ (analyse univariée), et de la distribution jointe de la hauteur significative et de la période moyenne $p(H_s, T_{02})$ (analyse bivariée). Nous avons également comparé, pour les deux types d'analyses, les résultats

obtenus en appliquant les modèles aux données de H_s et/ou T_{02} mesurées au large de l'Angola.

2.4.1.1 Analyse univariée : distribution marginale de la hauteur significative

Les deux modèles les plus utilisés pour la distribution marginale du H_s sont la distribution Log-normale et la distribution de Weibull à 2 ou 3 paramètres. Il n'y a pas de justification physique quant au choix du modèle. Les paramètres des modèles de distribution sont obtenus en ajustant ces derniers aux données réelles. Plusieurs techniques d'ajustements tels que la méthode des moments, la méthode des moindres carrés ou encore la méthode du maximum de vraisemblance peuvent être utilisées. Les ajustements sont finalement validés au moyen des tests statistiques notamment le test de χ^2 (Moore [1976]), le test de Kolmogorov-Smirnov (Massey Jr [1951]), ou encore le test de Anderson-Darling (Anderson, Darling [1954]).

La distribution Log-normale : En analysant des données de H_s de 11 stations dans l'Atlantique Nord fournies par le WBNA (Weather Bureau Data for North Atlantic) et d'autres données provenant du Pacifique, Jasper [1956], sous l'hypothèse d'une distribution de Rayleigh des hauteurs de vagues individuelles, montre que la fonction de densité de probabilité Log-normale peut être utilisée pour ajuster la distribution statistique à long-terme du H_s . La distribution Log-normale s'écrit

$$p_{H_s}(h) = \frac{1}{\sqrt{2\pi}\sigma h} \exp\left(-\frac{(\ln h - \mu)^2}{2\sigma^2}\right), 0 \leq h < \infty \quad (2.54)$$

où σ et μ sont les paramètres de la loi Log-normale classique déterminés à partir des données.

La distribution de Weibull : La fonction de distribution de Weibull à 3 paramètres s'écrit (Nordenström [1973]) :

$$p_{H_s}(h) = \frac{\beta(h-\gamma)^{\beta-1}}{\alpha^\beta} \exp\left[-\left(\frac{h-\gamma}{\alpha}\right)^\beta\right], 0 \leq h < \infty \quad (2.55)$$

où α est le paramètre d'échelle, β le paramètre de forme, et γ paramètre de position. La loi de Weibull à 2 paramètres correspond au cas $\gamma = 0$.

2.4.1.2 Étude comparative des modèles de distribution marginale du H_s appliqués aux données mesurées au large de l'Angola

À titre d'exemple sur l'étude de la sensibilité des modèles de distribution marginale, les trois modèles (Log-normale, Weibull à 2 paramètres et Weibull à 3 paramètres) ont été ajustés aux données de H_s obtenues au large des côtes Angolaises. Les données ont été recueillies à l'aide d'une bouée ancrée par 1450 m de profondeur entre mars 2001 et mars 2003. Les modèles ont été ajustés avec la méthode du maximum de vraisemblance, et la qualité des ajustements a été évaluée par jugement visuel et en utilisant les statistiques de χ^2 , et de Kolmogorov-Smirnov.

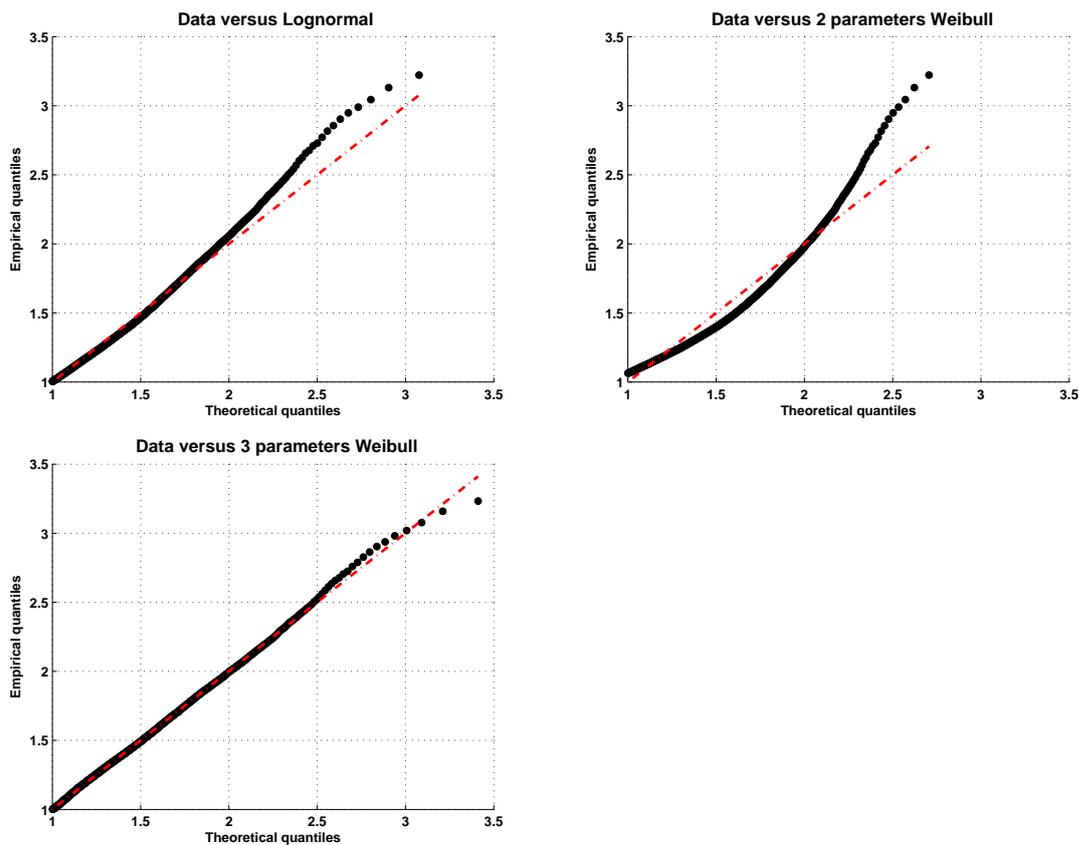


FIGURE 2.11 – Ajustement des modèles Log-normale, Weibull à 2 paramètres et Weibull à 3 paramètres aux données de H_s mesurées au large de l'Angola

La figure 2.11 montre les résultats des différents ajustements. La loi de Weibull à 2 paramètres présente le résultat le moins satisfaisant ; l'ensemble de l'ajustement est de mauvaise qualité. La qualité de l'ajustement est meilleure pour le modèle Log-normale, sauf pour les fortes valeurs du H_s qui sont nettement sous-estimées. Ce résultat a aussi été rapporté par des auteurs tels que Soares et al. [1988], Teng, Palao [2001]. La loi de Weibull à 3 paramètres donne le meilleur résultat ; elle ajuste mieux l'ensemble des valeurs du H_s comparée aux deux autres modèles. Ces résultats sont confirmés par les statistiques de χ^2 , et de Kolmogorov-Smirnov dont les valeurs sont contenues dans le tableau 2.2. On remarque en effet que pour l'ensemble des deux statistiques, le modèle de Weibull à trois paramètres affiche les valeurs les plus petites, traduisant ainsi un meilleur ajustement. Il faut, toutefois, noter que la qualité de l'ajustement n'est pas bonne pour les valeurs extrêmes de H_s . En effet, il est difficile de trouver un modèle qui ajuste correctement à la fois les valeurs extrêmes et les autres valeurs du H_s . Il existe des modèles spécifiques pour les valeurs extrêmes et certains de ces modèles seront présentés dans la sous-section 2.4.2.

	Log-normale	Weibull2	Weibull3
χ^2	112	291	7
$K-S$	0.0375	0.0962	0.0097

 TABLE 2.2 – Valeurs des statistiques de χ^2 , et de Kolmogorov-Smirnov (K-S)

2.4.1.3 Analyse bivariée : distribution jointe de la hauteur significative et de la période zero-crossing

La hauteur significative et la période zero-crossing ne sont pas souvent indépendantes et de nombreuses applications en génie océanique nécessitent de modéliser aussi la structure de dépendance entre les deux paramètres. Comme les lois marginales des deux paramètres, seules, ne permettent pas de tenir compte de leur dépendance mutuelle, on a souvent recours aux lois jointes. Il existe plusieurs méthodes qui permettent de déterminer la loi jointe de H_s et T_p :

- utilisation des lois bivariées
- utilisation de la loi marginale de H_s avec une loi conditionnelle de T_{02}
- utilisation des lois marginales de H_s et de la cambrure S combinée à un modèle physique de T_{02} comme une fonction de H_s et S
- utilisation des lois marginales de H_s et T_{02} et d'une transformation isoprobabiliste (transformation de Nataf, transformation de Rosenblatt)

Afin de simplifier les notations dans les équations qui vont suivre, nous désignerons par x la période zero-crossing T_{02} et par y la hauteur significative H_s .

Cas 1 : utilisation des lois bivariées : Les principales lois bivariées que l'on trouve dans la littérature sont : la loi Log-normale bivariée [Ochi \[1978\]](#) et sa forme modifiée [Fang, Hogben \[1982\]](#), et la loi de Weibull bivariée [Gran \[1992\]](#).

La loi Log-normale bivariée est définie par

$$p(x, y) = \frac{0.5}{xy\pi\sqrt{1 - \alpha_{xy}^2\epsilon_x\epsilon_y}} \exp \left\{ -\frac{0.5}{1 - \alpha_{xy}^2} \left[\frac{(\ln x - \xi_x)^2}{\xi_x^2} - \frac{2\alpha_{xy}(\ln x - \xi_x)(\ln y - \xi_y)}{\epsilon_x\epsilon_y} + \frac{\ln(y - \xi_y)^2}{\epsilon_y^2} \right] \right\} \quad (2.56)$$

où $\xi_x, \xi_y, \epsilon_x, \epsilon_y, \alpha_{xy}$ sont les paramètres de la distribution et sont définis par $\xi_x = E[\ln x]$, $\epsilon_x^2 = Var[\ln x]$, $\xi_y = E[\ln y]$, $\epsilon_y^2 = Var[\ln y]$, $\alpha_{xy} = Cov[\ln x, \ln y]/\epsilon_x\epsilon_y$. La forme modifiée de cette loi ([Fang, Hogben \[1982\]](#)) est identique à la fonction Log-normale bivariée à l'exception de l'addition d'un terme multiplicatif pour tenir compte de la dissymétrie (skewness) du H_s .

La loi de Weibull bivariée quant à elle s'écrit

$$p(x, y) = \frac{\gamma y^{\gamma-1}}{\eta^\gamma} \exp \left\{ -\left(\frac{y}{\eta}\right)^\gamma \right\} \frac{\beta x^{\beta-1}}{\alpha^\beta} \exp \left\{ -\left(\frac{x}{\alpha}\right)^\beta \right\} \quad (2.57)$$

avec $\alpha, \beta, \gamma, \eta$ les paramètres de la distribution.

Cas 2 : utilisation de la loi marginale de H_s avec une loi conditionnelle de T_{02} : Cette méthode a été proposée par [Mathisen, Bitner-Gregersen \[1990\]](#) et est aussi connue

sous le nom de "Conditional Modeling Approach (CMA)". Elle combine une loi de Weibull à 3 paramètres pour le H_s avec une loi conditionnelle pour T_{02}

$$p_{H_s}(y) = \frac{\beta_y}{\alpha_y} \left(\frac{y - \gamma_y}{\alpha_y} \right)^{\beta_y - 1} \exp \left\{ - \left(\frac{y - \gamma_y}{\alpha_y} \right)^{\beta_y} \right\}. \quad (2.58)$$

La loi conditionnelle est donnée par une fonction Log-normale

$$p_{T_{02}|H_s} = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left\{ - \frac{(\ln x - \mu)^2}{2\sigma^2} \right\} \quad (2.59)$$

où μ et σ sont fonctions de H_s :

$$\mu = E[\ln x] = a_0 + a_1 y^{a_2} \quad (2.60)$$

$$\sigma = std[\ln x] = b_0 + b_1 e^{b_2 y} \quad (2.61)$$

Les coefficients a_i, b_i ; $i = 0, 1, 2$ sont estimés à partir des données.

Cas 3 : utilisation des lois marginales de H_s et de la cambrure S combinée avec un modèle physique de T_{02} comme une fonction de H_s et S : La méthode a été proposée par [Vrijling, Bruinsma \[1980\]](#) et elle est basée sur l'hypothèse d'indépendance entre la hauteur significative et la cambrure. En eau profonde, d'après la théorie linéaire des vagues, la cambrure des vagues est définie par

$$S = \frac{2\pi H_s}{g T_{02}^2} \quad (2.62)$$

La densité de probabilité jointe de H_s et T_{02} est obtenue par transformation de la densité de probabilité jointe de H_s et de S . La fonction jointe est alors donnée par

$$p(T_{02}, H_s) = f(H_s, S) |J| = f(H_s) f(S) \frac{4\pi H_s}{g \left(\frac{2\pi H_s}{g S} \right)^{3/2}} \quad (2.63)$$

où J est le Jacobien de la transformation.

Cas 4 : utilisation des lois marginales de H_s et T_{02} et de la transformation de Nataf : Si les lois marginales de H_s et de T_{02} sont connues, nous pouvons alors construire une loi jointe en utilisant la transformation de [Nataf \[1962\]](#). La densité de probabilité jointe de H_s et de T_{02} peut s'écrire :

$$p(x, y) = \frac{f_x(x) f_y(y)}{\phi(\Phi^{-1}(F_x(x))) \phi(\Phi^{-1}(F_y(y)))} \phi_{xy} \left[\Phi^{-1}(F_x(x), F_y(y)), \rho \right] \quad (2.64)$$

où $f_x(x)$ et $f_y(y)$ sont respectivement les fonctions des densités marginales de T_{02} et H_s , $\Phi(\cdot)$ est la fonction de répartition normale, ϕ_{xy} est la densité jointe de la loi normale bivariée, et ρ est la matrice de corrélation définie positive estimée à partir des densités marginales.

En raison de la modélisation simplifiée de dépendance entre les variables cette approche doit être utilisée avec prudence (Nestegård et al. [2006], Bitner-Gregersen, Hagen [1999]). En effet, la dépendance entre les deux variables est uniquement modélisée par le coefficient de corrélation. Cependant, cette approche donne généralement des résultats satisfaisants et présente, surtout, l'avantage d'être facilement généralisable à plus de deux variables, ce qui est très difficile à faire avec les autres méthodes (Sagrilo et al. [2011]).

2.4.1.4 Étude comparative des modèles de distribution jointe de H_s et T_{02} appliqués aux données mesurées au large de l'Angola

Nous nous proposons de modéliser la distribution jointe de H_s et T_{02} au large de l'Angola en utilisant les différentes approches présentées et de comparer les résultats par rapport à la distribution jointe empirique. Les données utilisées sont les mêmes que celles présentées dans la Section 2.4.1.2 (données de bouée). Nous avons sélectionné les approches présentées dans les trois premiers cas :

- loi Log-normale bivariée
- loi marginale de H_s (Weibull à 3 paramètres) avec une loi conditionnelle de T_{02} (Log-normale) (CMA)
- loi marginale de H_s (Weibull) et de la cambrure (Weibull) combinée à un modèle physique

Les contours logarithmiques de la distribution jointe empirique de H_s et T_{02} et des distributions jointes obtenues avec les trois approches sont présentées dans la figure 2.12. La distribution jointe obtenue avec le CMA est celle qui se rapproche le plus de la distribution jointe empirique. Exceptée une légère sous-estimation de la probabilité jointe des couples (H_s, T_{02}) aux fortes valeurs de H_s , le CMA restitue bien la distribution jointe de H_s et de T_{02} . De précédentes études (Mathisen, Bitner-Gregersen [1990] et Teng, Palao [2001]) utilisant respectivement des données provenant de la mer du Nord et du Nord Pacifique ont également montré que le CMA donne le meilleur résultat. Le CMA est aussi la méthode recommandée par le Det Norske Veritas (DNV) (Nestegård et al. [2006]). La loi Log-normale bivariée surestime clairement les probabilités jointes des états de mer les plus cambrés. En conséquence, elle n'est recommandée pour des applications liées aux vagues extrêmes. La dernière approche, qui utilise les lois marginales de H_s et de la cambrure, présente l'avantage d'être plus facile à utiliser. Cependant, les autres approches donnent, nettement, de meilleurs résultats. En effet, dans cette approche, on fait l'hypothèse que H_s et la cambrure sont indépendants, ce qui n'est généralement pas vérifié.

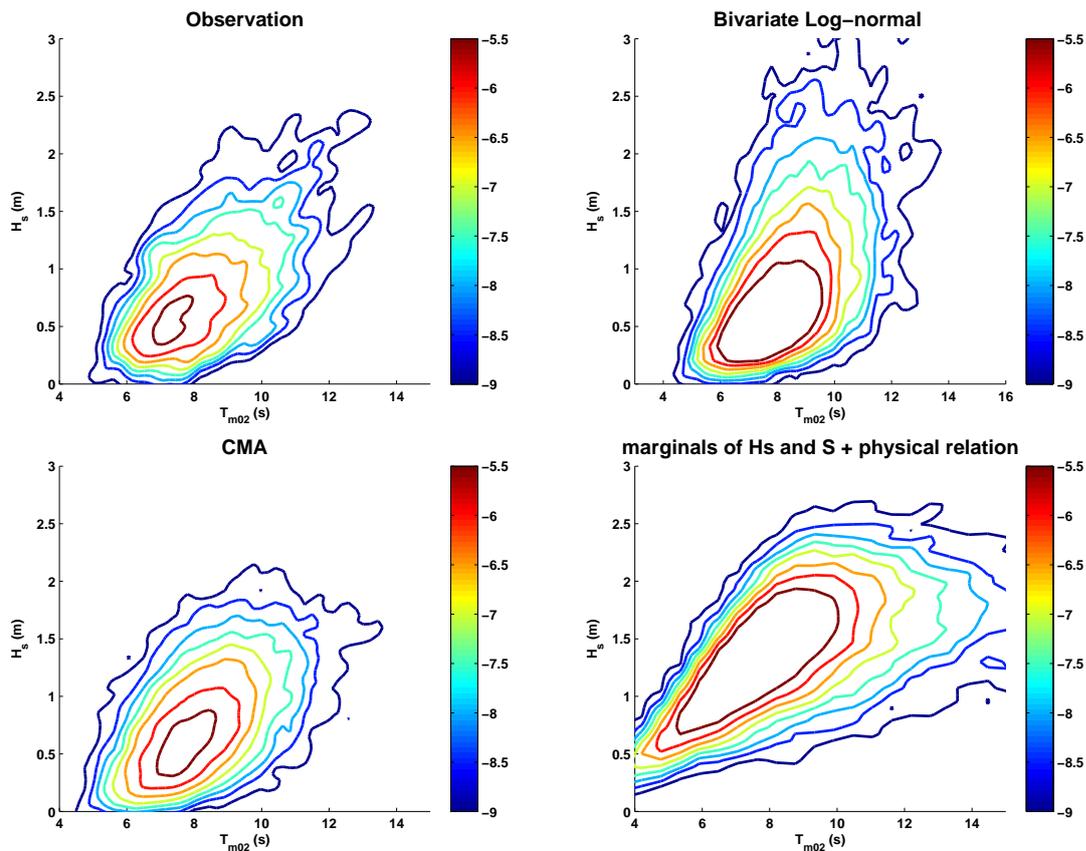


FIGURE 2.12 – Contours logarithmiques de distribution jointe de H_s et de T_{02} au large de l'Angola.

2.4.2 Estimation des extrêmes : approche par événement ou Peak Over Threshold method (POT)

Cette méthode consiste d'abord à identifier des tempêtes à partir de la série-temporelle de H_s , puis à analyser statistiquement ces tempêtes comme des événements indépendants. C'est la méthode qui est généralement utilisée dans l'analyse des états de mer extrêmes et est recommandée par Mathiesen et al. [1994] comme la méthode la plus adaptée pour déterminer les hauteurs significatives extrêmes de vagues. Par rapport à la première méthode qui utilise la distribution globale, le POT présente deux avantages quant à la description statistique à long-terme des états de mer extrêmes :

- si les données affichent plus d'une distribution statistique due à la présence de différents régimes physiques, la sélection des maxima de H_s uniquement dans les tempêtes permet de concentrer l'analyse sur le régime physique qui domine dans les états de mer extrêmes.
- les tempêtes sont statistiquement des événements indépendants, cela procure une base théorique solide et simplifie l'interprétation des résultats.

Les événements sur lesquels nous baserons notre modélisation stochastique de la climatologie seront conformes à la définition d'événements indépendants utilisée dans les méthodes POT à la différence qu'ils seront identifiés en distinguant les houles des mers

du vents (chapitre 3).

2.4.2.1 Définition d'une tempête

Une tempête est généralement définie comme une séquence ininterrompue d'états de mer dans lesquelles la hauteur significative des vagues (H_s) reste au dessus d'un seuil fixe h_{crit} (Boccotti [1986]). En général, le début de la tempête est considéré comme l'instant où le H_s passe au dessus du seuil, et l'instant où le H_s repasse en deçà du seuil correspond à la fin de la tempête. La valeur du seuil h_{crit} dépend du climat de la zone étudiée. Même si

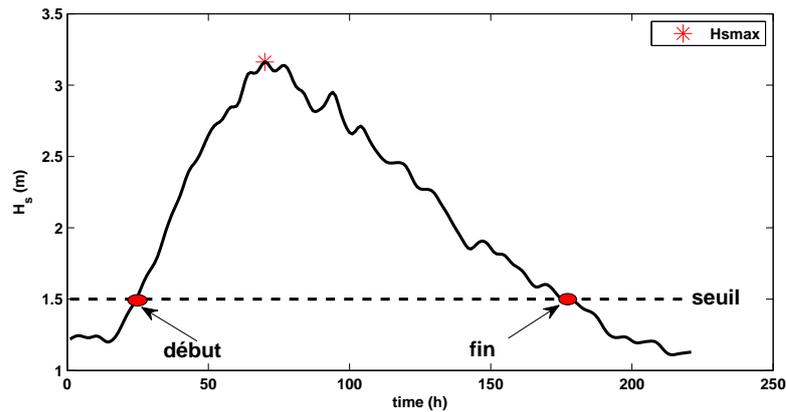


FIGURE 2.13 – Définition d'une tempête

cette valeur reste arbitraire, elle est généralement choisie proportionnellement à la hauteur significative moyenne : $h_{crit} = k\overline{H_s}$. Il est évident que la valeur du paramètre k détermine le nombre de tempêtes ainsi que leurs durées et a donc une influence sur les résultats statistiques. Par exemple, Boccotti [2000], propose de prendre $k = 1.5$ pour la Méditerranée. En règle générale, il faut choisir un seuil permettant, au regard des données, d'avoir un nombre suffisant de tempêtes (plusieurs dizaines de préférence, Holthuijsen [2007]). Le seuil ne doit cependant pas couper de tempête unique en deux tempêtes indépendantes. Le choix du seuil peut aussi prendre en compte la saisonnalité en imposant une variation de ce dernier avec la saison.

Une fois les tempêtes identifiées, deux types d'analyses sont possibles : l'analyse des maxima du H_s des tempêtes ou l'analyse globale de l'évolution du H_s dans la tempête qui permet de prendre aussi en compte la durée de la tempête.

2.4.2.2 Analyse des maxima de H_s des tempêtes

Distributions des maxima de H_s : Mathiesen et al. [1994] recommandent l'utilisation de la distribution Weibull ajustée avec la méthode du maximum de vraisemblance pour les maxima de H_s . Plus tard, plusieurs auteurs (Coles et al. [2001], Castillo [2012]) introduisent la distribution de Pareto généralisée (GPD : generalised Pareto distribution) qui est la distribution théorique des maxima de H_s de tempêtes indépendantes. En effet, sous l'hypothèse que les tempêtes sont indépendantes les unes des autres et que leur fréquence suit une loi de Poisson, la distribution asymptotique théorique des maxima est une

GPD. La distribution GPD s'écrit

$$\begin{cases} G_{y;k,\sigma}(y) = 1 - \left(1 + k \frac{y}{\sigma}\right)^{-\frac{1}{k}} & \text{si } k \neq 0 \\ G_{y;\sigma}(y) = 1 - \exp\left(-\frac{y}{\sigma}\right) & \text{si } k = 0 \end{cases} \quad (2.65)$$

où k est le paramètre de forme et σ le paramètre d'échelle. Cette distribution est recommandée par plusieurs auteurs (Vledder van et al. [1993], Thompson et al. [2009]), mais beaucoup d'autres préfèrent les distributions classiques des valeurs extrêmes (GEV, Weibull et Gumbel). La distribution GEV est donnée par

$$G_{y;\mu,\sigma,\xi}(y) = \exp\left\{-\left[1 + \xi \left(\frac{y - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (2.66)$$

où μ est le paramètre de position, σ le paramètre de forme et ξ le paramètre d'échelle. La distribution de Gumbel correspond au cas $\xi = 0$ et la distribution de Weibull correspond au cas $\xi < 0$.

Période de retour des hauteurs significatives : Les résultats de l'analyse des valeurs extrêmes sont traduits généralement en période de retour. La période de retour, ou temps de retour, caractérise le temps statistique entre deux occurrences d'un événement naturel d'une intensité donnée. Si $1 - F(H_s > h)$ est la probabilité que H_s dépasse un seuil h , t le nombre d'années d'observations du H_s , N le nombre d'observations pendant t , et $\lambda = N/t$ la fréquence d'échantillonnage, le nombre total d'observations pour lesquelles $H_s > h$ est,

$$K = N(1 - F(H_s > h)) = t\lambda(1 - F(H_s > h)) \quad (2.67)$$

La période de retour T de h est définie alors par

$$T = \frac{1}{\lambda(1 - F(H_s > h))} \quad (2.68)$$

i.e. h sera dépassée en moyenne une fois au moins chaque T années.

Pour une fonction de distribution GEV avec $\xi \neq 0$, la valeur du H_s correspondant à une période de retour T est obtenue par

$$\hat{H}_{sT} = \mu + \frac{\sigma}{\xi} \left[1 - \left(-\log\left(1 - \frac{1}{T}\right)\right)^\xi\right]. \quad (2.69)$$

Étude comparative des modèles de distributions d'extrêmes de H_s appliqués aux données mesurées au large de l'Angola : La figure 2.14 montre les ajustements des différentes distributions (GEV, Gumbel, Weibull et GPD) sur les maxima de H_s au large de l'Angola. Les tempêtes ont été identifiées, à partir de 21 ans de données de Hindcast, avec un seuil fixé à 2.25 m. Un total de 61 tempêtes ont été identifiées avec ce seuil. Les observations visuelles confirmées par les tests statistiques montrent que les modèles GEV et Gumbel donnent les meilleurs résultats avec un léger avantage pour le modèle GEV. Il faut remarquer qu'avec ces données, le GPD donne le résultat le moins satisfaisant. Cela peut être dû au seuil choisi pour identifier les tempêtes. En effet, l'utilisation du GPD

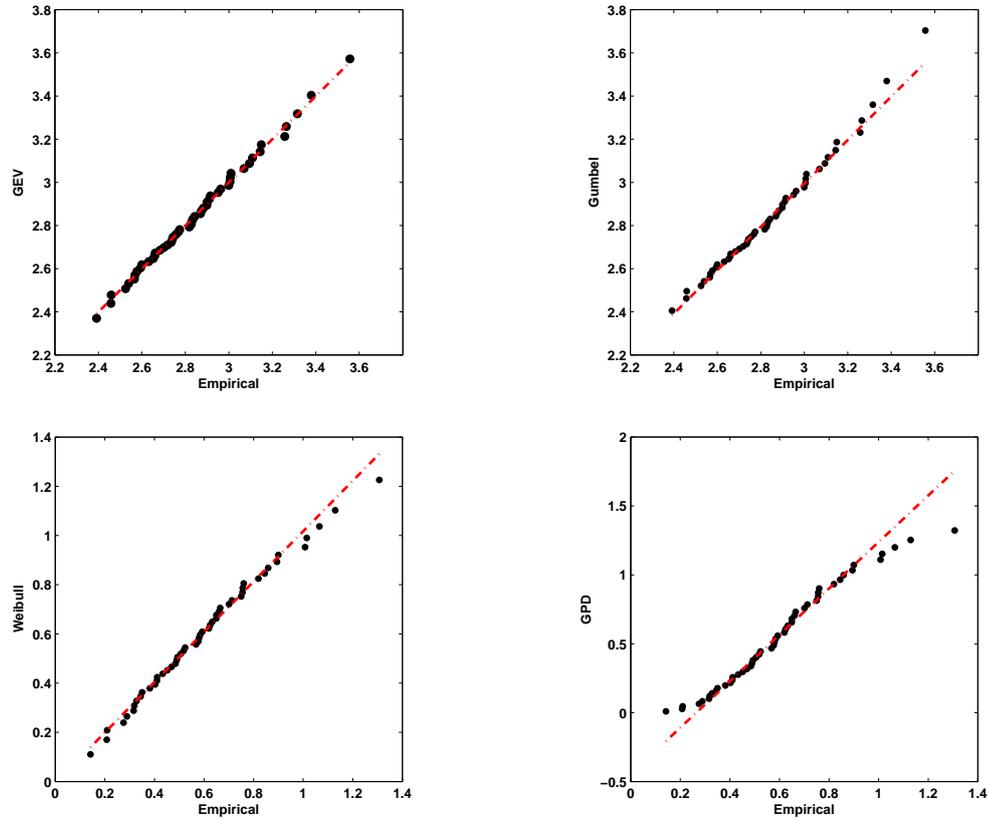


FIGURE 2.14 – Ajustement des maxima de H_s des tempêtes identifiées à partir des données de Hindcast aux large de l’Angola. Les ajustements obtenus par le GEV et le Gumbel sont les meilleurs avec un léger avantage pour le GEV. Les p -value obtenues pour les modèles GEV, Gumbel, Weibull et GPD sont respectivement 0.86, 0.85, 0.80 et 0.16.

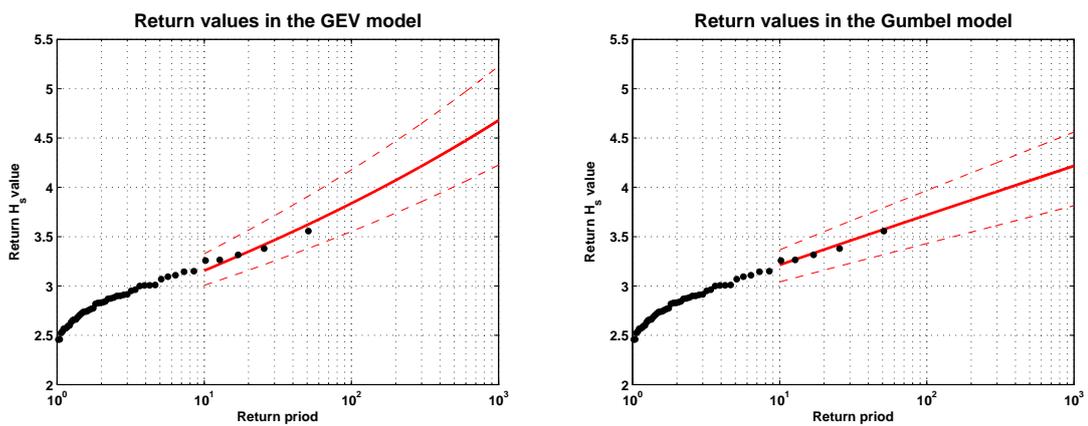


FIGURE 2.15 – Valeurs de retour du H_s en fonction des périodes de retour pour les modèles GEV et Gumbel. Les intervalles de confiance (95%) sont représentés en trait pointillé.

suppose que les événements soient indépendants, ce qui impose un seuil suffisamment haut. Cependant, un seuil haut implique un petit nombre de tempêtes qui peut être insuffisant pour la modélisation statistique.

La figure 2.15 montre les valeurs de retour des hauteurs significatives en fonction des périodes de retour obtenues avec les modèles GEV et Gumbel. Pour une période de retour donnée, les valeurs du H_s obtenues avec le modèle GEV sont légèrement supérieures à celles obtenues avec le modèle Gumbel. On trouve, par exemple, pour la période de retour de 100 ans, $H_s = 3.75 m$ avec le GEV et $H_s = 3.5 m$ avec le Gumbel. La différence entre les valeurs de retour données par les deux modèles augmente avec la période de retour. Même si le test statistique a révélé que la qualité de l'ajustement obtenu avec le GEV est meilleure que celle de l'ajustement obtenu avec le modèle Gumbel, il n'est pas évident de choisir entre les deux modèles en observant la figure 2.15. Néanmoins, le modèle Gumbel a un degré de liberté de moins et peut être par conséquent préférable au GEV.

2.4.2.3 Analyse de l'évolution du H_s dans les tempêtes

Généralement, la tempête réelle est modélisée par une fonction paramétrique qui présente l'avantage d'impliquer très peu de paramètres et de tenir compte de l'évolution du H_s au sein de la tempête. Les statistiques à long-terme sont alors déterminées par analyse des paramètres du modèle.

Equivalent Triangular Storm (ETS) (Boccotti [2000]) : Le modèle ETS associe à chaque tempête identifiée à partir de la série temporelle du H_s un triangle de hauteur a et de base b . La hauteur a du triangle représente l'intensité de la tempête et est égale à la hauteur significative maximale (H_{smax}) dans la tempête. La base b du triangle représente la durée de la tempête et est calculée en sorte que la hauteur maximale de vague individuelle attendue dans la tempête soit égale à celle obtenue avec le modèle ETS (figure 2.16). Le modèle s'écrit :

$$H_{sETS}(t) = a \left(1 - \left| \frac{2t}{b} \right| \right), \quad -\frac{b}{2} \leq t \leq \frac{b}{2} \quad (2.70)$$

Dans la pratique, la hauteur du triangle a est directement déduite de la série temporelle de H_s alors que la base b est déduite à partir d'une procédure itérative. Le modèle ETS a été utilisé par Arena, Pavone [2009], Arena, Pavone [2006] et Martín-Hidalgo et al. [2014] pour l'analyse des vagues extrêmes. Ce modèle permet d'obtenir une solution analytique pour les périodes de retours des hauteurs de vagues.

Equivalent Triangular Duration Storm (ETDS) (Corbella, Stretch [2013]) : Ce modèle associe à chaque tempête, un triangle de hauteur H égale à la hauteur significative maximale (H_{smax}), et de base D égale à la durée réelle de la tempête (figure 2.16). La climatologie des états de mer est obtenue par une simulation multivariée des paramètres du modèle. Ce modèle a aussi été utilisé par Martín-Hidalgo et al. [2014] pour évaluer la progression de fissure dans une structure mécanique marine.

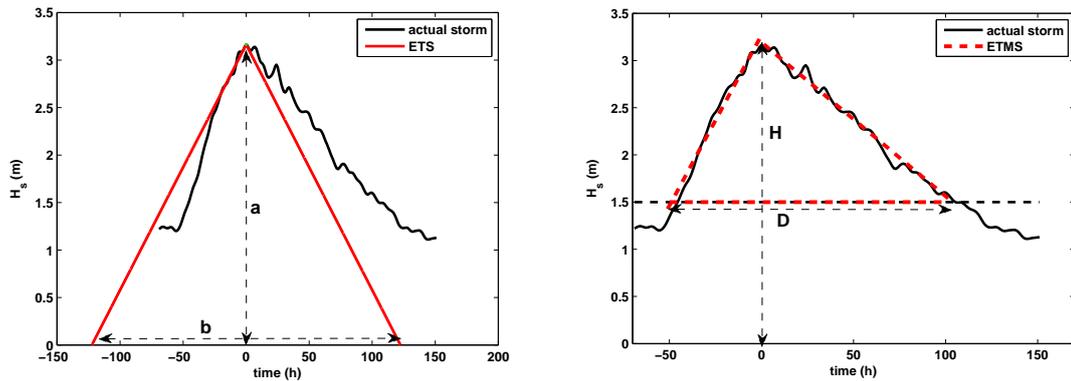


FIGURE 2.16 – Modèles triangulaires de tempêtes : ETS (gauche), ETMS (droite)

2.5 Conclusion

Dans ce chapitre nous avons d'abord présenté les différents moyens d'observations des états de mer puis la caractérisation d'un état de mer par l'analyse vague par vague ou par l'analyse spectrale. Nous avons ensuite abordé la modélisation statistique des vagues à court-terme (échelle de temps allant de 20 min à 6 h environ), notamment les distributions statistiques des hauteurs et des crêtes de vagues linéaires et non-linéaires. La modélisation statistique des états de mer à long-terme a aussi été présentée avec une application aux données d'états de mer en Afrique de l'Ouest. Les deux approches les plus utilisées dans la modélisation de la climatologie des états de mer ont été présentées :

- l'approche utilisant l'ensemble des données (initial distribution method) ;
- l'approche par événements indépendants qui utilise les données de H_s uniquement à partir d'un seuil (Peak Over Treshold method) et donc plus adaptée à la modélisation des extrêmes de H_s .

Dans les deux approches de modélisation de la climatologie des états de mer, les modèles qui ont été proposés sont généralement ajustés aux paramètres d'états de mer estimés à partir de la densité spectrale totale. Or, un état de mer résulte généralement de la superposition de deux ou plusieurs systèmes de vagues (houles et mer du vent), d'intensités, de fréquences et de directions différentes. Il apparaît clairement qu'une description des états de mer à l'échelle des systèmes de vagues (houles et mer du vent) permettrait d'augmenter la précision dans la modélisation de la climatologie.

De plus, les paramètres caractéristiques des houles et des mers du vent affichent une cohérence temporelle qui reflète leurs dynamiques. Cependant, cette cohérence temporelle n'est généralement pas prise en compte dans la modélisation de la climatologie. L'approche par événements utilisée dans nos travaux et décrite dans les chapitres suivants permet une représentation de la climatologie des états de mers à l'échelle des systèmes de vagues tout en intégrant leur cohérence temporelle.

Identification des événements de houle et de mer du vent

L'identification des événements constitue la première étape de notre démarche de modélisation de la climatologie utilisant l'approche par événements. En effet, les événements étant les éléments de base de cette approche, il est primordial en premier lieu de les identifier au sein d'une série temporelle d'états de mer caractérisés par leurs spectres directionnels. Notre objectif est de partir d'une base de données constituée de spectres directionnels, couvrant la période où les données sont disponibles, et d'élaborer une nouvelle base de données constituée par les événements de houle et de mer du vent.

Les méthodes d'identification d'événements de houle et de mer du vent utilisent généralement deux étapes. Chaque spectre individuel est d'abord partitionné. Ensuite, les événements sont identifiés par un suivi temporel des paramètres caractéristiques des systèmes de vagues. L'utilisation de ces méthodes nécessite le calibrage d'un nombre important de critères empiriques, rendant souvent difficile le traitement automatique des données. Suite à des difficultés rencontrées dans l'utilisation des méthodes existantes que nous détaillerons dans ce chapitre, nous avons développé une nouvelle méthode (TESP : Time-domain Expansion of Spectral Partitionning technique) qui intègre la dimension temporelle directement dans le partitionnement. Cette méthode permet une identification d'événements de houle et de mer du vent cohérente avec leurs dynamiques et garantit la continuité des événements, mal respectée par les méthodes existantes. Ce chapitre présente cette nouvelle méthode d'identification d'événements de houle et de mer du vent et introduit en préambule, le concept d'"événement" de houle et de mer du vent.

Sommaire

3.1	Concept d'événement de houle et de mer du vent.	39
3.1.1	Dynamique des systèmes de houle et événement de houle	40
3.1.2	Dynamique des systèmes de mer du vent et événement de mer du vent	41
3.2	Méthodes existantes pour l'identification des événements de systèmes de vagues	42
3.2.1	Partitionnement	42
3.2.2	Classification	45
3.2.3	Suivi temporel des systèmes de vagues	46
3.2.4	Spectral Partitioning for Operational Parameters identification (SPOP)	47
3.2.5	Limites des méthodes existantes	48
3.3	Méthode proposée : Time-domain Expansion of Spectral Partitionning technique (TESP)	49
3.3.1	Motivations	49
3.3.2	Filtrage	51
3.3.3	Extraction	52
3.3.4	Regroupement	52
3.3.5	Estimation des paramètres caractéristiques	54
3.3.6	Classification	54
3.4	Résultats.	55
3.4.1	Données <i>in situ</i> d'Afrique de l'Ouest	55
3.4.2	Données de Hindcast de la mer de Gascogne	57
3.4.3	Comparaison de TESP avec SPOP	58
3.5	Conclusion	59

3.1 Concept d'événement de houle et de mer du vent

De nombreuses études de spectres directionnels, qu'ils soient issus des mesures (bouées ou SAR) ou de modèles numériques de vagues, ont montré que la surface de la mer en un endroit donné est fréquemment la superposition de plusieurs systèmes de vagues (houles et mers du vent), généralement d'intensités, de fréquences et de directions différentes (figure 3.1).

Il est possible de décomposer le spectre caractéristique d'un état de mer en systèmes de vagues, et des études précédentes ont montré l'importance d'une description fine (à l'échelle des systèmes de vagues) des états de mer, notamment, pour des applications aussi variées que le dimensionnement des structures (Bitner-Gregersen [2005]) ou l'estimation du potentiel récupérable de l'énergie des vagues. (Kerbiriou et al. [2007]). Chaque système de vagues identifié peut être ensuite caractérisé par des paramètres synthétiques (H_s , T_p , θ_p , etc.).

De plus, les phénomènes météorologiques (e.g. tempête, dépressions, etc.) qui sont à l'origine des systèmes de vagues ont généralement une durée nettement supérieure à la durée de définition d'un état de mer stationnaire (30 min à 3h). Par un suivi temporel des systèmes de vagues, il est donc possible d'identifier les débuts et les fins de composants associés à un phénomène météorologique unique et à sa propagation.

Nous désignons par "événement" un système de vagues (houles ou mers du vent) en évolution au cours du temps, observable pendant une durée significative et que l'on peut relier à un unique phénomène météorologique source. Nous distinguons les événements de houle, conséquence de phénomènes météorologiques lointains, des événements de mer du vent qui sont, à l'opposé, la conséquence de phénomènes météorologiques locaux. L'identification des événements de houle et de mer du vent peut être réalisée à partir d'une série temporelle de spectres directionnels d'états de mer, connaissant les dynamiques théoriques suivant lesquelles évoluent un système de houle ou un système de mer du vent dans un événement.

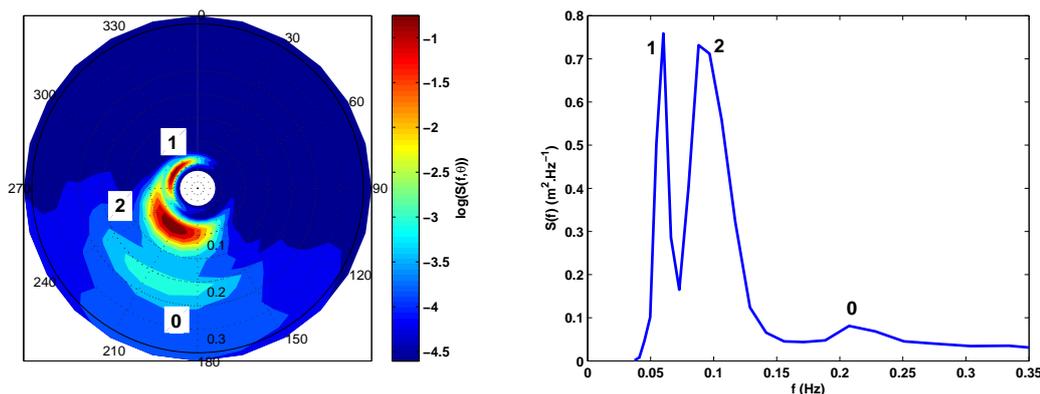


FIGURE 3.1 – Spectre d'état de mer obtenu au large de l'Angola (Girassol) le 04 janvier 2003 à 21h00. Représentation du spectre directionnel 2D (à gauche) et représentation du spectre omnidirectionnel 1D (à droite). On note la présence de trois systèmes de vagues : une houle primaire (1), une houle secondaire (2) et une mer du vent (0).

3.1.1 Dynamique des systèmes de houle et événement de houle

Les houles désignent les systèmes de vagues qui ont été générés par des phénomènes océano-météorologiques en une région éloignée du site d'observation. Elles sont généralement de grande période, peuvent parcourir des milliers de kilomètres, et traverser les océans. Par exemple, les houles observées en Afrique de l'Ouest sont, pour la plupart, générées par des tempêtes répétitives localisées entre le cap Horn ou le large des côtes argentines et le cap de Bonne-Espérance.

Les houles se propagent le long des géodésiques encore appelées grands cercles : cercles sur des plans passant par le centre de la Terre, dont les méridiens. Il est possible de mettre en évidence une relation entre la distance parcourue par les houles et l'évolution de leur fréquence de pic. En effet, les houles sont caractérisées par une forte dispersion : on observe d'abord l'arrivée des vagues de faibles fréquences (grandes périodes) suivies par les vagues de plus grandes fréquences (courtes périodes). Cela s'explique par le fait que les tempêtes contiennent une assez grande variété de fréquences d'une part et d'autre part que la vitesse de groupe dépend fortement de la fréquence, ($Cg = g/(4\pi f)$ en eau profonde). Dans la limite où le point d'observation est très loin de la tempête source, celle-ci peut alors être considérée comme ponctuelle dans le temps et dans l'espace. L'évolution de la fréquence de pic en un point d'observation fixe, est donnée par

$$f_p(t) = \frac{g(t - t_0)}{4\pi R\alpha} \quad (3.1)$$

où α est la distance sphérique entre la tempête et le point d'observation, R est le rayon de la Terre, et t_0 est la date de génération. Cette évolution typique de la fréquence de pic permet ainsi, d'identifier un système de houle, généré par une même tempête, en évolution au cours du temps. C'est en cela que la définition d'un événement de houle prend tout son sens. Malgré l'idéalisation sous-jacente à la relation 3.1, de nombreuses travaux ont confirmé sur la base d'observations *in situ* qu'elle permet de bien prédire l'évolution des systèmes de houle. (Barber, Ursell [1948], Munk et al. [1963], Snodgrass et al. [1966], Aarnes, Krogstad [2001], Hanson, Phillips. [2001], Collard et al. [2009], Delpy et al. [2010]).

Pour identifier le passage d'un système de houle en un point on utilise souvent la méthode dite *Ridge Line Method* dont une illustration est présentée dans la figure 3.2. Employée pour la première fois par Barber, Ursell [1948], puis par Munk et al. [1963] et Snodgrass et al. [1966], elle constitue la méthode de référence dans l'identification des systèmes de houle confirmant également la pertinence de la théorie linéaire en eau profonde pour l'étude des systèmes de houle. La figure 3.2 présente l'historique de spectres mesurés à partir d'une bouée et pendant le mois de décembre 2002 au large de l'Angola. Ces données sont celles qui seront principalement utilisées dans ce chapitre. Sur cette figure, les contours espacés régulièrement de -2.4 à 3.4 indiquent le logarithme du spectre $S(f)$, et les couleurs donnent la direction moyenne pour chaque fréquence. Les événements de houle sont reconnaissables dans les basses fréquences par l'évolution quasi-linéaire de leurs fréquences de pic. La ligne oblique pointillée du 23 décembre à 0.05 Hz au 29 décembre à 0.11 Hz correspond à la variation théorique de la fréquence de pic dans cet événement de houle.

Notons également que l'équation 3.1 est utilisée pour estimer la distance de génération des houles. Grâce à l'information directionnelle des systèmes, il est possible d'évaluer la

zone géographique où la tempête qui est en à l'origine s'est produite (Hasselmann et al. [1996], Aarnes, Krogstad [2001], Hanson, Phillips. [2001], Collard et al. [2009], Delpy et al. [2010]).

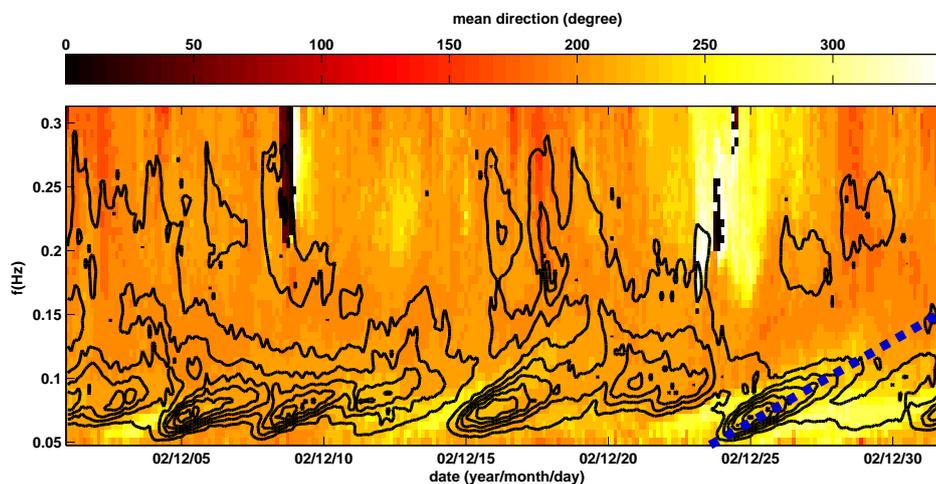


FIGURE 3.2 – Historique de spectres mesurés pendant le mois de décembre 2002 au large de l'Angola (adapté à partir de Collard et al. [2009]).

3.1.2 Dynamique des systèmes de mer du vent et événement de mer du vent

Les mers du vent sont des systèmes de vagues qui se forment sous l'action du vent local. Elles dépendent de plusieurs facteurs, 3 principalement : le fetch, c'est à dire la distance d'action du vent, la durée d'action du vent et la vitesse du vent. La hauteur significative des vagues augmente avec l'amplification de ces facteurs. En effet, en théorie linéaire, les vagues de haute fréquence générées par le vent se dispersent selon l'équation 2.7 sur une ouverture angulaire relativement importante (crêtes courtes), et ce d'autant plus que ces vagues sont directement influencées par les fluctuations rapides d'intensité et de direction du vent local. L'énergie transmise aux vagues par le vent s'accumule au cours du temps, les vagues étant elle-mêmes soumises à des interactions non linéaires, de telle sorte que d'une part, la fréquence de pic des trains d'ondes générées diminue, et d'autre part, leur hauteur significative augmente (Price, Bishop [1974]).

Comme dans le cas des houles, il est possible, de définir un événement de mer du vent comme une mer du vent en évolution au cours du temps et que l'on peut relier à une même tempête (ou dépression) source. Théoriquement, la fréquence de pic des systèmes de mer du vent diminue au cours d'un événement. Cependant, cette évolution de la fréquence de pic n'est généralement pas aussi régulière que celle des houles à cause des fluctuations locales du vent.

3.2 Méthodes existantes pour l'identification des événements de systèmes de vagues

Avant de présenter la méthode qui a été développée dans ces travaux, il nous a paru utile de présenter brièvement les méthodes existantes d'identification d'événements ainsi que leurs limites.

Dans les méthodes qui existent, l'identification des événements de houle et de mer du vent utilise généralement deux étapes. La première, souvent appelée *partitionnement*, consiste à extraire dans un spectre d'état de mer, les différents systèmes de vagues (partitions) qui le constituent, et à les classer en type houle ou mer du vent, suivant leurs caractéristiques. La seconde étape est le *suivi temporel* des systèmes de vagues, et consiste à déterminer, entre deux instants de mesure, quelles partitions relèvent d'un même système en évolution au cours du temps (événement).

3.2.1 Partitionnement

Le partitionnement se fait généralement en trois étapes :

- l'extraction de systèmes qui consiste à être capable de déceler automatiquement les composantes spectrales pertinentes dans un jeu de données fréquence-direction représentatif d'un état de mer complexe ;
- le regroupement qui consiste à fusionner suivant des critères bien définis plusieurs composantes de systèmes extraits de caractéristiques pas suffisamment différentes pour être considérées comme distinctes (élimination d'artefact ou de bruit de mesure).
- la classification qui consiste à affecter à chaque système de vagues, le type houle ou mer du vent.

3.2.1.1 L'extraction : Algorithme de ligne de partage des eaux (Watershed algorithm)

[Gerling \[1992\]](#) a été le premier à proposer une méthode d'extraction. Sa méthode est restée finalement très peu utilisée, principalement en raison de sa lourdeur. De nos jours, la méthode d'extraction la plus utilisée est basée sur l'algorithme de la ligne de partage des eaux, utilisée pour la première fois par [Hasselmann et al. \[1996\]](#).

En morphologie mathématique, l'algorithme de la ligne de partage des eaux, proposé par [Digabel, Lantuéjoul \[1978\]](#), est un algorithme de segmentation d'images qui permet de décomposer une image en régions homogènes. Si nous comparons l'image constituée par un spectre directionnel renversé ($-S(f, \theta)$) à un relief terrestre, cette image peut se diviser en plusieurs zones. Chaque zone sera définie comme un bassin versant. Un bassin versant¹ correspond à l'espace géographique dans lequel toutes les eaux convergent vers le même minimum. Les différents bassins versants sont délimités par une ligne de partage des eaux ou ligne de crêtes des massifs bordant le bassin versant [Coster, Chermant \[1985\]](#). Pour l'obtenir, il faut imaginer l'immersion d'un relief dans de l'eau, en précisant que l'eau ne peut pénétrer dans les vallées que par ses minima. La ligne de partage des eaux est représentée par les points où deux lacs disjoints se rejoignent au cours de l'immersion.

1. On peut par exemple utiliser dans MATLAB la routine de recherche des bassins versants WATER-SHED.m basée sur les travaux de [Meyer \[1994\]](#).

Pour empêcher le mélange des eaux venant des minimums différents, on crée un barrage élémentaire en chaque point de contact (figure 3.3).

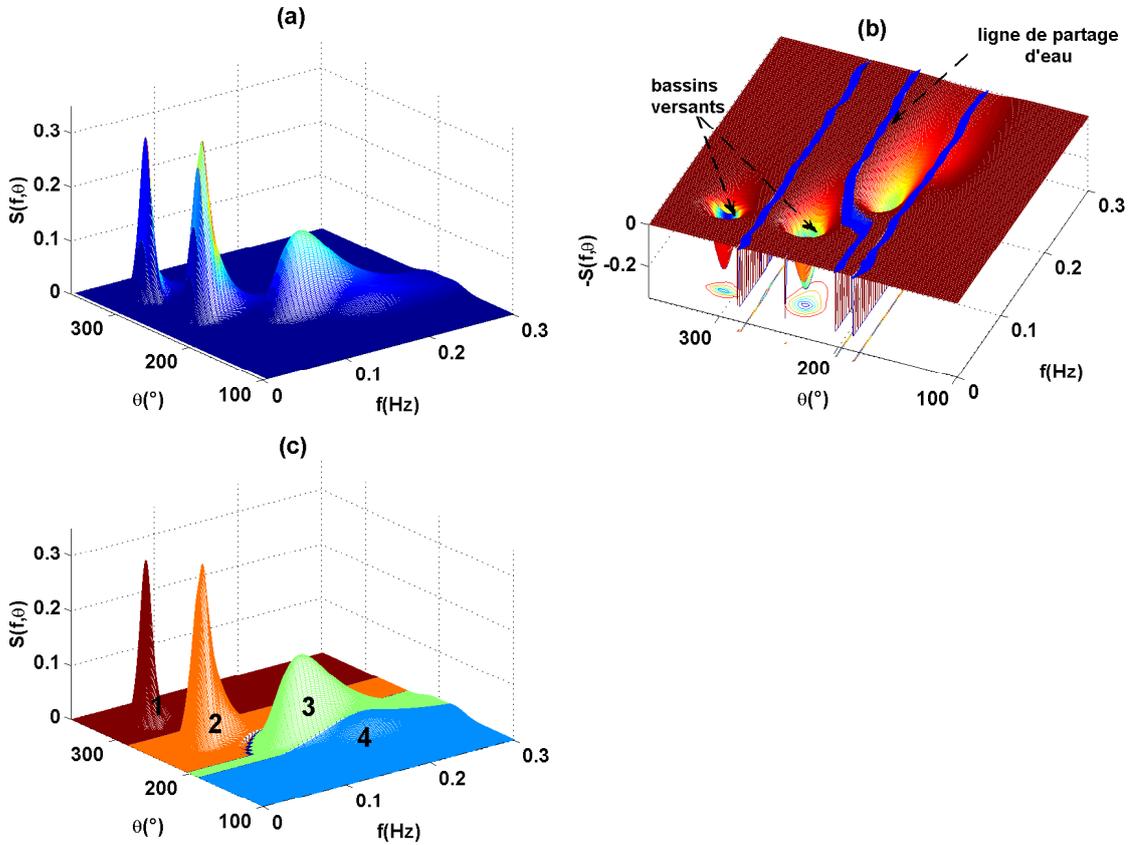


FIGURE 3.3 – Illustration de l'extraction par l'algorithme de la ligne de partage des eaux en 2D : (a) spectre directionnel ; (b) spectre directionnel renversé, traité comme un relief topographique et recherche des lignes de partage des eaux ; (c) les bassins versants identifiés, numérotés de 1 à 4, correspondent aux partitions.

3.2.1.2 Le regroupement

Souvent, les partitions obtenues de l'extraction n'ont pas toutes de signification physique. En effet, à cause du bruit dans le spectre et de la sensibilité de l'algorithme de la ligne de partage des eaux, le nombre de partitions issues de l'extraction est généralement grand. Le regroupement consiste donc à fusionner les composantes de caractéristiques insuffisamment différentes pour être considérés comme distinctes. Les critères utilisés pour le regroupement sont, pour la plupart, basés sur la distance quadratique entre pics (comparée ou non à l'étalement spectral des deux pics), le contraste minimal et l'énergie minimale (Aarnes, Krogstad [2001], Hanson, Phillips. [2001]).

Distance entre les pics : Deux partitions $p(1)$ et $p(2)$ sont regroupées si la distance quadratique entre les pics Δf^2 et l'étalement moyen en fréquence $\bar{\delta f^2}$ remplissent la condi-

tion suivante :

$$\Delta f^2 \leq \kappa \overline{\delta f^2}, \quad (3.2)$$

$$\Delta f^2 = (f_{px,1} - f_{px,2})^2 + (f_{py,1} - f_{py,2})^2, \quad (3.3)$$

avec $f_{px} = f_p \cos \theta_p$, et $f_{py} = f_p \sin \theta_p$. Le facteur d'étalement κ est ajusté empiriquement de manière à optimiser la performance du partitionnement du spectre.

Contraste entre les pics : Ce critère évalue la hauteur du point de selle (qu'on note S_s) entre deux pics. S'il est trop haut les deux pics ne sont pas suffisamment démarqués l'un de l'autre. La condition de regroupement de deux pics $p(1)$ et $p(2)$ s'écrit :

$$S_s > \nu \min(\max(p(1)), \max(p(2))), \quad (3.4)$$

avec ν une constante fixée empiriquement.

Énergie faible : Les partitions pour lesquelles l'énergie totale est en dessous d'un seuil d'énergie sont négligées,

$$e \leq \frac{A}{f_p^4 + B} \quad (3.5)$$

où A et B sont des constantes empiriques.

Les choix des valeurs des constantes qui interviennent dans les critères de regroupement ne sont pas ou peu justifiés dans la littérature. Il s'agit généralement de consensus empiriques ou d'adaptation par essais successifs d'un paramètre optimum à un jeu de données. La différence entre les algorithmes proposés dans la littérature repose essentiellement sur le choix des constantes. Le tableau 3.1 résume les différentes valeurs prises par les paramètres (κ , ν et A et B) suivant les auteurs. Un exemple de regroupement de partitions est montré dans la figure 3.4.

	distance entre les pics (κ)	contraste entre les pics (ν)	énergie faible (A, B)
Hasselmann et al. [1996]	1	0.85	-
Voorrips et al. [1997]	0.5	0.7	-
Hanson, Phillips. [2001]	0.4 – 0.5	0.65 – 0.75	$(2.10^{-6}, 2.10^{-3})-$ $(6.10^{-5}, 3.10^{-2})$

TABLE 3.1 – Tableau résumant les valeurs des constantes utilisées pour le regroupement dans la littérature.

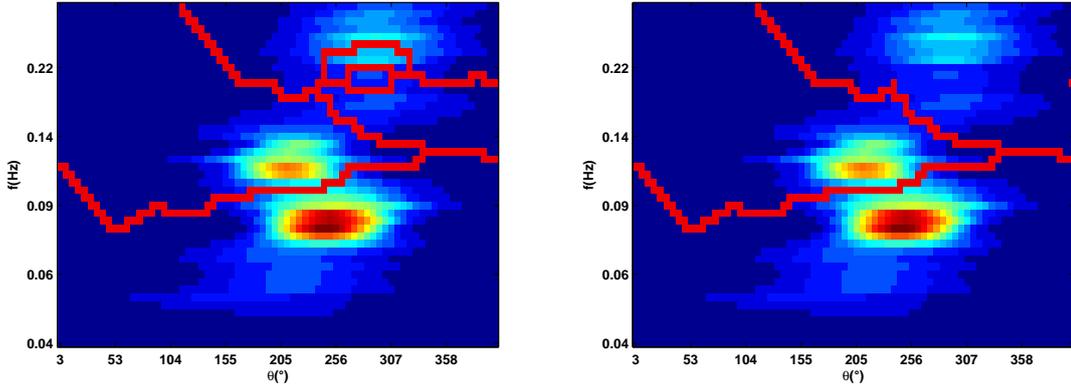


FIGURE 3.4 – Exemple de partitionnement du spectre de la figure 3.1 : extraction par l'algorithme des lignes de partage des eaux (gauche), puis regroupement des systèmes (droite)

3.2.2 Classification

Après le regroupement, chaque partition significative est caractérisée par ses paramètres synthétiques (H_s , f_p , θ_p , etc) qui sont obtenus à partir des moments spectraux calculés avec la portion du spectre correspondante. Chaque partition peut être ensuite classée en type houle ou mer du vent suivant ses caractéristiques. La classification est généralement basée sur la connaissance des données de vent. En effet, les mers du vent, étant des systèmes de vagues créées par le vent à l'endroit où il souffle, ont des caractéristiques qui sont fortement liées à celles du vent. Ainsi, une partition dont les caractéristiques sont proches de celles de ce que le vent observé aurait généré sera donc considérée comme une mer du vent : c'est ce qui est généralement traduit dans le critère dit de "l'âge des vagues" (Hanson, Phillips. [2001]).

Ce critère stipule que le pic de la mer du vent est contenu dans une région du spectre délimitée par une parabole telle que :

$$f_{pws} \leq \frac{g}{2\pi\beta U_{10} \cos(\delta)} \quad (3.6)$$

avec $\delta = \theta_w - \theta_m$, f_{pws} = fréquence de pic de la mer du vent, U_{10} = module de la vitesse du vent à 10m, θ_w = direction du vent, θ_m = direction moyenne du pic, β = valeur empirique. Le paramètre β de l'équation 3.6 est une constante empirique supérieure ou égale à 1, dont la valeur est réglée par l'observateur et n'est pas justifiée par des critères physiques mais permet d'ajuster la décomposition des états de mer aux spécificités locales. Par exemple Hanson, Phillips. [2001] choisissent $\beta = 1.5$ alors que Kerbiriou et al. [2007] utilisent $\beta = 4/3$.

Ce critère ne peut être utilisé que si les données de vent sont disponibles. En l'absence de données du vent, une fréquence de séparation constante est souvent fixée, en regard des spécificités locales, pour distinguer la houle de la mer du vent Kerbiriou et al. [2007].

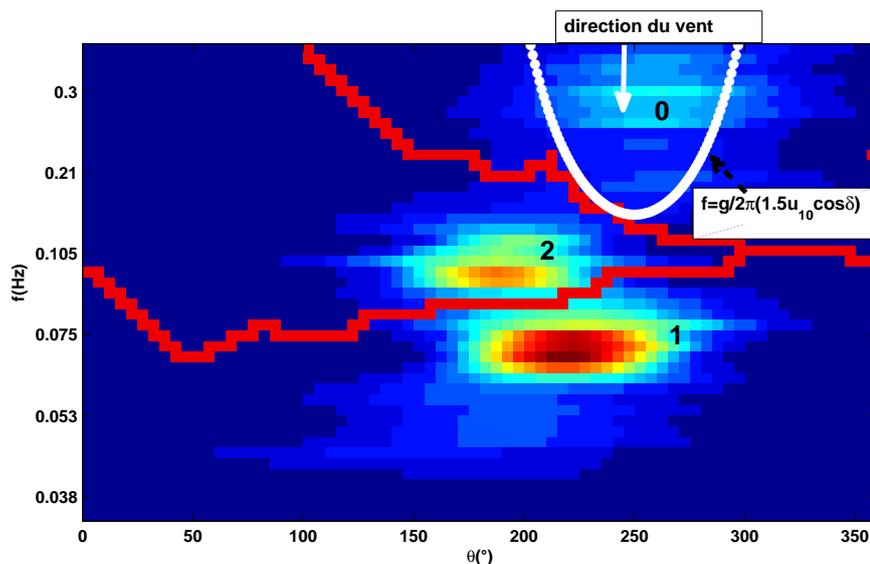


FIGURE 3.5 – Illustration de la classification des systèmes de vagues : la mer du vent (système 0) est localisée dans la parabole définie par le critère de l'âge des vagues $\{H_s = 0.54m, f_p = 0.26Hz, \theta_p = 257^\circ\}$, le système 2 correspond à la houle secondaire $\{H_s = 0.46m, f_p = 0.11Hz, \theta_p = 187^\circ\}$, et le système 1 correspond à la houle primaire $\{H_s = 0.62m, f_p = 0.08Hz, \theta_p = 217^\circ\}$.

3.2.3 Suivi temporel des systèmes de vagues

À partir des critères (critères d'affectation temporelle) basés sur la dynamique des houles et des mers du vent, il est possible d'associer des systèmes de vagues de même type obtenus à des pas de temps consécutifs pour les faire correspondre à un même événement.

Ainsi, Voorrips et al. [1997], et Aarnes, Krogstad [2001] définissent des critères basés sur la variation de l'énergie et la proximité dans le plan fréquence et direction. Pour faire correspondre une partition i d'un spectre A à l'instant t à une partition j d'un spectre B à l'instant $t + \Delta t$ il faut qu'elles :

- soient de même type (houle ou mer du vent),
- aient des énergies semblables :

$$\nu^{-1}e(i) \leq e(j) \leq \nu e(i), \quad (3.7)$$

- soient proches dans le plan spectro-directionnel :

$$|f_m(i) - f_m(j)| \leq \eta_f f_m(i), \quad |\theta_m(i) - \theta_m(j)| \leq \eta_\theta, \quad (3.8)$$

pour des valeurs appropriées des constantes ν , η_f et η_θ . Les constantes utilisées par Voorrips et al. [1997], et Aarnes, Krogstad [2001] sont récapitulées dans le tableau 3.2. De toute évidence, il est possible que chaque partition de A ne soit pas affectée à une partition de B . Si les deux spectres représentent l'état de la mer au même endroit, mais à deux instants différents, la non affectation peut être interprétée comme le début d'un nouvel événement ou la fin d'un événement ancien.

	ν	η_f	η_θ
Voorrips et al. [1997]	20	0.4 si $\theta_m(i) =$ $\theta_m(j)$	50° si $f_m(i) =$ $f_m(j)$
Aarnes, Krogstad [2001]	[10 – 20]	0.5	60°

TABLE 3.2 – Tableau résumant les valeurs des constantes utilisées pour le suivi temporel

Les algorithmes de partitionnement de spectres et du suivi temporel des systèmes de vagues sont, généralement, implémentés et présentés sous forme d'applications dans le but d'un traitement automatique des spectres d'états de mer. Une application de ce type dénommée SPOP (Spectral Partitioning for Operational Parameters identification) a été développé dans l'environnement Matlab par IFREMER et ACTIMAR (Kerbiou et al. [2007]).

3.2.4 Spectral Partitioning for Operational Parameters identification (SPOP)

SPOP est une application développée, qui permet le partitionnement et le suivi temporel des systèmes de vagues automatiques jusqu'à la production de statistiques climatologiques complètes.

SPOP reprend dans son implémentation, les étapes décrites ci-dessus avec quelques spécificités, notamment le choix des constantes empiriques intervenant dans les algorithmes. Il utilise pour l'extraction des systèmes, la fonction Watershed.m de Matlab, et pour le regroupement des systèmes, des critères basés d'une part sur les conditions jointes de proximité en fréquence et direction, et d'autre part sur l'énergie des systèmes. La classification des systèmes est faite grâce au critère de l'âge des vagues lorsque les données de vent sont disponibles. En absence des données de vent, une fréquence de séparation constante est fixée. SPOP reconstruit aussi la densité spectrale des systèmes de vagues identifiés en ajustant des modèles paramétriques. Ainsi, la mer du vent est ajustée par un Jonswap, et pour mieux représenter la forme souvent étroite du spectre de la houle, un Jonswap modifié est utilisé (Kerbiou et al. [2007]). La fonction de répartition directionnelle quant à elle est modélisée par la fonction \cos^{2s} dépendant de la fréquence. Il est également intégré dans SPOP une étape de correction mutuelle, qui permet de corriger l'énergie relative de chaque système et donc de reconstruire la densité spectrale de l'état de mer. SPOP effectue finalement un suivi temporel des systèmes de vague en utilisant des critères d'affectation temporelle sur les grandeurs physiques, H_s , f_p , et θ_m . La figure 3.6 montre l'identification des événements de houle et de mer du vent à partir des données *in situ* (large de l'Angola) du mois de décembre 2002 avec SPOP.

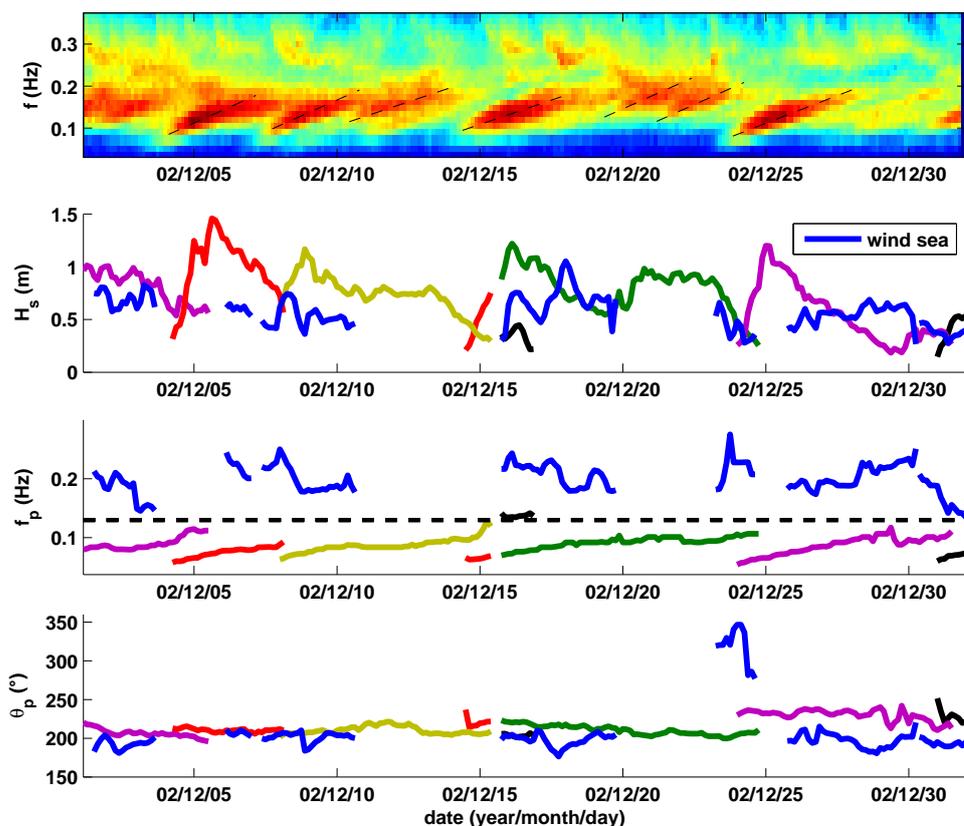


FIGURE 3.6 – Identification des événements avec SPOP (données *in situ*, décembre 2002). De haut en bas, les fenêtres correspondent respectivement à la série-temporelle des densités spectrales $S(f)$, la hauteur significative, la fréquence de pic et la direction de pic. Chaque événement est représenté par une couleur différente.

3.2.5 Limites des méthodes existantes

L'étape de regroupement, comme mentionné plus haut, est nécessaire pour fusionner les partitions non significatives issues de l'extraction. Mais, il est souvent difficile de fixer les valeurs des constantes empiriques, qui interviennent dans les critères de regroupement des partitions. Un mauvais choix de la valeur d'une constante peut entraîner, soit le regroupement abusif de deux partitions complètement indépendantes, soit le non regroupement de deux partitions qui n'en sont en réalité qu'une.

Le grand nombre de partitions identifiées lors de l'extraction est généralement dû au bruit dans le spectre. Portilla et al. [2009] proposent de réduire le bruit en lissant le spectre et de s'affranchir ainsi de l'étape du regroupement. Ainsi, se donnant un nombre de partitions significatives probable au vu de la climatologie du site, l'opération de lissage est répétée jusqu'à l'obtention, après extraction, du nombre de partitions espéré. Cette méthode présente l'inconvénient, pour l'utilisateur, de devoir savoir a priori le nombre de partitions finales à obtenir et donc le degré du lissage à effectuer, alors même que le

nombre de partitions significatives peut varier d'un spectre à un autre.

Dans le cas où les données de vent sont inexistantes, ce qui est le cas sur une période des données *in situ* utilisées pour nos travaux, la classification des systèmes de vagues se fait en considérant une fréquence de séparation constante. Or, la fréquence de séparation évolue dans le temps suivant les caractéristiques du vent et imposer qu'elle soit fixe pourrait introduire une erreur dans la classification. Une houle de fréquence élevée, parce qu'elle a été générée à proximité du point d'observation ou en fin de train de houle, peut donc se retrouver arbitrairement classée en type mer du vent. Réciproquement, une mer de vent de fréquence faible, au point d'être inférieure à la fréquence de séparation fixée, sera classée comme houle par erreur.

Dans l'étape du suivi temporel des systèmes de vagues, se pose également le problème du choix des constantes empiriques, notamment dans les critères d'affectation temporelle. C'est ainsi qu'un événement peut se retrouver arbitrairement séparé en deux, ou à l'inverse, deux événements complètement différents se retrouver fusionnés. Sur la figure 3.6 nous pouvons observer que l'événement de houle identifié par SPOP entre le 08 décembre et le 15 décembre (représenté en couleur jaune) est anormalement et excessivement long. En effet, on s'aperçoit lorsqu'on observe le graphe du haut, où est représentée la série-temporelle des densités spectrales que ce sont deux événements distincts qui ont été recollés.

L'identification automatique des événements de houle et de mer du vent est très difficile à opérer avec ces méthodes. SPOP par exemple utilise une dizaine de constantes empiriques et il est très difficile de les régler toutes simultanément. De plus, pour un même site, il arrive que la valeur d'une constante fonctionne bien sur des spectres d'un intervalle de temps donné, et que pour l'intervalle de temps suivant cette même valeur ne fonctionne plus.

Les différents problèmes énumérés ci-dessus ont rendu l'identification automatique des événements de houle et de mer du vent très difficile avec SPOP. Il faut souligner que SPOP n'a pas été, à la base, développé pour identifier directement les événements de houle et de mer du vent, mais pour faire le partitionnement de spectres individuels. Le suivi temporel de systèmes de vagues qui y est implémenté sert surtout à améliorer le partitionnement individuel des spectres. Afin de permettre la conservation de la cohérence temporelle des systèmes de vagues, nous avons donc développé une nouvelle méthode d'identification d'événements de houle et de mer du vent que nous présentons dans la section suivante.

3.3 Méthode proposée : Time-domain Expansion of Spectral Partitionning technique (TESP)

3.3.1 Motivations

La méthode que nous proposons est une extension du partitionnement à la dimension temporelle. Les méthodes existantes partitionnent le spectre $S(f, \theta)$ à chaque pas de temps, avant de faire un suivi temporel des paramètres (H_s, f_p, θ_p) des systèmes de vagues obtenus et d'identifier ainsi les événements de houle et de mer du vent. Notre approche permet de faire ces deux opérations en une fois, en incluant la dimension temporelle directement dans le processus du partitionnement. Le partitionnement est alors, effectué en 3D (temps, fréquence, direction), c'est-à-dire directement sur l'ensemble de la série-temporelle de spectres directionnels $S(t, f, \theta)$.

En effet, les événements de houle et de mer du vent sont identifiables directement dans la série-temporelle des densités spectrales comme le montre la figure 3.7. Sur cette figure, qui est une représentation dans le plan (temps-fréquence) de la densité spectrale, sont observables dans les basses fréquences, des structures quasi-linéaires qui sont la signature des événements de houle. Dans les hautes fréquences, il est également possible d’observer d’autres structures, aux allures moins régulières que les premières, et qui sont la signature des événements de mer du vent. La cohérence temporelle qu’affichent les systèmes de houle et de mer du vent, peut être directement utilisée dans la segmentation pour extraire des partitions plus cohérentes, et notre objectif est de construire une méthode automatisée à cette fin.

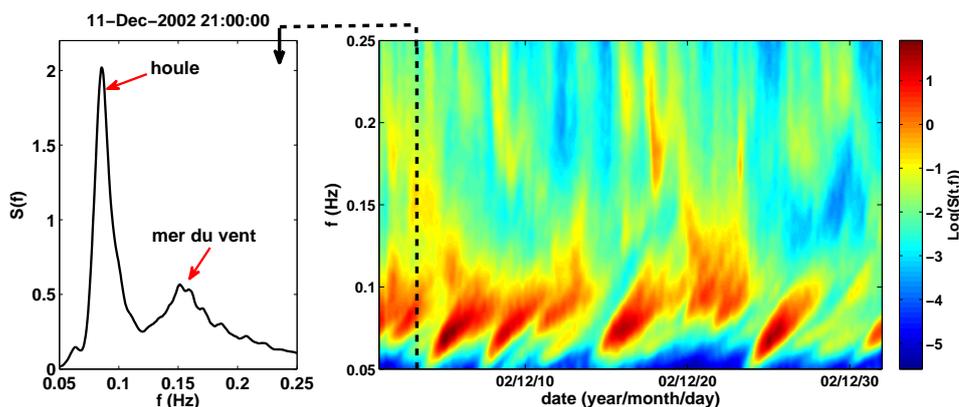


FIGURE 3.7 – Spectre $S(f)$ instantané (gauche) et série-temporelle de spectres ($S(t, f)$) en échelle logarithmique) des états de mer observés un mois de décembre 2002 aux large de l’Angola (droite).

Par l’inclusion de la dimension temporelle dans l’extraction, la méthode proposée, présente l’avantage de réduire considérablement le nombre de post-traitements, et donc de critères empiriques nécessaires à l’identification des événements. Cela favorise, par conséquent, le traitement automatique des données et réduit de même le temps de calcul.

Avec la méthode que nous proposons, il est possible d’extraire les événements en se limitant à la série temporelle des spectres fréquentiels ($S(t, f)$). Cependant, il serait impossible de séparer la fin d’un événement de houle du début d’un événement de mer du vent ou encore deux événements de houle très proches en fréquence mais de directions différentes. Or, ce cas de figure se présente parfois dans les régions où coexistent plusieurs systèmes de houle. En Afrique de l’Ouest par exemple, il est courant d’observer en hiver des systèmes de houle provenant de l’Atlantique Nord qui coexistent avec les systèmes de houle générées dans l’Atlantique Sud (Forristall et al. [2013]). La figure 3.8 montre à gauche deux coupes de la série-temporelle de spectres directionnels $S(t, f, \theta)$ du mois de janvier 2003 : la première à 210° , $S(t, f, 210^\circ)$ (direction Sud-Ouest), et la deuxième à 300° , $S(t, f, 300^\circ)$ (direction Nord-Ouest). Il est possible d’observer sur la coupe à 300° la présence de systèmes correspondant aux houles générées dans l’Atlantique Nord, et sur l’autre coupe, des événements de houle provenant de l’Atlantique Sud en plus des mers du vent. Le spectre directionnel obtenu le 21 janvier à 3h00 (à droite) confirme que les deux systèmes de houle évoluent à des fréquences très proches. Il apparaît clairement que l’information relative aux directions des deux systèmes de houles est indispensable pour les

séparer. D'où l'importance d'utiliser la série-temporelle de spectres directionnels $S(t, f, \theta)$ pour extraire les événements.

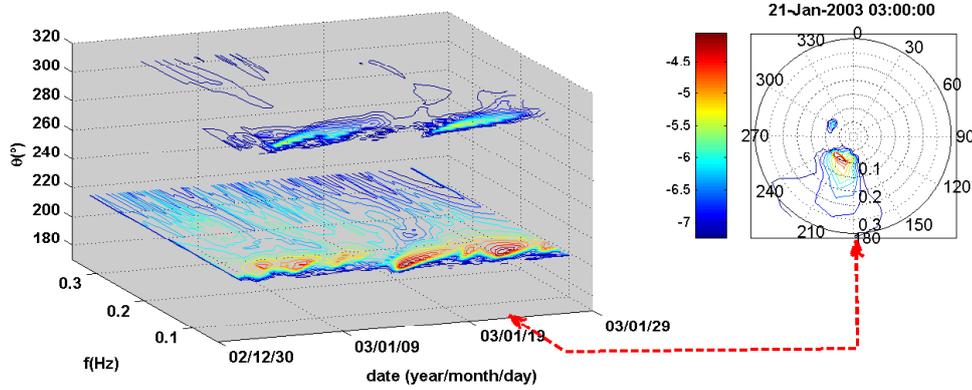


FIGURE 3.8 – Coupes dans les plans passant par 210° et 300° de la série-temporelle de spectres directionnels $S(t, f, \theta)$ du mois de janvier 2003 (gauche); spectre directionnel obtenu le 21 janvier à 3h00 (droite)

La méthode proposée comporte 5 étapes :

1. la filtrage : consiste à réduire le bruit dans l'image constituée par la série temporelle de la densité spectrale $S(t, f, \theta)$;
2. l'extraction : consiste à identifier dans l'image 3D constituée par $(S(t, f, \theta))$, les bassins versants qui correspondent aux événements ;
3. le regroupement : permet de regrouper les morceaux d'un événement qui a été arbitrairement séparé lors de l'extraction ;
4. l'estimation des paramètres : permet de caractériser chaque événement par ses paramètres synthétiques ;
5. la classification : consiste à affecter le type houle ou mer du vent à chaque événement identifié.

3.3.2 Filtrage

Les spectres issus des données *in situ* sont souvent bruités. Afin de réduire le bruit, et donc d'éviter l'identification d'un nombre important de bassins versants, $S(t, f, \theta)$ est d'abord filtré :

$$\hat{S}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n |H|^{-1/2} \mathbf{K} \left\{ |H|^{-1/2} (\mathbf{x} - \mathbf{x}_i) \right\} S(\mathbf{x}), \quad (3.9)$$

où $\mathbf{x} = [t, f, \theta]^T$, n est le nombre de points, H est une matrice 3×3 correspondant aux fenêtres de lissage suivant les 3 dimensions et \mathbf{K} le noyau tridimensionnel choisi comme gaussien,

$$\mathbf{K}(\mathbf{x}) = (2\pi)^{-3/2} \exp\left(\frac{-1}{2} \mathbf{x}^T \mathbf{x}\right). \quad (3.10)$$

Notons que la série-temporelle de spectres directionnels lissée sert uniquement à identifier les bassins versants ; l'estimation des paramètres se fait avec la série-temporelle de spectres directionnels initiale, afin de ne pas lisser l'évolution des paramètres H_s , f_p , θ_p , etc, au sein des événements.

3.3.3 Extraction

L'algorithme de la ligne de partage des eaux peut s'étendre facilement aux images 3D (ici l'espace temps, fréquence et direction), et a donc été utilisé pour extraire directement les événements de la série-temporelle de spectres directionnels.

La recherche des lignes de partage des eaux en 3D est difficile à conceptualiser. Nous nous servons de cet exemple extrait de [Agnus et al. \[2000\]](#) pour l'expliquer : considérons le volume formé par la série-temporelle de spectres directionnels $(-S(t, f, \theta))$ comme une roche poreuse. Chaque minimum est considéré comme un lac souterrain. On chauffe progressivement la roche. L'eau issue des lacs commence à occuper tout le volume, et se propage d'autant plus vite que la densité de la roche est faible. Lorsqu'en un endroit les eaux de deux lacs différents se rencontrent, il y a cristallisation au niveau du point de contact. Cette cristallisation crée une roche de densité infinie : l'eau ne peut plus se propager en ce point. Lorsque les eaux ont envahi toute la roche, on arrête le processus de chauffage : les lignes de partage des eaux sont l'ensemble des points où les cristaux se sont formés.

Pour obtenir les bassins versants, nous avons également utilisé la fonction *Watershed.m* de Matlab. On récupère en sortie une matrice 3D de même dimension que $S(t, f, \theta)$ dans laquelle les lignes de partage des eaux sont marquées par 0, et chaque bassin versant identifié est marqué par un nombre entier supérieur à 0.

Pour des raisons liées à la difficulté de visualisation de la série temporelle de spectres directionnels $S(t, f, \theta)$, nous illustrerons la méthode d'extraction proposée en 2D (figure 3.9). Sur la partie haute est représentée la série-temporelle des densités spectrales $(S(t, f))$ traitée comme un relief topographique renversé. La fenêtre du centre de la figure 3.9 montre les bassins versants identifiés après extraction. Chaque bassin versant porte un numéro qui lui est propre et peut correspondre, à cette étape, soit directement à un événement, soit à un fragment d'événement.

3.3.4 Regroupement

Il peut arriver, malgré le filtrage, qu'une ligne de partage des eaux passe au travers d'un événement et le sépare arbitrairement en deux. Comme par exemple les fractions d'événements 15 et 18 de la figure 3.9 (fenêtre du milieu), cela arrive généralement près du pic de l'événement où le gradient entre les pixels qui constituent l'image est important. Un post-traitement s'avère donc nécessaire pour supprimer, lorsqu'elles existent, ces lignes de partage des eaux qui coupent arbitrairement les événements.

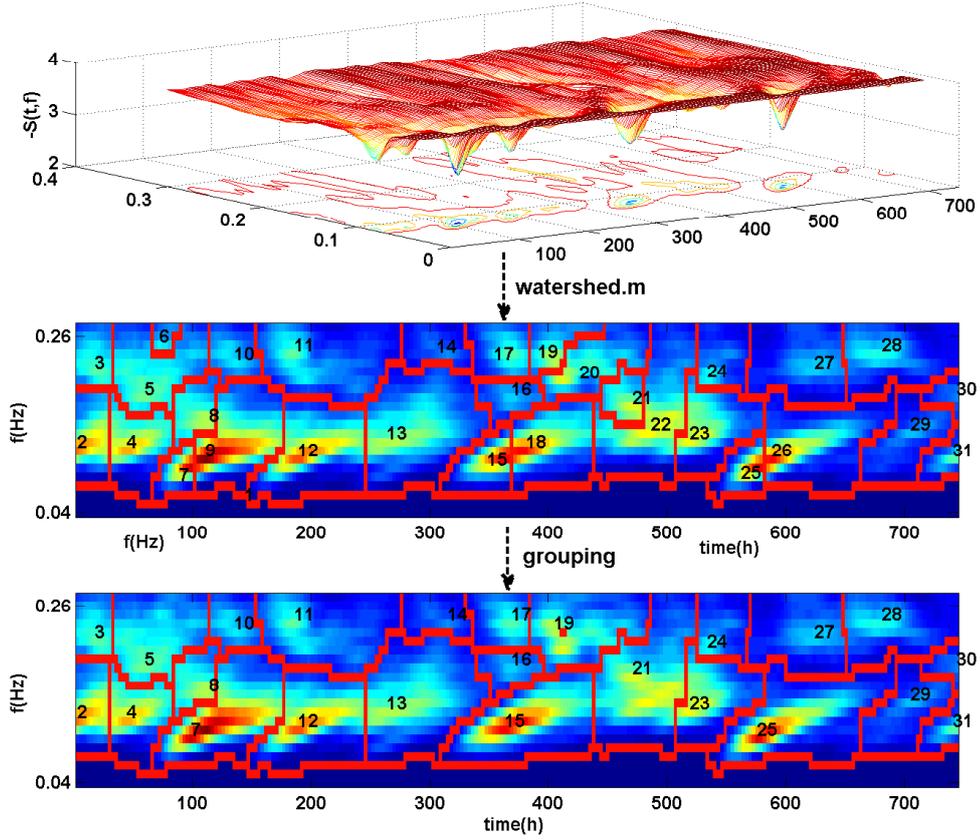


FIGURE 3.9 – Extraction des événements avec l’algorithme des lignes de partage des eaux : (haut) la série-temporelle $S(t, f)$ traitée comme une topographie renversée ; (milieu) identification des bassins versants ; (bas) regroupement des séquences de partitions.

Nous avons mis en place deux critères joints qui permettent de supprimer la ligne de partage des eaux lorsque cette dernière sépare arbitrairement un événement. Deux fractions A et B , dont les pics sont repérés respectivement par les coordonnées $(t_A, f_{p_A}, \theta_{p_A})$ et $(t_B, f_{p_B}, \theta_{p_B})$, sont regroupées si les deux critères suivants sont respectés :

$$\sqrt{(f_{p_B} \cos \theta_{p_B} - f_{p_A} \cos \theta_{p_A})^2 + (f_{p_B} \sin \theta_{p_B} - f_{p_A} \sin \theta_{p_A})^2} \leq D_0, \quad (3.11)$$

et

$$\left| \frac{e_B - e_A}{e_B} \right| \leq \Delta e, \quad (3.12)$$

où, e_i et e_j sont respectivement les énergies des pics des fractions d’événement i et j . L’énergie du pic de la fraction d’événement i à un instant t est

$$e_A = \int_{f_{min,t}}^{f_{max,t}} \int_{\theta_{min,t}}^{\theta_{max,t}} S(t, f, \theta) df d\theta. \quad (3.13)$$

$f_{min,t}$ et $f_{max,t}$ sont respectivement les limites inférieure et supérieure en fréquence et $\theta_{min,t}$ et $\theta_{max,t}$, les limites inférieure et supérieure en direction. D_0 et Δ_e sont des constantes à

fixer aux regards des spécificités du site. Dans la pratique, nous avons obtenu des résultats satisfaisants avec $D_0 = 0.01$ et $\Delta_e = 0.5$.

3.3.5 Estimation des paramètres caractéristiques

Chaque événement identifié peut être caractérisé par les paramètres synthétiques classiques d'états de mer. Le moment d'ordre p est calculé par

$$m_p(t) = \int_{f_{min,t}}^{f_{max,t}} \int_{\theta_{min,t}}^{\theta_{max,t}} f^p S(t, f, \theta) df d\theta \quad \text{pour} \quad t \in [t_i, t_f] \quad (3.14)$$

avec t_i et t_f les instants correspondants respectivement au début et à la fin de l'événement. La hauteur significative s'obtient donc par

$$H_s(t) = 4\sqrt{m_0(t)} \quad \text{pour} \quad t \in [t_i, t_f], \quad (3.15)$$

la fréquence d'énergie par

$$f_{-10}(t) = \frac{m_0(t)}{m_{-1}(t)} \quad \text{pour} \quad t \in [t_i, t_f], \quad (3.16)$$

et la direction moyenne par

$$\theta_m(t) = \arctan \left(\int_0^\infty b_1(t, f) S(t, f) df / \int_0^\infty a_1(t, f) S(t, f) df \right) \quad \text{pour} \quad t \in [t_i, t_f], \quad (3.17)$$

où les coefficients de Fourier a_1 et b_1 sont définis par les équations 2.24 et 2.25.

D'autres paramètres comme la puissance, la fréquence moyenne zero-crossing ou encore le paramètre d'acuité spectrale peuvent également être calculés.

3.3.6 Classification

Le type houle est affecté à un événement si sa fréquence de pic reste 75% du temps en deçà de la fréquence de séparation.

Lorsque les données de vent sont disponibles, le critère de l'âge des vagues (3.5) est utilisé pour définir la fréquence de séparation entre la houle et la mer du vent.

En absence des données de vent, considérant que le choix d'une fréquence de séparation constante peut induire une erreur dans la classification, parce qu'elle évolue en fonction des caractéristiques du vent, nous avons utilisé une fréquence de séparation qui s'exprime en fonction de la hauteur significative. Cette fréquence de séparation est basée sur les travaux de [Torsethaugen, Haver \[2004\]](#). En effet, sous l'hypothèse d'un état de mer complètement développé, la fréquence du pic spectral f_{pf} peut être écrite comme une fonction de la hauteur significative,

$$f_{pf} = \frac{1}{a_f H_s^{1/3}}. \quad (3.18)$$

Si $f_p \geq f_{pf}$, le système est classé comme une mer du vent et si $f_p < f_{pf}$, il est classé comme une houle. L'équation 3.18 est en fait obtenue comme suit : de l'expérience JONSWAP [Hasselmann et al. \[1973\]](#), les relations suivantes ont été établies :

$$H = k_1 F^{1/2} \quad (3.19)$$

$$T = k_2 F^{1/3} \quad (3.20)$$

$H = \frac{gH_s}{u^2}$, $T = \frac{gT_p}{u}$, et $F = \frac{gF_e}{u^2}$ sont respectivement les valeurs adimensionnelles des paramètres H_s , T_p et du fetch F_e , et u la vitesse du vent. Durant l'expérience JONSWAP, Hasselmann et al. [1973] ont utilisé la vitesse du vent à 10m au dessus de la surface de la mer (U_{10}) et trouvé que $k_1 = 1.6 \times 10^{-3}$ et $k_2 = 0.286$.

À partir de ces relations, il vient que :

$$\frac{T}{H^{1/3}} = k_3 F^{1/6} \quad k_3 = \frac{k_2}{k_1^{1/3}}. \quad (3.21)$$

En remplaçant les paramètres adimensionnels par leurs expressions, on obtient :

$$f_{pf} = \frac{1}{k_f F_e^{1/6} H_s^{1/3}} \quad (3.22)$$

où f_{pf} correspond à la fréquence de pic de l'état de mer complètement développé, et $k_f = k_3 g^{-1/2} = 0.78 sm^{-1/2}$. L'équation 3.18 s'obtient finalement en posant,

$$a_f = k_f F_e^{1/6}. \quad (3.23)$$

Le facteur a_f varie légèrement avec le fetch. Par exemple, pour un fetch de 370km, $a_f = 6.6 sm^{-1/3}$ et pour un fetch de 100km, $a_f = 5.3 sm^{-1/3}$.

3.4 Résultats

Afin d'évaluer la pertinence du TESP, nous l'avons utilisé pour identifier les événements de houle et de mer du vent sur deux sites à climatologies d'état de mer différentes. Le premier site est notre zone d'étude, l'Afrique de l'Ouest où les états de mer sont dominés par des houles. Le second site est le golfe de Gascogne où les mers du vent sont beaucoup plus intenses et où les données de vent disponibles sont fiables. Nous avons utilisé des données *in situ* pour l'Afrique de l'Ouest et des données de hindcast pour le golfe de Gascogne ; le but étant d'évaluer également la méthode sur deux différentes sources de données. Il est aussi intéressant de comparer, les événements identifiés des données *in situ* et de hindcast sur un même site notamment en Afrique de l'Ouest. Le chapitre 4 sera dédié à cette fin. Cette section se terminera par une comparaison entre SPOP et TESP.

3.4.1 Données *in situ* d'Afrique de l'Ouest

Nous présentons ici les résultats (figure 3.10) obtenus pour le mois de décembre 2002 à partir des données *in situ*. Sur cette figure, de haut en bas, sont représentées respectivement, la série-temporelle des densités spectrales $S(t, f)$, la hauteur significative, la fréquence de pic et la direction moyenne du pic. Il est possible d'observer globalement une certaine cohérence entre les événements repérables sur l'image de la série-temporelle des densités spectrales, et l'évolution des paramètres H_s , f_p et θ_m .

Il est facile de reconnaître les événements de houle, par leurs fréquences de pic se trouvant en deçà de la fréquence de séparation et dont les évolutions au cours du temps sont

croissantes. Suivant les directions de provenance, nous pouvons distinguer les événements de houle du Sud-Ouest (θ_m autour de 210°) et un événement de Nord-Ouest survenu à la fin du mois (θ_m autour de 300° , couleur turquoise). Les états de mer, pour cette période et pour la quasi-totalité du temps, sont dominés par les houles du Sud-Ouest. Les houles du Nord-Ouest sont quant à elles, de faible énergie avec des hauteurs significatives inférieures à $0.5m$. Nous pouvons également remarquer une évolution typique de la hauteur significative au cours d'un événement de houle. Le H_s commence par croître avec l'arrivée des premiers trains de houle, atteint un maximum, puis décroît. Cette évolution typique a été déjà soulignée par d'autres études notamment [Snodgrass et al. \[1966\]](#), [Gjevik et al. \[1988\]](#), [Husson \[2012\]](#).

Les événements de mer du vent sont les événements dont les fréquences sont au dessus de la fréquence de séparation (couleur bleue). Contrairement aux événements de houle, aucune évolution typique ne peut être identifiée avec certitude. Notons que leur fréquence de pic a tendance à décroître pendant l'événement.

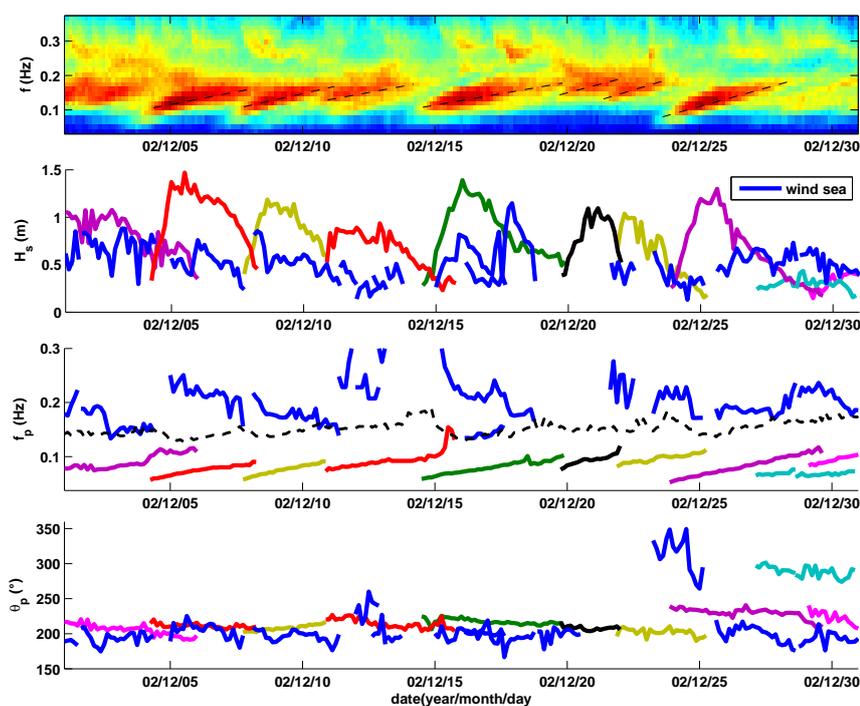


FIGURE 3.10 – Identification des événements : série-temporelle de la densité spectrale (fenêtre du haut) et paramètres synthétiques des événements. Les événements de mer du vent sont représentés en couleur bleue, et les événements de houle sont chacun représentés avec une couleur différente. Sur la troisième fenêtre, la ligne pointillée en couleur noire représente la fréquence de séparation.

3.4.2 Données de Hindcast de la mer de Gascogne

TESP a été également testée sur des données de HOMERE en mer de Gascogne. L'objectif est d'évaluer la pertinence de la méthode sur un autre site, où le vent est de forte amplitude et où les données de vent sont fiables. La base de données HOMERE a été construite avec le modèle numérique de vagues Wave Watch III®, et couvre la Manche, la mer d'Iroise et le golfe de Gascogne. Le point choisi pour cette étude a pour coordonnées ($45^{\circ}18'N$ et $1^{\circ}19'W$) et la profondeur de l'eau est de 21m. Les données, échantillonnées au pas d'une heure sont constituées de spectres directionnels, de la vitesse et de la direction du vent. La figure 3.11 montre le résultat pour la période allant du 19 février au 14 mars

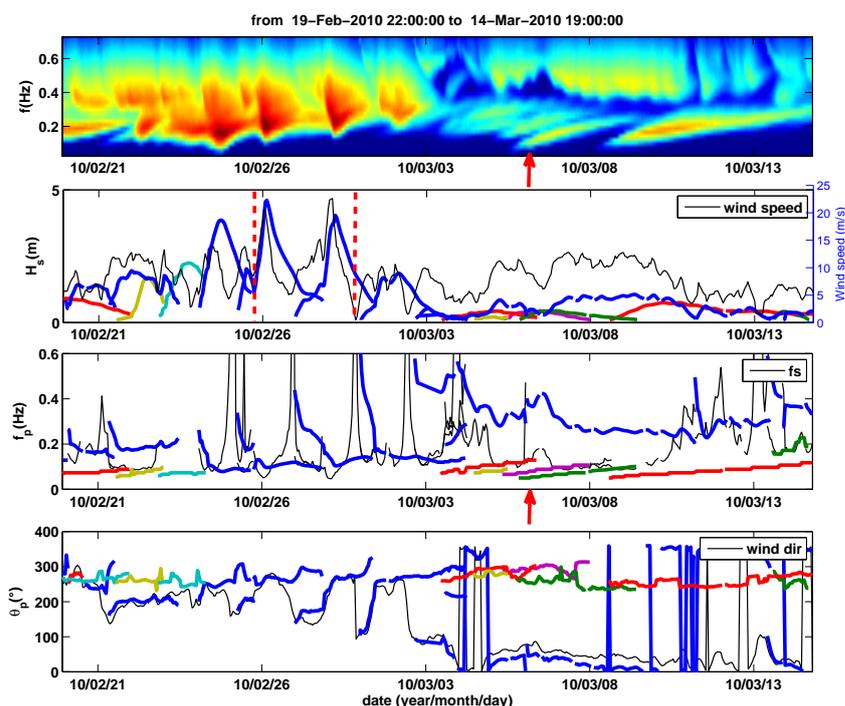


FIGURE 3.11 – Identification des événements à partir des données de Hindcast en mer de Gascogne. L'intensité et la direction du vent sont également représentées : série-temporelle de la densité spectrale (fenêtre du haut) et paramètres synthétiques des événements. Les événements de mer du vent sont représentés en couleur bleue, et les événements de houle sont chacun représentés avec une couleur différente. L'intensité et la direction du vent sont également représentées.

2010. La classification a été effectuée avec le critère de l'âge des vagues. Les événements de mer du vent sont représentés en bleu, et les événements de houle sont représentés chacun avec une couleur différente. L'intensité et la direction du vent sont également représentées.

Nous observons une bonne corrélation entre les caractéristiques de la mer du vent et le vent : sur la deuxième fenêtre de la figure, on remarque que l'intensité du vent conditionne la hauteur significative des mers du vent. Le H_s des mers du vent atteint généralement son maximum quasiment au même moment que le maximum de l'intensité

du vent. On note toutefois un léger retard qui n'est en réalité que le temps de réponse de la mer sous l'action du vent. De même la direction moyenne des mers du vent est très proche de la direction du vent. Il faut préciser qu'un écart maximum de 90° entre la direction de la mer du vent et celle du vent est acceptable selon le critère de l'âge des vagues. Remarquons, enfin, les forts événements de mer du vent survenus entre la période 26 février et le 1 mars 2010, repérables par deux traits pointillés en rouge sur la deuxième fenêtre graphique de la figure. Les fréquences de pic de ces événements de mer du vent sont très basses avec des vitesses du vent atteignant la valeur de 22 m/s. On note également l'absence d'événements de houle pendant cette période. En effet, cette période correspond au passage de la tempête Xynthia qui avait frappé plusieurs pays européens, et qui a causé des pertes de vies humaines et de nombreux dégâts matériels.

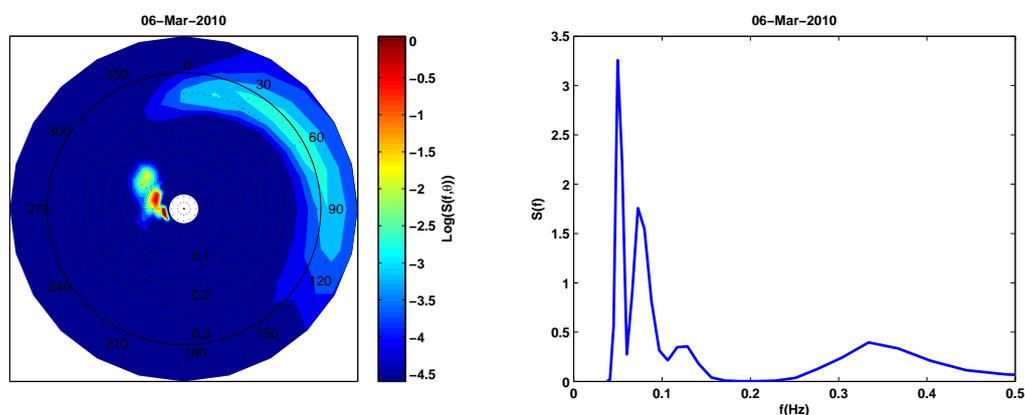


FIGURE 3.12 – Spectre directionnel (gauche) et omnidirectionnel (droite) du 6 mars 2010 à 3h00 :

Les fréquences de pic des événements de houle identifiés évoluent de façon quasi-linéaire comme le prévoit la théorie. Focalisons nous sur la date du 6 mars 2010, indiquée sur la figure 3.11 par la flèche rouge. En plus de la mer du vent, 3 systèmes de houle très proches en fréquences, ont été identifiées. L'analyse du spectre à cette date (figure 3.12) montre la présence effective de ces trois systèmes de houle. Cela montre que TESP permet d'obtenir un partitionnement cohérent des états de mer, même si ces derniers sont constitués de systèmes de vagues évoluant à des fréquences voisines, avec des hauteurs significatives faibles.

3.4.3 Comparaison de TESP avec SPOP

La figure 3.13 présente les événements identifiés avec SPOP (à gauche) et avec TESP (à droite), à partir des données *in situ*. Sur les parties hautes, l'image correspondant à la série temporelle des densités spectrales avant le partitionnement est représentée, dans le but de voir si les événements identifiés correspondent à ce que l'on peut observer sur l'image. En se référant à l'image de la série-temporelle des densités spectrales (fenêtre du haut), nous pouvons globalement dire que les événements identifiés avec TESP sont plus cohérents. Pour ce cas, certains événements de houle ont été arbitrairement recollés par SPOP ; c'est le cas pour l'événement entre le 15 décembre et le 25 décembre (couleur verte).

Comme mentionné plus haut, la difficulté principale dans l'utilisation de SPOP réside dans le choix des constantes qui interviennent dans les critères empiriques de regroupement et de suivi temporel. En absence de données de vent, la fréquence de séparation constante qui est utilisée dans SPOP peut aussi induire des erreurs dans la classification des systèmes. Dans ces critères de suivi temporel, SPOP autorise, d'un pas de temps au suivant, que le H_s soit multiplié ou divisé par 1.5. Or, en Afrique de l'Ouest, les hauteurs significatives des systèmes de vagues sont généralement faibles comme on peut le voir sur la figure 3.13, où le H_s maximum est 1.5m. Cela pourrait expliquer le fait que pour ce mois, SPOP fusionne des événements normalement distincts. Cependant, notons que le suivi temporel des systèmes de vagues qui est fait dans SPOP a pour premier objectif d'améliorer le partitionnement des spectres individuels.

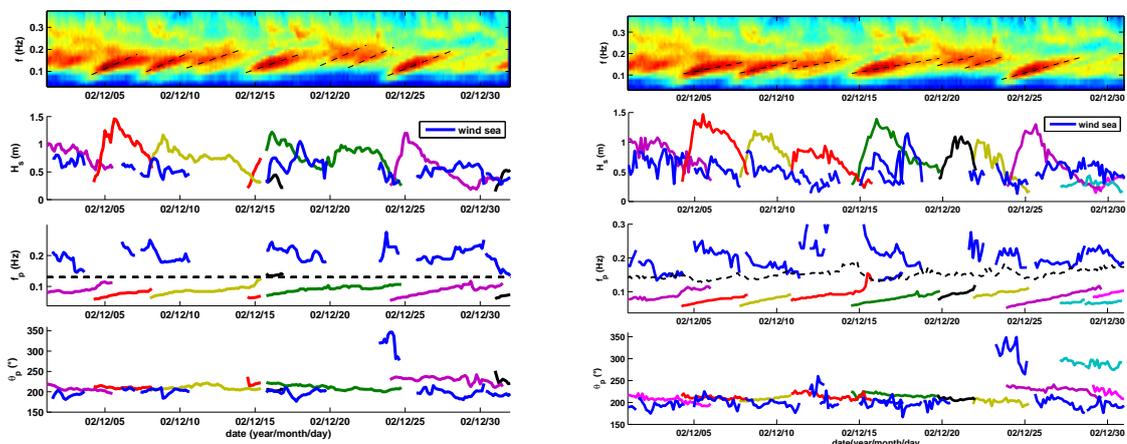


FIGURE 3.13 – Identification des événements avec SPOP (gauche) et TESP (droite)

3.5 Conclusion

Ce chapitre nous a permis, dans un premier temps, de définir le concept “événement”, élément de base de l’approche qu’on se propose d’utiliser dans ces travaux.

Nous avons ensuite présenté une nouvelle méthode d’identification d’événements de houle et de mer du vent (TESP), développée dans le cadre de ces travaux pour extraire les événements de houle et de mer du vent. Cette méthode, comme la plupart de celles qui existent, utilise l’algorithme de la ligne de partage des eaux. Mais, elle diffère des autres par l’inclusion de la dimension temporelle dans la segmentation : l’algorithme de la ligne de partage des eaux est appliqué en 3D, sur la série temporelle de spectres directionnels. TESP présente l’avantage de réduire considérablement le besoin de critères empiriques, souvent utilisés dans les méthodes existantes et favorise en conséquence le traitement automatique et la réduction du temps de calcul. Nous avons montré que TESP permet d’identifier des événements cohérents, avec des données d’Afrique de l’Ouest, mais aussi, avec des données d’un autre site (golfe de Gascogne) à dynamique plus complexe. La bonne corrélation entre les caractéristiques de la mer du vent et celles du vent, l’évolution quasi-linéaire de la fréquence de pic des événements de houle identifiés, sont les raisons pour lesquelles nous

pouvons conclure que la méthode proposée est pertinente. La comparaison avec SPOP a montré que TESP garantit, mieux qu'elle, l'extraction des événements cohérents avec la physique des systèmes de vagues.

Enfin, soulignons que l'implémentation de cette nouvelle méthode d'identification d'événements est d'une grande importance dans la mise en œuvre de l'approche que nous proposons d'utiliser. Elle a permis de constituer une base de données d'événements de houle et de mer du vent fiables qui a servi dans la mise en œuvre de notre approche.

Validation des données de hindcast à l'échelle des événements

Dans ce chapitre, nous comparons les événements de houle et de mer du vent identifiés à partir des données *in situ* avec ceux identifiés à partir des données en sortie de modèle (hindcast). L'objectif principal de cette étude est de valider les données de hindcast utilisées, par la suite, dans ces travaux. En effet, même s'il est préférable d'utiliser des données *in situ* (mesures de bouée par exemple) pour la modélisation statistique à long-terme des états de mer, nous ne disposons ici que de 3 ans de données sur le site Girassol (situé au large de l'Angola) et seulement d'une année de données sur le site Akpo (situé au large du Nigéria). Les durées que couvrent ces données d'observations sont trop courtes pour reproduire par modélisation stochastique leurs propriétés statistiques, notamment la dépendance entre les paramètres ainsi que l'alternance des saisons. Or, sur ces deux sites, nous disposons de 21 ans de données de hindcast dont l'utilisation peut permettre une modélisation stochastique plus pertinente et intégrant la variabilité saisonnière. Mais avant d'être utilisées, ces données de hindcast doivent être validées par rapport aux observations. Puisque nous proposons une représentation par événements de la climatologie des états de mer, cette validation doit s'opérer à l'échelle des événements afin de vérifier si la cohérence temporelle des systèmes de vagues est bien reproduite dans les données de hindcast.

Ce chapitre présente également une classification des systèmes de vagues en fonction de leur direction de provenance et leur contribution à l'énergie totale des états de mer sur les sites Girassol et Akpo.

Sommaire

4.1	Données utilisées	63
4.2	Comparaison des hauteurs significatives des états de mer	64
4.3	Comparaison, à l'échelle des événements de houle et de mer du vent, des données <i>in situ</i> et des données de hindcast	66
4.3.1	Classification des événements	66
4.3.2	Comparaison des événements de houle	69
4.3.3	Comparaison des événements de mer du vent.	73
4.4	Conclusion	75

4.1 Données utilisées

Cette étude considère deux sites situés au large des côtes d’Afrique de l’Ouest : Girassol localisé au large de l’Angola et Akpo localisé au large du Nigéria (voir figure 4.1). Sur le site Girassol, nous disposons des données de bouée couvrant la période de mars 2001 à mars 2004. Il faut souligner qu’à cause des problèmes de maintenance, beaucoup de données sont manquantes. La proportion de données manquantes est sensiblement égale à 25%. Sur Akpo, nous disposons des données de bouée d’avril 2003 à mai 2004.

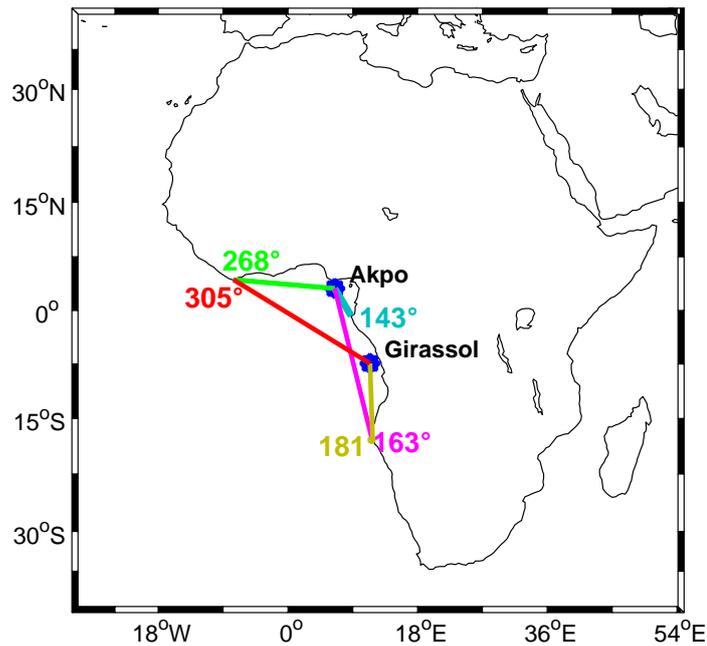


FIGURE 4.1 – Positions des sites Girassol et Akpo avec les principaux secteurs possibles de provenances des houles

Les données *in situ* sont en effet constituées de mesures de déplacements de la surface de la mer $\{x(t), y(t), z(t)\}$ obtenus à partir de bouées instrumentées ancrées par 1450 m de profondeur sur le site Girassol et par 1307 m de profondeur sur le site Akpo. À partir de ces mesures de déplacements, des spectres directionnels ont été estimés, à raison d’un spectre toutes les 30 min, puis moyennés sur 3 h afin de les rendre comparables aux données de hindcast, échantillonnées par pas de 3 h. La technique utilisée pour estimer les spectres directionnels est celle décrite dans (Kpogo-Nuwoklo et al. [2015a], voir Annexe B). Cette technique, développée dans le cadre de ces travaux, utilise la cohérence temporelle des systèmes de vagues pour améliorer l’estimation des spectres d’état de mer. Elle permet également d’interpoler correctement jusqu’à 12 h les spectres manquants et surtout de mieux identifier, délimiter et séparer les événements de houle et de mer du vent.

Quant aux données de hindcast disponibles, elles sont constituées de spectres directionnels et couvrent pour ces deux sites la période de 1991 à 2012. Elles proviennent de la

base de données IWOAGA¹ et ont été obtenues à partir du modèle numérique de vagues Wave Watch III ©(Ardhuin et al. [2011]).

4.2 Comparaison des hauteurs significatives des états de mer

La validation des données de hindcast se fait souvent sur les paramètres globaux, c'est à dire sans distinction de la houle et de la mer du vent (Swail et al. [2000], Cox et al. [1999], Boudière et al. [2013]). La hauteur significative est généralement utilisée comme paramètre primaire pour la validation des données en sortie de modèles. Avant de faire une comparaison à l'échelle des événements de houle et de mer du vent, il nous a semblé nécessaire de comparer d'abord les paramètres globaux, notamment la hauteur significative.

Les figures 4.2 et 4.3 présentent respectivement pour les sites Girassol et Akpo les résultats de la comparaison des hauteurs significatives. Sur Girassol, nous observons dans l'ensemble un bon accord entre les H_s avec un biais quasiment nul (-0.06 m) et un RMSE=0.2 m. Le diagramme quantile/quantile (en bas, à droite) montre que l'accord est excellent jusqu'à une hauteur significative de 2.5m, à partir de laquelle le modèle sous-estime les H_s . Le nuage des points (en bas, à gauche) montre une dispersion relativement faible des points autour de la première bissectrice avec un coefficient de corrélation linéaire de 0.87 entre le H_s .

En ce qui concerne Akpo, le biais est également très faible (0.02 m) et le *RMSE* est égal 0.18 m. Le diagramme quantile/quantile montre cependant une légère sur-estimation des faibles valeurs de H_s ($H_s \leq 0.8$ m) et une légère sous-estimation des $H_s > 2$ m. Toutefois, dans l'ensemble, les H_s issus des données de hindcast sont très proches de ceux observés avec coefficient de corrélation entre eux égal à 0.83.

En résumé, cette comparaison a montré que les hauteurs significatives globales calculées à partir des données de hindcast sont très proches dans l'ensemble de celles calculées à partir des observations, aussi bien sur le site Girassol que sur le site Akpo. Les données de hindcast sont donc d'une qualité globale suffisante. Notre objectif est de voir si ces résultats se confirment à l'échelle des événements de houle et de mer du vent.

1. IWOAGA : Integrated Ocean waves for Geophysical and other Applications (<http://wwz.ifremer.fr/iowaga>)

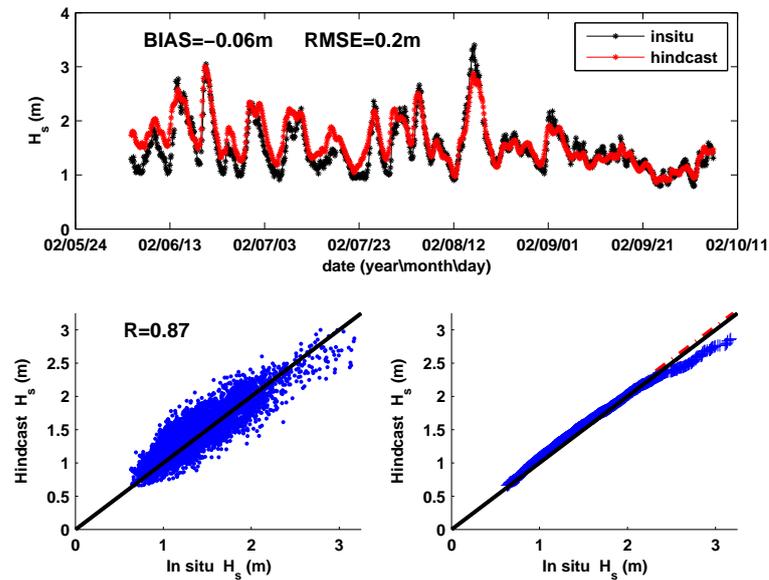


FIGURE 4.2 – Comparaison du H_s global sur le site Girassol : Série temporelle (en haut), nuage des points (en bas, à gauche) et diagramme quantile/quantile (en bas, à droite).

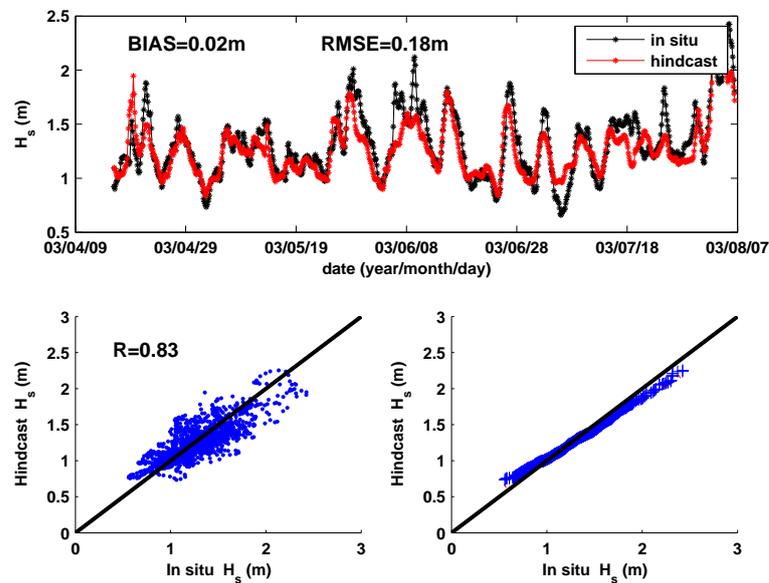


FIGURE 4.3 – Comparaison du H_s global sur le site Girassol : Série temporelle (en haut), nuage des points (en bas, à gauche) et diagramme quantile/quantile (en bas, à droite).

4.3 Comparaison, à l'échelle des événements de houle et de mer du vent, des données *in situ* et des données de hindcast

La comparaison entre les H_s des deux sources de données a montré que le modèle représente globalement bien les états de mer sur l'ensemble des deux sites. Cependant, comme mentionné dans les chapitres précédents, les houles et les mers du vent ont différentes sources et évoluent suivant des dynamiques différentes. Il est possible qu'une mauvaise reproduction des houles par le modèle soit arbitrairement compensée par les mers du vent (et vice versa). La comparaison des paramètres globaux ne permet pas de déceler ces erreurs, si elles existent. De plus, la calibration des modèles numériques s'opère souvent avec les paramètres globaux alors que les sources d'erreurs de la modélisation de la houle et de la mer du vent peuvent être complètement différentes : une mauvaise représentation de la mer du vent dans le modèle peut être due, par exemple, à la mauvaise connaissance du vent local, alors qu'une mauvaise représentation de la houle peut être liée, par exemple, à un mauvais paramétrage de sa dissipation. Il est donc nécessaire de faire une comparaison à l'échelle des systèmes de vagues, afin de s'assurer que le modèle reproduit correctement à la fois les systèmes de houle et de mer du vent. De plus, si nous voulons utiliser les données de hindcast en substitution des données d'observations, dans le cadre de notre démarche de la représentation par événements de la climatologie, il est indispensable de vérifier que la cohérence temporelle des systèmes de vagues est bien reproduite dans les données de hindcast.

Pour ce faire, les événements de houle et de mer du vent ont été d'abord identifiés à partir des séries temporelles de spectres directionnels des données *in situ* et des données de hindcast, avec la méthode décrite dans le chapitre 3. Pour réaliser la comparaison des événements, une classification de ces derniers est nécessaire afin de distinguer les événements de houle qui peuvent coexister au même instant.

4.3.1 Classification des événements

Généralement, les systèmes de vagues sont classés en catégories mer du vent, houle primaire, houle secondaire, etc. La houle primaire se distingue de la houle secondaire uniquement par comparaison de leur H_s ou T_p ou encore θ_p . Par exemple pour Nerzic et al. [2007], entre deux houles présentes au même instant, la houle primaire est celle dont la hauteur significative est la plus grande. En suivant cette classification, un système de houle en évolution (événement de houle), peut se retrouver classé dans des catégories différentes en fonction des valeurs prises par le H_s . Cette classification n'intègre donc pas assez de considérations physiques au sens où, par définition, un événement est une entité reflétant la dynamique d'un système de vague en évolution au cours du temps et devrait être classé dans une seule catégorie.

Afin de permettre une analyse cohérente avec la définition des événements, il nous a paru utile de classer les événements de houles en fonction de leur direction de provenance.

4.3.1.1 Site Girassol

Les diagrammes polaires montrant la hauteur significative et la période des systèmes de vagues en fonction de leur direction de provenance, pour les données *in-situ* (à gauche) et

pour les données de hindcast (à droite), sont présentés dans la (figure 4.4). La répartition angulaire des systèmes de vagues est globalement similaire dans les deux cas même. Notons tout de même que les systèmes de vagues sont beaucoup plus étalés dans cas des données d'observations, ce qui est dû à l'aspect bruité des données d'observations. On distingue, en plus des événements de mer du vent qui proviennent préférentiellement du secteur Sud-Ouest, deux catégories d'événements de houle se distinguent suivant leur direction de provenance :

- Les événements de houle du Sud-Ouest ($181^\circ \leq \theta_m < 270^\circ$), générée par des tempêtes répétitives, localisées dans l'Atlantique Sud. Ce sont les systèmes de vagues dominants en terme d'énergie et les plus fréquents sur ce site.
- Les événements de houle de Nord-Ouest ($270^\circ \leq \theta_m < 305^\circ$), présents uniquement pendant l'hiver boréal (décembre à février), à raison de 2 à 3 événements par mois.

Un avantage de cette classification par direction est qu'elle permet d'évaluer, suivant leur zone de génération, la contribution moyenne (en pourcentage) à l'énergie totale (H_s^2) ou à la puissance totale des états de mer des différents systèmes de vagues présents sur le site. Les contributions des systèmes de vagues ont été estimées et les résultats sont montrés dans le tableau 4.1. Que ce soit avec les données *in situ* ou avec les données de hindcast, les houles de Sud-Ouest sont les systèmes dominants en terme de contribution à l'énergie totale et la puissance puissance totale des états de mer. Elles contribuent à plus de 80% à l'énergie et à près de 90% à la puissance des états de mer. Les mers du vent contribuent faiblement à l'énergie totale ou à la puissance des états de mer, ce qui est dû à l'intensité modérée du vent au large de l'Angola (vitesse moyenne du vent égale 5 m/s). Quant aux houles de Nord-Ouest, leur contribution à l'énergie et à la puissance des états de mer est quasiment nulle (inférieure à 1%). Du fait des grandes distances parcourues depuis leur zone de génération (Atlantique Nord) jusqu'au point d'observation (au large de l'Angola), les houles de Nord-Ouest perdent une grande partie de leur énergie au travers des phénomènes de dispersion et dissipation. Ces houles ont généralement des hauteurs significatives très faibles, souvent inférieures à 0.5 m.

		mer du vent	houle du SW	houle du NW
énergie (H_s^2)	<i>in situ</i>	14.04%	85.57%	0.38%
	hindcast	17.76%	81.94%	0.30%
puissance (P_w)	<i>in situ</i>	7.28%	92.24%	0.46%
	hindcast	9.82%	89.92%	0.25%

TABLE 4.1 – Contribution à l'énergie totale et à puissance totale des systèmes de vagues (Girassol)

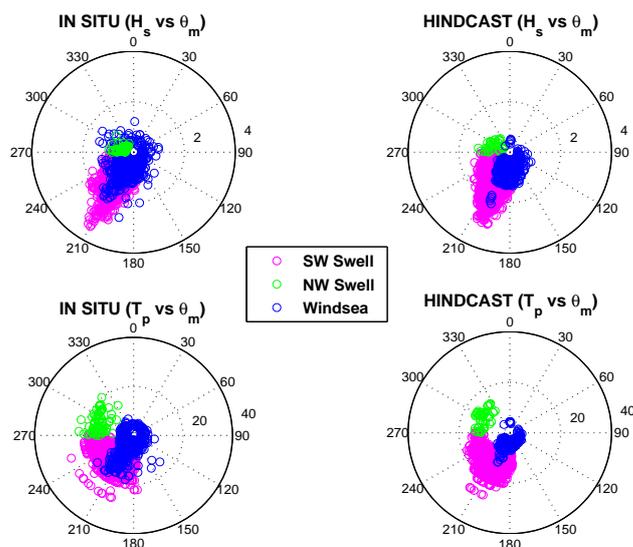


FIGURE 4.4 – Diagramme polaire de la hauteur significative et de la période de pic suivant la direction moyenne des systèmes de vagues (GIRASSOL).

4.3.1.2 Site Akpo

Sur le site Akpo, il n'est pas possible d'avoir des systèmes de houle provenant du secteur compris entre 268° et 143° dans le sens géographique (voir Figure 4.1). Le site Akpo est donc abrité des houles de Nord-Ouest. Sur les diagrammes polaires de la figure 4.5 on peut distinguer, en plus des événements de mer du vent :

- Les événements de houle provenant du secteur Sud-Est ($143^\circ \leq \theta_m < 163^\circ$), en faible proportion et généralement de hauteurs significatives faibles ($H_s < 1$ m).
- Les événements de houle de Sud et Sud-Ouest ($163^\circ \leq \theta_m < 268^\circ$), dominants en terme d'énergie et plus fréquents.

Nous avons également évalué pour le site Akpo la contribution à l'énergie et à la puissance des états de mers des systèmes en présence. Les résultats, présentés dans le tableau 4.2, montrent que les systèmes de vagues dominants sont également les houles du Sud-Ouest avec une contribution à plus de 70%. Comparées au site Girassol, les mers du vent sur le site Akpo ont une contribution supérieure à l'énergie et à la puissance des états de mer. Les houles de Sud-Est sont les systèmes dont la contribution est la plus faible. Ces houles ne sont pas fréquentes et leur énergie est relativement faible ($H_s < 1$ m).

En définitive, la présente classification, basée sur la direction de provenance des houles, permet une analyse cohérente des événements. Par rapport à une classification basée sur les valeurs de H_s ou de T_p , elle garantit le fait qu'un même système de houle en évolution soit classé dans une même catégorie. Cette classification a également permis de constater que les houles de Sud-Ouest contribuent majoritairement à l'énergie et à la puissance totale des états de mer en Afrique de l'Ouest.

4.3. Comparaison, à l'échelle des événements de houle et de mer du vent, des données *in situ* et des données de hindcast

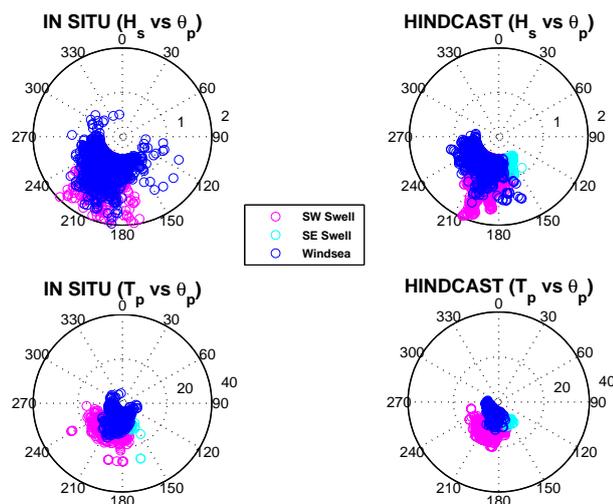


FIGURE 4.5 – Diagramme polaire de la hauteur significative et de la période de pic suivant la direction moyenne des systèmes de vagues (AKPO).

		mer du vent	houle du SW	houle du SE
énergie (H_s^2)	<i>in situ</i>	26.42%	73.01%	0.47%
	hindcast	27.73%	70.72%	1.54%
puissance (P_w)	<i>in situ</i>	17.39%	81.97%	0.62%
	hindcast	20.42%	77.06%	2.5%

TABLE 4.2 – Contribution à l'énergie totale et à puissance totale des systèmes de vagues (Akpo)

4.3.2 Comparaison des événements de houle

On se propose de comparer les événements de houle sur les deux sites en suivant la classification présentée dans 4.3.1.

4.3.2.1 Site Girassol

La comparaison des événements de houle sur le site Girassol a été effectuée sur la période que couvrent les 3 ans de données *in situ*. La figure 4.6 montre les événements de houle identifiés à partir des données *in situ* et des données de hindcast pour le mois de janvier 2003. Nous avons choisi de montrer les résultats de ce mois car il correspond à la période où sont présents à la fois les événements de houle de Sud-Ouest et de Nord-Ouest. Notons que pour ce mois, des données *in situ* sont manquantes entre le 14 janvier et le 17 janvier.

Nous observons un bon accord entre les événements de houle de Sud-Ouest identifiés à partir des deux sources de données aussi bien pour les valeurs prises par les paramètres que pour la cohérence temporelle des systèmes de houle. Les hauteurs significatives affichent dans les deux cas, la même évolution : au début de l'événement, la hauteur significative

croît, atteint un maximum puis décroît. Les dates du maximum de H_s des événements coïncident quasiment. Sur les fréquences de pic, la coïncidence est encore notable alors qu'elle l'est moins en ce qui concerne les directions moyennes, ce qui dû en grande partie à l'aspect bruité dans l'évolution des directions moyennes estimées à partir des spectres d'observations.

Quant aux événements de houle de Nord-Ouest (Girassol), ils ne sont pas toujours présents dans les données *in situ*. Par exemple l'événement identifié entre le 18 janvier et le 24 janvier 2003 des données de hindcast (repérable par deux traits pointillés en rouge, (figure 4.6), n'est pas présent dans les données *in situ*. L'absence de cet événement n'est pas due à la méthode d'identification d'événement utilisée, car les spectres directionnels issus des données *in situ* pendant cette période montrent l'absence effective de systèmes de houle de Nord-Ouest (figure 4.7). Une explication plausible est que la bouée n'as pas pu mesurer cet événement de houle car ce dernier a une énergie très faible. Néanmoins, lorsqu'elles sont identifiables dans les données *in situ*, les événements de houle de Nord-Ouest concordent globalement bien avec les événements de houle identifiés à partir des données de hindcast.

Pour l'ensemble des événements de houle identifiés des 3 ans de données, la figure 4.8 montre les diagrammes quantile/quantile des hauteurs significatives pour les événements de houles de Sud-Ouest (à gauche) et les événements de houles de Nord-Ouest (à droite). Pour les houles de Sud-Ouest, on note un bon accord entre entre les quantiles des H_s jusqu'à 2.5 m où les H_s sont sous-estimés dans les données de hindcast. Le coefficient de corrélation entre les H_s est de 0.81. Malgré leur faibles H_s , les houles de Nord-Ouest présentent également une bonne corrélation apparente. Notons néanmoins une sur-estimation des fortes valeurs du H par le modèle.

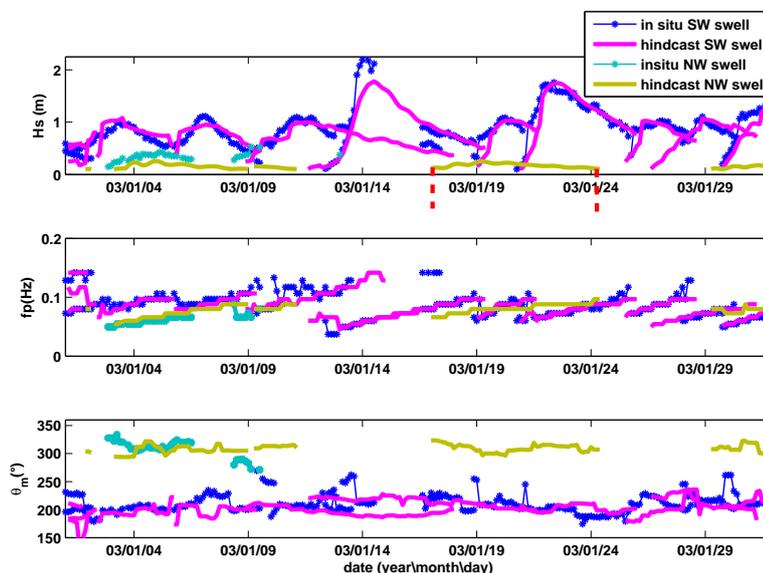


FIGURE 4.6 – Comparaison des événements de houle sur Girassol (janvier 2003)

4.3. Comparaison, à l'échelle des événements de houle et de mer du vent, des données *in situ* et des données de hindcast

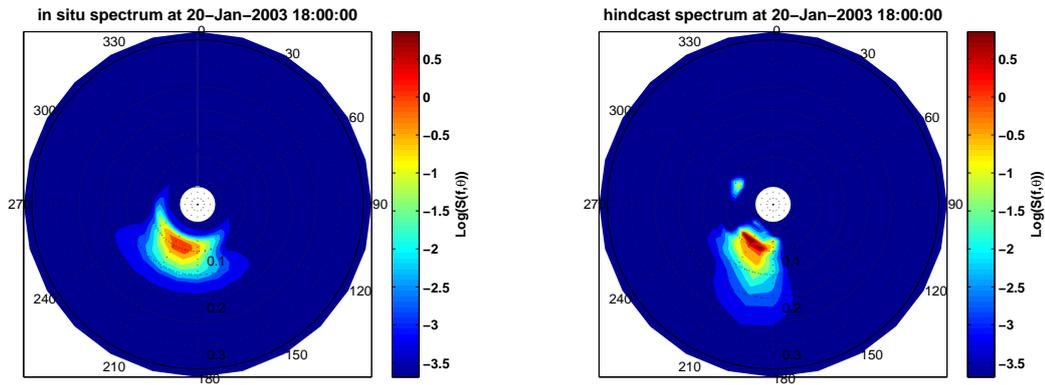


FIGURE 4.7 – Spectres directionnels obtenus le 20 janvier 2003 à 18h00 : *in situ* (à gauche) et hindcast (droite).

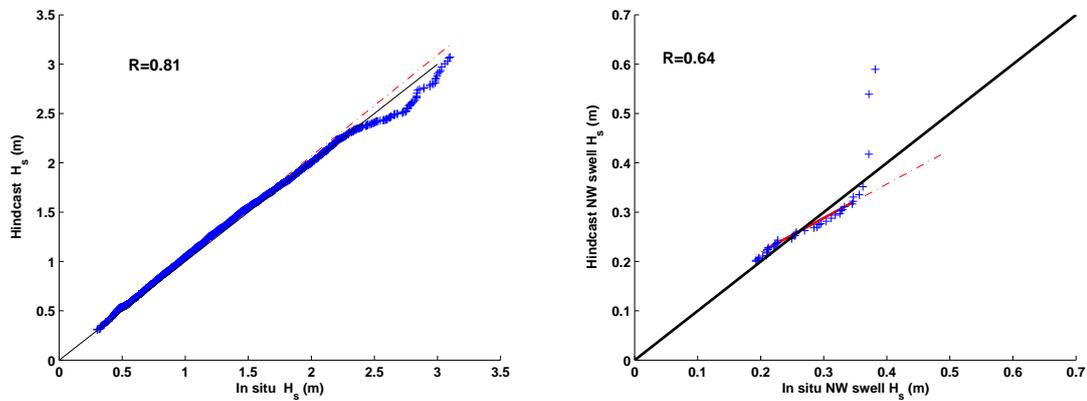


FIGURE 4.8 – Diagramme quantile/quantile des systèmes de houles : Sud-Ouest (à gauche), Nord-Ouest (à droite)

4.3.2.2 Site Akpo

Sur le site Akpo la comparaison a été réalisée sur la période d'une année que couvrent les données *in situ* disponibles. Les figures 4.9 montre les événements de houle identifiés à partir des données *in situ* (décembre 2003), comparés aux événements de houle identifiés à partir des données de hindcast. Entre les H_s , on note dans l'ensemble un bon accord même s'ils sont modérément sous-estimés dans les données de hindcast. Les débuts, les fins des événements ainsi que le maximum de H_s des événements coïncident quasiment ; on retrouve également la même cohérence temporelle des systèmes de houle. En ce qui concerne les fréquences de pic, leurs évolutions au sein des événements sont similaires dans les deux cas. L'accord est moins bon entre les directions de pic, en raison d'une part d'une large discrétisation en direction des spectres de hindcast (15°) et d'autre part de l'aspect bruité de l'évolution des directions de pic issues des données *in situ*.

Pour tous les événements de houle identifiés, la figure 4.10 montre les diagrammes quantile/quantile entre les H_s des houles du Sud-Ouest (à gauche) et des houles de Sud-

Est (à droite). Nous notons une bonne corrélation entre les quantiles des H_s des houles Sud-Ouest jusqu'à 1.15 m où ils sont sous-estimés dans les données de hindcast. En ce qui concerne les houles du Sud-Est, la corrélation est faible comparée à celle obtenue avec les houles du Sud-Ouest, mais reste apparente.

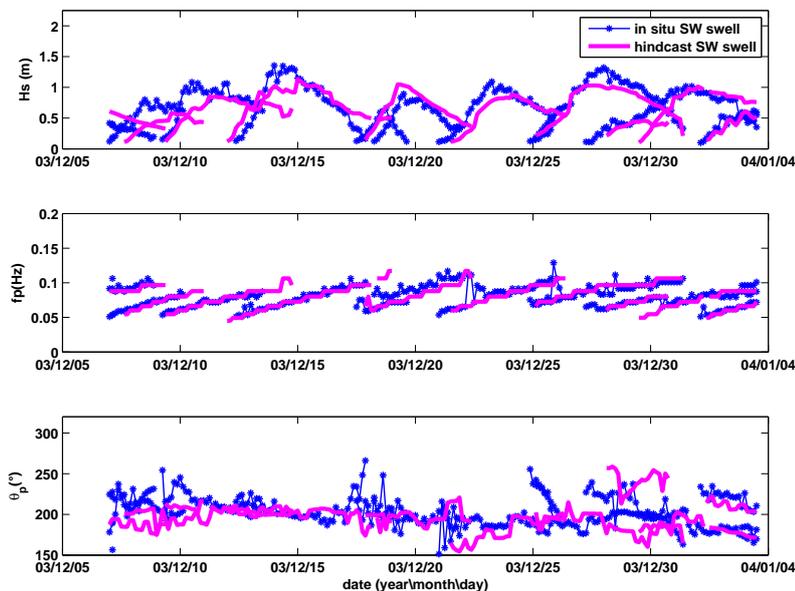


FIGURE 4.9 – Comparaison des événements de houle sur Akpo (décembre 2003)

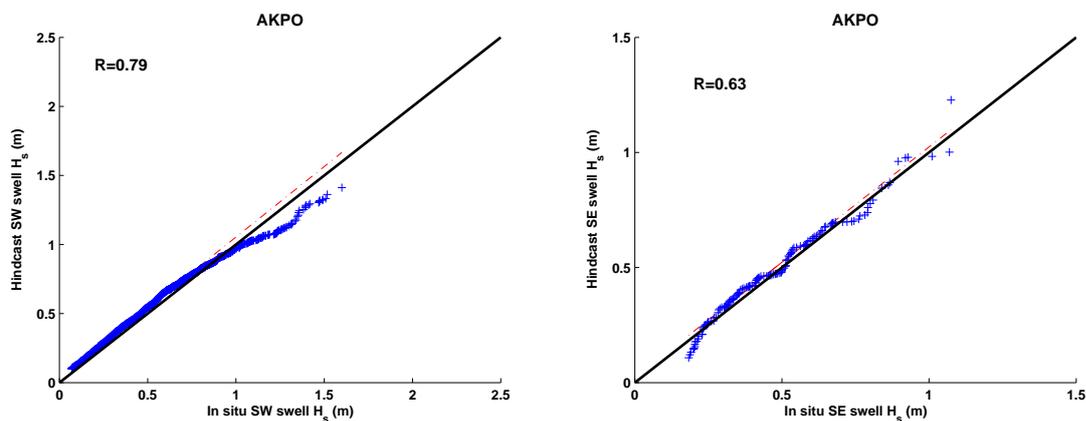


FIGURE 4.10 – Diagramme quantile/quantile des systèmes de houles : Sud-Ouest (à gauche), Sud-Est (à droite)

Au regard de cette étude comparative, nous pouvons conclure que les événements de houle issus des données de hindcast correspondent dans l'ensemble aux événements observés même si souvent, les valeurs extrêmes des H_s sont sous-estimées. Toutefois, ces valeurs correspondent à moins de 10% des valeurs prises par le H_s et n'ont de ce fait pas d'influence significative sur la qualité globale de l'accord entre les événements de houles issus des deux sources de données. De plus, dans ces travaux notre objectif est de représenter au mieux toute la distribution des états de mer et non pas exclusivement les états de mers extrêmes. Néanmoins, l'utilisation de ces données de hindcast pour des applications sensibles aux états de mer extrêmes doit se faire avec précaution. Soulignons également que la cohérence temporelle des événements de houle est globalement reproduite dans les données de hindcast.

4.3.3 Comparaison des événements de mer du vent

On s'intéresse maintenant aux événements de mer du vent sur les sites Girassol et Akpo. La figure 4.11 présente pour les sites Girassol et Akpo, des séquences d'événements de mer du vent identifiés à partir des données *in situ* et des données de hindcast. Ces séquences correspondent aux mêmes périodes que dans le cas des houles. Contrairement aux événements de houle, l'accord paraît moins bon quant à la cohérence temporelle des événements. Nous pouvons observer des instants où des mers du vent sont identifiées à partir des données *in situ* alors qu'elles sont absentes des données de hindcast. Quand les mers du vent sont identifiées dans les deux cas, nous remarquons une différence entre les valeurs des paramètres, notamment la hauteur significative.

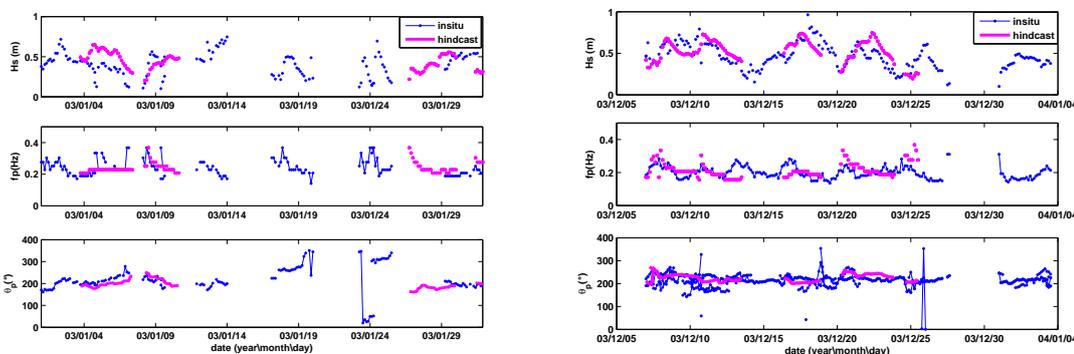


FIGURE 4.11 – H_s Comparaison des événements de mer du vent sur Girassol pour janvier 2003 (à gauche) et sur Akpo décembre 2003 (à droite)

Pour l'ensemble des événements de mers du vent identifiés, la figure 4.12 montre les diagrammes quantile/quantile des H_s pour le site Girassol (à gauche), et pour le site Akpo (à droite). Sur le site Girassol, les quantiles de H_s sont pour la plupart sur-estimés dans les données de hindcast. En ce qui concerne le site Akpo, même si on note une meilleure corrélation que par rapport à Girassol, elle reste cependant faible.

Il est plausible que la différence entre les systèmes de mer du vent issus des données *in situ* et des données de hindcast soit due à un mauvais accord entre le vent observé et le vent réanalysé, utilisé en forçage du modèle qui génère les données de hindcast. En effet, les caractéristiques des mers du vent (H_s , T_p , θ_p) sont fortement dépendantes des

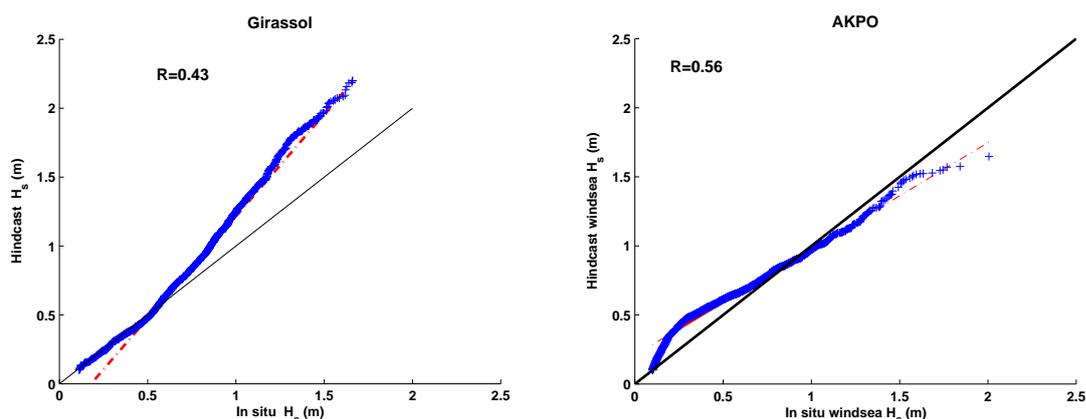


FIGURE 4.12 – Diagramme quantile/quantile des H_s des événements de mer du vent : sur le site Girassol (à gauche) et sur le site Akpo (à droite).

caractéristiques du vent local (vitesse et direction), du fetch et du temps d'action du vent. Il est donc prévisible que lorsque les données du vent utilisées dans le modèle (hindcast) diffèrent des données de vent observées (*in situ*), les mers du vent qu'elles génèrent diffèrent aussi par leurs caractéristiques. De plus, la hauteur significative et la période de pic de la mer du vent sont proportionnelles au carré de la vitesse du vent. En conséquence, une légère différence entre les vitesses du vent *in situ* et du vent de hindcast peut se traduire par une grande différence entre les H_s et T_p .

Nous avons comparé, les données du vent *in situ* et les données du vent utilisées dans le hindcast pour la période où les observations sont disponibles (figure 4.13). Le coefficient de corrélation linéaire entre les vitesses du vent (U_{10}) des deux sources de données est seulement de 0.42. Les différences entre les deux sources de données du vent s'expliquent par une mauvaise connaissance de la climatologie du vent en Afrique de l'Ouest. Cette zone dispose de très peu de données d'observations, ce qui limite considérablement la calibration des modèles numériques de vent. Cependant, le diagramme quantile/quantile (en bas, à droite de la figure 4.13) montre que l'intensité du vent est plutôt sous-estimée dans les données de vent réanalysées alors les H_s des mers du vent sont quant à eux sur-estimés. Nous pensons que cela peut être dû au paramétrage utilisé dans le modèle numérique, pour les termes relatifs aux mers du vent, dans le cas d'un vent de faible intensité.

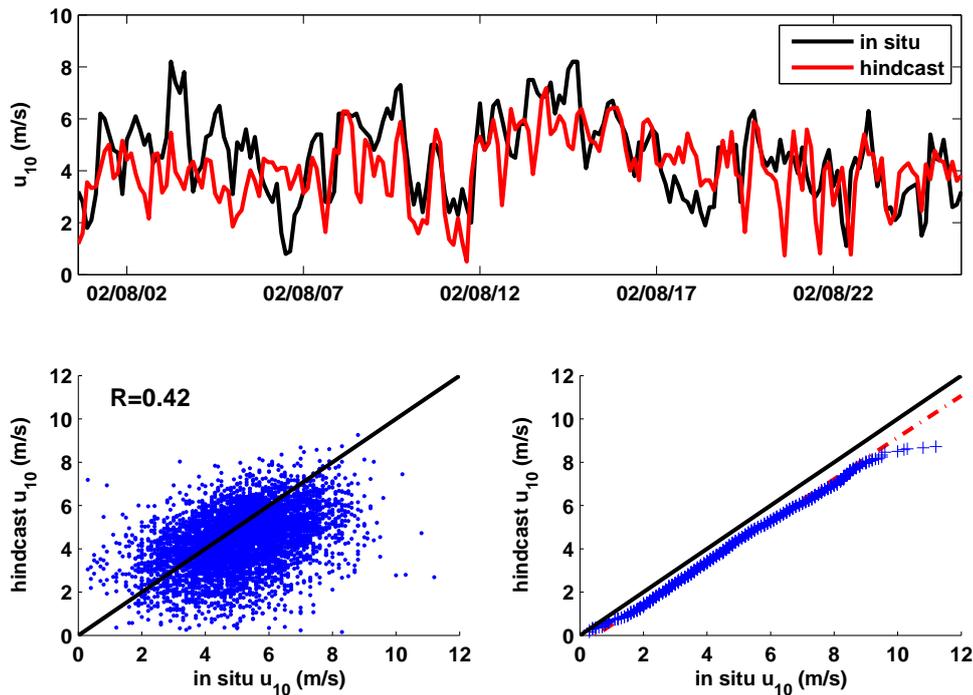


FIGURE 4.13 – Comparaison des vitesses du vent sur Girassol : série temporelle (haut) ; scatter plot (en bas, à gauche) et quantile/quantile plot (en bas, à droite).

4.4 Conclusion

Dans ce chapitre, nous avons proposé une classification des événements en fonction de leur zone de génération. Cette classification a permis, d'une part, d'évaluer pour les systèmes présents, leurs contributions à l'énergie ou à la puissance totale des états de mer, et d'autre part, d'effectuer une analyse comparative plus raisonnée des événements.

Ainsi, les résultats ont montré, aussi bien sur le site Girassol que sur le site Akpo, que les houles de Sud-Ouest constituent les systèmes dominants. Elles contribuent à plus de 80% sur Girassol et près de 70% de l'énergie totale des états de mer. La contribution des mers de vent reste faible (moins de 20% en général).

La comparaison des événements a montré que les événements de houle identifiés à partir des données de hindcast sont globalement similaires à ceux observés, malgré une sous-estimation des valeurs extrêmes. Quant aux mers du vent, nous avons noté une différence explicable par la différence constatée entre les données de vent *in situ* et celles utilisées pour le forçage du modèle.

Au regard de cette étude et pour les objectifs qui sont les nôtres, notamment celui de représenter le mieux possible tous les états de mers et non exclusivement les extrêmes, nous pouvons conclure que les données de hindcast peuvent servir de données de référence pour la modélisation de la climatologie des houles. L'utilisation de ces données permettra d'intégrer la variabilité saisonnière dans le générateur stochastique d'événements de houle qui sera présenté dans le chapitre 6.

Modélisation d'un événement de houle

La modélisation des événements individuels de houle, qui constitue la deuxième étape de la démarche adoptée, est nécessaire à la caractérisation d'une climatologie de vagues par événements. Elle a pour but de représenter chacun des événements de houle par un nombre réduit de paramètres à partir desquels les événements peuvent être entièrement reconstruits.

La modélisation des événements de houle consiste à représenter avec des formes paramétriques, l'évolution des paramètres caractéristiques d'états de mer (H_s , T_p , θ_m , etc.) dans un événement de houle, comme dans le cas des modèles triangulaires de tempête présentés dans 2.4.2.3. Les événements de houle observés en Afrique de l'Ouest proviennent de tempêtes répétitives d'une période moyenne d'occurrence de 3 à 4 jours et majoritairement localisées dans l'Atlantique Sud. Au sein de ces événements de houle, les paramètres caractéristiques évoluent quasiment suivant les mêmes configurations comme nous l'avons déjà souligné dans les chapitres 3 et 4, ce qui rend possible leur modélisation.

Ce chapitre présente le modèle d'événement de houle développé dans le cadre de ces travaux. Le modèle est constitué de trois formes paramétriques correspondant aux évolutions respectivement de la puissance, de la fréquence d'énergie et de la direction moyenne dans l'événement. Il permet de caractériser entièrement les événements de houle, chacun avec son étendue temporelle, par un ensemble de sept paramètres. Ces sept paramètres suffisent à synthétiser, par exemple, pour un événement de houle de durée 144 heures (48 fois trois heures stationnaires), l'information des 48 valeurs instantanées des puissances, 48 valeurs instantanées des fréquences et 48 valeurs instantanées des directions.

Sommaire

5.1	Choix des paramètres synthétiques utilisés pour la caractérisation d'un événement de houle	79
5.2	Modèle d'évolution de la puissance au sein d'un événement de houle	81
5.2.1	Les différentes fonctions paramétriques étudiées.	82
5.2.2	Analyse des résultats de la modélisation de l'évolution de la puissance avec les 3 fonctions paramétriques.	83
5.3	Modèle d'évolution de la fréquence moyenne d'énergie au sein d'un événement de houle	86
5.4	Modèle d'évolution de la direction moyenne au sein d'un événement de houle	87
5.5	Récapitulatif	89
5.6	Résultats et discussions sur la modélisation des événements de houle avec le SEM	91
5.6.1	Modélisation d'une série temporelle d'événements de houle.	91
5.6.2	Comparaison de la distribution jointe des couples (H_s, T_e) issus des événements de houle observés et ceux issus du SEM.	93
5.6.3	Validation pratique : évaluation du Potentiel Houlomoteur Extractible	95
5.7	Conclusion	99

5.1 Choix des paramètres synthétiques utilisés pour la caractérisation d'un événement de houle

Comme introduit dans le chapitre 3, un événement de houle désigne l'évolution au cours du temps d'un système de houles que l'on peut relier à un même phénomène météorologique source (dépression ou tempête lointaines). Il est caractérisé par l'évolution de ses paramètres caractéristiques (H_s , T_p , θ_p , etc.). La modélisation des événements de houle consiste donc à identifier des modèles d'évolution de ces paramètres caractéristiques. Avant d'aborder la modélisation des événements de houle, nous allons d'abord expliquer les choix des paramètres avec lesquels les événements de houles ont été caractérisés dans le cadre cette étude.

Les événements de houle peuvent être caractérisés par une variété de paramètres, les mêmes qui permettent de caractériser un état de mer, et présentés dans 2.2.2.2. Couramment, l'état de mer est caractérisé par un paramètre d'amplitude (H_s) ou d'intensité (P_w), une période ou fréquence caractéristique (f_p , f_{m-10} , f_{m02} , etc.), un paramètre de direction (θ_p , θ_m) et éventuellement des paramètres de largeur spectrale et d'étalement directionnel. Dans le cadre de cette thèse, nous avons choisi de caractériser un événement de houle par la puissance des vagues par mètre de largeur de crête (P_w), la fréquence moyenne d'énergie (f_{m-10} que nous désignerons par la suite par f_e) et la direction moyenne (θ_m).

Le choix s'est porté sur la puissance (équation 2.30) plutôt que sur la hauteur significative parce que d'une part, la puissance des vagues permet de représenter l'inter-dépendance entre la hauteur significative et la fréquence moyenne d'énergie en un seul paramètre. D'autre part, son évolution au sein d'un événement de houle présente une allure plus lisse que celle de la hauteur significative et se prête donc mieux à la modélisation. De plus, la puissance présente l'avantage d'être sommable alors que la somme des hauteurs significatives n'a aucune signification physique. Le choix de la puissance permet donc d'obtenir pour son modèle d'évolution, des paramètres qui ont un sens physique. Toutefois, notons que la hauteur significative se déduit aisément de la puissance connaissant la fréquence moyenne d'énergie.

Afin d'être cohérent dans le choix des paramètres caractéristiques des événements de houle, la fréquence moyenne d'énergie (f_e) a été choisie car c'est elle qui intervient dans le calcul de la puissance. Par ailleurs, il faut souligner que pour les systèmes de houle, les différents types de fréquences (f_p , f_e , f_{m02}) ont généralement des valeurs voisines. En effet, à cause des grandes distances parcourues par les houles, leurs spectres ont généralement une largeur de bande en fréquence étroite. Le spectre d'un système de houle est d'autant plus étroit que la distance parcourue par les houles est grande et ceci est lié à la relation de dispersion des vagues (Gjevik et al. [1988]). La figure 5.1 montre les corrélations qui existent entre la fréquence moyenne d'énergie, la fréquence de pic et la fréquence moyenne zero-crossing pour les sites Girassol et Akpo. Pour ces deux sites, les coefficients qui relient les différentes fréquences sont proches de 1, montrant ainsi que les fréquences moyennes sont proches de la fréquence de pic. Les relations linéaires empiriques qui existent entre les fréquences peuvent servir à calculer f_p ou f_{m02} lorsque f_e est connue.

La direction moyenne est la direction pondérée par l'énergie des vagues alors que la direction de pic correspond à la direction de la fréquence de pic. Lorsque l'on travaille sur des spectres discrétisés, la direction de pic ne peut progresser que par sauts discrets, alors que la direction moyenne a une variation continue. Cette raison motive le choix de

la direction moyenne pour la caractérisation des événements de houle.

Soulignons par ailleurs que la largeur spectrale et l'étalement directionnel des houles observées en Afrique de l'Ouest sont faibles en raison des grandes distances qui séparent leur zone de génération et le point d'observation. De plus, les variations de la largeur spectrale et de l'étalement directionnel sont fortement corrélées à la fréquence et la connaissance de cette dernière suffit souvent à les déterminer. En conséquence, dans cette étude, ces paramètres n'ont pas été utilisés dans la caractérisation des événements de houles. Cependant, ils peuvent être déduits en fonction de la fréquence à partir des relations établies lors du projet WASP¹ (West Africa Swell Project) Forristall et al. [2013].

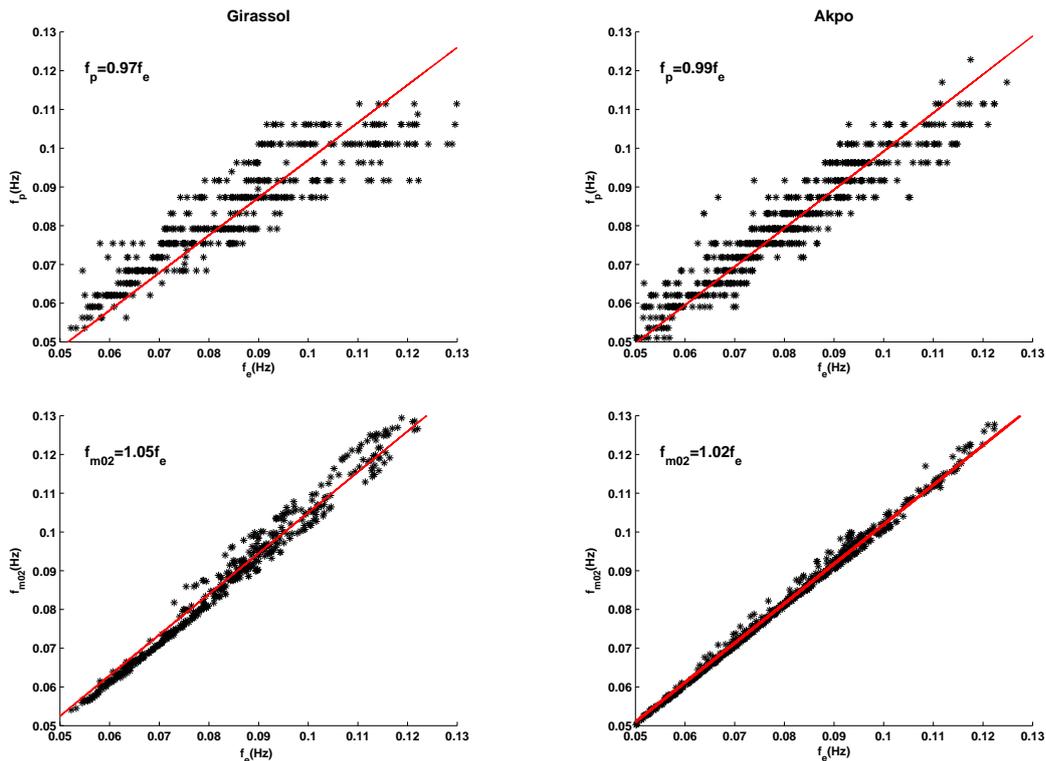


FIGURE 5.1 – Fréquence de pic (f_p) et fréquence zero-crossing (f_{m02}) en fonction de la fréquence d'énergie (f_e) : Girassol (gauche) et Akpo (droite)

La figure 5.2 (fenêtre à gauche) montre l'ensemble des événements de houle extraits de la base de données *in situ* du site Girassol en utilisant la méthode décrite dans 3.3. Les événements y sont représentés en faisant correspondre à l'instant zéro la date où leur puissance atteint son maximum. Il faut noter qu'au total, 201 événements ont été extraits de cette base de données, après avoir écarté ceux pour lesquels la durée est inférieure à 24 h et qui sont en réalité des fragments d'événements, conséquence de discontinuité dans les données. Afin d'harmoniser les événements de houle extraits et de les rendre comparables les uns aux autres, nous avons considéré qu'un événement de houle commence lorsque sa puissance passe au dessus d'un seuil de 0.5 kW.m^{-1} et se termine lorsque cette

1. Le projet WASP est un projet commun (Join Industry Project) qui a été mené par Shell, Ifremer, Chevron, Marathon et Total pour analyser les données disponibles sur les houles en Afrique de l'Ouest (du Nigeria à la Namibie).

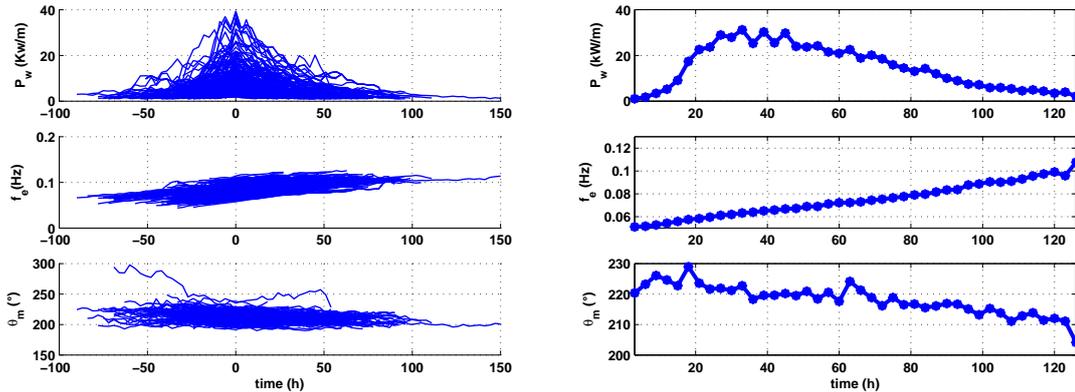


FIGURE 5.2 – Ensemble des 201 événements de houle extraits des données Girassol (gauche) ; événement de houle typique (droite)

dernière passe en deçà de ce seuil. Pour avoir une idée d'équivalence de ce seuil en hauteur significative, l'exemple suivant est donné : au début d'un événement dont les premiers trains de houles ont une fréquence de 0.05 Hz (longueur d'onde = 624 m, en eau profonde), ce seuil correspond à un H_s de 0.22 m et à la fin de l'événement, il correspond à un H_s de 0.3 m lorsque la fréquence des houles est 0.1 Hz (longueur d'onde = 156 m, en eau profonde).

Les évolutions typiques de la puissance, de la fréquence d'énergie et de la direction moyenne d'un événement de houle sont montrées dans la figure 5.2 (à droite). La puissance croît au début de l'événement, atteint un maximum puis décroît. La fréquence d'énergie affiche une évolution quasiment linéaire comme l'indique la théorie (équation 3.1). La direction moyenne quant à elle décroît quasi-linéairement, indiquant ainsi que la tempête source s'est déplacée de l'Ouest vers l'Est. Notre objectif est maintenant d'identifier des fonctions paramétriques pour modéliser ces évolutions typiques de la puissance, de la fréquence d'énergie et de la direction moyenne au sein d'un événement de houle.

5.2 Modèle d'évolution de la puissance au sein d'un événement de houle

Le but est d'identifier une fonction paramétrique qui permette de modéliser au mieux l'évolution de la puissance au sein d'un événement de houle. Le problème est comparable, par exemple à la recherche de formes paramétriques de spectres d'état de mer (JONSWAP, Pierson-Moskowitz, etc.). L'évolution typique de la puissance au sein d'un événement de houle est semblable à celle de la hauteur significative, comme soulignée par des études antérieures (Gjevik et al. [1988], Hanson, Phillips. [2001], Husson [2012]). Comme pour le H_s , au début de l'événement, la puissance croît avec l'arrivée des premiers trains de houle, atteint un maximum, puis décroît. Il faut également remarquer que la phase ascendante est généralement plus pentue que la phase descendante. Ce comportement ne peut pas être modélisé par une fonction symétrique. Sur la base des travaux existants (Boccotti [2000], Arena, Pavone [2009], Martín-Hidalgo et al. [2014], Olagnon et al. [2014]), une fonction triangulaire a été initialement proposée pour modéliser l'évolution de la puissance au sein

d'un événement. Il a été observé par la suite que la fonction triangulaire ne modélise pas toujours correctement l'ensemble des valeurs prises par la puissance au sein des événements et que d'autres fonctions, notamment la fonction de densité d'une loi Log-normale et la fonction de densité d'une loi Gamma donnent de meilleurs résultats (Kpogo-Nuwoklo et al. [2015b]). Nous présentons ci-après les différentes fonctions paramétriques étudiées avant de conclure sur celle qui modélise au mieux l'évolution de la puissance au sein d'un événement de houle.

5.2.1 Les différentes fonctions paramétriques étudiées

Pour modéliser l'évolution de la puissance au sein d'un événement de houle, 3 fonctions paramétriques ont été successivement étudiées : la fonction triangulaire, la fonction de densité d'une loi Log-normale et la fonction de densité d'une loi Gamma.

5.2.1.1 Fonction triangulaire

L'équation de la fonction triangulaire proposée s'écrit :

$$P_w(t) = \begin{cases} E\alpha_1(t - t_0) + P_{w_{max}} & \text{si } t_i \leq t < t_0 \\ E\alpha_2(t - t_0) + P_{w_{max}} & \text{si } t_0 \leq t \leq t_f \end{cases} \quad (5.1)$$

où α_1 et α_2 sont respectivement les pentes ascendante et descendante, t_i et t_f respectivement les dates du début et de fin de l'événement, t_0 la date à laquelle la puissance atteint son maximum, et $P_{w_{max}}$ la puissance maximale dans l'événement de houle (figure 5.3). Le paramètre E défini par

$$E = \int_{t_i}^{t_f} P_w(t) dt, \quad (5.2)$$

en $kWh.m^{-1}$ est l'énergie par mètre de largeur de crête de l'événement de houle pris dans sa totalité. Il intervient comme un facteur de normalisation, et permet d'obtenir pour le modèle la même énergie que dans l'événement observé.

5.2.1.2 Loi Log-normale

L'idée est de représenter l'évolution de la puissance, normalisée par l'énergie de l'événement par la fonction de densité d'une loi Log-normale (Jobe [1989], Weisstein [2010]) (voir figure 5.4) :

$$P_w(t) = \frac{E}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(t) - \mu)^2}{2\sigma^2}\right), t_i \leq t \leq t_f \quad (5.3)$$

où le paramètre E est le même que dans l'équation (5.2), et les paramètres σ et μ sont exprimés en fonction de la puissance maximale ($P_{w_{max}}$) et de l'écart type (sd) par :

$$\sigma = \sqrt{\ln\left(\frac{sd^2}{P_{w_{max}}^2} + 1\right)} \quad (5.4)$$

$$\mu = \ln(P_{w_{max}}) + \sigma^2 \quad (5.5)$$

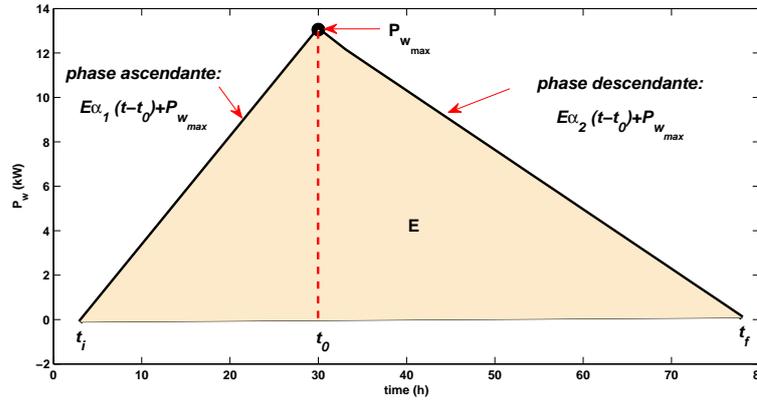


FIGURE 5.3 – Modèle triangulaire pour l'évolution de la puissance au sein d'un événement de houle. La partie colorée correspond à l'énergie de l'événement (E).

5.2.1.3 Loi Gamma

Comme dans le cas de la loi Log-normale, nous cherchons à paramétrer l'évolution de la puissance normalisée par l'énergie totale de l'événement avec la fonction de densité d'une loi Gamma (Johnson, Kotz [1969]) (figure 5.5). La loi Gamma offre de par la variation de ses paramètres, une grande variété de formes (Ferreira, Guedes Soares [2000]), laquelle permet de l'adapter à la modélisation de l'évolution de la puissance au sein d'un événement de houle. Le modèle s'écrit :

$$P_w(t) = \frac{E}{\tau^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\tau}\right), t_i \leq t \leq t_f \quad (5.6)$$

où $\Gamma(\alpha)$ est une fonction Gamma incomplète généralisée définie par :

$$\Gamma(\alpha) = \int_{t_i}^{t_f} t^{\alpha-1} \exp(-t) dt. \quad (5.7)$$

Le paramètre de forme α et le paramètre d'échelle τ sont tels que $\alpha > 1$ et $\tau > 0$. Le paramètre E est l'énergie totale de l'événement (équation 5.2). La date du maximum de puissance t_0 peut être obtenue en fonction des paramètres α et τ par

$$t_0 = (\alpha - 1)\tau \quad (5.8)$$

5.2.2 Analyse des résultats de la modélisation de l'évolution de la puissance avec les 3 fonctions paramétriques

Pour obtenir les paramètres, les fonctions sont ajustées à l'évolution de la puissance au sein des événements de houle observés (issus des données *in situ*) par la méthode des moindres carrés pondérés. Parce que le début et la fin des événements de houle correspondent aux états de mer de faible énergie et ne sont pas toujours bien estimés, nous avons

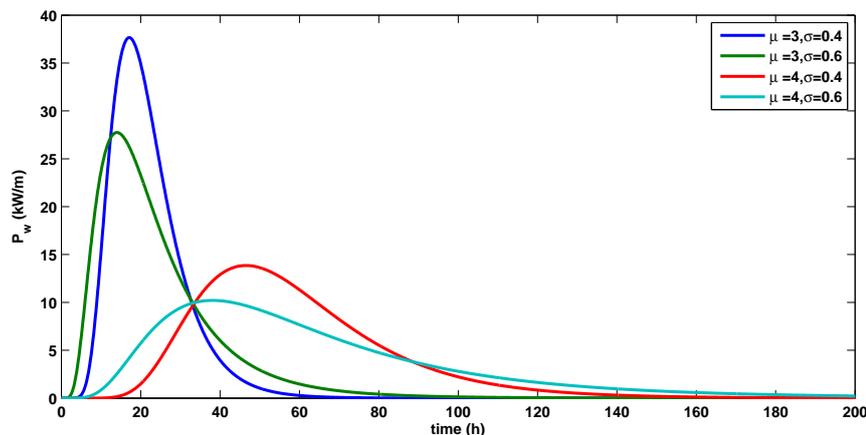


FIGURE 5.4 – Loi Log-normale pour $E = 700 \text{ kWh/m}$ et pour différentes valeurs de μ et σ .

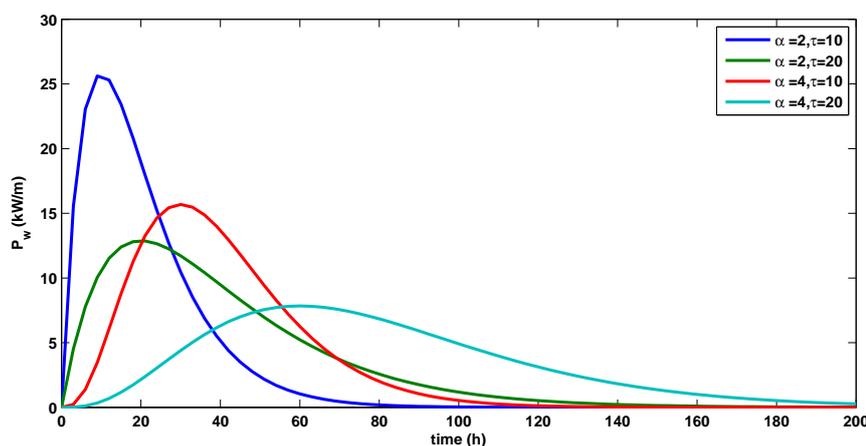


FIGURE 5.5 – Loi Gamma pour $E = 700 \text{ kWh/m}$ et pour différentes valeurs de α et τ .

choisi de donner plus de poids aux états de mer d'intensités fortes en utilisant comme poids les valeurs de la puissance observée.

La figure 5.6 montre, pour un événement de houle, l'ajustement des 3 fonctions paramétriques (partie haute de la figure). Par rapport à la fonction triangulaire, les lois Log-normale et Gamma modélisent mieux l'évolution de la puissance au sein de cet événement de houle. Ce résultat est également confirmé par le graphe du bas de la figure 5.6 où est représentée l'erreur relative en valeur absolue définie par :

$$MAE = \left| \frac{P_w^{obs} - P_w^{mod}}{P_w^{obs}} \right| \quad (5.9)$$

où P_w^{obs} la puissance observée et P_w^{mod} la puissance obtenue avec les fonctions paramétriques. Les lois Log-normale et Gamma présentent en effet les MAE les plus faibles,

moins de 15% pour les puissances supérieures à 5 kW.m^{-1} . Il faut remarquer que pour les puissances très faibles ($P_w < 5 \text{ kW.m}^{-1}$) les erreurs sont élevées pour les 3 fonctions. La fonction triangulaire présente des erreurs faibles près du maximum de puissance, mais qui augmentent dès qu'on s'en éloigne. Le modèle triangulaire ne permet pas de modéliser correctement l'ensemble des valeurs de la puissance, alors que certaines applications, par exemple le dimensionnement en fatigue, peuvent être sensible à des erreurs sur les houles de faible énergie notamment. On pourrait en effet imaginer qu'une réduction de la base du triangle permettrait de mieux modéliser les faibles valeurs de la puissance, mais ce serait aux dépens des fortes valeurs de puissance. Ce problème ne se pose pas avec les lois Log-normale et Gamma qui représentent globalement bien l'ensemble des valeurs de la puissance.

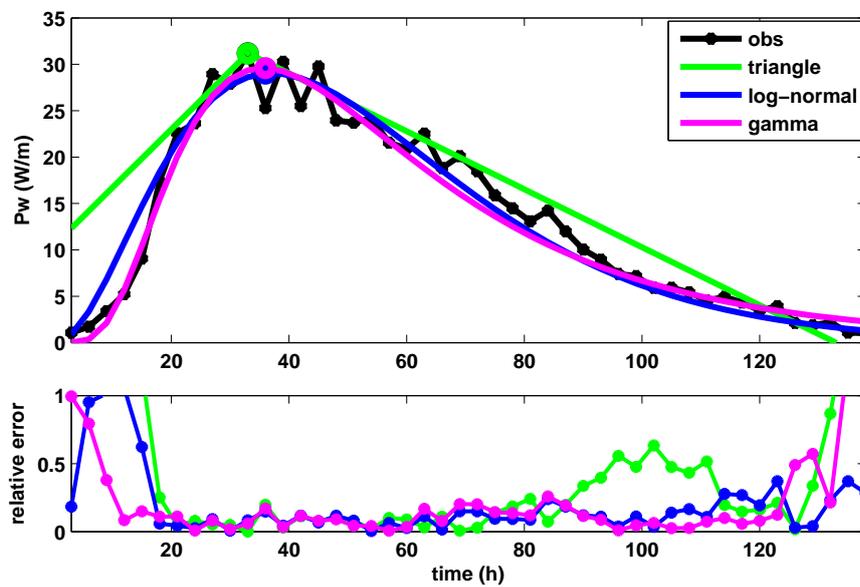


FIGURE 5.6 – Les différentes fonctions paramétriques ajustées à un événement de houle (haut); erreur relative associée (bas)

La qualité des ajustements par les 3 modèles a été évaluée sur l'ensemble des événements extraits des données *in situ* Girassol. Nous avons calculé pour chaque événement, l'erreur quadratique moyenne normalisée définie par :

$$NRMSE = \frac{\left\langle \left(P_w^{obs} - P_w^{mod} \right)^2 \right\rangle^{1/2}}{\langle P_w^{obs} \rangle} = \frac{RMSE}{\langle P_w^{obs} \rangle} \quad (5.10)$$

L'erreur quadratique moyenne a été normalisée pour rendre les événements comparables entre eux. La figure 5.7 montre les boîtes à moustache² de la distribution des NRMSE,

2. Pour chaque boîte à moustache, le trait qui se trouve à l'intérieur de la boîte correspond à la médiane, les limites inférieure et supérieure de la boîte correspondent respectivement aux quantiles 25% et 75%, et les barres inférieure et supérieure qui sont reliées par des traits en pointillés à la boîte correspondent respectivement aux quantiles 10% et 90%.

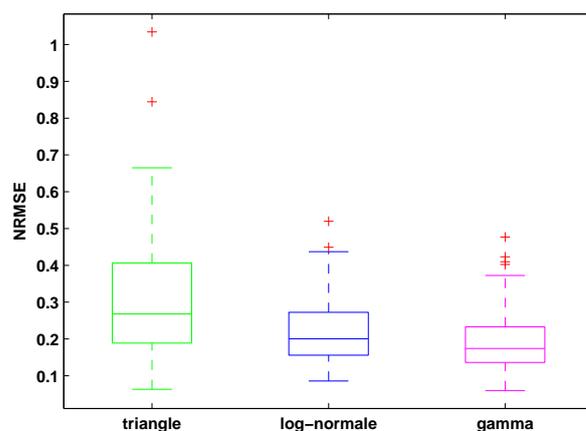


FIGURE 5.7 – Boîtes à moustache de la distribution des NRMSE.

pour chaque paramétrique, de l'ensemble des 201 événements de houle extraits. Le triangle présente pour l'ensemble des événements les NRMSE les plus élevées. La loi Log-normale et la loi Gamma affichent des erreurs plus faibles avec un NRMSE médian inférieur à 20%. Entre ces deux lois paramétriques, la loi Gamma affiche les plus faibles erreurs. De plus, la loi Gamma de par sa flexibilité, permet de modéliser une grande variété de formes en comparaison aux deux autres fonctions.

Dans une étude approfondie réalisée dans le cadre de ces travaux (voir Annexe C), un critère autre que le NRMSE a été utilisé pour évaluer les trois fonctions paramétriques étudiées : la distribution des crêtes de vagues individuelles. En effet, dans des applications comme la résistance des structures offshore, il est essentiel de prendre en compte la distribution des vagues individuelles extrêmes, d'où l'intérêt de valider les fonctions paramétriques étudiées par rapport à leur aptitude à reproduire la distribution des vagues extrêmes individuelles. Pour cela, nous avons comparé la distribution des crêtes de vagues non-linéaires obtenue à partir de l'événement de houle observé, et celle obtenue à partir de l'événement de houle modélisé par les 3 fonctions paramétriques. Les résultats ont montré que la loi Gamma modélise bien tant le déroulement de l'événement (problèmes de fatigue, d'érosion, de ressources énergétiques, d'opérations à la mer) que les extrêmes de vagues individuelles qu'on en déduit (dimensionnement aux extrêmes). L'étude a fait l'objet d'un article de conférence (European Wave and Tidal Energy Conference, EWTEC 2015, Annexe C).

5.3 Modèle d'évolution de la fréquence moyenne d'énergie au sein d'un événement de houle

Les houles se propagent à la vitesse de groupe qui est inversement proportionnelle à la fréquence, $Cg = g/(4\pi f)$ en eau profonde. À la limite où le point d'observation est très loin de la tempête source, celle-ci peut alors être considérée comme ponctuelle dans le temps et dans l'espace et un observateur placé en ce point voit d'abord arriver les houles de faible

fréquence suivies des houles de fréquence plus élevée. Lorsque les distances parcourues sont suffisamment grandes, la fréquence de pic évolue linéairement avec le temps (Sverdrup, Munk [1947], Snodgrass et al. [1966], Gjevik et al. [1988], Hanson, Phillips. [2001], Delpy et al. [2010]). Cette considération physique motive le choix d'un modèle linéaire pour l'évolution de la fréquence moyenne dans un événement de houle. Le modèle s'écrit :

$$f_e(t) = \frac{g}{4\pi d}t + \frac{1}{T_{max}} \quad (5.11)$$

Le modèle fait intervenir 2 paramètres :

- d : distance indicative du point d'observation à la zone de génération des houles ;
- T_{max} : période maximale de houle dans l'événement.

Remarquons que la distance réellement parcourue par les houles peut être théoriquement obtenue en remplaçant f_e dans l'équation 5.11 par la fréquence de pic. Le résultat obtenu permet, à titre indicatif, de remonter à la zone de génération de la tempête, mais n'est généralement pas très précis parce que (Kerbirou [2007]) :

- Une petite différence entre fréquences entraîne des écarts importants sur les distances estimées.
- La zone de génération n'est pas ponctuelle, ce qui implique que, d'une extrémité à l'autre, les vagues ont une plus ou moins grande distance à parcourir jusqu'au point d'observation.
- Les composantes basse fréquence sont plus rapides mais sont générées après les composantes haute fréquence, d'où l'influence de la durée de l'événement.
- La tempête génératrice des vagues se déplace. Les vagues produites peuvent être propagées dans toutes les directions (même si la plus grande partie de l'énergie est propagée dans la direction de génération de la tempête). Le déplacement de la tempête modifie la distance d à parcourir.

5.4 Modèle d'évolution de la direction moyenne au sein d'un événement de houle

Les houles observées en Afrique de l'Ouest sont majoritairement générées dans l'Atlantique Sud, entre le large des côtes argentines ou le cap Horn, et le cap de Bonne Espérance (Prevosto et al. [2013], Nerzic et al. [2007], Olagnon et al. [2014]). Ces houles sont générées par des dépressions ou des tempêtes répétitives qui, suivant la circulation atmosphérique générale, se déplacent du cap Horn vers le cap de Bonne Espérance. Deux constats ont motivé le choix du modèle d'évolution de la direction moyenne.

Nous avons d'abord remarqué que les vitesses angulaires de déplacement des dépressions génératrices des événements de houles sont relativement constantes. En effet, la figure 5.8 montre, en guise d'exemple, le suivi dans le temps d'une dépression dans l'Atlantique Sud. Les données de vent ECMWF (ERA-Interim) ont été utilisées pour identifier et suivre la dépression. Dans cette dépression, les vitesses de vent sont comprises entre 10 m/s et 17 m/s . Sur les fenêtres en dessous des cartes, est représentée l'évolution dans le temps de l'angle sous lequel un observateur situé sur le site Girassol pourrait voir arriver les trains de houle générés par cette dépression, sous l'hypothèse que ces dernières se pro-

pagent suivant les grands cercles³. Pour cet exemple, l'évolution de la direction moyenne de provenance des houles générées par la tempête est quasi linéaire, excepté pour les deux derniers pas de temps. La variation plus importante de la vitesse angulaire pour ces deux derniers pas de temps est due à l'étalement de la dépression. Somme toute, la vitesse angulaire moyenne de déplacement de cette dépression relative au point d'observation est quasiment constante en première approximation.

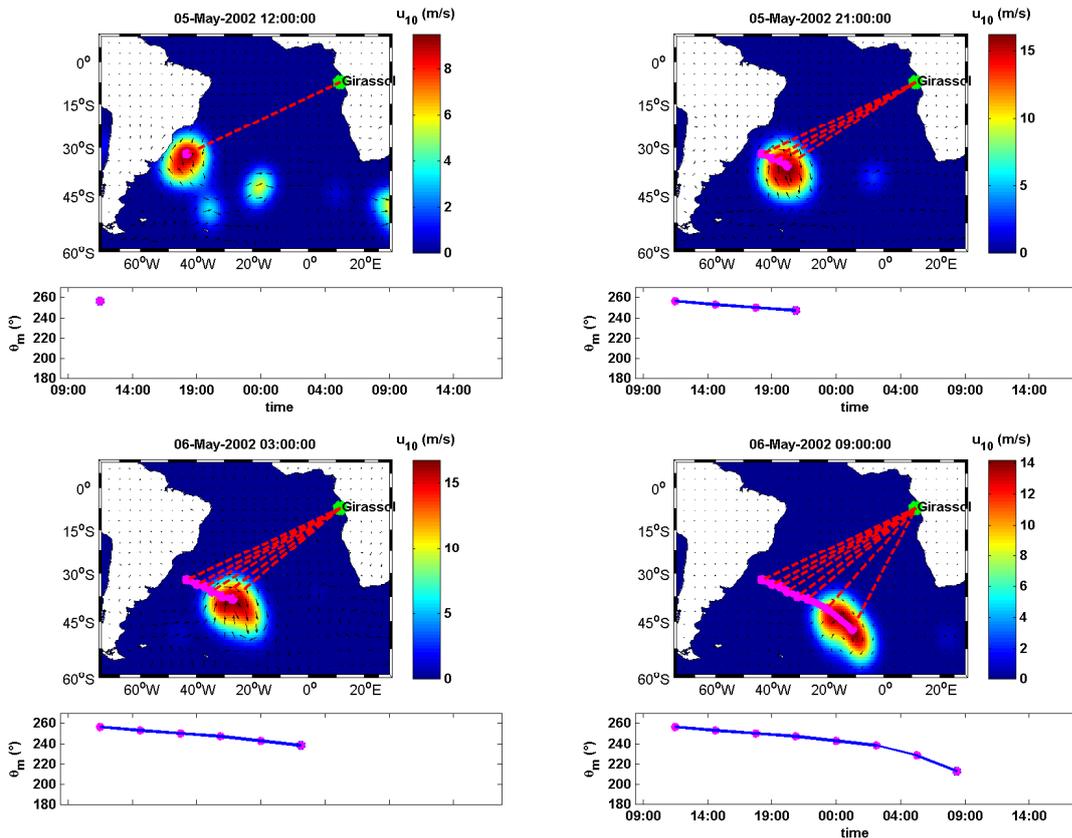


FIGURE 5.8 – Suivi dans le temps d'une dépression : sur la fenêtre en dessous de chaque carte est représentée l'évolution dans le temps de l'angle sous lequel, un observateur situé sur le site Girassol voit arriver les trains de houle.

Nous avons ensuite remarqué que pour l'ensemble des événements de houle extraits, les variations d'angle entre le début et la fin des événements sont faibles. Sur la figure 5.9 est représenté l'histogramme des écarts angulaires définis comme la différence d'angle en valeur absolue entre la direction des houles en début d'événement et la direction des houles en fin d'événement :

$$\Delta\theta = |\theta_m(t_f) - \theta_m(t_i)|. \quad (5.12)$$

3. Les grands cercles sont les géodésiques d'une sphère, c'est-à-dire les chemins possédant la plus petite courbure et les arcs de grands cercles sont par conséquent les plus courts chemins reliant deux points à la surface d'une sphère. Sur Terre, les méridiens et l'équateur sont des grands cercles ; les autres parallèles n'en sont pas puisque leur taille varie en fonction de la latitude. Les houles qui se propagent en suivant une droite sur un océan plan, se propagent en réalité le long des géodésiques dans le cas général

Les écarts d'angle sont généralement inférieurs à 25° pour les événements de houle. Ils peuvent être considérés comme faibles si on sait que les modèles numériques de vagues considèrent généralement une discrétisation par pas de 15° en direction.

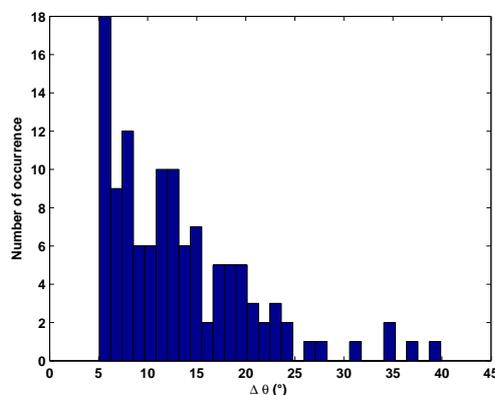


FIGURE 5.9 – Histogramme des variations d'angle entre le début et la fin des événements de houle

Au regard de ces deux remarques, il nous a paru raisonnable de modéliser l'évolution de la direction moyenne au sein d'un événement de houle par un modèle linéaire :

$$\theta_m(t) = \omega t + \theta_0, \quad (5.13)$$

où la pente ω , supposée constante, est assimilable à la vitesse angulaire moyenne de déplacement de la tempête relativement au point d'observation, et θ_0 l'angle sous lequel on voit arriver au point d'observation le premier train de houle. Les paramètres du modèle sont obtenus par régression linéaire.

5.5 Récapitulatif

En résumé, la loi Gamma pour l'évolution de la puissance, les modèles linéaires pour la fréquence d'énergie et la direction moyenne, forment le modèle d'événement de houle que nous désignerons par la suite par un acronyme en anglais "SEM" pour "Swell Event Model". Le SEM se compose donc de 7 paramètres au total résumés dans le tableau 5.1. La figure 5.10 montre un exemple d'événement de houle et sa représentation modélisée. Cette modélisation est nécessaire à la caractérisation d'une climatologie par événements. En plus de permettre la conservation de la cohérence temporelle des systèmes de houle, un autre avantage de la modélisation des événements est la réduction du nombre de paramètres nécessaires pour représenter l'événement de houle. Par exemple, l'événement de houle de la figure 5.10 est représenté par 42 valeurs instantanées de P_w , 42 valeurs instantanées de f_e et 42 valeurs instantanées de θ_m soit au total 126 valeurs d'observations. Avec le SEM, seuls 7 paramètres sont nécessaires pour le reconstruire entièrement.

Paramètres synthétiques	Modèle	Paramètres du SEM	Signification
puissance	$P_w(t) = \frac{E}{\tau^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(\frac{-t}{\tau}\right)$	E	énergie de l'événement par mètre de largeur de crête de vagues
		α	paramètre de forme
		τ	paramètre d'échelle
fréquence d'énergie	$f_e(t) = \frac{g}{4\pi d} t + \frac{1}{T_{max}}$	d	distance indicative parcourue par les houles
		T_{max}	période maximale de houle dans l'événement
direction moyenne	$\theta_m(t) = \omega t + \theta_0$	ω	penne, i.e. vitesse angulaire moyenne, relative au point d'observation, de la tempête
		θ_0	direction de provenance du premier train de houle

TABLE 5.1 – Tableau récapitulatif du SEM

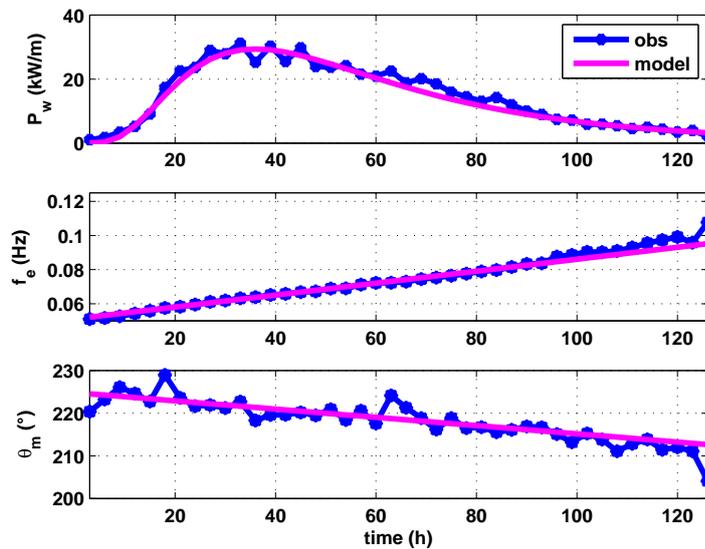


FIGURE 5.10 – Un événement et sa représentation modélisée

5.6 Résultats et discussions sur la modélisation des événements de houle avec le SEM

Cette section, subdivisée en trois parties, présente les résultats de la modélisation des événements de houle avec le SEM sur le site Girassol et sur le site Akpo. La première partie présente les résultats de la modélisation d'une série temporelle d'événements de houle par le SEM. Dans la deuxième partie, on compare la distribution jointe de H_s et T_e issus des événements de houle modélisés avec celle issue des événements de houle observés. La troisième partie présente les résultats obtenus avec le SEM pour un cas test d'application : l'évaluation du potentiel houlomoteur extractible d'un convertisseur d'énergie des vagues.

5.6.1 Modélisation d'une série temporelle d'événements de houle

Nous présentons ici les résultats obtenus quant à la modélisation d'une série temporelle d'événements de houle avec le SEM.

Les figures 5.11 et 5.12 présentent l'ajustement du SEM aux événements de houle identifiés à partir des données *in situ*, respectivement, pendant le mois de décembre 2012 pour le site Girassol, et pendant le mois de décembre 2003 pour le site Akpo. Sur les deux figures, chaque événement de houle est indiqué par un numéro qui lui est propre. L'observation de ces deux figures permet de dire que le modèle se comporte globalement bien quand il est ajusté aux événements observés tant sur le site Girassol que sur le site Akpo.

Sur la puissance, nous avons quantifié l'erreur en calculant pour chaque événement l'erreur quadratique moyenne normalisée (NRMSE, équation 5.10). Le NRMSE a été calculé pour un seuil de puissance de $0.5 kW/m$, puis pour un seuil de $3 kW/m$. Le tableau 5.2 contient les NRMSE calculés pour les événements de la figure 5.11 (Girassol), et le tableau 5.3 pour événements de la figure 5.12 (Akpo). Pour un seuil de puissance de $0.5 kW/m$, les NRMSE sont généralement inférieurs à 20%, excepté pour les événements qui commencent et qui terminent les historiques. Par exemple le premier événement de la figure 5.11 est en réalité un événement incomplet, ce qui explique la mauvaise modélisation de sa puissance par la loi Gamma. Quant à l'événement indiqué par le numéro 9 sur la figure 5.11, il correspond à un événement de houle de Nord-Ouest (direction autour de 300°) et sa puissance est très faible; cela explique la valeur élevée du NRMSE. Cette explication est soutenue par les valeurs des NRMSE calculés avec des puissances supérieures à $3 kW/m$ (dernière ligne des tableaux). En effet, les NRMSE obtenus avec un seuil de $3 kW/m$ sont inférieurs à ceux obtenus avec un seuil de $0.5 kW/m$ signifiant ainsi que les erreurs relatives sont plus importantes lorsque les puissances deviennent faibles. Le choix des poids, lors de l'ajustement du modèle est une explication plausible à ce résultat. En prenant comme poids les valeurs de la puissance observée, on choisit de bien représenter les états de mer d'intensité forte et il est donc attendu que ces états de mer soient mieux modélisés, comparés aux états de mer de faible intensité. Il nous a paru raisonnable de faire ce choix d'une part parce que les états de mer de faible énergie ne sont pas toujours bien estimés, et d'autre part parce que pour la plupart des applications, l'érosion côtière par exemple, les états de mer d'intensité faible sont moins importants.

Le modèle linéaire se comporte également bien dans la modélisation des fréquences moyennes et des directions moyennes. Remarquons au passage que les modèles linéaires ajustés aux directions moyennes sont majoritairement décroissants, indiquant ainsi que les

tempêtes génératrices des houles se déplacent de l'Ouest (Cap Horn) vers l'Est (Cap de Bonne Espérance).

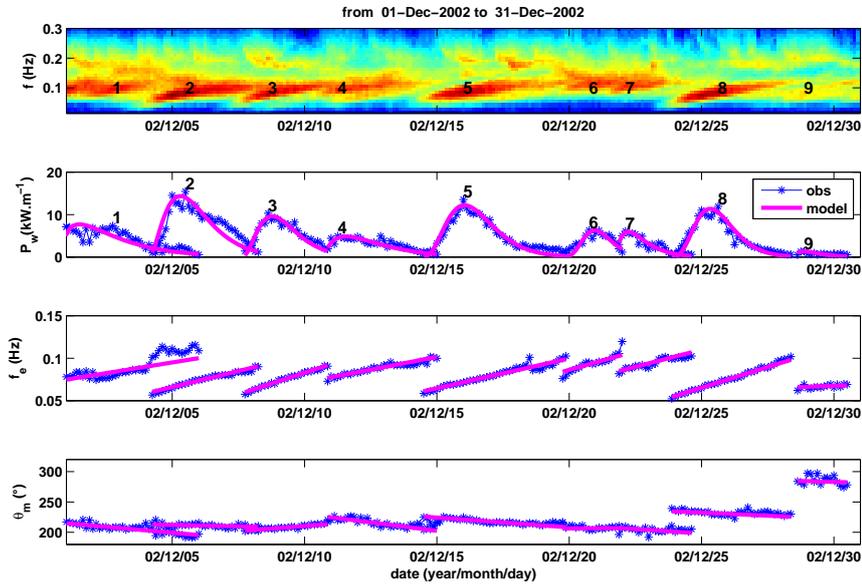


FIGURE 5.11 – Historique d'événements de houle et leurs modélisations (données *in situ*, Girassol) ; les événements de houle portent chacun un numéro différent.

N° d'événement	1	2	3	4	5	6	7	8	9
NRMSE (%) - $0.5kW/m$	27	17	8	15	13	18	19	15	37
NRMSE (%) - $3kW/m$	21	15	5	9	4	8	7	5	-

TABLE 5.2 – NRMSE associés aux événements de houle de la figure 5.11 (Girassol)

N° d'événement	1	2	3	4	5	6	7	8
NRMSE (%) - $0.5kW/m$	24	22	17	13	16	19	18	28
NRMSE (%) - $3kW/m$	10	15	10	7	8	12	7	-

TABLE 5.3 – NRMSE associés aux événements de houle de la figure 5.12 (Akpo)

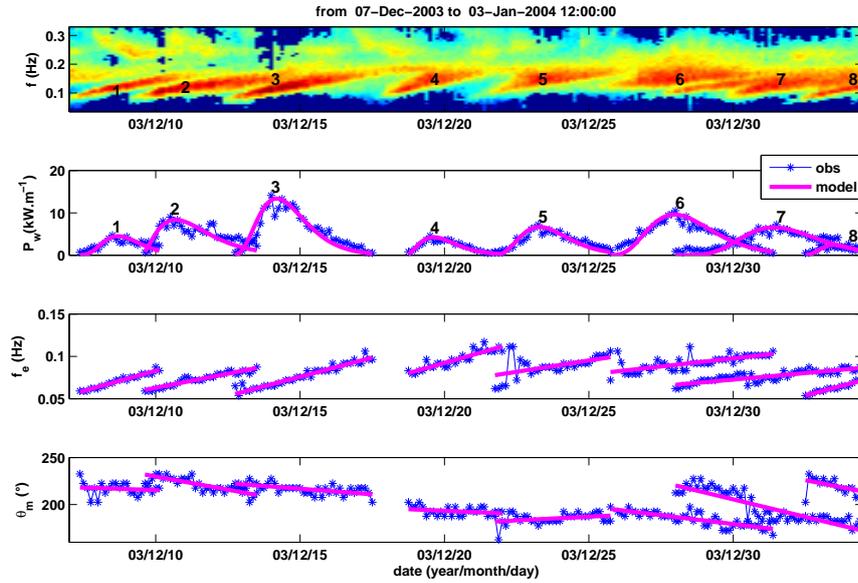


FIGURE 5.12 – Historique d'événements de houle et leurs modélisations (données *in situ*, Akpo) ; les événements de houle portent chacun un numéro différent.

5.6.2 Comparaison de la distribution jointe des couples (H_s, T_e) issus des événements de houle observés et ceux issus du SEM.

Pour les calculs en fatigue des structures offshore et pour d'autres types d'applications, on s'intéresse prioritairement aux paramètres H_s et T_e et à leur distribution jointe. Il est donc intéressant d'analyser la distribution jointe des paramètres d'états de mer à partir des événements de houle modélisés. On se propose donc de comparer la distribution jointe du H_s et T_e issue des observations à celle obtenue en remplaçant les événements de houle observés par leurs SEM. Le but est de répondre à la question suivante : remplacer les événements observés par leurs SEM introduit-il une différence significative dans la distribution jointe du H_s et T_e ? Une réponse négative à cette question permettrait de conclure que la distribution jointe du H_s et T_e dérivant du SEM peut être utilisée avec confiance dans les applications en génie océanique.

Dans le SEM, les hauteurs significatives se déduisent du modèle de la puissance, sous l'hypothèse de profondeur infinie par :

$$H_s^{mod} = \sqrt{\frac{64\pi P_w^{mod}}{\rho g^2 T_e^{mod}}} \quad (5.14)$$

où P_w^{mod} , et $T_e^{mod} = 1/f_e^{mod}$ sont respectivement la puissance et la période moyenne issues du modèle d'événements de houle.

La figure 5.13 compare la distribution jointe du H_s et T_e issus des événements observés et la distribution jointe du H_s^{mod} et T_e^{mod} . Les niveaux de contours représentés sont les mêmes pour les deux distributions jointes et sont régulièrement espacés de 0.00475

entre 0.001 et 0.02. Nous observons globalement un bon accord entre les deux distributions jointes. Les lignes de contours sont très peu différentes aussi bien pour les faibles probabilités que pour les probabilités élevées.

De même, les distributions marginales de la hauteur significative et de la période moyenne observées sont proches de celles obtenues avec le modèle. Les modes des distributions marginales de H_s et T_e coïncident quasiment. Afin de valider ces observations, nous avons utilisé un test statistique, le test de Kolmogorov-Smirnov qui permet de déterminer si les H_s et T_e issus des observations et du modèle suivent la même loi marginale. La statistique de ce test est :

$$KS_{stat} = \max_x |F_{obs}(x) - F_{mod}(x)|. \quad (5.15)$$

où F_{obs} est la fonction de répartition de H_s ou de T_e issus des observations et F_{mod} est la fonction de répartition de H_s^{mod} ou de T_e^{mod} issus du modèle. Le tableau 5.4 présente les résultats de la statistique du test de Kolmogorov-Smirnov ainsi que les $p-value$ ⁴ associées. Même si les $p-value$ sont faibles, surtout pour le T_e , elles restent supérieures à 0.05 signifiant ainsi qu'avec un risque de 5% de se tromper, on ne peut pas rejeter l'hypothèse selon laquelle les H_s et T_e issus des observations et du modèle dérivent des mêmes lois marginales.

À partir de ce résultat, nous pouvons conclure que remplacer l'historique des événements de houle issus des données d'observations par l'historique de leurs modèles d'événements, n'introduit pas de différences significatives dans la distribution marginale des hauteurs significatives et des périodes d'énergie ainsi que dans leur distribution jointe.

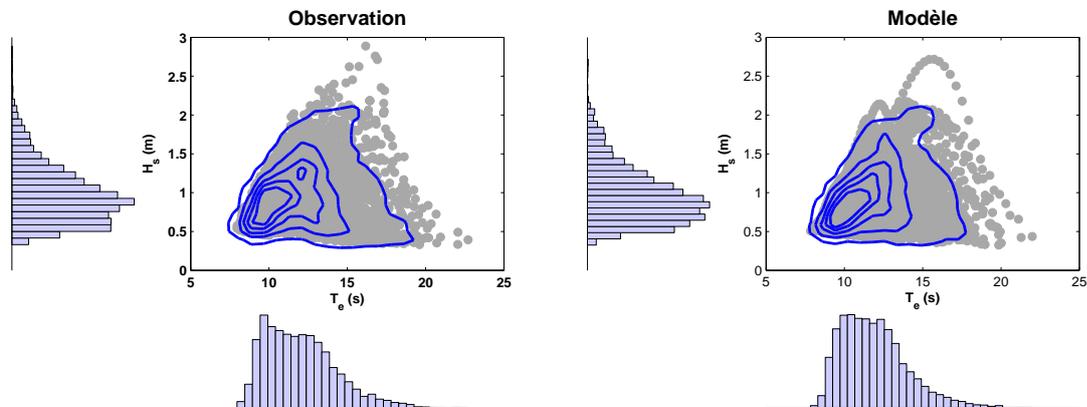


FIGURE 5.13 – Distributions jointes des hauteurs significatives et des périodes moyennes (Données *in situ*, Girassol). Les contours sont régulièrement espacés de 0.00475 entre 0.001 et 0.02.

4. La $p-value$ représente la probabilité qu'on aurait, si l'hypothèse nulle est vraie, d'avoir tort et de la rejeter néanmoins. Ici notre hypothèse nulle est que deux vecteurs dérivent d'une même loi marginale. On peut rejeter l'hypothèse nulle si la $p-value$ est inférieure à 0,05.

	H_s	T_e
KS_{stat}	0.0174	0.0240
p -value	0.415	0.089

TABLE 5.4 – Résultats du test de Kolmogorov-Smirnov appliqués aux distributions de H_s et T_e

5.6.3 Validation pratique : évaluation du Potentiel Houlomoteur Extractible

On propose dans cette partie de valider le SEM sur un cas d'application : l'évaluation du Potentiel Houlomoteur Extractible par un convertisseur d'énergie des vagues. Plus précisément, on cherche à quantifier l'erreur qu'on commet dans l'évaluation du Potentiel Houlomoteur Extractible en remplaçant les événements de houle observés par leurs modèles. Pour cela, nous allons utiliser un exemple de convertisseur de type uni-directionnel pour lequel la récupération de l'énergie se fait suivant une direction privilégiée, et dont la réponse est spectro-directionnelle (Li, Yu [2012], O. Falcão de [2010]). Par rapport à un convertisseur omnidirectionnel, le convertisseur uni-directionnel permet d'évaluer également le modèle linéaire de l'évolution de la direction moyenne.

5.6.3.1 Convertisseur uni-directionnel à orientation adaptable

On considère un convertisseur houlomoteur uni-directionnel, c'est à dire un système dont la fonction de transfert d'extraction dépend à la fois de la fréquence et de la direction relative par rapport au champ de vagues incident. On considère également que ce convertisseur est orienté à chaque pas de temps en sorte que sa direction principale soit alignée suivant la direction moyenne à la fréquence de pic instantanée des trains de vagues. Il existe un grand nombre de concepts, reposant sur des technologies variées pour de tels types de convertisseurs (O. Falcão de [2010]).

La fonction de transfert R_{pd} générique du type de convertisseur uni-directionnel utilisé a été construite sur la base de la fonction de transfert d'un convertisseur omnidirectionnelle (type absorbeur ponctuel) R_p , modulée par une fonction d'étalement directionnel. La fonction d'étalement directionnel utilisée est de type \cos^{2s} , normalisée en sorte que :

$$\int R_p df = \int \int R_{pd} df d\theta \quad (5.16)$$

De fait, la fonction de transfert directionnelle est donnée par :

$$R_{pd} = R_p \frac{2^{2s-1} \Gamma(s+1)^2}{\pi \Gamma(2s+1)} \cos^{2s}(\theta - \theta_m) \quad (5.17)$$

La figure 5.14 montre la fonction de transfert $R_p(f)$ utilisée (à gauche) et sa représentation modulée en fonction de l'étalement directionnel $R_{pd}(f, \theta)$ (à droite). La fonction de transfert $R_{pd}(f, \theta)$ met en évidence une réponse bimodale dont les pics sont suivant deux directions opposées. En fréquence, les pics sont localisés autour $[0.085Hz - 0.1Hz]$ avec une largeur spectrale s'étendant sur la bande $[0.05Hz - 0.33Hz]$.

Le potentiel extractible est évalué par la relation intégrale :

$$P = 2 \int \int R_p(f, \theta_m) S(f, \theta) df d\theta \quad (5.18)$$

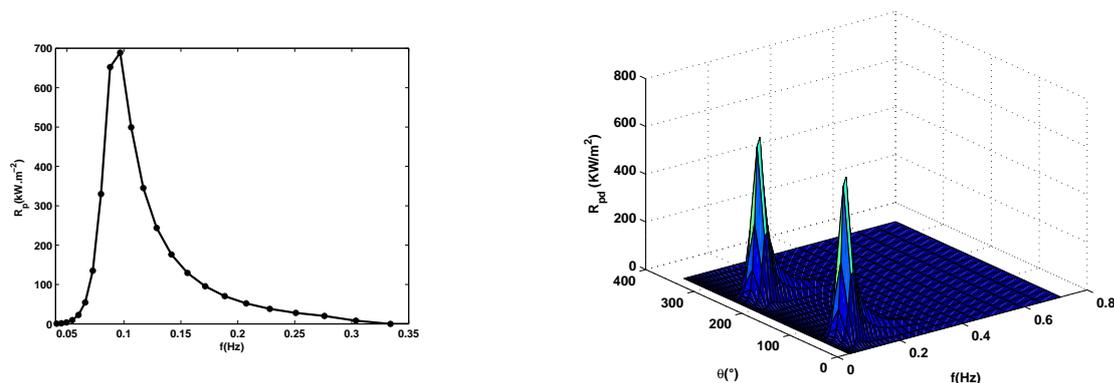


FIGURE 5.14 – Fonction de transferts de puissance omnidirectionnel (à gauche) et unidirectionnel (à droite)

Notre but est d'évaluer la puissance produite par le convertisseur en utilisant d'abord les événements de houle observés, puis les événements de houle modélisés et comparer les résultats. Mais préalablement, comme on peut remarquer dans l'équation 5.18, l'estimation de la puissance produite par le convertisseur nécessite la reconstitution du spectre de houle $S(f, \theta)$:

$$S(f, \theta) = S(f) \cdot D(f, \theta). \quad (5.19)$$

La fonction \cos^{2s} a été choisi pour $D(f, \theta)$ avec s dépendant de la fréquence et pour $S(f)$, nous avons utilisé la forme spectrale Log-normale plutôt que la forme Jonswap communément utilisée. En effet, les houles observées en Afrique de l'Ouest ont généralement une largeur de bande très étroite et les résultats du projet WASP (West Africa Swell Project, Forristall et al. [2013], Olagnon et al. [2013]) ont montré qu'une forme spectrale Log-normale permet une meilleure modélisation des spectres de houle que la forme spectrale Jonswap. Pour illustrer ce fait, nous montrons dans la figure 5.15 l'ajustement d'un spectre de houle issu des données du site Girassol avec la forme spectrale Jonswap et le Log-normal. Nous pouvons observer que la forme spectrale Jonswap ne modélise pas bien ce spectre à cause d'une faible pente dans sa phase croissante et surtout dans sa phase décroissance. Le spectre Log-normal est défini par :

$$S(f) = \frac{m_0}{f \sigma_0 \sqrt{2\pi}} \exp\left(-\frac{(\ln(f) - \mu_0)^2}{2\sigma_0^2}\right) \quad (5.20)$$

où m_0 est le moment d'ordre 0 et les paramètres σ_0 et μ_0 sont exprimés en fonction de la fréquence de pic (f_p) et de l'écart type de la distribution normale (sd) par :

$$\sigma_0 = \sqrt{\ln\left(\frac{sd^2}{f_p^2} + 1\right)} \quad (5.21)$$

$$\mu_0 = \ln(f_p) + \sigma_0^2 \quad (5.22)$$

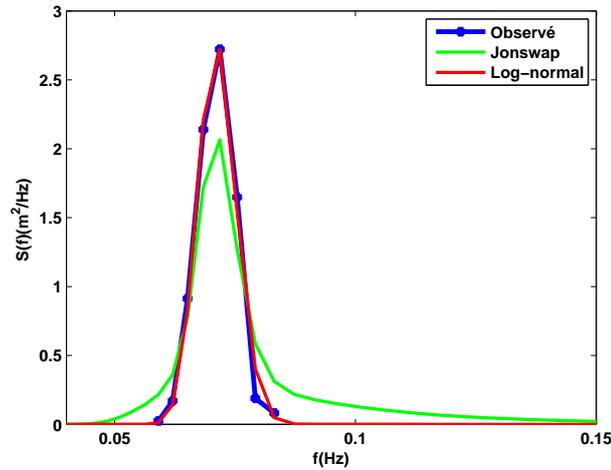


FIGURE 5.15 – Modèles Jonswap et Log-normal ajustés à un spectre de houle

Précisons que la reconstitution des spectres concerne uniquement les événements de houle modélisés. Pour les événements de houle observés, nous avons utilisé les spectres qui ont été estimés avec les données *in situ*.

5.6.3.2 Résultats

Nous présentons ici les résultats de la comparaison entre la puissance extraite estimée à partir des événements de houle observés et modélisés sur les deux sites (figures 5.16 et 5.17).

La fenêtre du haut montre la puissance instantanée extraite par le convertisseur pour les événements de houles observés en comparaison avec la puissance instantanée produite en utilisant les événements de houles modélisés. Sur la fenêtre du bas est représentée la différence entre les puissances instantanées extraites. Sauf pour quelques instants où l'écart entre les deux puissances instantanées est grand, on note dans l'ensemble un bon accord avec des biais quasiment nuls (0.98 kW pour Girassol et 0.32 kW pour Akpo) et des RMSE inférieurs à 5% (3.8 kW pour Girassol et 3.4 kW pour Akpo). Les instants où la différence entre les deux puissances extraites est importante, correspondent à ceux où on note une variation abrupte de la puissance des houles observées due au bruit dans le signal correspondant à la puissance des houles observées. En effet, dans l'évaluation de la puissance extraite, les spectres utilisés pour les houles observées étant ceux estimés directement à partir des données brutes, sont moins lisses que les spectres reconstruits pour les houles modélisées, ce qui explique ces variations abruptes par moment dans le signal de la puissance extraite estimée avec les houles observées. On a choisi d'utiliser pour les houles observées les spectres estimés à partir des données brutes plutôt que des spectres reconstruits afin d'évaluer réellement l'erreur qu'un utilisateur du SEM commettrait dans l'évaluation de la puissance extraite par rapport à l'étalement en fréquence et en direction des spectres de houles observées.

Dans la figure 5.18 sont présentés les diagrammes quantile-quantile de la puissance extractible pour l'ensemble des événements de houles observés et les événements modélisés

sur les deux sites. On observe un bon accord entre l'estimation de la puissance produite à partir des événements observés et modélisés. On note tout de même une sous-estimation de la puissance produite sur le site Girassol pour des puissances produites supérieures à 40 kW. Toutefois avec des coefficients de corrélation de Pearson égales à 0.93, nous pouvons conclure qu'en modélisant les événements de houles par le SEM, on n'introduit pas de différence significative dans l'estimation de la puissance produite.

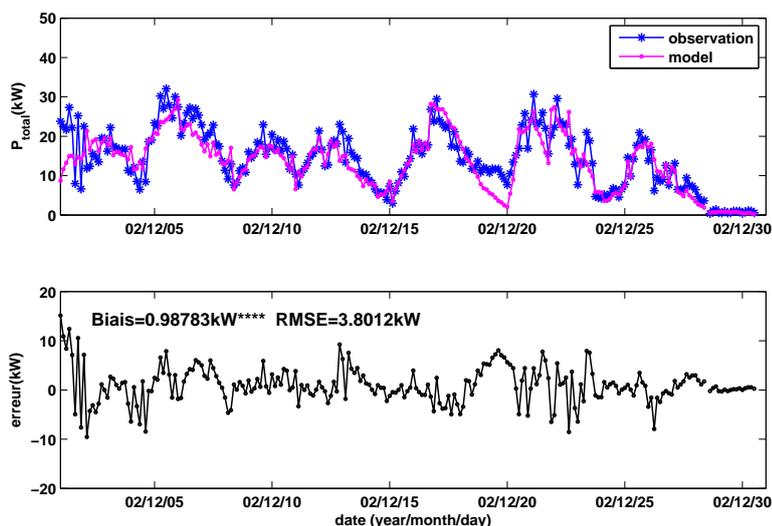


FIGURE 5.16 – Potentiel houlomoteur extractible avec le convertisseur uni-directionnel (Girassol)

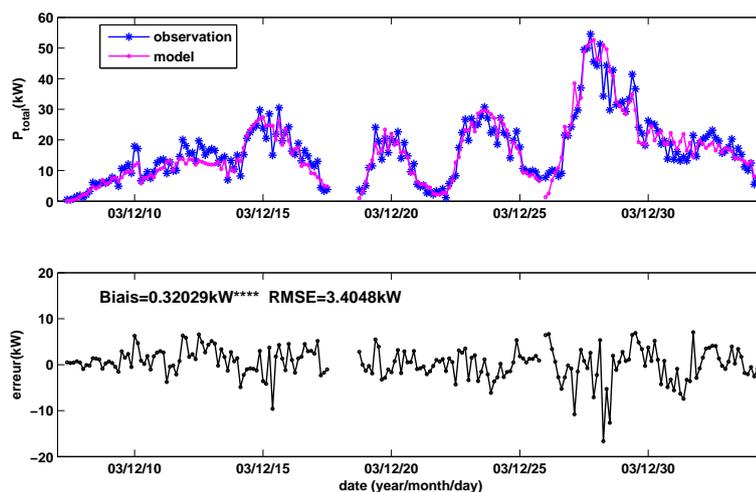


FIGURE 5.17 – Potentiel houlomoteur extractible avec le convertisseur uni-directionnel (Akpo)

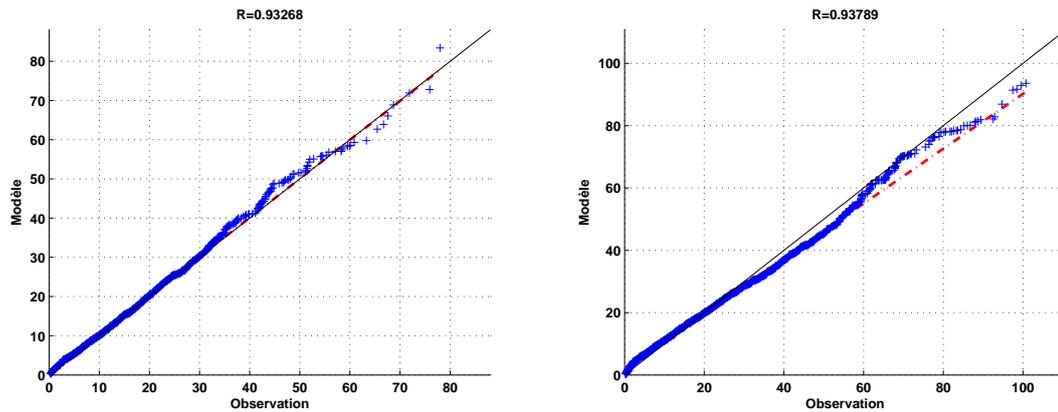


FIGURE 5.18 – Diagramme quantile-quantile comparant la puissance extractible sur le site Akpo (à gauche) et Girassol (à droite)

5.7 Conclusion

Nous avons présenté dans ce chapitre, le modèle d'événement de houle (SEM) qui permet de représenter chacun des événements de houle par un ensemble de 7 paramètres. Le SEM est composé de la loi Gamma ajustée à l'évolution de la puissance et de modèles linéaires ajustés aux évolutions de la fréquence d'énergie et de la direction moyenne. Nos résultats ont montré que les événements de houles observés en Afrique de l'Ouest peuvent être modélisés par le SEM sans différence significative dans la distribution jointe de la hauteur significative et de la période d'énergie et pour les applications inhérentes au génie océanique.

La modélisation des événements de houle est une étape essentielle dans notre démarche de la modélisation de la climatologie des houles avec l'approche par événements. Le SEM permet de réduire le nombre de paramètres nécessaires à la caractérisation d'une climatologie par rapport à une représentation par série temporelle d'états de mer. Il permet également de conserver la cohérence temporelle des systèmes de houles que nous désirons reproduire dans la simulation statistique d'événements de houle. Dès lors que chacun des événements de houle peut être complètement caractérisé par les valeurs des 7 paramètres du SEM (E , α , τ , d , T_{max} , ω et θ_0), il est possible, par modélisation stochastique de la distribution jointe de ces 7 paramètres, de simuler des événements de houles individuels. Mais, pour reconstituer la climatologie des houles, il est indispensable de modéliser également le processus de leurs occurrences. Le chapitre 6 sera dédié à cette fin.

Générateur stochastique d'événements de houle

Le dimensionnement des structures ou l'estimation de l'évolution morphodynamique du trait de côte ou d'autres applications en génie océanique nécessitent de connaître les conditions d'états de mer sur de très longues durées (jusqu'à 100 ans ou plus) alors que les données d'observations souvent disponibles, ne couvrent qu'une durée très limitée (par exemple 3 ans seulement pour le site Girassol). Même les données de hindcast (21 ans pour le site Girassol) sont aussi limitées dans le temps car les modèles numériques de vagues qui servent à les générer, utilisent en forçage des données de vent, généralement indisponibles sur de longues durées. La modélisation stochastique apparaît comme une alternative privilégiée, car elle permet, d'une part, de simuler des données ayant les mêmes propriétés statistiques que les données d'observations (ou de hindcast), et d'autre part, de dépasser la limite imposée, en durée, par ces observations (ou le hindcast).

Dans ce chapitre, nous présentons un générateur stochastique d'événements de houle, développé dans le cadre de cette étude, qui permet de simuler des séries temporelles d'événements de houle utilisables pour des applications en génie océanique. La mise en œuvre de ce générateur stochastique comprend deux étapes :

- La modélisation des distributions des paramètres du SEM caractéristiques d'un événement de houles individuel. En effet, nous avons montré dans le chapitre 5 que chaque événement de houle peut être caractérisé par les 7 paramètres du SEM. La modélisation des distributions de ces paramètres permet donc la simulation d'un grand nombre d'événements de houle ayant les mêmes propriétés statistiques que ceux qui ont été observés.
- La modélisation de l'occurrence des événements de houles : il s'agit de modéliser la manière dont surviennent les événements, ce qui permet de reconstruire des séries temporelles d'événements de houle sur des périodes plus longues que celles pour lesquelles on dispose de données.

Dans l'optique de simuler des séries temporelles d'événements de houles qui soient les plus proches possibles des données de référence, un certain nombre de propriétés statistiques notamment la saisonnalité ou encore la dépendance entre les paramètres du SEM sont prises en compte dans notre générateur. Ce chapitre commence par quelques analyses statistiques dont les résultats seront intégrés dans la mise en œuvre du générateur. Ces analyses statistiques comportent l'étude de la variabilité saisonnière et de la variabilité

interannuelle ainsi que l'étude de la dépendance entre les paramètres du SEM. Par la suite, nous présenterons les deux étapes qui constituent le générateur stochastique d'événement de houle.

Dans tout le chapitre, les données utilisées sont celles de la base de données de hindcast du site Girassol. Comparées aux données *in situ*, ces données de hindcast couvrent une période plus longue et offrent de ce fait la possibilité d'étudier et d'inclure dans notre générateur la variabilité saisonnière. De plus, nous avons montré dans le chapitre 4 que les événements de houles sont bien reproduits dans les données de hindcast qui peuvent donc être utilisées en substitution des données d'observations.

Sommaire

6.1	Analyses statistiques des données de paramètres du SEM	103
6.1.1	Variabilité saisonnière	103
6.1.2	Variabilité interannuelle	105
6.1.3	Étude de la dépendance entre les paramètres du SEM	106
6.2	Distributions jointes des paramètres du SEM	109
6.2.1	Présentation du schéma de la modélisation des distributions jointes des paramètres du SEM	109
6.2.2	Modélisation des distributions jointes de $\{E, d\}$ et $\{\omega, \theta_0\}$ par utilisation des copules	111
6.2.3	Distribution conditionnelle des paramètres α, τ et T_{max} par utilisation des GLM	128
6.3	Modélisation de l'occurrence des événements de houles	135
6.3.1	Temps d'inter-arrivée des événements de houle	135
6.3.2	Modélisation de la distribution du temps d'inter-arrivée des événements de houle	136
6.4	Reconstruction et validation des séries temporelles d'événements de houle	139
6.4.1	Reconstruction de séries temporelles d'événements de houle	139
6.4.2	Validation des séries temporelles d'événements de houle simulés	140
6.5	Conclusion	147
7.1	Synthèse des principales conclusions	149
7.2	Perspectives	151

6.1 Analyses statistiques des données de paramètres du SEM

On se propose d'étudier d'une part les variabilités saisonnière et interannuelle des paramètres du SEM (E , α , τ , d , T_{max} , ω et θ_0) et d'autre part la dépendance entre ces paramètres.

En effet, l'intensité des événements de houle ainsi que leur fréquence varient suivant les saisons et d'une année à une autre. L'hiver correspond généralement à la période où l'activité des tempêtes devient plus intense et plus fréquente. Les houles générées en hiver dans leur zone de génération ont généralement plus d'intensité que celles qui y sont générées en été. En outre, suivant les années, il est également possible d'observer pour une même saison, une variabilité dans l'intensité ou dans la périodicité des houles. Afin de simuler des événements de houles réalistes par rapport aux observations, il est nécessaire de tenir compte des variabilités saisonnière et interannuelle si elles existent. D'où l'importance de cette étude.

Une autre propriété statistique importante à reproduire dans la simulation est la dépendance entre les paramètres du SEM. Les paramètres caractéristiques des états de mer ne sont généralement pas indépendants et on s'attend à ce qu'il en soit de même pour les paramètres du SEM. La modélisation de la dépendance entre les paramètres est primordiale si l'on veut que notre générateur simule une climatologie de houles dont les propriétés soient statistiquement équivalentes à celles de référence (hindcast) et qui puisse être utilisée avec confiance dans les applications en génie océanique.

6.1.1 Variabilité saisonnière

L'intensité des houles observées en Afrique de l'Ouest affiche une forte variabilité saisonnière comme souligné par des études antérieures (Nerzic et al. [2007], Forristall et al. [2013]). Les houles de plus fortes amplitudes sont généralement enregistrées en été boréal (correspondant à l'hiver dans l'hémisphère sud) comme le montre la figure 6.1 où est représenté un an d'événements de houle caractérisés par leur puissance et identifiés à partir des données de hindcast du site Girassol (année 2008). Nous pouvons en effet y remarquer que les événements de forte puissance sont enregistrés entre mai et août. Ce résultat est quasiment le même pour toutes les autres années.

Il est fort probable que les paramètres du SEM (E , α , τ , d , T_{max} , ω et θ_0), avec lesquels nous caractérisons désormais les événements de houles, affichent également cette saisonnalité. Pour le vérifier, les événements de houles issus des 21 années de données de hindcast ont été modélisés par le SEM. Les données des paramètres du SEM ont été ensuite regroupées suivant chaque mois de l'année.

La figure 6.2 montre les distributions mensuelles de l'énergie des événements E , de la distance parcourue par les houles d , de la période maximale T_{max} et de la direction de provenance des premiers trains de houle θ_0 . La saisonnalité est perceptible dans les distributions des paramètres E et T_{max} . Les événements de houle enregistrés entre mai et août affichent des valeurs médianes de E proches de 900 kWh.m^{-1} et des valeurs médianes de T_{max} autour des 18 s. Pendant cette période de l'année, l'énergie des événements de houle peut atteindre 3000 kWh.m^{-1} avec des périodes maximales de houle atteignant les 27 s. Pendant les autres périodes de l'année, les valeurs médianes de E sont généralement inférieures à 500 kWh.m^{-1} et les valeurs médianes de T_{max} sont souvent inférieures à 16s.

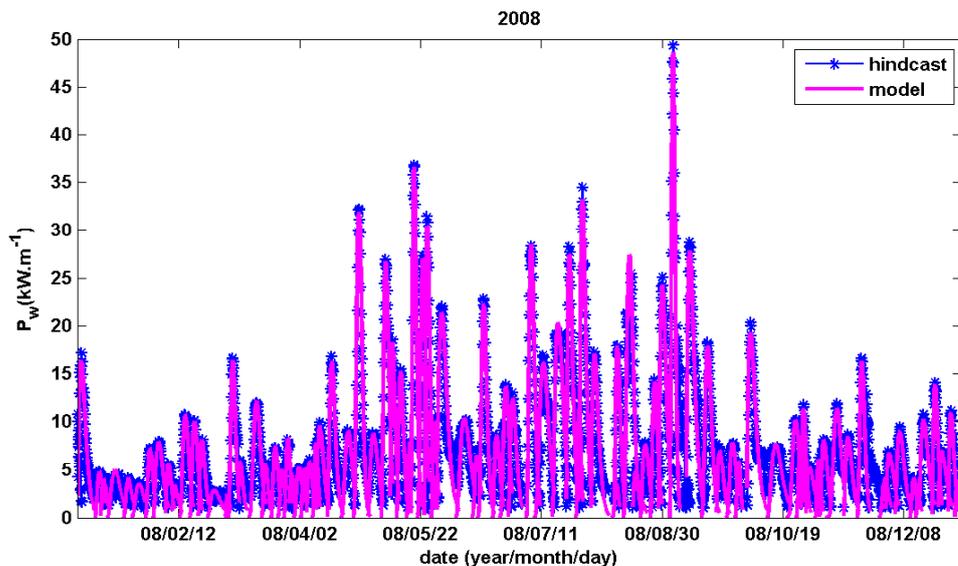


FIGURE 6.1 – Puissance des événements de houle identifiés à partir des données hindcast (Girassol) de l'année 2008 et SEM associés

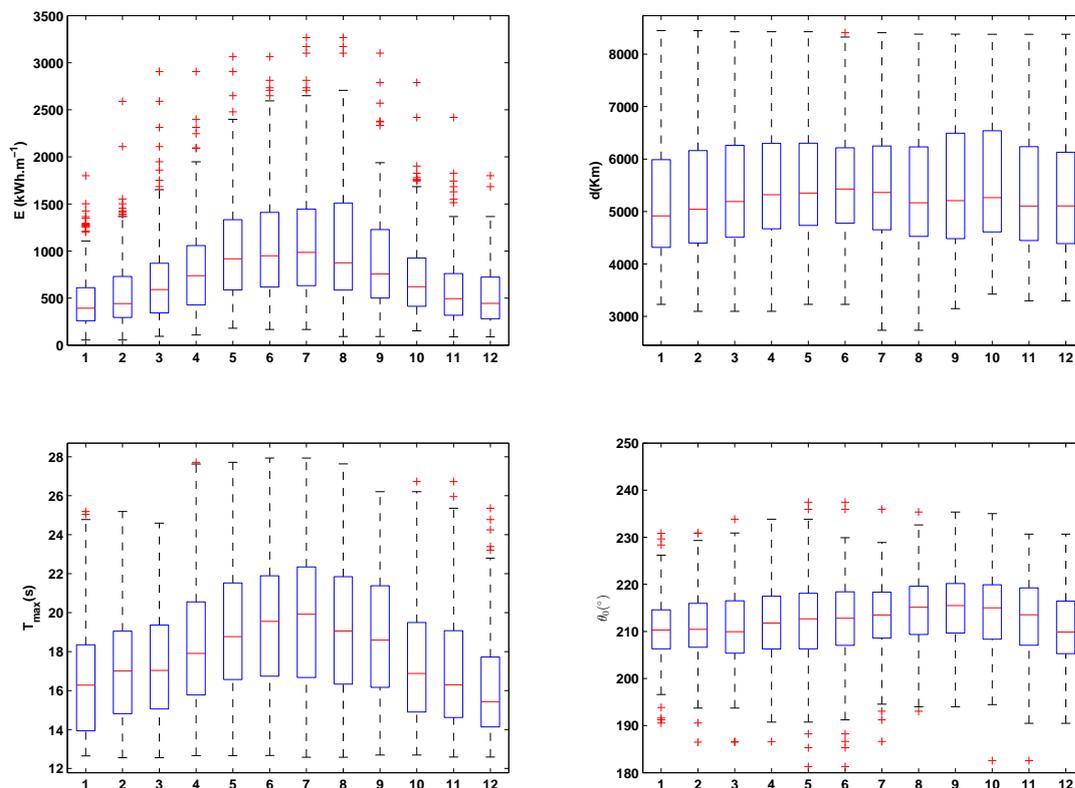


FIGURE 6.2 – Boxplot des distributions mensuelles des paramètres E , d , T_{max} et θ_0 .

On note également une corrélation entre E et T_{max} : les événements de houle de forte intensité sont ceux qui comptent en leur sein les houles de grandes périodes. En effet, la période de mai à août correspond à l’hiver austral où dans l’Atlantique Sud, zone de génération prépondérante des houles observées en Afrique de l’Ouest, les tempêtes deviennent plus fréquentes et plus intenses (Braun [2008], Guishard et al. [2009]), favorisant ainsi la naissance de vagues de grandes périodes et de grandes amplitudes. Précisons que les distributions du paramètre d’échelle τ et du paramètre de forme α , non représentées ici, montrent également une variabilité saisonnière.

Les distributions de la distance parcourue par les houles (d) et de la direction de provenance des premiers trains de houle dans l’événement (θ_0) ne montrent pas quant à elles, de variabilités saisonnières très marquées comme dans le cas de E et T_{max} . Les valeurs médianes de d sont proches de 5000 km et celles de θ_0 sont approximativement égales 200° tout au long de l’année. Nous pouvons conclure en conséquence que les saisons ont très peu d’influence sur la zone de génération des houles.

En résumé, cette étude a révélé que les saisons influent beaucoup sur la distribution de l’énergie des événements ainsi que les périodes qu’ils contiennent. Même si pour des raisons de simplicité, la saisonnalité est rarement prise en compte dans la simulation statistique des paramètres d’états de mer (Corbella, Stretch [2013], Callaghan et al. [2008], Li et al. [2014]), on se propose d’en tenir compte dans cette thèse dans le but de simuler une climatologie de houle réaliste par rapport à la référence (hindcast). Il existe plusieurs façons de prendre en compte la variabilité saisonnière des données. Une approche parfois utilisée consiste à “stationnariser” la série de référence c’est-à-dire retrancher la composante saisonnière et la modéliser séparément. Cette approche n’a pas été retenue car elle induit la plupart du temps des biais (Monbet et al. [2007]). Dans cette thèse nous avons préféré utiliser des modèles dont les paramètres changent en fonction de chaque mois de l’année.

6.1.2 Variabilité interannuelle

On s’intéresse maintenant à la variabilité interannuelle des événements de houle. Pour cela on considère deux variables : le nombre total d’événements de houle par année et l’énergie totale des événements de houle cumulée par année. La figure 6.3 montre l’évolution de ces deux variables entre 1991 et 2012. La fréquence des événements de houle varie peu d’une année à une autre. L’année au cours de laquelle le plus grand nombre d’événements de houle a été enregistré (91 événements) est 2009 et l’année 1996 constitue celle pour laquelle le nombre d’événements de houle enregistrés est le plus petit (74 événements). En moyenne, le nombre d’événements par an est égal à 80 et d’une année à une autre, l’écart par rapport à cette valeur moyenne est faible.

L’évolution de l’énergie totale annuelle des événements de houle montre, en revanche, une forte variabilité entre 1991 et 2012. L’énergie totale des événements de houle est de 95000 kWh.m⁻¹ en 1991, l’année où elle est maximale et sa valeur minimale, enregistrée en 2004, est de 65000 kWh.m⁻¹. Cependant, la série dont nous disposons n’est pas assez longue pour en dégager un cycle de la variation interannuelle de l’énergie totale des événements de houle. Sur la série représentée à la figure 6.3, aucun cycle de variation ne se dégage a priori. En conséquence, la variabilité interannuelle n’a pas été prise en compte dans notre générateur stochastique d’événements de houle.

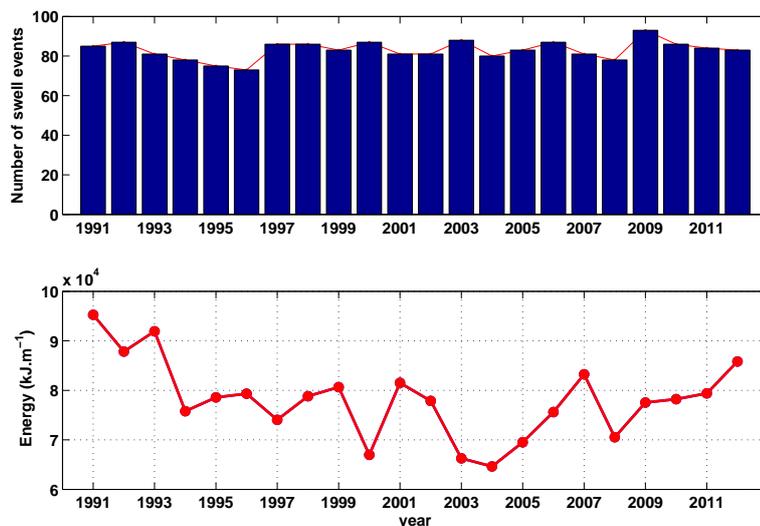


FIGURE 6.3 – Nombre annuel d'événements de houle (haut) et énergie totale annuelle (bas)

6.1.3 Étude de la dépendance entre les paramètres du SEM

La reconstitution d'une climatologie fiable de houle passe par le réalisme de la simulation des événements de houles. Cela impose de modéliser non seulement les distributions marginales des paramètres du SEM, mais aussi les dépendances entre ces derniers. Dans cette section, est proposée une étude de la dépendance entre les paramètres du SEM dont le but est de déterminer quels sont les paramètres corrélés les uns aux autres pour en tenir compte dans la mise en œuvre de notre générateur.

La dépendance entre deux variables aléatoires se mesure généralement par l'analyse graphique du nuage de points de ces deux variables, ou par le biais du coefficient de corrélation de Pearson défini par :

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (6.1)$$

avec n le nombre d'observations. Ce coefficient, par sa définition, permet d'analyser uniquement la relation linéaire supposée entre deux variables aléatoires. Néanmoins, il n'est pas approprié dans les circonstances suivantes (Scherrer [2007]) :

- Si l'on veut mettre en évidence toute autre relation monotone (croissante ou décroissante) entre deux variables quantitatives, et pas uniquement une relation linéaire.
- Si l'on désire tester le degré de la corrélation dans le cas où les variables ne sont pas distribuées normalement. C'est le cas pour la plupart des paramètres du SEM.
- Lorsque le nombre d'observations est très faible.

L'utilisation d'un coefficient de corrélation non-paramétrique tel que le Tau de Kendall (Kendall [1938]) permet de s'affranchir de ces problèmes. C'est pourquoi le Tau de Kendall a été adopté dans cette thèse comme mesure de la dépendance entre les paramètres SEM.

6.1.3.1 Tau de Kendall (corrélation de rang)

Le Tau de Kendall ou coefficient de corrélation de Kendall ($\hat{\tau}_k$) est une statistique qui mesure la corrélation de rang et repose sur la notion de paires d'observations concordantes et discordantes. Soit $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ un ensemble d'observations des variables jointes X et Y . Les paires d'observations (x_i, y_i) et (x_j, y_j) sont dites concordantes si $x_i < x_j$ et $y_i < y_j$ ou si $x_i > x_j$ et $y_i > y_j$. Elles sont dites discordantes si $x_i < x_j$ et $y_i > y_j$ ou si $x_i > x_j$ et $y_i < y_j$. Dans le cas où $x_i = x_j$ ou $y_i = y_j$, la paire n'est ni concordante ni discordante. Le Tau de Kendall est égal à la différence entre la probabilité de concordance et la probabilité de non concordance :

$$\hat{\tau}_k = \frac{n_c - n_d}{n_c + n_d} \quad (6.2)$$

où, n_c est le nombre de paires d'observations concordantes et n_d le nombre de paires d'observations discordantes. Le Tau de Kendall possède les propriétés suivantes :

- Si X et Y sont indépendants alors $\hat{\tau}_k(X, Y) = 0$.
- Il est symétrique, c'est à dire : $\hat{\tau}_k(X, Y) = \hat{\tau}_k(Y, X)$.
- Sa valeur est comprise entre -1 et 1 .
- Si Y est une fonction croissante de X , alors $\hat{\tau}_k(X, Y) = 1$.
- Si Y est une fonction décroissante de X , alors $\hat{\tau}_k(X, Y) = -1$.
- Si a et b sont des fonctions strictement croissantes, alors $\hat{\tau}_k(a(X), b(Y)) = \hat{\tau}_k(X, Y)$

6.1.3.2 Résultats et discussions

Pour tenir compte de la saisonnalité, cette étude a été réalisée séparément pour chaque mois. Nous présentons ici les résultats du mois d'août où les événements de fortes intensités sont enregistrés. Les résultats obtenus sont semblables pour les tous mois.

La figure 6.4 montre les nuages de points entre les couples de paramètres du SEM. L'analyse visuelle révèle une dépendance entre l'énergie E et certains paramètres, notamment la distance d (6.4 (a)) et la période maximale T_{max} (6.4 (b)). Les événements de forte énergie sont ceux qui parcourent généralement les plus grandes distances et contiennent les houles de plus grande période. En effet, par rapport à la zone d'observation des houles (au large de l'Angola) il est plausible que plus on s'en éloigne vers le Sud, plus on a de chance d'avoir des tempêtes de forte intensité qui génèreraient des vagues de forte amplitude et de grande période. On peut également noter une dépendance entre α et τ (6.4 (g)) et ω et θ_0 (6.4 (i)). Pour les autres couples de paramètres l'observation visuelle ne permet pas a priori de déterminer de dépendance.

Pour confirmer ces observations et évaluer statistiquement la dépendance entre les paramètres du SEM, le Tau de Kendall a été calculé pour chaque couple. Les résultats sont présentés sous forme de matrice dans le tableau 6.1. Les valeurs entre les parenthèses correspondent aux p -value associés à chaque valeur du Tau de Kendall. Une p -value < 0.05 signifie que le Tau de Kendall est significativement non nul et les paramètres concernés ne peuvent pas être considérés comme indépendants. Les valeurs des p -value obtenues confirment des dépendances entre les couples $\{E, d\}$, $\{E, T_{max}\}$, $\{\alpha, \tau\}$, $\{\tau, d\}$, $\{\tau, T_{max}\}$, ainsi que $\{d, T_{max}\}$. Les paramètres du SEM relatifs à la direction moyenne des événements de houles $\{\omega, \theta_0\}$ sont dépendants entre eux, mais ils sont indépendants des autres paramètres du SEM.

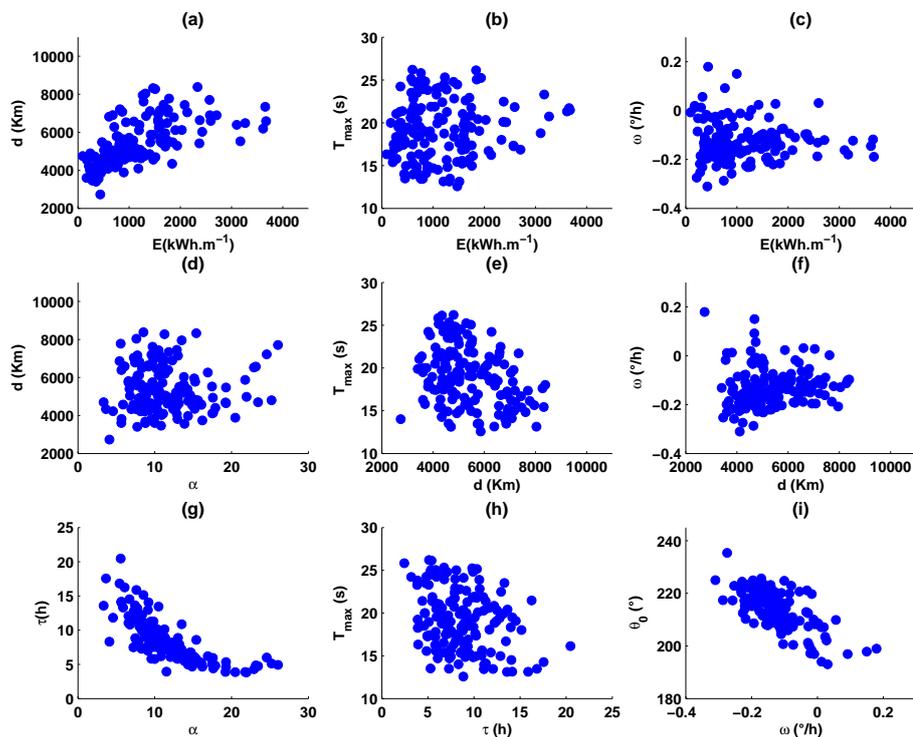


FIGURE 6.4 – Nuages de points des paramètres du SEM

En définitive cette étude de dépendance a montré que les paramètres E , α , τ , d et T_{\max} sont plus ou moins dépendants les uns des autres. Cela signifie, pour le générateur d'événements de houle, qu'il faut modéliser non seulement leurs distributions marginales mais aussi leurs structures de dépendance. Nous pouvons également conclure que les paramètres du modèle de l'évolution de la direction moyenne des houles $\{\omega, \theta_0\}$ sont d'une part dépendants entre eux et d'autre part indépendants des autres paramètres. En conséquence, dans le générateur, la dépendance entre ces deux paramètres doit être modélisée mais leur distribution jointe peut être modélisée indépendamment des 5 autres paramètres que sont E , α , τ , d et T_{\max} .

	E	α	τ	d	T_{max}	ω	θ_0
E	1 -	0.10 (0.1876)	-0.01 (0.9213)	0.47 (3.10^{-11})	0.21 (0.0049)	-0.01 (0.8543)	0.04 (0.5182)
α	0.10 (0.1876)	1 -	-0.75 (10^{-33})	-0.14 0.07	0.47 (4.10^{-11})	0.01 (0.8716)	-0.01 (0.8108)
τ	-0.01 (0.9213)	-0.75 (10^{-33})	1 -	0.42 (3.10^{-9})	-0.45 (2.10^{-11})	-0.04 (0.5210)	0.07 (0.3207)
d	0.47 (3.10^{-11})	-0.13 (0.07)	0.42 (3.10^{-9})	1 -	-0.40 (4.10^{-8})	0.06 (0.3672)	0.08 (0.2605)
T_{max}	0.21 (0.0049)	0.47 (4.10^{-11})	-0.45 (2.10^{-11})	-0.40 (4.10^{-8})	1 -	0.01 (0.8312)	-0.01 (0.8994)
ω	-0.01 (0.8543)	0.01 (0.8716)	-0.04 (0.5210)	0.06 (0.3672)	0.01 (0.8312)	1 -	-0.70 (7.10^{-28})
θ_0	0.04 (0.5182)	-0.01 (0.8108)	0.07 (0.3207)	0.08 (0.2605)	-0.01 (0.8994)	-0.70 (7.10^{-28})	1 -

TABLE 6.1 – Tau de Kendall entre les paramètres du SEM et p -value associée (cas du mois d'août). Les couples de paramètres pour lesquels le Tau de Kendall est significativement non nul sont indiqués en gras.

6.2 Distributions jointes des paramètres du SEM

La première étape dans la mise en œuvre de notre générateur stochastique d'événements de houle est la modélisation des distributions jointes des paramètres du SEM. Les analyses statistiques effectuées sur les paramètres du SEM ont révélé, d'une part, qu'ils sont marqués par une forte variabilité saisonnière (Section 6.1.1) et, d'autre part, qu'ils sont plus ou moins dépendants les uns des autres (Section 6.1.3). Les modèles de distributions jointes sont obtenus en modélisant à la fois les lois marginales des paramètres et les structures de dépendance entre eux. Pour tenir compte de la saisonnalité, les modèles de distributions jointes sont déterminées séparément pour chaque mois. Ces modèles permettront ensuite de simuler des paramètres du SEM ayant les mêmes propriétés statistiques que ceux de référence. À partir des paramètres du SEM simulés, on peut reconstituer un grand nombre d'événements de houles individuels.

6.2.1 Présentation du schéma de la modélisation des distributions jointes des paramètres du SEM

L'étude de la dépendance a révélé que les paramètres du modèle d'évolution de la direction moyenne des houles au sein d'un événement $\{\omega, \theta_0\}$ sont indépendants des autres paramètres, mais corrélés entre eux. Pour cette raison, la distribution jointe de ces deux paramètres est modélisée indépendamment des 5 autres paramètres du SEM.

Les autres paramètres, à savoir E , α , τ , d et T_{max} , sont plus ou moins dépendants entre eux et doivent être idéalement simulés à partir d'une distribution de probabilité jointe à 5

dimensions. Cependant, de telles distributions jointes sont difficiles à obtenir sauf, dans le cas où, pour chaque dimension, les variables aléatoires suivent une loi normale. En effet, la loi normale peut être facilement généralisée à n dimensions avec comme paramètres, le vecteur des moyennes et la matrice de variance-covariance. Par contre, la généralisation à n dimensions des autres lois de probabilité est difficile à obtenir mathématiquement. De plus, pour inférer les paramètres d'une distribution de probabilité jointe à 5 dimensions, il faut disposer d'une grande taille d'échantillons. Dans notre cas, nous ne disposons que de 160 échantillons en moyenne pour chaque mois, ce qui est insuffisant pour inférer les paramètres d'une loi jointe à 5 dimensions avec confiance.

Une méthode souvent utilisée pour déterminer la probabilité jointe de plusieurs variables caractéristiques des états de mer est l'approche de modélisation conditionnelle (Conditional Modelling Approach (CMA), dans laquelle, la distribution de probabilité jointe est obtenue à partir des lois marginales et d'une série de lois conditionnelles (Bitner-Gregersen et al. [1991], Bitner-Gregersen [2005], Callaghan et al. [2008]). En suivant cette même démarche, nous avons fait le choix de modéliser d'abord la loi jointe du couple de paramètres $\{E, d\}$ puis de déterminer les distributions des autres paramètres conditionnellement à $\{E, d\}$. Le couple $\{E, d\}$ a été choisi plutôt qu'un autre couple de paramètres parce que, d'une part, E qui représente l'énergie de l'événement est la même dans le modèle que dans l'événement observé, alors que les autres paramètres découlent des modèles d'évolution de la puissance et de la fréquence moyenne au sein des événements de houle. D'autre part, parmi les 4 autres paramètres, d est celui dont la corrélation avec E est la plus forte (voir tableau 6.1).

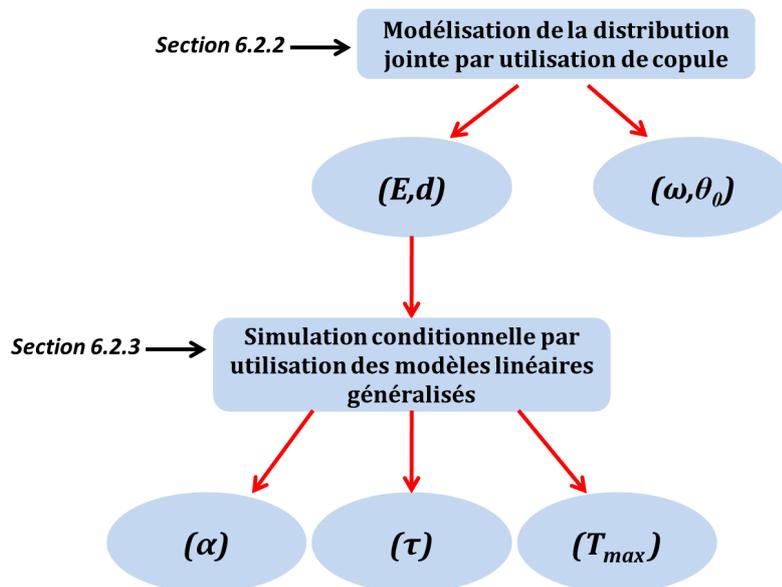


FIGURE 6.5 – Diagramme montrant le schéma utilisé pour déterminer les distributions des paramètres du SEM

Le schéma de la démarche utilisée pour déterminer les distributions jointes des paramètres du SEM est présenté dans la figure 6.5. Pour modéliser les distributions de probabilités jointes de $\{\omega, \theta_0\}$ et de $\{E, d\}$ nous avons utilisé une approche basée sur les copules (Section 6.2.2). Une copule est un objet mathématique qui permet de caractériser

la dépendance entre deux ou plusieurs variables aléatoires sans se préoccuper de leurs lois marginales. Les distributions des paramètres α , τ , et T_{max} sont ensuite déterminées conditionnellement à celle de $\{E, d\}$ par utilisation des modèles linéaires généralisés (generalized linear model (GLM)) (Section 6.2.3).

6.2.2 Modélisation des distributions jointes de $\{E, d\}$ et $\{\omega, \theta_0\}$ par utilisation des copules

Les résultats de l'étude de la dépendance (6.1.3) entre les paramètres du SEM ont montré que les paramètres E et d d'une part, et d'autre part ω , θ_0 ne sont pas indépendants. Il existe de nombreuses méthodes pour modéliser la dépendance entre les paramètres océano-météorologiques pour des applications aussi variées que le dimensionnement des structures, l'érosion côtière, l'énergie des vagues, etc. Parmi ces méthodes, nous pouvons citer :

- le modèle logistique, utilisé par Morton, Bowers [1996] pour modéliser la dépendance entre les extrêmes de vent et de hauteur significative, ou encore par Galiatsatou, Prinós [2011] pour estimer la distribution jointe de la hauteur des vagues et de la surcote.
- les transformations isoprobabilistes de Rosenblatt (Rosenblatt [1952], utilisée par exemple par Nerzic et al. [2007]) et de Nataf (Der Kiureghian, Pei-Ling [1986], utilisée par exemple par Sagrilo et al. [2011]).
- les lois de distributions bivariées (voir 2.4.1.3).

Une méthode alternative est l'utilisation des copules dont la définition et les propriétés sont présentées dans la sous-section 6.2.2.2. Les copules sont des outils relativement innovants de modélisation de la structure de dépendance de plusieurs variables aléatoires. Même si elles ont été introduites dès 1959 par Sklar [1959], leur utilisation est récente et est due à de récentes études qui ont montré combien elles sont des outils performants pour modéliser la dépendance entre les variables aléatoires (Joe [1997], Nelsen [1999], Salvadori, De Michele [2007]). Les copules constituent un outil statistique qui présente de nombreux avantages. Outre une grande souplesse dans la mise en œuvre de l'analyse multivariée, les copules autorisent une sélection plus étendue des distributions conjointes des variables. Les copules permettent une représentation moins naïve de la dépendance statistique fondée sur la mesure traditionnelle de corrélation linéaire qui présente des limites dans l'étude de l'interdépendance entre deux variables (Embrechts et al. [1999]). Elles permettent donc de modéliser des dépendances non-linéaires et non-monotones entre variables. En outre, elles autorisent des distributions de probabilités jointes moins restrictives, prenant mieux en compte certains comportements tels que l'asymétrie ou la dépendance des queues. Elles permettent la construction de distributions multidimensionnelles assez générales et ce, indépendamment des lois marginales qui peuvent être différentes et quelconques. Par conséquent, elles permettent de s'affranchir de certaines hypothèses peu réalistes faites dans les études empiriques.

L'engouement envers l'utilisation des copules a connu ces dernières décennies un essor considérable surtout dans la gestion des risques financiers, l'économétrie ou l'assurance (Cherubini et al. [2004], Jaworski et al. [2010]). L'utilisation des copules dans les sciences environnementales a commencé en hydrologie (Favre et al. [2004], Salvadori, De Michele [2007]). Plus récemment, la modélisation par les copules a été introduite et utilisée dans

de nombreuses études pour estimer la distribution jointe des variables en génie océanique et côtier : De Michele et al. [2007] utilisent les copules pour modéliser la dépendance entre la hauteur significative, la direction, la durée et le temps d'inter-arrivée des tempêtes ; De Waal, Gelder van [2005] modélisent la dépendance entre les extrêmes de H_s et T_p par les copules ; Corbella, Stretch [2012], Corbella, Stretch [2013], Li et al. [2014], Wahl et al. [2012], Masina et al. [2015] ont également utilisé les copules pour la simulation statistique des paramètres d'état de mer pour des applications liées à l'érosion côtière.

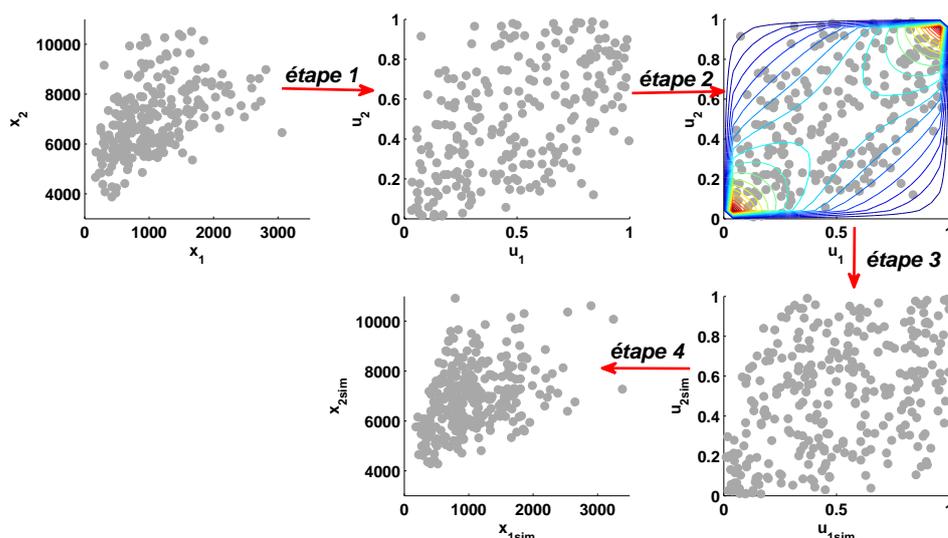


FIGURE 6.6 – Illustration des différentes étapes pour la simulation d'un couple de variables aléatoires $\{x_1, x_2\}$ par utilisation d'une copule

Les différentes étapes de la simulation d'un couple de variables aléatoires $\{x_1, x_2\}$, par utilisation d'une copule sont illustrées dans la figure 6.6 et décrites comme suit :

- étape 1 : À partir des observations de x_1 et x_2 , on calcule les fonctions de répartition empiriques $u_1 = F(x_1)$ et $u_2 = F(x_2)$ en utilisant un estimateur par noyau.
- étape 2 : On ajuste la copule à la distribution jointe empirique de u_1 et u_2 soit $F(u_1, u_2)$.
- étape 3 : Avec la copule ajustée, on simule des couples de variables aléatoires uniformes : $(u_{1sim}, u_{2sim}) \in [0, 1]^2$.
- étape 4 : Les variables uniformes simulées avec la copule sont transformées dans l'échelle physique des variables x_1 et x_2 en utilisant les inverses des fonctions de répartition marginales théoriques, préalablement ajustées aux fonctions de répartition empiriques de x_1 et x_2 .

La mise en œuvre de ces 4 étapes nécessite 2 outils fondamentaux :

- les lois marginales des paramètres x_1 et x_2 et
- la copule qui permet de modéliser la structure de dépendance entre x_1 et x_2 .

Ces deux outils peuvent être déterminés indépendamment l'un de l'autre. De plus, la validation de la loi jointe peut également s'effectuer à deux niveaux : sur les lois marginales et sur la structure de dépendance qu'est la copule. C'est ce deuxième schéma de validation qui a été adopté dans cette étude. Dans ce qui va suivre, nous présenterons d'abord les

lois marginales utilisées pour les paramètres E , d , ω et θ_0 , puis nous nous intéresserons à la modélisation de la dépendance par les copules.

6.2.2.1 Distribution marginale des paramètres (E , d , ω et θ_0)

On cherche à modéliser la distribution marginale empirique des paramètres du SEM notamment E , d , ω et θ_0 par des lois de distributions de probabilités théoriques existantes. En effet, une distribution empirique ne peut permettre de simuler que des valeurs qui ont été observées. Remplacer la distribution empirique par une loi théorique qui s'y ajuste bien permet de dépasser la limite imposée par les observations.

Il n'y a pas de raison physique justifiant le choix d'une loi de distribution théorique par rapport à une autre. Lorsqu'on s'intéresse, par exemple, à l'estimation de la probabilité d'occurrence des états de mer pour le dimensionnement des structures marines en fonction de chargements extrêmes, il existe des lois de distributions (GPD, GEV) qui sont spécialement dédiées aux valeurs extrêmes (voir 2.14). Cependant, pour de nombreuses autres applications, notamment le dimensionnement en réponse à la fatigue ou encore l'estimation du potentiel d'énergie récupérable des vagues, on cherche plutôt à bien représenter la probabilité d'occurrence non seulement des états de mer extrêmes, mais aussi des états de mer de faible et moyenne intensités. Notre objectif est de bien modéliser l'ensemble de la distribution des paramètres du SEM et non pas exclusivement les valeurs extrêmes.

La distribution marginale de E a été modélisée par la loi de probabilité Gamma¹. C'est la loi qui modélise au mieux la distribution empirique de E . De plus, elle est flexible et permet de mieux prendre en compte l'ensemble de la distribution de E (Ferreira, Guedes Soares [2000]). Pour tenir compte de la saisonnalité, on impose que les paramètres de la loi prennent des valeurs différentes en fonction des mois. La fonction de répartition de la loi Gamma s'écrit alors :

$$F(x|a(k), b(k)) = \frac{1}{b(k)^{a(k)}\Gamma(a(k))} \int_0^x t^{a(k)-1} \exp\left(\frac{-t}{b(k)}\right) dt, k \in \{1, 2, \dots, 12\} \quad (6.3)$$

où k symbolise les mois de janvier à décembre. Pour chaque mois, les paramètres de forme $a(k)$ et d'échelle $b(k)$ de la loi sont obtenus en maximisant la Log-vraisemblance définie par :

$$L(a(k), b(k)) = (a(k) - 1) \sum_{i=1}^n \log(x_i) - \frac{1}{b(k)} \sum_{i=1}^n x_i - na(k) \log(b(k)) - n \log(\Gamma(a(k))) \quad (6.4)$$

avec n le nombre total d'observations. La figure 6.7 montre l'évolution mensuelle des paramètres $a(k)$ et $b(k)$. On retrouve la signature de la saisonnalité dans l'évolution des deux paramètres : les fortes valeurs de $b(k)$, associées aux faibles valeurs de $a(k)$ sont obtenues en été boréal.

1. La loi de probabilité Gamma est utilisée ici comme la distribution théorique marginale de E . Dans le SEM nous avons utilisé la fonction correspondant à la densité d'une loi Gamma pour modéliser l'évolution de la puissance au sein d'un événement. Il est important de distinguer ces deux différentes utilisations de la distribution Gamma

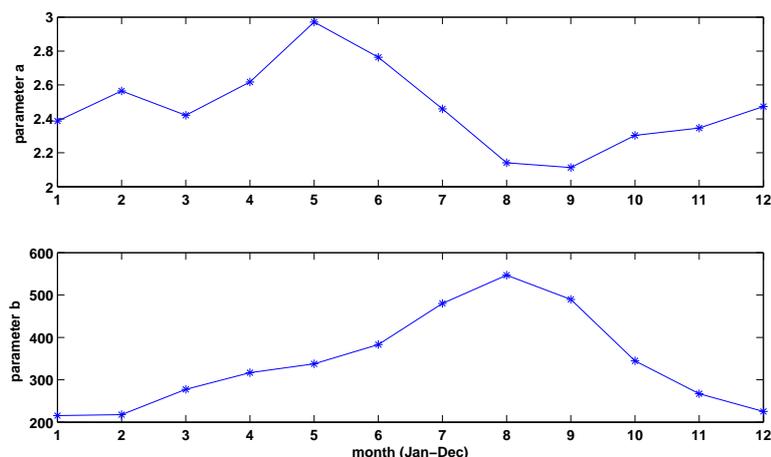


FIGURE 6.7 – Évolution mensuelle du paramètre de forme $a(k)$ et d'échelle $b(k)$ de la loi Gamma ajustée aux observations de E .

La qualité des ajustements a été validée par le test de Kolmogorov-Smirnov qui est un test d'hypothèse non-paramétrique utilisé pour déterminer si un échantillon suit bien une loi donnée connue par sa fonction de répartition continue (Massey Jr [1951]). La statistique de ce test est le maximum de la différence absolue entre la fonction de répartition empirique $G(x)$ et la fonction de répartition théorique $F(x)$:

$$D^* = \max_x |G(x) - F(x)|. \quad (6.5)$$

Les résultats des ajustements sont présentés dans la figure 6.8. On note globalement un bon accord entre la loi théorique et la distribution empirique de E même si pour certains mois (février, juillet, août, octobre et novembre) les valeurs extrêmes ne sont pas bien modélisées par la loi Gamma. Rappelons que notre objectif est de modéliser au mieux l'ensemble de la distribution non pas exclusivement les valeurs extrêmes. Il est difficile de trouver une loi qui prenne en compte à la fois les valeurs extrêmes et les autres valeurs de la distribution. C'est pourquoi, comme souligné plus haut, dans le cas où on s'intéresse aux valeurs extrêmes, des lois spécifiques telles que les GEV et les GPD doivent être privilégiées. Toutefois, les p -value du test de Kolmogorov-Smirnov (tableau 6.2) affichent des valeurs nettement supérieures à 0.05, signifiant ainsi qu'avec un risque de 5% de se tromper, la distribution de E dérive d'une loi Gamma. En se référant aux résultats du test de Kolmogorov-Smirnov, nous pouvons conclure que la loi Gamma ajuste bien la distribution marginale du paramètre E .

	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
a	2.38	2.56	2.42	2.61	2.97	2.76	2.45	2.14	2.11	2.30	2.34	2.47
b	215	217	277	316	337	383	480	546	489	344	266	225
p -value	0.93	0.75	0.50	0.67	0.74	0.89	0.98	0.26	0.88	0.98	0.93	0.94

TABLE 6.2 – Valeurs mensuelles des paramètres $a(k)$ et $b(k)$ et p -value du test de Kolmogorov-Smirnov

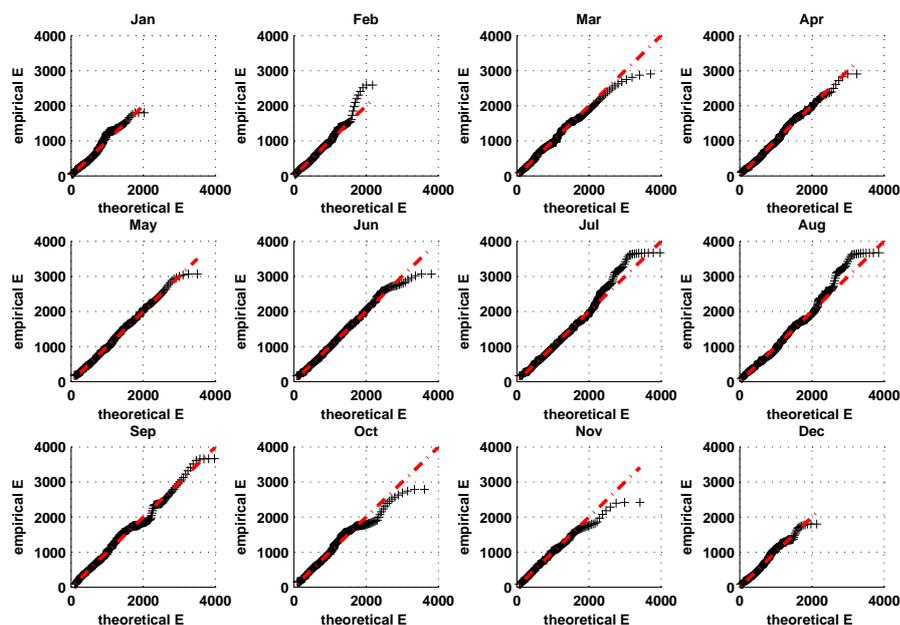


FIGURE 6.8 – Diagramme quantile-quantile de la comparaison entre la distribution marginale empirique de E et celle donnée par la loi Gamma

En ce qui concerne les autres paramètres, suite à des tests réalisés avec plusieurs lois, les meilleurs résultats ont été obtenus avec la loi log-normale pour d et les lois normales pour ω et θ_0 . La qualité des ajustements a été également validée par le test Kolmogorov-Smirnov. Le tableau 6.3 présente, pour chaque mois, les paramètres des lois ajustées ainsi que les p -value du test de Kolmogorov-Smirnov. Les résultats des ajustements sont montrés dans les diagrammes quantile-quantile des figures 6.9, 6.10 et 6.11. Les p -value des tests de Kolmogorov-Smirnov obtenus sont toutes supérieures à 0.05 permettant de conclure ainsi que les lois théoriques choisies ajustent bien les distributions empiriques.

$$\text{paramètre } d \leftrightarrow \text{loi log-normale : } F(x|\mu(k), \sigma(k)) = \frac{1}{\sigma(k)\sqrt{2\pi}} \int_0^x \frac{\exp\left(\frac{-(\ln(t)-\mu(k))^2}{2\sigma(k)^2}\right)}{t} dt$$

	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
μ	8.84	8.87	8.87	8.88	8.87	8.88	8.89	8.87	8.87	8.90	8.87	8.86
σ	0.31	0.34	0.32	0.31	0.31	0.31	0.35	0.37	0.36	0.36	0.33	0.31
$p\text{-value}$	0.21	0.25	0.30	0.28	0.55	0.46	0.37	0.40	0.50	0.51	0.36	0.24

$$\text{paramètre } \omega \leftrightarrow \text{loi normale : } F(x|\mu(k), \sigma(k)) = \frac{1}{\sigma(k)\sqrt{2\pi}} \int_{-\infty}^x \exp\left(\frac{-(t - \mu(k))^2}{2\sigma(k)^2}\right) dt$$

	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
$-\mu$	0.058	0.059	0.059	0.053	0.043	0.045	0.052	0.067	0.081	0.090	0.083	0.065
σ	0.080	0.083	0.098	0.107	0.109	0.108	0.106	0.105	0.099	0.093	0.096	0.094
$p\text{-value}$	0.83	0.67	0.51	0.38	0.69	0.91	0.83	0.86	0.72	0.64	0.67	0.92

$$\text{paramètre } \theta_0 \leftrightarrow \text{loi normale : } F(x|\mu(k), \sigma(k)) = \frac{1}{\sigma(k)\sqrt{2\pi}} \int_{-\infty}^x \exp\left(\frac{-(t - \mu(k))^2}{2\sigma(k)^2}\right) dt$$

	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
μ	208	208	208	209	210	210	210	211	211	211	211	208
σ	7.58	7.87	8.81	9.25	9.84	9.49	8.81	9.03	8.95	8.76	8.68	8.10
$p\text{-value}$	0.96	0.99	0.69	0.76	0.37	0.63	0.84	0.87	0.50	0.71	0.36	0.58

TABLE 6.3 – Valeurs mensuelles des paramètres des lois marginales ajustées aux paramètres (d , ω et θ_0) et $p\text{-value}$ des tests de Kolmogorov-Smirnov

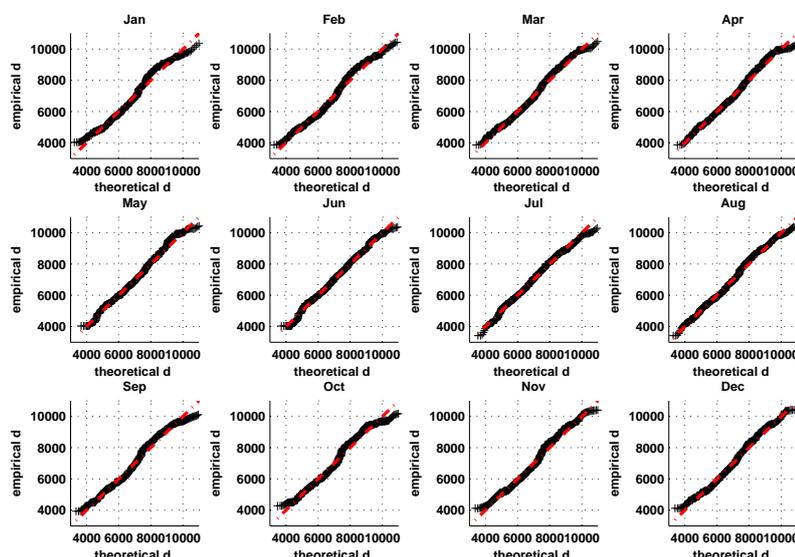


FIGURE 6.9 – Diagramme quantile-quantile de la comparaison entre la distribution de marginale empirique de d et celle de la loi log-normale

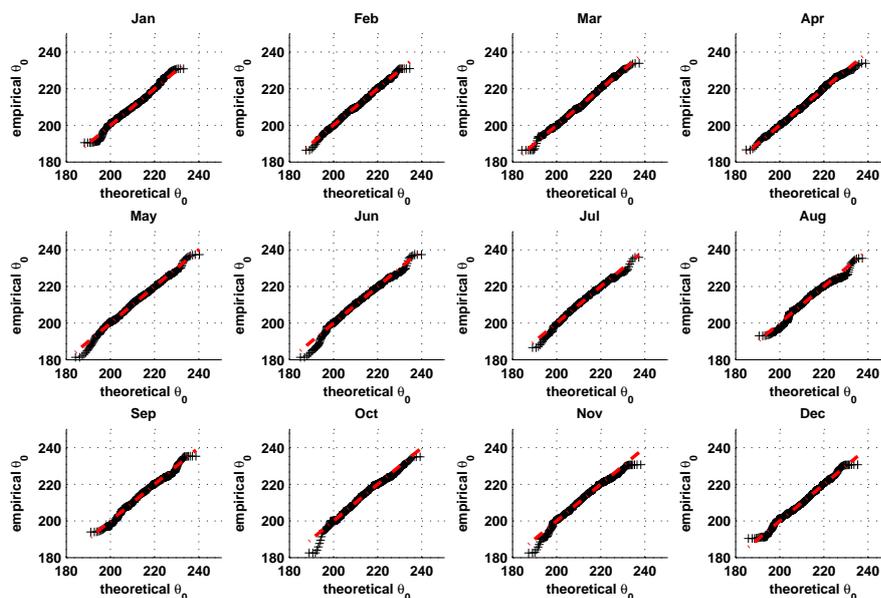


FIGURE 6.10 – Diagramme quantile-quantile de la comparaison entre la distribution de marginale empirique de θ_0 et celle de la loi normale

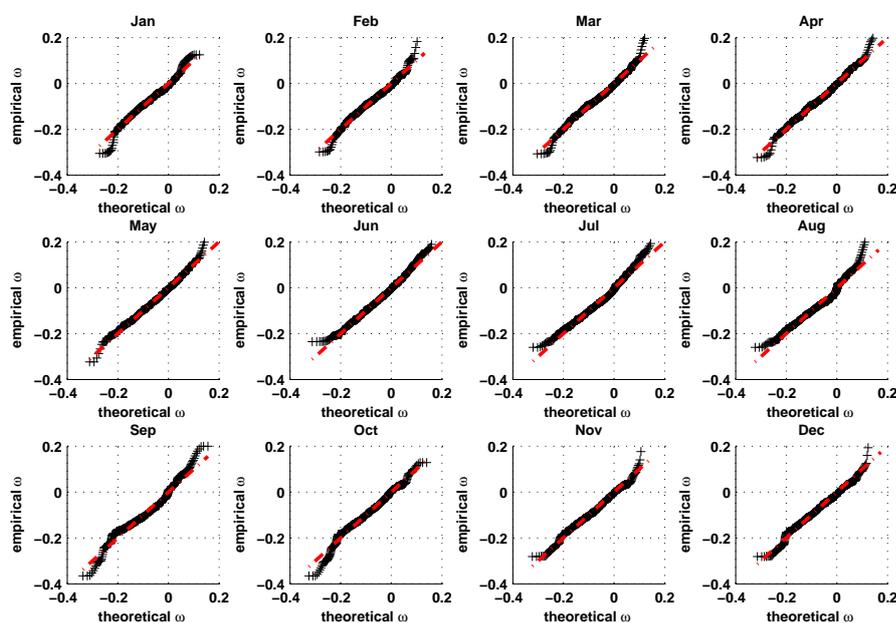


FIGURE 6.11 – Diagramme quantile-quantile de la comparaison entre la distribution de marginale empirique de ω et celle de la loi normale

6.2.2.2 Modélisation de la dépendance entre $\{E, d\}$ et $\{\omega, \theta_0\}$ par les copules

Définition et propriétés générales des copules bivariées

Pour une définition complète des copules, le lecteur peut se référer aux ouvrages de référence tels que Joe [1997] et Nelsen [1999]. Nous introduisons ici, les notions de bases nécessaires à l'utilisation des copules qui est faite dans cette thèse.

Définition : Une copule bivariée C est une fonction de $[0, 1]^2 \rightarrow [0, 1]$ et est définie par les caractéristiques suivantes :

- (i) $C(u, 0) = C(0, u) = 0 \forall u \in [0, 1]$
- (ii) $C(u, 1) = C(1, u) = u \forall u \in [0, 1]$
- (iii) $C(v_1, v_2) - C(v_1, u_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0 \forall (u_1, u_2) \in [0, 1]^2, (v_1, v_2) \in [0, 1]^2$
tel que $0 \leq u_1 \leq v_1 \leq 1$ et $0 \leq u_2 \leq v_2 \leq 1$.

Soient U_1 et U_2 deux variables aléatoires uniformes sur $[0, 1]$, alors on a $C(u_1, u_2) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \forall (u_1, u_2) \in [0, 1]^2$. Cette définition assure donc que la copule est une distribution de probabilité avec des marges uniformes.

Le théorème de Sklar [1959] précise le lien défini par la copule C , déterminée à partir de la distribution jointe F , entre les fonctions de répartition marginales univariées F_1 et F_2 et la distribution complète bivariée F .

Soit F une distribution bivariée de marges F_1 et F_2 . La copule C associée à F s'écrit :

$$C(u_1, u_2) = C(F_1(x_1), F_2(x_2)) \quad (6.6)$$

$$= F(F_1^{-1}(u_1), F_2^{-1}(u_2)) \quad (6.7)$$

$$= F(x_1, x_2) \quad (6.8)$$

C est unique lorsque les marges F_1 et F_2 sont continues. La densité f d'une loi bivariée peut s'écrire aussi en fonction de la densité c de la copule associée et des densités des marginales f_1 et f_2 :

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \times f_1(x_1) \times f_2(x_2) \quad (6.9)$$

Copules étudiées

On présente dans cette partie quelques exemples de copules qui ont été utilisées dans le cadre de la thèse. Il existe 4 grandes familles de copules : la famille des copules elliptiques (copule Gaussienne et copule de Student), la famille des copules Archimédiennes (copule de Clayton, copule de Gumbel et copule de Frank, etc.), la familles des copules de valeurs extrêmes et la famille des copules Archimax. La famille Archimax est une généralisation des copules Archimédiennes et des copules de valeurs extrêmes. Les copules de valeurs extrêmes, comme l'indique leur nom, sont exclusivement dédiées à la modélisation des valeurs extrêmes. Puisqu'on cherche à modéliser l'ensemble des distributions des paramètres, elles ne sont pas considérées dans cette thèse.

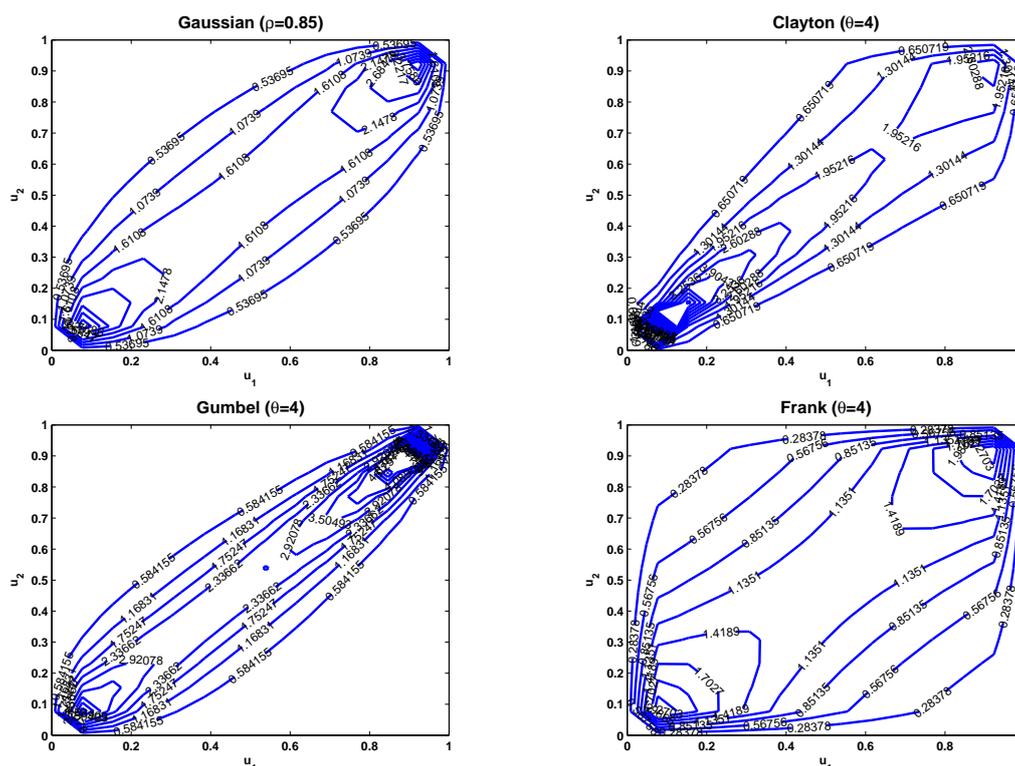


FIGURE 6.12 – Fonctions densités de probabilité des copules.

Copule Gaussienne : La copule Gaussienne bivariee est définie de la façon suivante :

$$C(u_1, u_2; \rho) = \Phi_\rho \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2) \right) \quad (6.10)$$

ρ est le coefficient de corrélation et Φ_ρ la distribution normale bivariee standard de matrice de corrélation fonction de ρ . Le Tau de Kendall τ_k est relié à ρ par :

$$\tau_k = \frac{2}{\pi} \arcsin \rho \quad (6.11)$$

Copule Archimédiennes : Les copules Archimédiennes sont définies de la manière suivante :

$$C(u_1, u_2) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) & \text{si } \varphi(u_1) + \varphi(u_2) \leq \varphi(0) \\ 0 & \text{sinon} \end{cases} \quad (6.12)$$

avec $\varphi(1) = 0$, $\varphi'(u) < 0$ et $\varphi''(u) > 0$ pour tout $0 \leq u \leq 1$. φ est appelée la fonction génératrice de la copule. Le Tau de Kendall pour les copules Archimédiennes est :

$$\tau_k = 1 + 4 \int_0^1 \frac{\varphi(u)}{\varphi'(u)} du \quad (6.13)$$

Quelques exemples de copules Archimédiennes bivariees sont présentés dans le tableau 6.4.

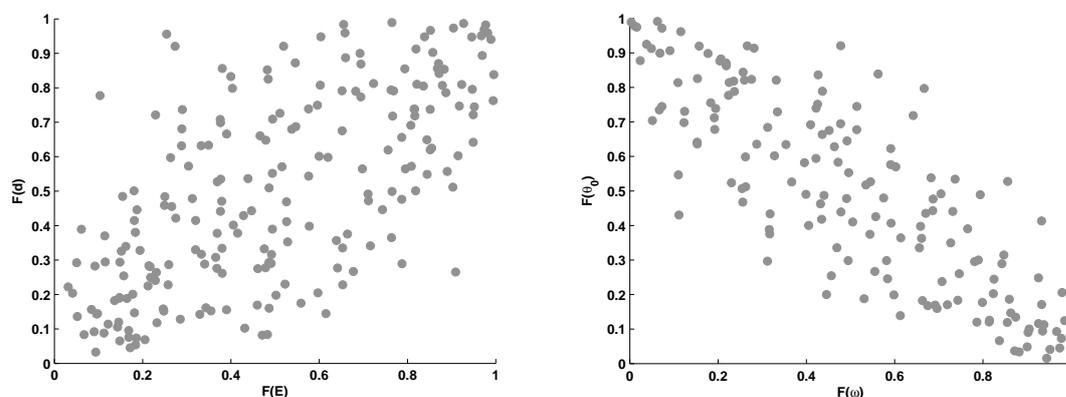
La figure 6.12 montre les fonctions densités de probabilité des différentes copules évoquées ci-dessus.

Nom	Générateur	Copule bivariee
Clayton ($\theta > 0$)	$u^{-\theta} - 1$	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}$
Gumbel ($\theta \geq 1$)	$(-\ln u)^\theta$	$\exp\left(-\left((\ln u_1)^\theta + (\ln u_2)^\theta\right)^{1/\theta}\right)$
Franck ($\theta \neq 0$)	$-\ln \frac{\exp(-\theta u) - 1}{\exp(-\theta) - 1}$	$-\frac{1}{\theta} \ln \left(1 + \frac{\exp((- \theta u_1) - 1)(\exp((- \theta u_2) - 1))}{\exp(-\theta) - 1}\right)$

TABLE 6.4 – Exemples de copules Archimédiennes bivariées

Sélection de la copule adaptée à la modélisation de la dépendance de $\{E, d\}$ et $\{\omega, \theta_0\}$

On cherche parmi les différentes copules présentées dans 6.2.2.2 et dont les fonctions de densités sont représentées à la figure 6.12, celles qui modélisent le mieux la dépendance de $\{E, d\}$ et $\{\omega, \theta_0\}$ i.e. la distribution jointe de leurs fonctions de répartition (figure 6.13).


 FIGURE 6.13 – Distributions jointes des fonctions de répartition empiriques de E et d (à gauche) et de ω et θ_0 (à droite) (Cas du mois d'août)

Pour $\{E, d\}$, nous avons évalué les 4 copules (Gaussienne, Gumbel, Clayton et Frank). En ce qui concerne $\{\omega, \theta_0\}$, la dépendance est négative entre les fonctions de répartition (figure 6.13). Étant donné que les copules de Gumbel et de Clayton ne sont pas définies pour les dépendances négatives, seules les copules Gaussienne et de Frank ont été évaluées pour la modélisation de la dépendance de $\{\omega, \theta_0\}$.

Les copules ont été ajustées par la méthode du maximum de vraisemblance. La log-vraisemblance s'écrit sous la forme :

$$l(\psi) = \sum_{i=1}^n \log [c_\psi F_1(x_{i1}; \psi), F_2(x_{i2}; \psi)] + \sum_{j=1}^2 \sum_{i=1}^n \log (f_j(x_{ij}; \psi)) \quad (6.14)$$

où, c_ψ est la fonction de densité de la copule C_ψ , x_{ij} la i ème observation pour la variable $j \in \{1, 2\}$, F_i la loi paramétrique choisie pour la variable j . Les paramètres ψ sont obtenus par maximisation de la log-vraisemblance.

Pour choisir la meilleure copule, on considère deux critères de comparaison et de sélection de modèles généralement utilisés en statistique : le critère d'information bayésien

(bayesian information criterion ou *BIC* Schwarz, others [1978]) et le critère d'information d'Akaike corrigé (corrected Akaike information criterion ou *AICc* Akaike [1998]). Ces deux critères sont beaucoup utilisés dans la sélection de copules (Smith [2003], Santos Silva dos, Lopes [2008], Masina et al. [2015]). Le *BIC* et le *AICc* sont respectivement définis par :

$$BIC = -2 \ln(L) + K \ln(n) \quad (6.15)$$

$$AICc = 2K - 2 \ln(L) + \frac{2K(K+1)}{n-K-1} \quad (6.16)$$

avec n le nombre d'échantillons, K le nombre de paramètres à inférer et L le maximum de la fonction de vraisemblance. La meilleure copule est celle qui présente le *BIC* et/ou le *AICc* les plus petits.

Le tableau 6.5 résume pour chaque mois, les différentes valeurs du *BIC* et du *AICc* obtenues pour les copules utilisées (Gaussienne, Gumbel, Clayton et Frank) pour la modélisation de la dépendance entre E et d . La copule Gaussienne affiche, pour la quasi totalité des mois, les plus petits *BIC* et *AICc*. Seul le mois de janvier voit la copule de Frank afficher un *BIC* et un *AICc* plus faibles que ceux obtenus avec la copule Gaussienne. Les copules de Gumbel et de Clayton présentent, les *BIC* et *AICc* les plus élevés. La figure 6.14 montre les contours de probabilité de la copule empirique en comparaison avec les contours de probabilité des 4 copules. La copule empirique bivariée est définie par (Deheuvels [1979]) :

$$C_n(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^2 \mathbf{1}\{F_{nj}(X_{ij} \leq u_j)\}, \quad (6.17)$$

où

$$F_{nj}(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_{ij}u_j \leq x), 1 \leq j \leq d \quad (6.18)$$

sont les fonctions de répartitions empiriques marginales. On peut observer que la copule Gaussienne présente les contours de probabilités les plus proches de ceux de la copule empirique aussi bien pour les faibles probabilités que pour les probabilités élevées.

De même, en comparant les *BIC* et *AICc* obtenus avec la copule Gaussienne et la copule de Frank pour le cas du couple (ω, θ_0) (tableau 6.6), les plus faibles valeurs sont obtenues avec la copule Gaussienne.

En définitive, aussi bien pour $\{E, d\}$ que pour $\{\omega, \theta_0\}$, la copule Gaussienne comparée aux autres, apparaît comme la copule la mieux adaptée à la modélisation de la dépendance entre les paramètres, d'après les critères du *BIC* et du *AICc*. En conséquence, elle sera dans la suite la copule considérée pour modéliser la structure de dépendance entre les paramètres.

copule Gaussienne												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (ρ)	0.57	0.61	0.62	0.54	0.50	0.54	0.60	0.66	0.53	0.53	0.57	0.57
<i>BIC</i>	-52.4	-91.4	-86.2	-63.8	-56.2	-67.5	-80.5	-97.0	-48.8	-41.2	-63.1	-56.5
<i>AICc</i>	-55.7	-94.7	-89.9	-67.2	-59.6	-70.9	-83.8	-100	-51.9	-44.2	-66.1	-59.5

copule de Gumbel												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (θ)	1.51	1.53	1.62	1.46	1.39	1.44	1.55	1.68	1.462	1.5	1.59	1.57
<i>BIC</i>	-40.1	-60.8	-72.2	-50.0	-39.4	-44.8	-58.3	-74.9	-38.7	-38.7	-58.4	-49.5
<i>AICc</i>	-43.1	-64.2	-75.5	-53.3	-42.8	-48.2	-61.6	-78.2	-41.9	-41.7	-61.5	-52.6

copule de Clayton												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (θ)	0.92	1.12	1.01	0.81	0.75	0.85	0.96	1.14	0.89	0.91	1.05	0.96
<i>BIC</i>	-43.4	-87.7	-66.7	-53.1	-50.8	-62.8	-66.1	-76.3	-47.1	-41.4	-62.7	-48.8
<i>AICc</i>	-46.5	-91.1	-70.0	-56.5	-54.2	-66.2	-69.4	-79.5	-50.2	-44.4	-65.9	-51.9

copule de Frank												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (θ)	4.07	4.31	4.54	3.51	3.17	3.63	4.47	5.18	3.51	3.59	4.11	4.25
<i>BIC</i>	-53.6	-85.0	-82.7	-57.8	-51.2	-66.0	-78.7	-95.1	-44.6	-39.4	-60.8	-54.4
<i>AICc</i>	-56.6	-88.4	-86.0	-61.2	-54.7	-69.4	-82.0	-98.0	-47.8	-42.5	-64.0	-58.5

TABLE 6.5 – *BIC* et *AICc* pour les différentes copules utilisées pour modéliser la dépendance entre E et d .

copule Gaussienne												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (ρ)	-0.7	-0.8	-0.7	-0.8	-0.7	-0.8	-0.7	-0.7	-0.8	-0.7	-0.8	-0.7
<i>BIC</i>	-107	-151	-82	-124	-134	-176	-111	-104	-156	-100	-144	-99
<i>AICc</i>	-110	-154	-85	-126	-137	-179	-114	-107	-159	-103	-147	-102

copule de Frank												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
paramètre (θ)	-6.3	-8.7	-5.8	-7.6	-7.1	-6.4	-7.2	-7.0	-8.4	-7.2	-8.4	-6.7
<i>BIC</i>	-100	-148	-82	-122	-124	-163	-105	-111	-147	-105	-137	-96
<i>AICc</i>	-103	-151	-85	-125	-127	-166	-108	-114	-150	-108	-140	-99

TABLE 6.6 – *BIC* et *AICc* pour les différentes copules utilisées pour modéliser la dépendance entre ω et θ_0 .

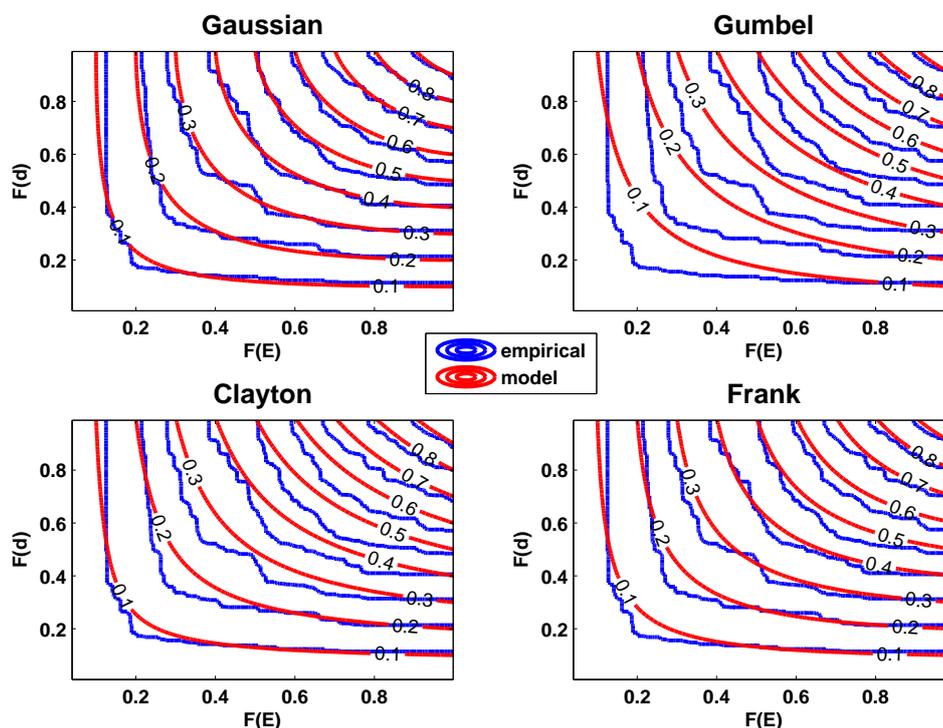


FIGURE 6.14 – Comparaison des contours de probabilités de la copule empirique de $\{E, d\}$ avec ceux des 4 copules (Cas du mois d’août).

Validation de la modélisation de la dépendance entre $\{E, d\}$ et $\{\omega, \theta_0\}$ avec la copule Gaussienne

Le *BIC* et le *AICc* ont permis de sélectionner parmi les copules étudiées, la copule Gaussienne comme la copule la mieux adaptée à la modélisation de la dépendance entre $\{E, d\}$ et $\{\omega, \theta_0\}$. Cependant, ces deux critères ne nous permettent pas de valider le fait que la copule Gaussienne peut être valablement utilisée en substitution de la copule empirique. Afin de s’assurer que la copule Gaussienne que nous utilisons reproduit bien la dépendance entre les paramètres, il est important de procéder à des tests statistiques d’adéquation.

Comme par exemple dans le cas d’une loi marginale, il existe des tests statistiques permettant de valider le fait que la copule empirique appartient à une famille paramétrique de copules ; l’objectif étant d’utiliser la copule paramétrique pour reproduire statistiquement la même dépendance entre les variables. Les tests statistiques utilisés pour les lois marginales (test de Kolmogorov-Smirnov, test de χ^2) peuvent a priori être généralisés en 2 dimensions et utilisés pour valider les ajustements des copules. Cependant, des études ont montré que le test de χ^2 et le test de Kolmogorov-Smirnov ne sont pas performants pour juger de la validité de l’ajustement d’une copule (Genest, Rémillard [2008], Genest et al. [2009], Berg [2009]) : le premier dépend fortement du nombre de classes en lequel est subdivisé l’espace tandis que la second est très sensible aux déviations dans la majeure partie

de la distribution. En comparant la puissance d'un certain nombre de tests, Berg [2009] et Genest et al. [2009] ont montré que les tests les plus performants, quant à la validation de l'ajustement d'une copule, sont basés sur le processus de copule empirique et reposent sur la statistique de Anderson-Darling et surtout sur la statistique de Cramér-von Mises.

Dans le test basé sur le processus de copule empirique mis en place, on cherche à vérifier l'hypothèse H_0 suivant laquelle la copule empirique C_n peut être considérée comme une copule Gaussienne :

$$H_0 : C_n \in C_\rho \quad \text{et} \quad H_1 : C_n \notin C_\rho \quad (6.19)$$

Les statistiques de Cramér-von Mises S_n et de Anderson-Darling A_n sont respectivement définies par :

$$S_n = \sum_{j=1}^n \{C_n(\mathbf{u}_j) - C_\rho(\mathbf{u}_j)\}^2 \quad (6.20)$$

avec $\mathbf{u}_j = (u_{1j}, u_{2j})$,

$$A_n = \max_{i,j} \frac{C_n(\mathbf{u}_j) - C_\rho(\mathbf{u}_j)}{\sqrt{C_\rho(\mathbf{u}_j) [1 - C_\rho(\mathbf{u}_j)]}} \quad (6.21)$$

Les deux statistiques comparent de manière différente, la copule Gaussienne à la copule empirique et conduisent à accepter l'hypothèse H_0 lorsque leurs valeurs sont faibles par rapport à un seuil fixé. Il faut souligner que la statistique de Anderson-Darling donne plus d'importance à la dépendance dans les queues des distributions et permet, par conséquent, d'évaluer la modélisation de la dépendance entre les valeurs extrêmes par la copule Gaussienne. De plus, pour ces deux tests il est possible d'obtenir une *p-value* approximative par la méthode d'auto-échantillonnage ou de *bootstrap* paramétrique² (Genest, Rémillard [2008]). Avec un seuil fixé à 5%, l'hypothèse H_0 selon laquelle la copule empirique appartient à la famille Gaussienne est acceptée si *p-value* > 0.05. Dans ce cas, nous pouvons alors conclure que la copule Gaussienne peut être utilisée en substitution de la copule empirique.

Le tableau 6.7 présente les résultats des tests statistiques d'ajustement de la copule Gaussienne pour $\{E, d\}$ et $\{\omega, \theta_0\}$. Les *p-value* sont indiquées entre les parenthèses et sont obtenues à partir de la méthode proposée par Genest, Rémillard [2008] avec 5000 échantillons du *bootstrap* paramétrique. Le choix du nombre d'échantillons du *bootstrap* résulte d'un compromis entre le temps de calcul et la précision de l'estimation de la *p-value*. Nous pouvons constater que les valeurs des *p-value* des deux tests statistiques affichent des valeurs supérieures à 0.05 aussi bien pour $\{E, d\}$ que pour $\{\omega, \theta_0\}$. L'hypothèse H_0 ne peut pas être rejetée et nous concluons ainsi que la copule Gaussienne est appropriée à la modélisation de la dépendance de $\{E, d\}$ et de $\{\omega, \theta_0\}$.

2. Il s'agit de générer N échantillons de taille n tirés de la loi C_ρ , et de calculer à chaque fois la statistique d'intérêt

Couple $\{E, d\}$												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
S_n	0.0178 (0.50)	0.0244 (0.21)	0.0092 (0.98)	0.0126 (0.86)	0.0112 (0.93)	0.0181 (0.47)	0.0231 (0.23)	0.0176 (0.44)	0.0220 (0.33)	0.017 (0.55)	0.0187 (0.44)	0.0245 (0.20)
A_n	0.7072 (0.82)	1.2097 (0.78)	1.3785 (0.66)	1.0772 (0.51)	0.4789 (0.90)	0.5737 (0.92)	1.4183 (0.56)	2.155 (0.34)	1.5297 (0.12)	1.3338 (0.14)	1.5551 (0.26)	1.081 (0.52)

Couple $\{\omega, \theta_0\}$												
	Jan	Fev	Mar	Avr	mai	Jun	Jul	Aou	Sep	Oct	Nov	Dec
S_n	0.0137 (0.89)	0.0198 (0.41)	0.0196 (0.61)	0.0177 (0.65)	0.0166 (0.67)	0.0145 (0.74)	0.0179 (0.66)	0.014 (0.90)	0.0214 (0.22)	0.0209 (0.50)	0.0143 (0.81)	0.0217 (0.43)
A_n	2.7302 (0.41)	2.9991 (0.87)	1.5862 (0.71)	2.9677 (0.56)	2.7191 (0.76)	4.3841 (0.52)	2.3686 (0.65)	2.8835 (0.30)	4.4022 (0.31)	2.1761 (0.65)	3.4414 (0.65)	2.7642 (0.32)

TABLE 6.7 – Statistiques de Cramér-von Mises S_n et de Anderson-Darling A_n calculées à partir de la copule Gaussienne pour $\{E, d\}$, et pour $\{\omega, \theta_0\}$. Les p -value sont indiquées dans les parenthèses et sont obtenues à partir d'un *bootstrap* à 5000 échantillons

6.2.2.3 Résultat de la simulation des couples $\{E, d\}$ et $\{\omega, \theta_0\}$ avec la copule Gaussienne

Nous avons d'une part déterminé et validé les lois marginales des paramètres E , d , ω et θ_0 (voir 6.2.2.1) et d'autre part, nous avons obtenu la validation que la copule Gaussienne est la copule adaptée à la modélisation de la dépendance de $\{E, d\}$ et de $\{\omega, \theta_0\}$ et peut être utilisée pour reproduire une dépendance statistiquement équivalente à celle entre les paramètres (voir 6.2.2.2). Nous disposons donc des deux outils évoqués dans l'introduction de la Section 6.2.2 pour maintenant simuler $\{E, d\}$ et $\{\omega, \theta_0\}$.

Conformément aux étapes 3 et 4 de la simulation d'un couple de variables aléatoires $\{x_1, x_2\}$ par utilisation d'une copule qui ont été présentées dans 6.2.2, nous avons simulé des variables aléatoires uniformes avec la copule Gaussienne obtenue. Ces variables ainsi simulées sont transformées dans l'échelle physique des paramètres en utilisant pour chaque paramètre, l'inverse de sa loi marginale.

La figure 6.15 montre pour comparaison, à gauche la distribution jointe de $\{E, d\}$ issue des données de hindcast (servant de référence) et à droite la distribution jointe de $\{E, d\}$ simulée avec un même nombre d'échantillons. Les histogrammes correspondant aux distributions marginales des paramètres y sont également représentés. Nous notons globalement un bon accord entre la distribution jointe de $\{E, d\}$ de référence et celle simulée. Nous pouvons également remarquer que les contours de probabilités sont assez proches. Soulignons toutefois quelques légères différences. En effet la distribution empirique comporte plus de points correspondants aux valeurs de E supérieures à 3000 kWh.m^{-1} que dans la simulation. Une explication plausible est que la loi marginale théorique (Gamma), ajustée à la distribution empirique de E sous-estime les valeurs extrêmes de E comme on peut le constater dans le figure (6.8) pour le cas du mois d'août. C'est l'inverse de cette loi marginale qui a été utilisée pour transformer les variables uniformes simulées par la copule, en échelle physique. D'où cette légère réduction du nombre de points simulés

correspondant aux valeurs extrêmes de E . Malgré cette légère sous-estimation apparente des valeurs extrêmes de E , les tests de Kolmogorov pour les lois marginales et les tests de Cramér-von Mises et d' Anderson-Darling pour la copule qui ont été effectués permettent de conclure que la distribution jointe de $\{E, d\}$ simulée est très proche de celle de référence dans sa globalité.

Pour ce qui concerne la simulation de $\{\omega, \theta_0\}$ (figure 6.16), l'accord est également notable entre la distribution jointe simulée et celle de référence.

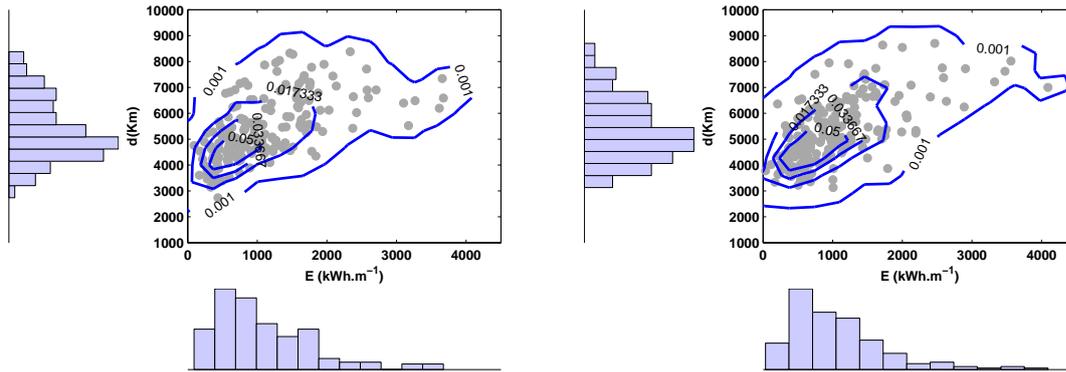


FIGURE 6.15 – Distributions jointes de $\{E, d\}$: référence (à gauche) et simulée (à droite). Les histogrammes correspondent aux distributions marginales des paramètres

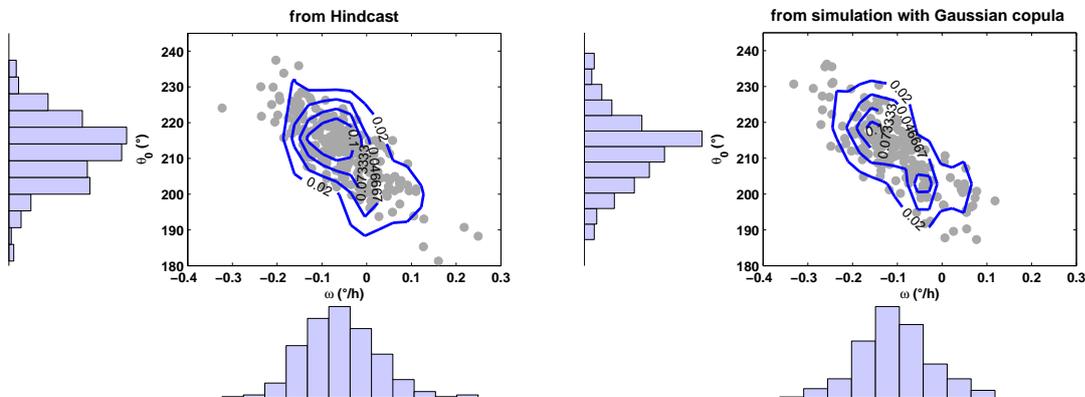


FIGURE 6.16 – Distributions jointes de $\{\omega, \theta_0\}$: référence (à gauche) et simulée (à droite). Les histogrammes correspondent aux distributions marginales des paramètres

À la lumière des résultats obtenus, nous pouvons conclure que la copule Gaussienne et les lois marginales ajustées permettent de bien modéliser la distribution jointe de $\{E, d\}$ et $\{\omega, \theta_0\}$. Les simulations réalisées avec ces distributions jointes donnent des résultats globalement semblables aux données de référence. Nous présentons pour compléter, dans les figures 6.17 et 6.18, les résultats des simulations en comparaison avec les données de référence pour chaque mois. Nous avons simulé pour chaque mois, 700 couples de $\{E, d\}$ et $\{\omega, \theta_0\}$ correspondant approximativement à 100 ans de séquence d'événements de houle

sur la base d'un nombre moyen de 7 événements de houle par mois. Pour l'ensemble des mois, les nuages de points des couples $\{E, d\}$ et $\{\omega, \theta_0\}$ simulés sont similaires aux nuages de points de référence.

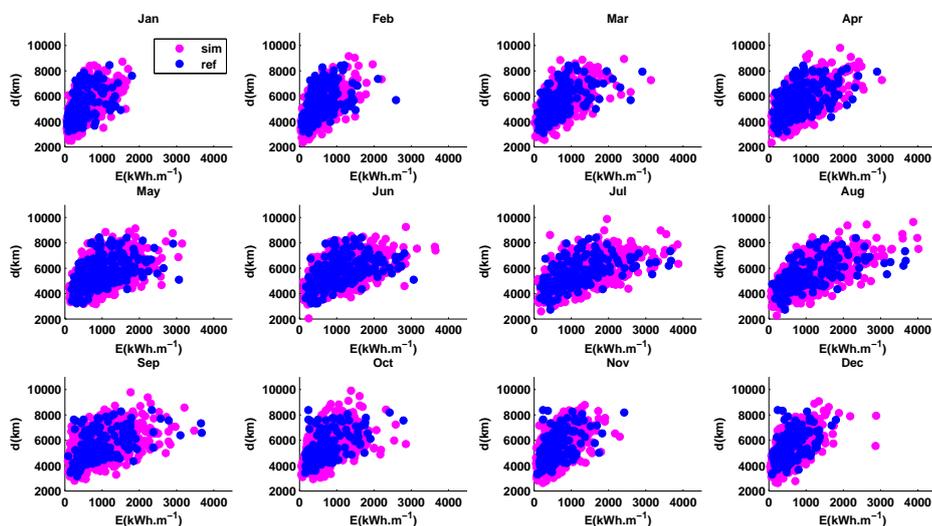


FIGURE 6.17 – Nuage de points de $\{E, d\}$ simulés en comparaison avec les données de référence.

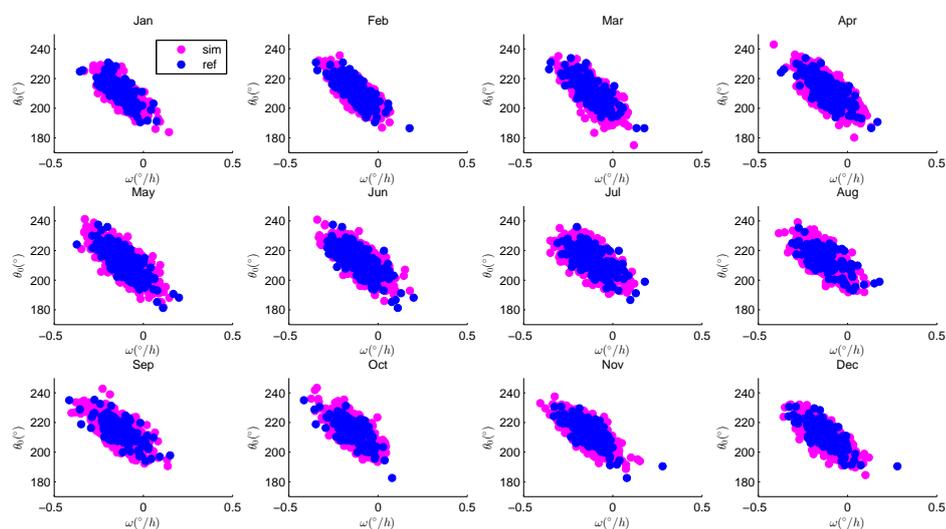


FIGURE 6.18 – Nuage de points de $\{\omega, \theta_0\}$ simulés en comparaison avec les données de référence.

6.2.3 Distribution conditionnelle des paramètres α , τ et T_{max} par utilisation des GLM

Conformément au schéma de simulation proposé (figure 6.5), on désire simuler les paramètres α , τ et T_{max} conditionnellement aux valeurs de E et d simulés par utilisation de modèles linéaires généralisés (GLM). Les GLM sont une généralisation souple de la régression linéaire en termes de loi. Ils permettent l'analyse de données discrètes mais aussi de données continues pour lesquelles la loi normale n'est pas des plus adaptées. Dans leur ouvrage, [McCullagh, Nelder \[1989\]](#), en font une présentation complète. Dans les sciences environnementales, le potentiel et la pertinence des GLM ont été démontrés dans de nombreuses études ([Coe, Stern \[1982\]](#), [Stern, Coe \[1984\]](#), [Yan et al. \[2002\]](#), [Chandler \[2005\]](#)).

6.2.3.1 Définition des GLM

Notons y le vecteur des observations de taille n , réalisation du vecteur aléatoire Y , variable à expliquer. Les GLM sont formés de trois composantes suivantes :

- la variable à expliquer Y , composante aléatoire à laquelle est associée une loi de probabilité.
- les variables explicatives $X_1 \dots X_k$ utilisées comme prédicteurs dans le modèle, qui définissent sous forme d'une combinaison linéaire la composante déterministe.
- le lien qui décrit la relation fonctionnelle entre la combinaison linéaire des variables $X_1 \dots X_k$ et l'espérance mathématique de la variable à expliquer Y .

Composante aléatoire : On suppose que les composantes $Y_i (i = 1, \dots, n)$ de Y sont indépendantes et distribuées selon une loi appartenant à la famille exponentielle, c'est-à-dire que la fonction de densité de la variable aléatoire Y_i s'écrit :

$$f(y_i; \theta_i, \phi) = \exp \left\{ \frac{y_i \theta_i b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\} \quad (6.22)$$

où, θ_i est un paramètre canonique et ϕ un paramètre de dispersion. Les fonctions a , b et c sont spécifiées en fonction du type de loi exponentielle. Les lois de probabilités telles que la loi normale, la loi binomiale, la loi de Poisson, la loi Gamma et la loi de Gauss inverse appartiennent à la famille exponentielle.

Composante déterministe : La composante déterministe, exprimée sous forme d'une combinaison linéaire $\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$ (appelée aussi prédicteur linéaire) précise quels sont les prédicteurs.

Lien : La troisième composante d'un modèle linéaire généralisé est le lien entre la composante aléatoire et la composante déterministe. Il spécifie comment l'espérance mathématique de Y notée μ est liée au prédicteur linéaire construit à partir des variables explicatives. On peut modéliser l'espérance μ directement (régression linéaire usuelle) ou modéliser une fonction monotone $g(\mu)$ de l'espérance :

$$g(\mu) = \beta \mathbf{X} = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad (6.23)$$

A toute loi de probabilité de la composante aléatoire est associée une fonction spécifique de l'espérance appelée paramètre canonique. Le tableau 6.8 résume pour quelques distributions continues appartenant à la famille exponentielle, les fonctions lien canoniques associées.

Distribution	Nom	Fonction lien
Normale	Identité	$\beta\mathbf{X} = \mu$
Exponentiel	Inverse	$\beta\mathbf{X} = \mu^{-1}$
Gamma	Inverse	$\beta\mathbf{X} = \mu^{-1}$
Gauss inverse	Inverse carré	$\beta\mathbf{X} = \mu^{-2}$

TABLE 6.8 – Exemples de distributions continues appartenant à la famille exponentielle et fonctions lien canoniques associées

L'estimation des paramètres β de dimension k , coefficients de la combinaison linéaire des covariables permettant d'expliquer le vecteur Y , s'effectue généralement en maximisant la log-vraisemblance du vecteur des paramètres canoniques θ qui s'écrit :

$$L(\theta; y) = \sum_{i=1}^n \left[\frac{y_i \theta_i b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right] \quad (6.24)$$

Quant au choix du meilleur modèle, il est souvent réalisé en comparant la déviance des différents modèles. La déviance Dev mesure la différence entre la log-vraisemblance du modèle L et la log-vraisemblance du modèle dit saturé L_{sat} c'est-à-dire le modèle possédant autant de paramètres que d'observations :

$$Dev = 2(L - L_{sat}) \quad (6.25)$$

On montre qu'asymptotiquement, Dev suit une loi du χ^2 à $n - k$ degrés de liberté ce qui permet de construire un test de rejet ou d'acceptation du modèle selon que la déviance est jugée significativement ou non importante.

Enfin, le diagnostic du modèle est souvent réalisé par l'analyse des résidus de Pearson qui comparent les valeurs observées y_i et la moyenne μ_i des valeurs prédites pondérées par leur précision estimée par l'écart-type σ_i :

$$r_i^{(k)} = \frac{Y_i - \mu_i}{\sigma_i} \quad (6.26)$$

Si le modèle ajusté est adéquat, tous les résidus de Pearson r_i auront une espérance nulle et une variance égale à 1.

6.2.3.2 Application des GLM pour la simulation de α , τ et T_{max}

On cherche à identifier le meilleur GLM pour les trois paramètres α , τ et T_{max} . Guidé par le degré de corrélation entre les différents paramètres, nous avons testé plusieurs combinaisons de variables explicatives pour en retenir la meilleure d'après le critère de la déviance. Les meilleurs modèles obtenus pour α , τ et T_{max} sont présentés ci-après :

Paramètre T_{max} :

- Variables explicatives : E et d
- Distribution : normale
- Fonction lien canonique : identité

Variables explicatives	Estimation
E	24.82×10^{-4}
d	-13.10×10^{-4}
Ordonnée à l'origine	28.71

TABLE 6.9 – Estimation des paramètre du GLM exprimant T_{max} en fonction de E et d (cas du mois d'août)

Paramètre τ :

- Variables explicatives : d et T_{max}
- Distribution : normale
- Fonction lien canonique : identité

Variables explicatives	Estimation
T_{max}	-0.17
d	8.2×10^{-4}
Ordonnée à l'origine	7.32

TABLE 6.10 – Estimation des paramètre du GLM exprimant τ en fonction de d et T_{max} (cas du mois d'août)

Paramètre α :

- Variables explicatives : E , τ , d et T_{max}
- Distribution : Gamma
- Fonction lien canonique : inverse

Variables explicatives	Estimation
E	9.10×10^{-6}
tau	0.997×10^{-4}
d	-11.86×10^{-4}
T_{max}	-0.361×10^{-4}
Ordonnée à l'origine	0.1655

TABLE 6.11 – Estimation du GLM exprimant α en fonction de E , τ , d et T_{max} (cas du mois d'août)

Les diagnostics des trois modèles sont présentés dans les figures 6.19, 6.20, 6.21 respectivement pour les paramètres T_{max} , τ et α . Ces trois figures s'interprètent de la manière suivante :

- Sur les parties gauches des figures sont représentés les résidus en fonction des valeurs prédites de Y . Si le modèle est parfait, l'espérance des résidus doit être égale zéro c'est à dire que les résidus doivent être distribués symétriquement autour de l'axe horizontal passant par zéro. La courbe rouge qui indique la ligne d'interpolation de l'ensemble des points doit être horizontale.
- À droite, est représentée la racine carrée des résidus divisés par leur écart type en fonction des valeurs prédites. Pour un modèle parfait, cette quantité doit être constante pour toutes les valeurs prédites. Les points doivent être alors symétriquement distribués et la courbe rouge horizontale.

Pour l'ensemble des trois figures, les résidus sont distribués quasi symétriquement autour de zéro. Même si les courbes rouges ne sont pas horizontales comme pour un modèle parfait, nous pouvons voir qu'elles s'en rapprochent. Il faut noter que cette courbe est très influencée par les quelques points, numérotés dans la figure, qui s'écartent relativement de l'horizontale passant par zéro. L'analyse des graphes en partie droite des figures montre également que la variance des résidus est quasiment constante en fonction des valeurs prédites pour l'ensemble des trois modèles. Les déviations qu'on observe, notamment dans la figure 6.20 à la fin de la courbe et dans la figure 6.21 au début de la courbe, sont également dues aux points numérotés. Cependant, pour les trois modèles, les effectifs de ces points sont trop faibles pour avoir un réel effet sur la précision des modèles. Dans les figures 6.22 et 6.23 nous montrons la comparaison avec les données de référence des nuages de points entre d et T_{max} simulés et entre α et τ simulés. On observe que les nuages de points simulés sont semblables à ceux de référence.

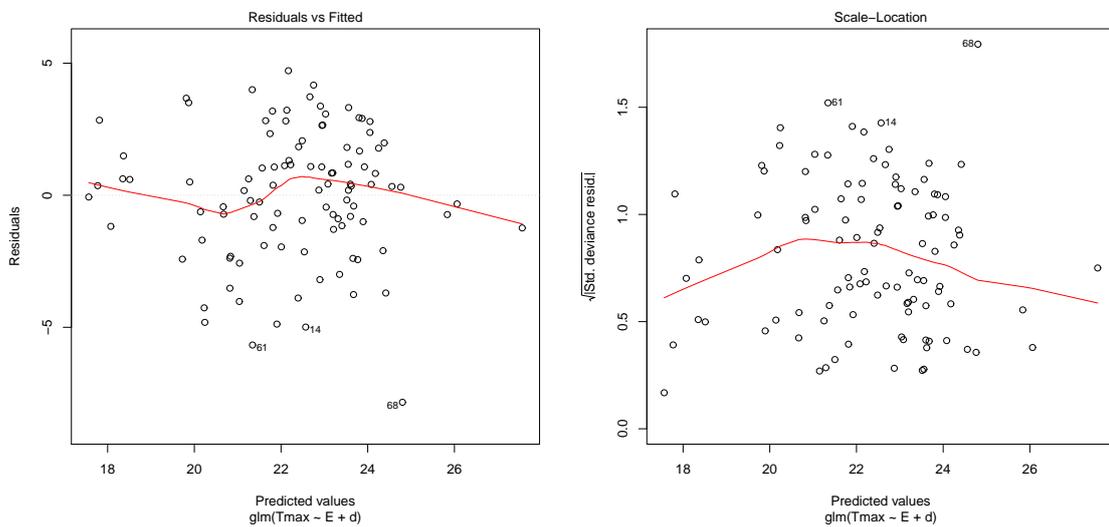


FIGURE 6.19 – Diagnostic du GLM exprimant T_{max} en fonction de E et d

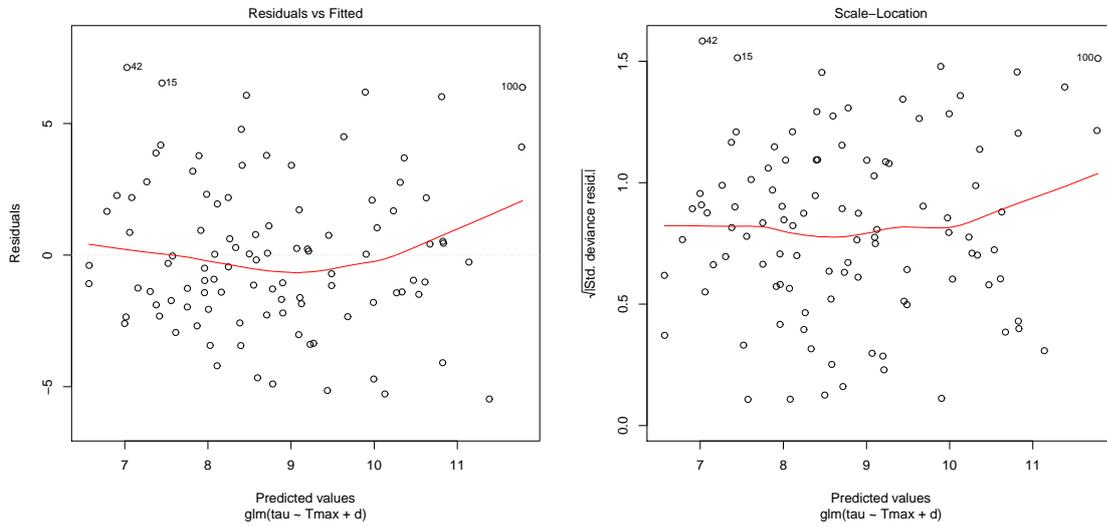


FIGURE 6.20 – Diagnostic du GLM exprimant τ en fonction de d et T_{max}

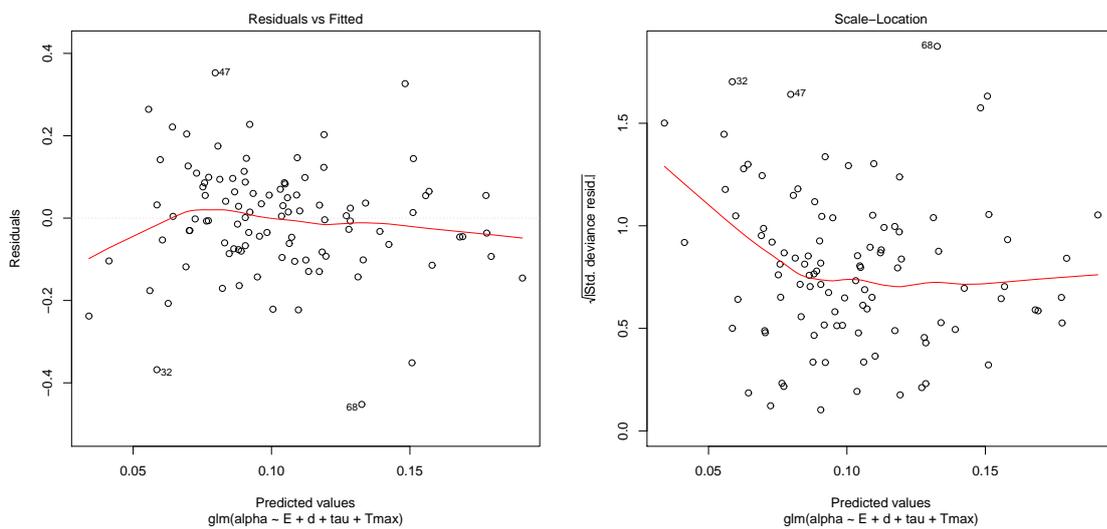


FIGURE 6.21 – Diagnostic du GLM exprimant α en fonction de E , τ , d et T_{max}

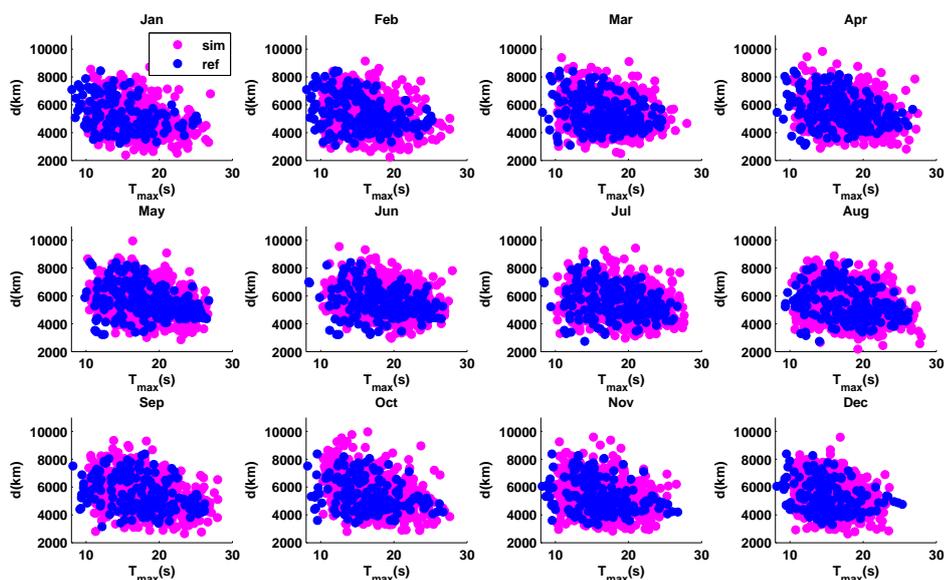


FIGURE 6.22 – Nuage de points entre d et T_{max} simulés en comparaison avec les données de référence.

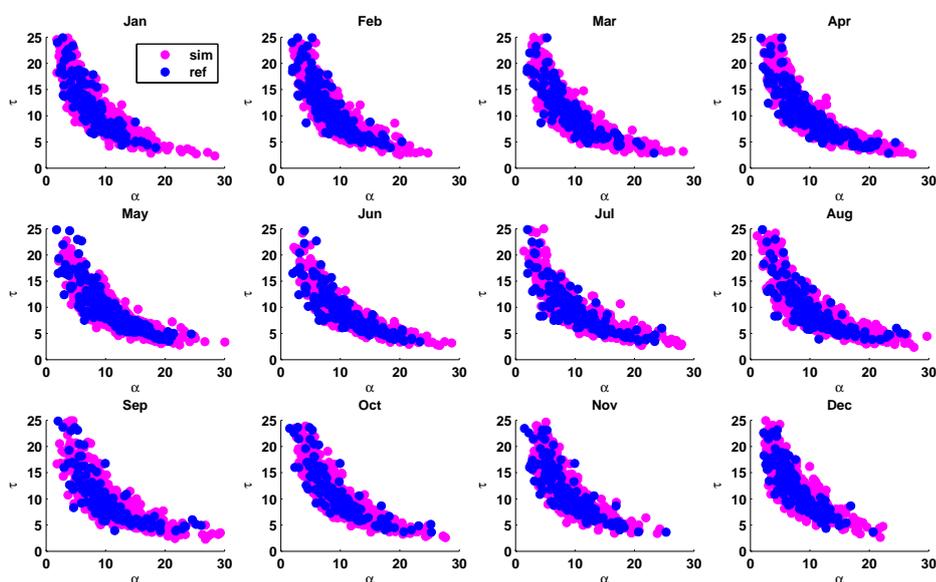


FIGURE 6.23 – Nuage de points de α et τ simulés en comparaison avec les données de référence.

6.2.3.3 Critères physiques de sélection des paramètres simulés

À partir des distributions jointes obtenues par utilisation des copules et des distributions conditionnelles utilisant les GLM, nous pouvons simuler un grand nombre de septuplets de paramètres $\{E, \tau, \alpha, d, T_{max}, \omega, \theta_0\}$ et donc reconstruire un grand nombre d'événements de houles. Les modèles de lois marginales et la copule gaussienne utilisés permettent de simuler des valeurs des paramètres autres que celles qui ont été observés. Ils peuvent cependant conduire à la simulation de paramètres ayant certes les mêmes propriétés statistiques que les observations, mais ne reflétant aucune réalité physique. Par exemple, les limites géographiques imposées par la zone de génération des houles, impliquent également une limite quant aux distances parcourues par les houles. Pour le cas du site Girassol, il est impossible que des houles du Sud-Ouest soient générées hors de la zone délimitée par les deux traits bleus et les côtes dans la figure 6.24. De même l'angle sous lequel on voit arriver les premiers trains de houle (θ_0) ne peut être compris que dans un intervalle angulaire défini par les limites de la zone géographique. En outre, la période maximale (T_{max}) des houles ne peut être physiquement plus grande qu'un seuil qui dépend des caractéristiques de la tempête génératrice.

Au regard de ces considérations physiques, il nous a paru raisonnable d'éliminer les septuplets de paramètres simulés qui ne reflètent aucune réalité physique. Ainsi, parmi les septuplets de paramètres simulés, sont éliminés ceux pour lesquels les critères suivants sur d , θ_0 et T_{max} ne sont pas respectés :

- La distance d doit être inférieure à un seuil fixé à $9000km$. Cette distance est sensiblement égale à la distance maximale qu'il est possible de parcourir par les houles (figure 6.24).
- La direction de provenance des premiers trains de houles θ_0 doit être comprise dans l'intervalle $[180^\circ, 270^\circ]$ délimité par les deux traits bleus dans la figure 6.24.
- La période maximale des houles T_{max} doit être inférieure ou égale à $28s$. Cette limite nous a paru raisonnable car même les plus longues houles observées dans le pacifique atteignent rarement cette période (Gjevik et al. [1988], Ardhuin et al. [2009]).

Même s'il y a très peu de septuplets de paramètres ne respectant pas ces conditions, les mettre en place garantit que les événements de houle simulés reflètent la réalité physique et la climatologie qui s'en déduit est donc plus réaliste.

À cette étape, nous sommes en mesure de simuler des événements de houles individuels statistiquement équivalents aux événements de référence. Pour reconstruire des séries temporelles d'événements de houle, il faudra également modéliser la manière dont surviennent les événements. L'étape suivante, dans notre démarche de générateur stochastique d'événements de houle, est donc la modélisation de l'occurrence de ces événements de houles.

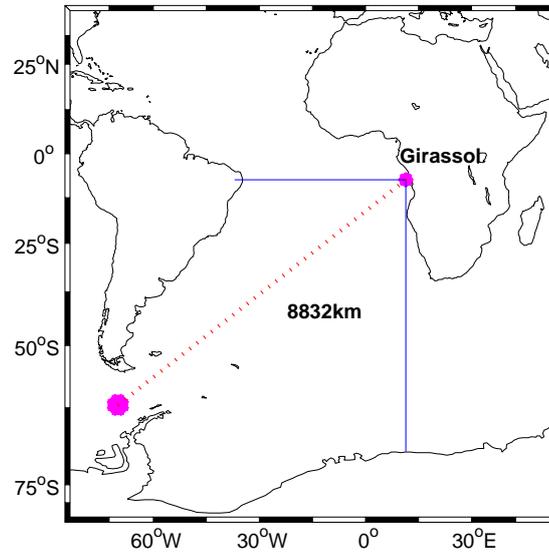


FIGURE 6.24 – Zone de génération possible des houles du Sud-Ouest observées sur le site Girassol délimitée par les deux lignes bleues et les abords des continents

6.3 Modélisation de l'occurrence des événements de houles

La deuxième étape dans la conception de notre générateur stochastique d'événements de houle consiste à modéliser l'occurrence des événements afin de reconstruire des séries temporelles réalistes d'événements de houle. En effet, de nombreuses applications comme l'estimation de l'évolution morphodynamique du trait de côte ou encore l'estimation de la probabilité d'occurrence des fenêtres météorologiques pour les opérations en mer, nécessitent de modéliser la manière dont s'enchaînent les événements. Par exemple, des événements de houle de forte intensité associées à des temps d'inter-arrivée faibles augmenteraient sans doute les risques d'érosion. Par ailleurs, une longue durée entre deux événements successifs favoriserait les opérations marines (installation des plates-formes offshore, des convertisseurs d'énergie de vagues, maintenance des systèmes en mer, etc.).

Pour modéliser l'occurrence des événements, il nous faut, dans un premier temps, définir un temps d'inter-arrivée entre les événements, puis modéliser la distribution empirique des données du temps d'inter-arrivée.

6.3.1 Temps d'inter-arrivée des événements de houle

Nous définissons le temps d'inter-arrivée noté I comme la durée entre les dates de début de deux événements successifs. Cette définition du temps d'inter-arrivée est illustrée dans la figure 6.25 (à gauche). Dans la même figure à droite est représentée la distribution mensuelle des données de I . Contrairement à certains paramètres du SEM notamment l'énergie des événements (E), il n'y a pas d'effet de saisonnalité notable dans les distributions de I ; la distribution de I est à peu près semblable pour tous les mois. Même si les

événements de houle les plus intenses surviennent pendant l'été, la fréquence d'occurrence des événements quant à elle est relativement constante au long de l'année.

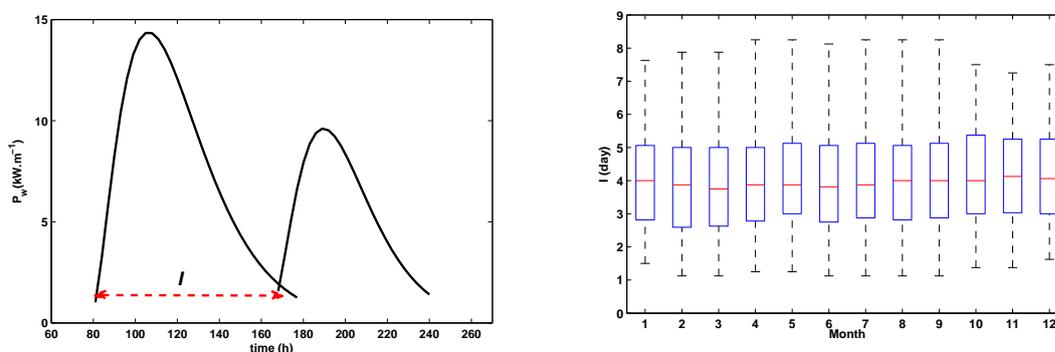


FIGURE 6.25 – Définition du temps d'inter-arrivée entre deux événements de houle successifs (à gauche); distribution mensuelle des données du temps d'inter-arrivée (droite)

6.3.2 Modélisation de la distribution du temps d'inter-arrivée des événements de houle

Généralement la fréquence d'événements aléatoires indépendants est modélisée par une loi de Poisson. Cela implique que le temps d'inter-arrivée entre événements successifs suit une loi exponentielle (Ross, others [1996]). Or, les distributions marginales de I obtenues à partir des données sont différentes de celle d'une loi exponentielle. La figure 6.26 (à gauche) montre l'histogramme de la distribution empirique de I pour le mois d'août. On peut y voir que cette distribution n'est pas celle d'une loi exponentielle. D'autres types de lois qui permettent de modéliser de façon plus générale la distribution du temps d'inter-arrivée, notamment la loi de Weibull (McShane et al. [2008]) et la loi Gamma (Sim [1992], Miller, Bhat [1997]) sont souvent utilisées. Entre la loi de Weibull (figure 6.26, à droite) et la loi Gamma (figure 6.26, au centre), la dernière modélise mieux les données de I et a donc été retenue pour la modélisation des temps d'inter-arrivée. La loi Gamma est très flexible dans la modélisation du temps d'inter-arrivée et il faut remarquer que la loi exponentielle est un cas particulier de la loi Gamma. Remarquons cependant que le modèle Gamma, pour ce mois, surestime légèrement les valeurs de I qui dépassent 7 jours, ce qui conduirait dans la simulation à des temps d'inter-arrivée légèrement plus longs. mais, cela ne risque pas d'affecter de manière conséquente la simulation parce que l'effectif de I pour lequel $I > 7\text{jours}$ est très faible.

Le temps d'inter-arrivée est corrélé avec la durée de l'événement qui le précède. En effet, tant qu'une tempête reste active dans la zone de génération des houles, il y a moins de chance d'en voir apparaître une autre dans la même zone. Ainsi, au point d'observation des houles, si un événement de forte intensité dure dans le temps, la probabilité de détecter un nouvel événement dans la même direction est faible, conduisant dans ce cas à de longs temps d'inter-arrivée. Or, la durée de l'événement est elle-même corrélée à ses paramètres caractéristiques notamment à son énergie (E) et à la distance parcourue par les houles (d). Par exemple, la durée de l'événement de houle est d'autant plus longue que cet événement a été généré loin du point d'observation. En conséquence, il est prévisible que I soit aussi

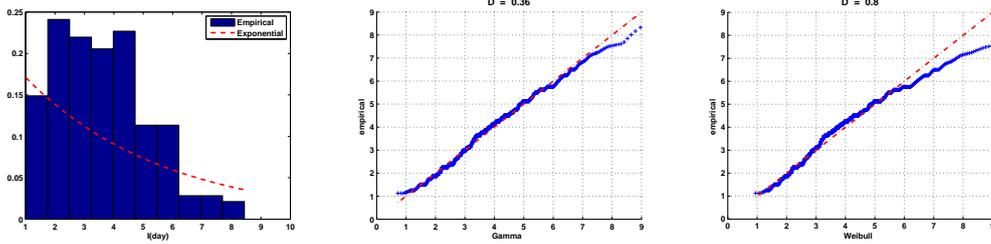


FIGURE 6.26 – Histogramme de I (à gauche) ; Diagramme quantile-quantile entre distribution empirique de I et les lois Gamma (au centre) et de Weibull (à droite) (cas du mois d'août)

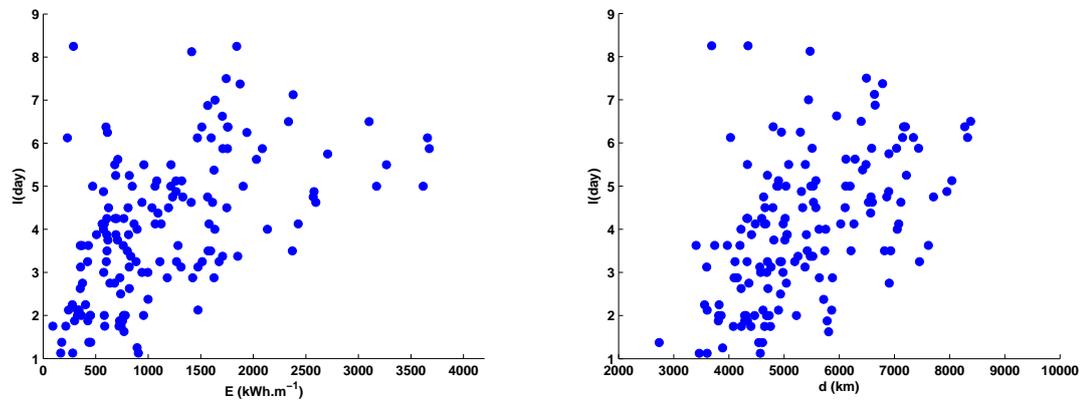


FIGURE 6.27 – Nuage de points entre I et E (à gauche) et entre I et d (à droite)(cas du mois d'août)

corrélé à E et à d comme le montrent les nuages de points de la figure 6.27. Si nous voulons simuler des temps d'inter-arrivée représentatifs des observations, il est nécessaire de prendre en compte cette corrélation entre I et les paramètres de l'événement qui le précède.

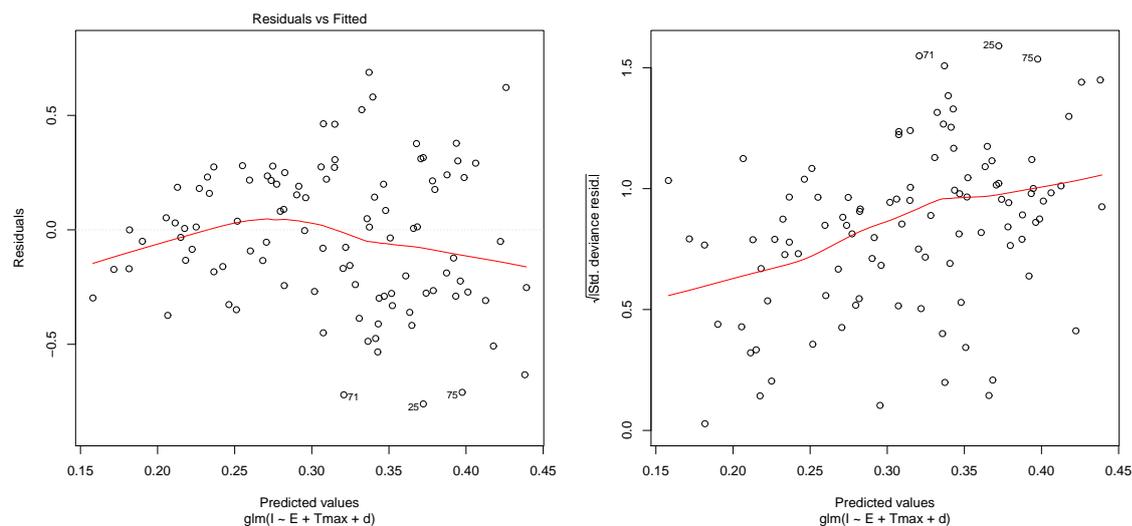
En tenant compte du fait que la loi marginale de I est bien modélisée par une loi Gamma et que I est corrélé aussi bien à E et d qu'avec T_{max} , il nous a paru raisonnable de simuler I conditionnellement à ces paramètres en utilisant un GLM (Section 6.2.3). Ce GLM utilise la loi Gamma comme distribution de I et une fonction lien canonique inverse :

$$g(\mathbb{E}(I)) = \beta_0 + \beta_1 E + \beta_2 d + \beta_3 T_{max} \quad (6.27)$$

où $\mathbb{E}(I)$ est l'espérance de I et g la fonction lien inverse. Les valeurs des coefficients β sont renseignées dans le tableau 6.12.

Le diagnostic du modèle est présenté dans la figure 6.28 qui s'interprète comme indiqué plus haut dans la section 6.2.3. Sur la partie gauche sont représentés les résidus en fonction des valeurs prédites de I . Les résidus sont quasiment distribués de manière symétrique autour de zéro. Sur la partie droite est représentée la racine carrée des résidus normalisés. La tendance n'est pas horizontale comme elle le serait si le modèle était parfait. On remarque que la variance croît légèrement avec les valeurs prédites, ce qui se traduirait

Variables explicatives	Estimation (β_i)
E	-1.44×10^{-6}
d	-0.51×10^{-6}
T_{max}	-1.57×10^{-2}
Ordonnée à l'origine	1.1

 TABLE 6.12 – Coefficients du GLM exprimant I en fonction de E , τ , d , T_{max} .

 FIGURE 6.28 – Diagnostic du GLM exprimant I en fonction de E , τ , d , T_{max} .

dans la simulation par une variabilité importante autour de la moyenne des grandes valeurs de I . Nous montrons comme résultat de simulation I dans la figure 6.29 la comparaison entre la distribution jointe du couple $\{d, I\}$ simulé et celle de référence. On note dans l'ensemble une bonne ressemblance même si on surestime par la loi Gamma des valeurs de $I > 7$ jours comme évoqué plus haut, et s'il existe une variabilité importante de la variance des résidus pour ces points. Toutefois, cela n'a pas d'influence significative sur la reconstruction de la climatologie des houles car très peu de points sont finalement concernés.

Avec la simulation des événements de houle individuels et une modélisation adaptée de l'occurrence des événements, nous disposons des deux éléments qui constituent notre générateur d'événements de houles. Il est maintenant possible de reconstruire des séries temporelles d'événements de houle sur des durées plus longues que les 21 années de données de hindcast. Même si nous avons statistiquement validé la simulation des événements de houles individuels et le modèle d'occurrence des événements, il est important de valider les séries temporelles d'événements reconstruites afin de s'assurer qu'elles peuvent être utilisées avec confiance pour des applications en génie océanique et côtier.

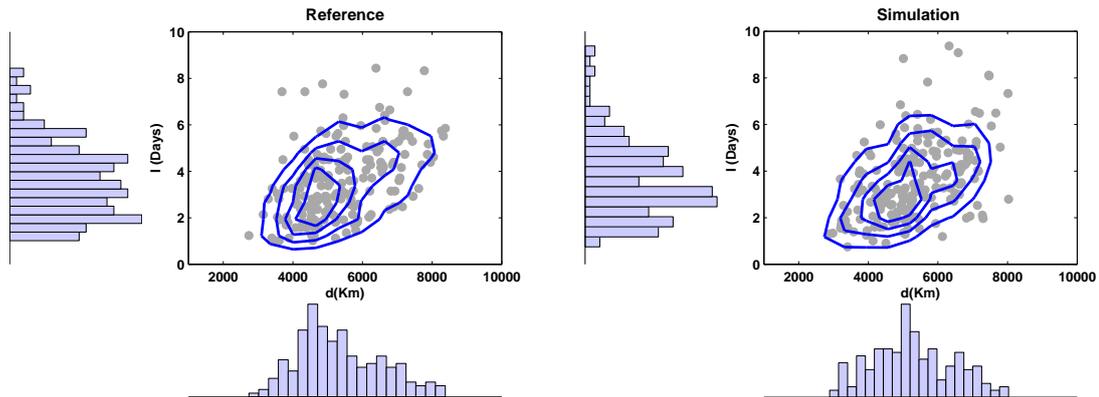


FIGURE 6.29 – Distributions jointes de $\{d, I\}$: référence (à gauche) et simulée (à droite) (Cas du mois d'août). Les contours sont régulièrement espacés de 0.01 entre 0.01 et 0.04

6.4 Reconstruction et validation des séries temporelles d'événements de houle

6.4.1 Reconstruction de séries temporelles d'événements de houle

On se propose de reconstruire des séries temporelles d'événements de houle sur de grandes périodes. On dispose pour cela des données simulées de septuplets de paramètres du SEM à partir desquelles on peut reconstruire individuellement un grand nombre d'événements de houle. On dispose également des données simulées du temps d'inter-arrivée entre les événements de houles associés à chaque septuplet de paramètres du SEM. Avec ces données du temps d'inter-arrivée, il est possible de faire suivre des événements au cours de chaque mois, pour reconstruire les séries temporelles d'événements de houles. Les événements se suivent en faisant l'hypothèse qu'au cours de chaque mois, ils surviennent de manière indépendante. Il nous a paru raisonnable de faire cette hypothèse car il n'y a pas de raison physique conduisant à penser qu'un événement de houle et son suivant soient corrélés alors même qu'ils ont été générés par deux phénomènes météorologiques différents. L'hypothèse a été vérifiée en étudiant la dépendance entre les paramètres E et d d'un événement de houle et son suivant. La figure 6.30 montre les nuages de points entre les paramètres E et d d'un événement et de son suivant. Il n'y a pas de corrélation notable et les Tau de Kendall calculés ($\tau_k = 0.009$ pour E et $\tau_k = 0.04$ pour d) peuvent significativement être considérés comme nuls, ce qui confirme que notre hypothèse de reconstruction est raisonnable.

À titre d'exemple, la figure 6.31 (à gauche) montre l'exemple des puissances simulées pour une série temporelle d'événements de houle d'une année et à droite l'exemple du mois d'août où l'on présente l'évolution de tous les paramètres. On peut remarquer que la saisonnalité est reproduite dans les séries temporelles d'événements de houle simulés. La prochaine étape de notre démarche consiste à valider la reconstruction des séries temporelles de houles simulées.

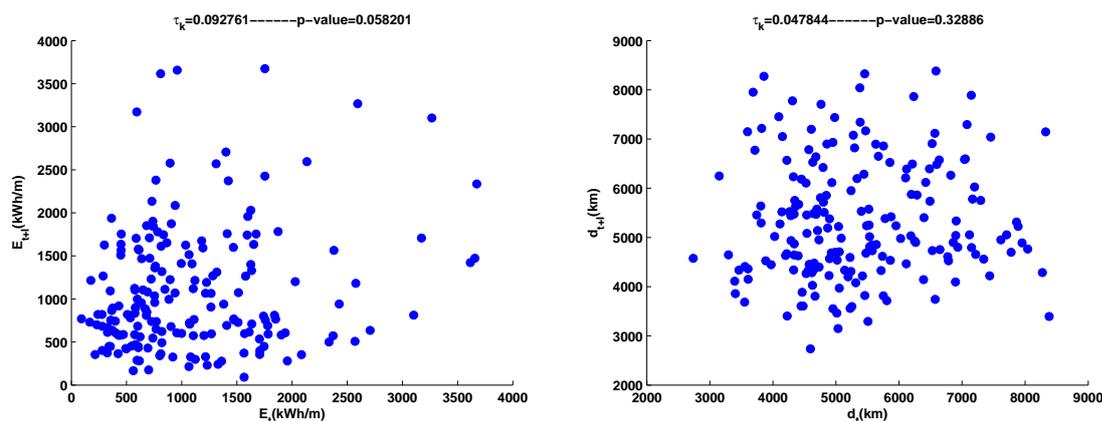


FIGURE 6.30 – Nuage de points entre les paramètres E et d d'un événement de houle et son suivant.

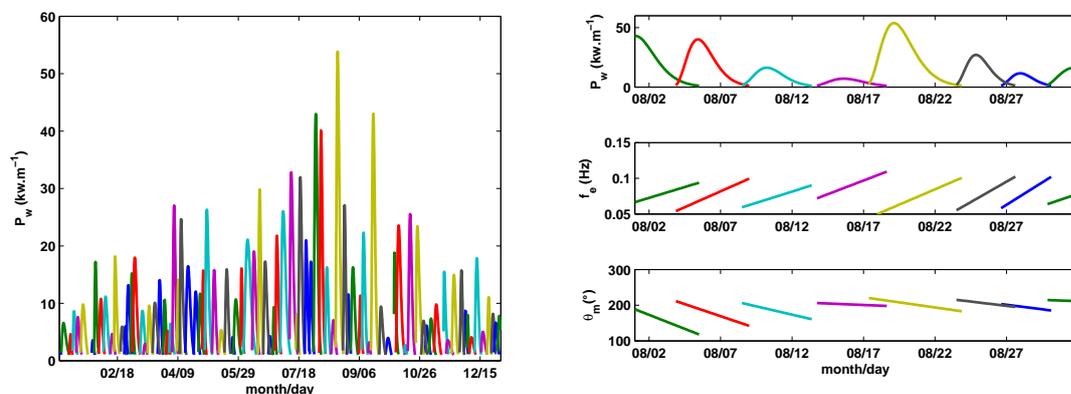


FIGURE 6.31 – Exemple de simulation d'une année d'événements de houle caractérisés par leur puissance (à gauche) ; Zoom sur le mois d'août pour les paramètres P_w , f_e et θ_m (à droite)

6.4.2 Validation des séries temporelles d'événements de houle simulés

Afin de pouvoir utiliser les séries temporelles d'événements de houle simulés, il est indispensable de vérifier qu'elles sont statistiquement équivalentes aux données de référence. Nous avons choisi de valider dans cette étude les données simulées sur deux caractéristiques statistiques :

- la distribution des hauteurs significatives globales,
- la durée de persistance des états de mer.

6.4.2.1 Hauteur significative globale

Le paramètre le plus important permettant de caractériser l'intensité des états de mer est la hauteur significative. Il nous a donc paru nécessaire de valider les séries temporelles d'événements de houles reconstruites sur la base de ce paramètre, d'autant plus que la modélisation ne s'est pas portée directement sur celui-ci.

On se propose donc de comparer la distribution mensuelle totale des H_s obtenus par simulation avec la distribution des H_s issus des données de référence. Parce que les données de hindcast servant de référence couvrent une durée de 21 ans, nous avons également simulé pour chaque mois, 21 séries temporelles d'événements de houle. Les hauteurs significatives se déduisent de la puissance et de la période d'énergie en utilisant l'équation (5.14). Il est possible de construire un intervalle de fluctuation autour de l'évolution des paramètres P_w , f_e et θ_m simulées à partir des intervalles de confiance obtenus avec les lois marginales ajustées. D'où la possibilité de construire également un intervalle de fluctuation autour du H_s indiquant la variabilité du générateur (voir figure 6.32). Le H_s total à un instant t est :

$$H_{st} = \sqrt{\sum_1^n H_{s_i}^2} \quad (6.28)$$

avec n le nombre de systèmes présents à l'instant t .

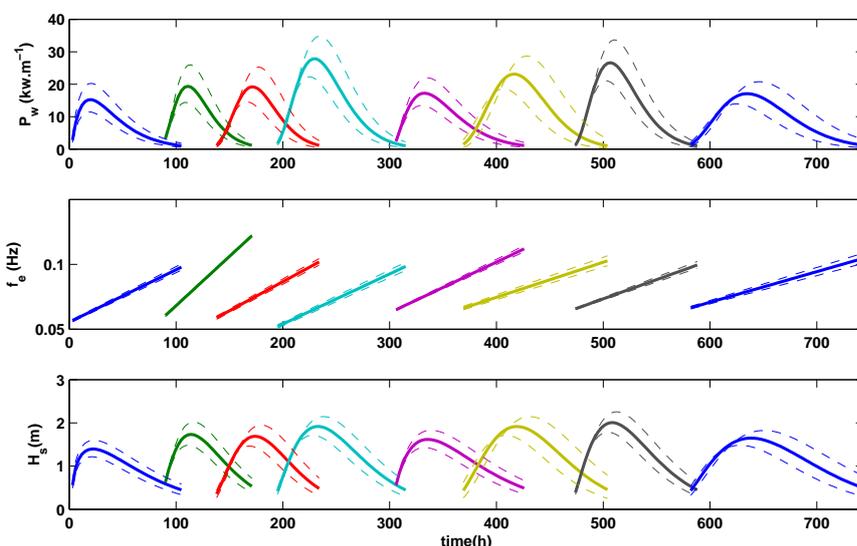


FIGURE 6.32 – Exemple d'un mois de séries temporelles d'événements de houle simulés. La Puissance, la période d'énergie et la hauteur significative déduite sont représentées avec les intervalles de fluctuation à 95% (traits en pointillés).

La figure 6.33 présente les diagrammes quantile-quantile comparant la distribution du H_s total simulé et du H_s total de référence. Les quantiles des H_s simulés correspondent dans l'ensemble aux quantiles des H_s de référence. On remarque néanmoins que le générateur ne reproduit pas bien les valeurs extrêmes du H_s . Excepté pour le mois de Juillet, les H_s extrêmes sont sous-estimés par le générateur. Les modèles de lois marginales des paramètres du SEM que nous avons utilisés dans le générateur ne sont pas adaptés aux valeurs extrêmes. C'est ce qui explique le fait que le générateur ne reproduit pas correctement les valeurs extrêmes de H_s . Toutefois notre objectif est de représenter au mieux l'ensemble des états de mer car de nombreuses applications, telles que la fatigue ou encore l'estimation du potentiel d'énergie récupérable des vagues, nécessitent que l'ensemble

des états de mer soit bien modélisé. Pour des applications qui sont uniquement sensibles aux états de mer extrêmes notamment le dimensionnement au chargement extrême, il est possible d'utiliser des lois marginales adaptées aux valeurs extrêmes en remplacement de celles qui sont introduites dans le générateur pour la présente étude.

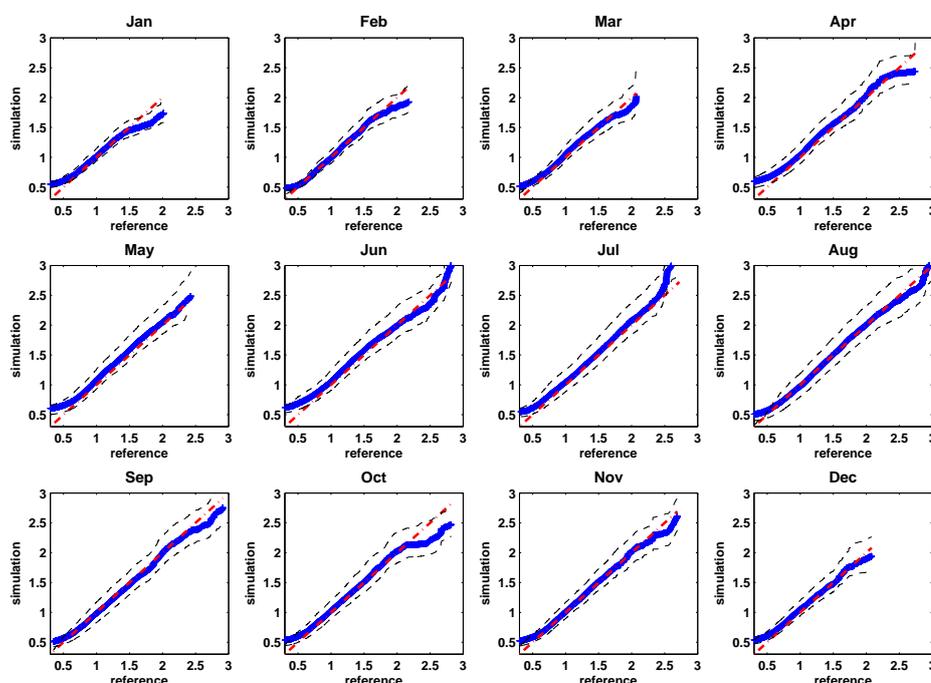


FIGURE 6.33 – Diagramme quantile-quantile de la distribution du H_s total simulé et du H_s total de référence.

6.4.2.2 Durée de persistance

Parmi les caractéristiques statistiques que l'on souhaite restaurer dans les simulations, on accorde une importance particulière aux durées de persistance qui sont déterminantes dans de nombreuses applications en génie océanique et côtier. Nous pouvons citer par exemple l'estimation de l'évolution morphodynamique du trait de côte où les durées de persistance des états de mer déterminent le transport de sédiments. Par ailleurs, l'installation d'une plateforme offshore, d'un convertisseur d'énergie de vagues ou encore la maintenance d'un système en mer ne pouvant se faire que sous certaines conditions favorables d'états de mer (temps calmes), il est nécessaire d'estimer la probabilité des fenêtres météorologiques (weather windows) favorables à ces opérations en mer à diverses périodes de l'année.

On considère une opération en mer au large de l'Afrique de l'Ouest (installation de plate-forme offshore, de convertisseur d'énergie de vagues, maintenance de système en mer, etc.), qui ne peut être effective que si la hauteur significative totale des houles reste en deçà d'un seuil H_{ac} donné pendant une durée τ_{ac} donnée. On cherche à déterminer de

manière statistique pour un mois de l'année, le nombre de jours favorables à une telle opération ou le nombre de jours d'accès au site (N_{ac}). Notre but est de comparer les durées de persistance ainsi que le nombre de jours d'accès au site calculés à partir des séries temporelles d'événements de houle simulés et de les comparer à ceux calculés à partir des données de référence (hindcast).

Calcul des durées de persistance et du nombre de jours d'accès au site

Les durées de persistance, les probabilités d'occurrence des fenêtres météorologiques ainsi que les nombres de jours d'accès au site ont été estimés à partir de la méthode décrite dans Walker et al. [2013] et utilisée dans Maisondieu et al. [2014] dont on rappelle les grandes étapes.

La première étape consiste à estimer la probabilité de dépassement d'un seuil donné H_{ac} de la hauteur significative en ajustant aux données empiriques une loi de Weibull à 3 paramètres (voir figure 6.34),

$$P(H > H_{ac}) = \exp\left(-\left(\frac{H_{ac} - x_0}{b}\right)^k\right) \quad (6.29)$$

où les paramètres x_0 , b et k sont estimés à partir des données.

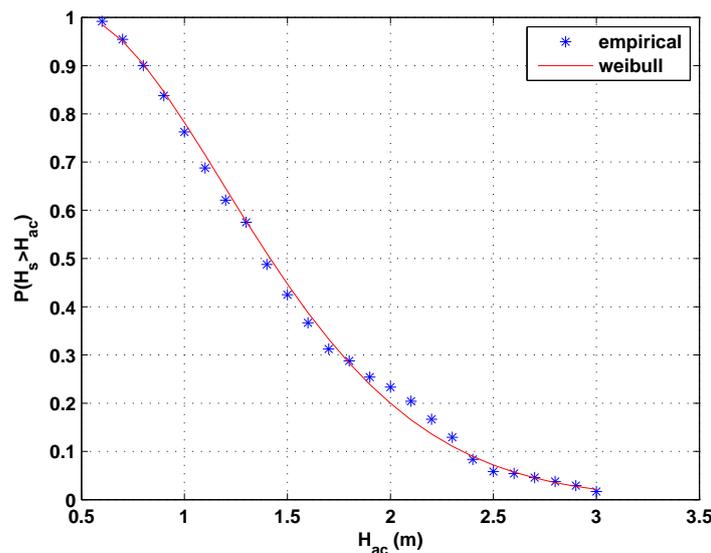


FIGURE 6.34 – Exemple d'ajustement de la loi de Weibull à 3 paramètres aux données empiriques de probabilités de dépassement de seuil

La durée moyenne pendant laquelle la hauteur significative reste continuellement en deçà du seuil H_{ac} est :

$$\tau_{ac} = \frac{1 - P(H > H_{ac})}{P(H > H_{ac})} \frac{A}{[-\ln(P(H > H_{ac}))]^\beta} \quad (6.30)$$

avec

$$A = \frac{35}{\sqrt{\Gamma}}$$

$$\beta = 0.6\Gamma^{0.287},$$

$$\Gamma = k + \frac{1.8x_0}{\bar{H} - x_0},$$

$$\bar{H} = b\Gamma \left(1 + \frac{1}{k}\right) + x_0.$$

La probabilité pour que persistent des conditions d'états de mer favorables pendant une durée normalisée X_i est donné par :

$$P(X_i > X_{ac}) = \exp(-C_{ac}(X_i)^{\alpha_{ac}}) \quad (6.31)$$

où

$$C_{ac} = \left[\Gamma \left(1 + \frac{1}{\alpha_{ac}}\right) \right]^{\alpha_{ac}} \quad (6.32)$$

$$\alpha_{ac} = 0.267\Gamma \left(\frac{H_{ac}}{\bar{H}}\right)^{-0.4} \quad (6.33)$$

La probabilité que la hauteur significative reste continuellement en deçà d'un seuil H_{ac} pendant une durée donnée est alors :

$$P(T > \tau_{ac}) = P(X_i > X_{ac}) \cdot (1 - P(H > H_{ac})). \quad (6.34)$$

Le nombre de jours de conditions favorables (ou d'accès) i.e. le nombre de jours où la hauteur significative reste continuellement en deçà d'un seuil H_{ac} pendant une durée donnée peut être alors obtenu par

$$N_{ac} = D \cdot P(T > \tau_{ac}) \quad (6.35)$$

où D est le nombre de jours correspondant à la période d'étude.

Résultats et discussions

On se propose de comparer les quantités suivantes, calculées à partir des données simulées avec celles calculées à partir des données de référence :

- La durée moyenne pendant laquelle la hauteur significative reste continuellement en deçà d'un seuil (équation 6.35) (durée de persistance).
- Le nombre de jours d'accès i.e le nombre de jours par mois où la hauteur significative reste continuellement en deçà d'un seuil pendant une durée donnée (équation 6.35).

Nous avons également simulé une série temporelle d'événements de houle d'une durée de 21 ans pour chaque mois de l'année. Les quantités que nous désirons comparer ont été calculées séparément pour chaque mois de l'année.

La figure 6.35 compare la durée de persistance en fonction du H_s calculée à partir des données simulées et à partir des données de référence. Nous présentons ici les résultats du mois de février et du mois d'août. Les durées de persistance calculées pour l'ensemble des 21 années sont montrées sur la partie du haut. Sur la partie du bas sont représentées les durées de persistance médianes estimées à partir des 21 années. Les traits en pointillés (couleur rose) indiquent un intervalle de fluctuation à 95%. Cet intervalle de fluctuation

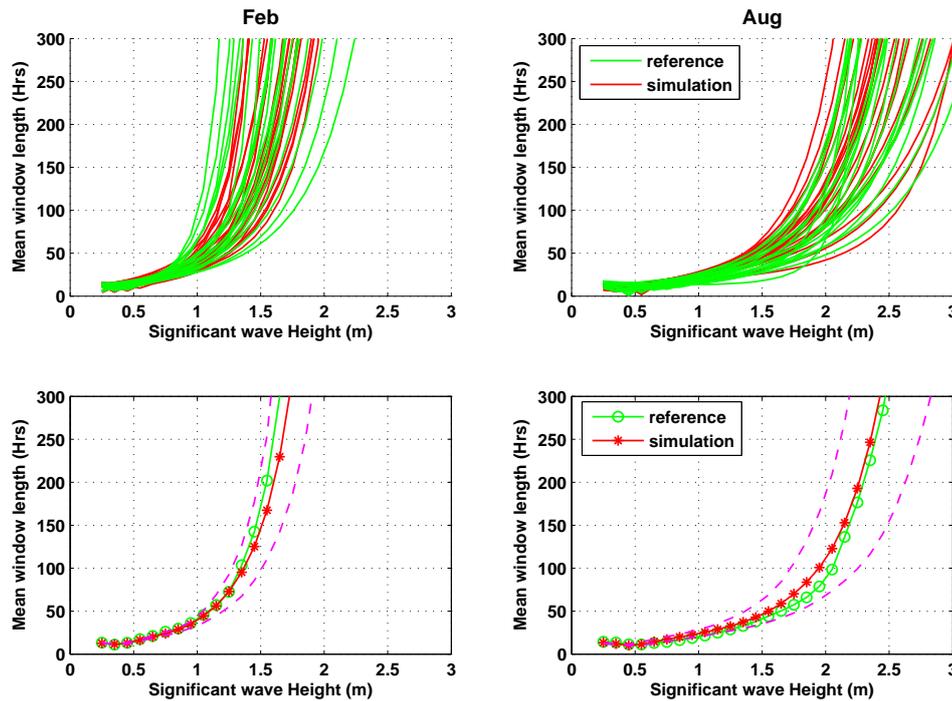


FIGURE 6.35 – Durée de persistance : pour l'ensemble des 21 années (en haut) ; valeur médiane (en bas)

représente la variabilité du modèle. On peut considérer que, si la courbe de couleur verte (représentant la durée de persistance de référence) est comprise entre les traits pointillés, la durée de persistance de référence aurait pu résulter de la simulation.

Dans l'ensemble, les durées de persistance calculées à partir des données simulées correspondent bien à celles calculées à partir des données de référence. Au mois de février, pour un H_s de $1m$ la durée de persistance médiane est de 50 heures et elle est de 200 heures pour un H_s de $1.5 m$. Pour ce même seuil de $1.5 m$, la durée de persistance est seulement de 50 heures au mois d'août qui correspond à la période où l'on enregistre les événements les plus forts en intensité. Remarquons néanmoins que l'intervalle de fluctuation est assez large pour les $H_s > 2m$, principalement dans le cas du mois d'août, ce qui pourrait conduire à un écart important entre la durée de persistance issue des simulations et celle de référence. Par exemple pour un $H_s = 2.5 m$, l'écart pourrait atteindre 150 heures. Cependant, 90% des valeurs de H_s sont inférieures à $2 m$ sur les 21 années et la méthode d'estimation de la durée de persistance utilisée n'est pas bien adaptée aux extrêmes. En effet, la loi de Weibull, ajustée aux probabilités de dépassement de seuil ne permet toujours pas de bien modéliser à la fois les valeurs extrêmes et les autres valeurs. Mais ceci ne constitue pas une limite de la méthode d'estimation des durées de persistance utilisée car les opérations d'installation ou de maintenance en mer ne sont généralement réalisées lorsque quand les conditions d'état de mer sont modérées.

La dernière statistique que nous avons étudiée est le nombre de jours d'accès par mois à un site. Cette comparaison a été réalisée pour trois cas de conditions favorables : la première définie par $H_s \leq 1 m$, la deuxième par $H_s \leq 1.5 m$ et la troisième pour un $H_s \leq 2 m$, toutes les trois pour une durée de 24 h. Les résultats correspondants aux nombres médians de jours d'accès au site pour chaque mois sont présentés dans la figure 6.36. Soulignons d'abord que le nombre de jours d'accès reflète la variabilité saisonnière de la climatologie des houles en Afrique de l'Ouest. Pendant le mois d'août, le nombre de jours favorables à une opération en mer, qui nécessite les conditions ($H_s \leq 1 m$ et $\tau_{ac} \geq 24 h$) n'est que de 5 jours contre 12 jours en janvier.

Pour les deux premiers cas ($H_s \leq 1 m$ et $\tau_{ac} \geq 24 h$) et ($H_s \leq 1.5 m$ et $\tau_{ac} \geq 24 h$), on peut remarquer que les nombres de jours d'accès sont dans l'ensemble similaires à ceux de référence. On note cependant que pour les mois d'août et de septembre correspondant au cas ($H_s \leq 1 m$ et $\tau_{ac} \geq 24 h$), la courbe de référence est très proche de la limite inférieure de l'intervalle de fluctuation ce qui signifie que le générateur tend à sur-estimer légèrement le nombre de jours d'accès. En revanche pour le mois de décembre et pour les deux cas, le générateur sous-estime le nombre de jours d'accès. Mais la courbe verte se situe globalement dans l'intervalle de fluctuation et est souvent proche de la courbe rouge signifiant que le générateur restitue bien pour ces seuils le nombre de jours d'accès au site.

Pour ce qui est du dernier cas ($H_s \leq 2 m$ et $\tau_{ac} \geq 24 h$), nous remarquons une différence plus importante que dans les deux premiers cas entre le nombre de jours d'accès obtenu par les données simulées par le générateur et le nombre de jours d'accès estimés à partir des données de références. Toutefois, l'accord reste toujours bon car il y a finalement très peu de mois pour lesquels la courbe verte est complètement en dehors de l'intervalle de fluctuation. Il est plausible que la différence plus importante pour ce cas soit due au fait que les valeurs extrêmes ne soient pas prises en compte par le générateur. En effet, un $H_s = 2 m$ correspond pour la plupart des mois, à un quantile supérieur à 95%. Un autre facteur qui contribue à cette différence est la loi de Weibull qui, pour les valeurs extrêmes de H_s , n'ajuste pas bien les probabilités de dépassement de seuil. La tendance générale pour ce dernier cas est que le nombre de jours d'accès est sous-estimé, ce qui vu sous un autre angle, peut être un avantage. En effet, pour une opération en mer nécessitant des ressources humaines et financières importantes, il est préférable de sous-estimer le nombre de jours d'accès et d'être sûr que l'opération se déroulera dans les conditions favorables que de les sur-estimer et d'être surpris pendant l'opération par des conditions d'états de mer indésirables.

En définitive, pour les conditions modérées au regard de la climatologie des houles sur le site ($H_s \leq 1 m$ et $\tau_{ac} \geq 24 h$) et ($H_s \leq 1.5 m$ et $\tau_{ac} \geq 24 h$) et même pour ($H_s \leq 2 m$ et $\tau_{ac} \geq 24 h$), le générateur stochastique d'événements de houle reproduit globalement bien les durées de persistance et les nombres de jours d'accès au site pour des opérations en mer.

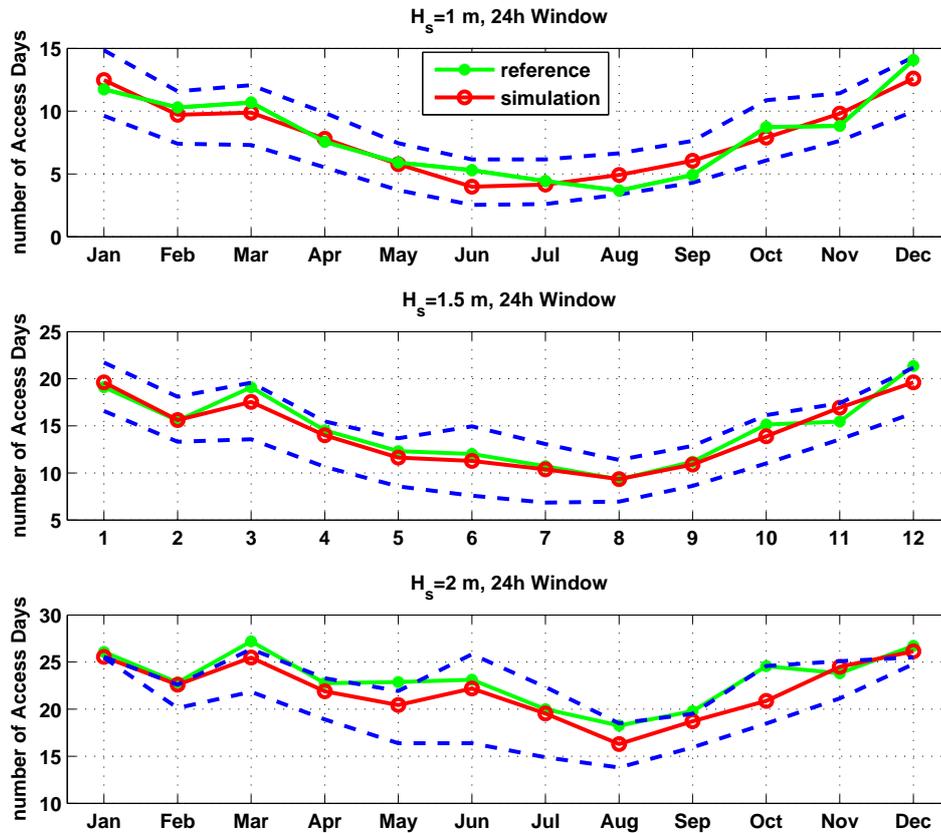


FIGURE 6.36 – Nombre de jours par mois d'accès au site

6.5 Conclusion

Dans ce chapitre nous avons présenté un générateur stochastique d'événements de houle qui permet de simuler des séries temporelles d'événements de houles sur des durées plus longues que celles pour lesquels les données sont disponibles. À cause de la limite en durée imposée par les données d'observations ou de hindcast, la modélisation stochastique s'offre comme une solution privilégiée pour simuler des données d'états de mer, de caractéristiques statistiquement équivalentes aux données observées, sur des durées plus longues. Le générateur stochastique d'événements de houle qui a été développé dans le cadre de ces travaux et présenté dans ce chapitre, intègre la variabilité saisonnière des houles ce qui permet de simuler des séries d'événements de houles réalistes pour les applications auxquelles elles seront dédiées. La mise en œuvre du générateur a nécessité deux étapes. La première étape a consisté en la modélisation des distributions jointes des paramètres du SEM permettant de simuler des événements de houles individuels. Dans la deuxième étape de la mise en œuvre du générateur nous avons modélisé l'enchaînement des occurrences des événements de houle, permettant la reconstitution de climatologie de houles.

En plus des validations statistiques qui ont été effectuées à chaque étape de la mise

œuvre du générateur, les séries temporelles d'événements de houles reconstruites ont été validées en comparant les distributions des hauteurs significatives totales et leurs durées de persistance issues de la simulation avec celles issues des données de références. Un bon accord a été obtenu dans les deux cas : la distribution des H_s ainsi que leur durée de persistance sont bien reproduites par le générateur. Cependant, les valeurs extrêmes de H_s sont sous-estimées par le générateur du fait que les lois marginales retenues ici ne permettent pas de modéliser à la fois les valeurs extrêmes et le corps de la distribution. Une amélioration de ce générateur consisterait à utiliser par exemple des lois hybrides pour modéliser la distribution marginale des paramètres. Une loi hybride est en effet une loi formée de deux types de distributions principales ; par exemple une loi Gamma et un GEV. De telles lois, si elles sont validées sur les données par des tests statistiques, permettraient de bien modéliser l'ensemble de la distribution des paramètres y compris les extrêmes et pourraient aider à l'amélioration de la performance du générateur.

En définitive, le générateur stochastique d'événements de houle présenté dans ce chapitre peut valablement servir à simuler des données de houles en Afrique de l'Ouest, utilisables pour des applications qui sont peu sensibles aux états de mer extrêmes : estimation des fenêtres météorologiques, dimensionnement à la fatigue, estimation du potentiel d'énergie récupérable etc.

Conclusions et perspectives

L'objectif principal de ces travaux était d'analyser et de représenter la climatologie de vagues en utilisant une nouvelle approche stochastique, qui s'appuie sur une représentation par événements des données d'états de mer. En effet, les événements sont des entités qui reflètent l'évolution au cours du temps des systèmes de vagues. L'utilisation de cette approche avait pour but de simplifier la structure des données par réduction du nombre de paramètres nécessaires à la caractérisation des états de mer complexes et de reproduire la cohérence temporelle des systèmes de vagues.

La démarche adoptée dans la mise en œuvre de cette approche se décompose en trois étapes. Dans la première étape, les événements de houles ont été identifiés à partir d'une série temporelle de spectres directionnels caractéristiques des états de mer. La deuxième étape a consisté en la modélisation des événements de houles. Dans la dernière étape, on s'est attaché à développer un générateur stochastique d'événements de houle permettant la reconstitution de climatologies de houles.

7.1 Synthèse des principales conclusions

Les principaux résultats obtenus à l'issue de ces travaux sont synthétisés ci-après :

- Dans un premier temps, nous avons proposé une nouvelle méthode qui permet d'effectuer à la fois le partitionnement des spectres d'états de mer et le suivi temporel des systèmes de houle ou de mer du vent : Time-domain Expansion of Spectral Partitionning technique (TESP). Cette méthode, basée sur l'algorithme de la ligne de partage des eaux appliqué en 3D, utilise la cohérence temporelle des événements de houle et de mer du vent dans la segmentation, permettant ainsi d'extraire des événements cohérents avec la dynamique des systèmes de vagues. Par rapport aux méthodes existantes, elle garantit la continuité des événements extraits et la réduction du nombre de critères empiriques, souvent utilisés dans les méthodes de partitionnement, réduisant également considérablement le temps de calcul. Nous avons montré que la méthode n'est pas uniquement adaptée aux données de l'Afrique de l'Ouest, mais qu'elle peut aussi s'utiliser sur des spectres issus d'une autre zone géographique ayant une dynamique différente, qu'ils proviennent des observations ou des modèles numériques de vagues. Les techniques d'estimation spectrale dévelop-

pées au cours de ces travaux et présentées dans l'Annexe B ont considérablement contribué à l'amélioration de l'identification des événements, principalement à partir de données *in situ*.

- Ces travaux ont également permis de proposer un modèle d'événement de houle : Swell Event Model (SEM). Ce modèle est constitué de formes paramétriques qui permettent de modéliser les évolutions de la puissance, de la fréquence d'énergie et de la direction moyenne d'un système de houle dans un événement. Le SEM utilise la forme paramétrique Gamma pour l'évolution de la puissance et des modèles linéaires pour les évolutions de la fréquence d'énergie et de la direction moyenne. Il permet de représenter chaque événement de houle individuel par un ensemble de seulement sept paramètres offrant ainsi l'avantage de synthétiser l'information des dizaines de valeurs instantanées des puissances, fréquences et directions observées. Le modèle a été validé, d'une part, sur la distribution jointe des hauteurs significatives et des périodes d'énergie, et d'autre part, sur un cas d'application : l'estimation de la puissance extraite par un convertisseur d'énergie des vagues. Dans les deux cas, les résultats obtenus avec le modèle sont très proches de ceux obtenus avec les événements de houles observés, permettant de valider ce modèle pour son utilisation dans des applications en génie océanique et côtier en Afrique de l'Ouest.
- Un autre résultat important de ces travaux est la réalisation d'un générateur stochastique d'événements de houles. Ce générateur qui prend en compte la dépendance entre les paramètres et la variabilité saisonnière, a été construit en deux étapes. Dans la première étape, nous avons modélisé la distribution jointe des paramètres du modèle SEM par utilisation de modèles de loi marginales, de copules et de modèles linéaires généralisés, ce qui permet de simuler des événements de houle en nombre arbitraire. La seconde étape a consisté en la modélisation de l'occurrence des événements, permettant la reconstruction de séries temporelles d'événements de houles et la reconstitution de la climatologie des houles. Ce générateur a d'abord été validé à chaque étape de son développement grâce à des tests statistiques, notamment sur les ajustements des lois marginales (test de Kolmogorov-Smirnov), des copules (tests de Cramér-von Mises et de Anderson-Darling) et des GLM (analyse de résidus de Pearson). Les séries temporelles de houles reconstruites ont été ensuite validées sur les distributions des hauteurs significatives et leur durée de persistance. Le générateur peut valablement servir à simuler des données de houles en Afrique de l'Ouest, utilisables pour des applications qui sont peu sensibles aux états de mer extrêmes : estimation des fenêtres météorologiques, dimensionnement à la fatigue, estimation du potentiel d'énergie récupérable etc. Même si cela n'a pas été notre objectif, soulignons que le générateur peut être adapté aux états de mer extrêmes en utilisant des lois marginales (GEV, GPD) et des copules adaptées à ces derniers. Enfin, le générateur a été développé avec les données du site Girassol, mais il peut facilement être adapté à un autre site en Afrique de l'Ouest en réajustant les paramètres des lois marginales, des copules et des GLM ajustés.

7.2 Perspectives

Ces travaux peuvent trouver des applications ou des prolongements dans plusieurs directions. On donne ci-après quelques éléments sur ces perspectives futures.

- Il est possible d'améliorer le générateur stochastique d'événements de houle, notamment par utilisation des lois hybrides dans la modélisation des distributions marginales empiriques. Les lois hybrides sont en effet des lois formées de deux principaux types de distributions. L'idée est d'associer les lois marginales utilisées dans le générateur avec des lois de valeurs extrêmes (GEV ou GPD). De telles lois, une fois validées par des test statistiques, permettraient de bien modéliser tout le corps de la distribution des paramètres, conduisant ainsi à l'amélioration de la performance du générateur. Cependant, l'utilisation des lois hybrides fait intervenir d'autres difficultés : elles augmentent le nombre de paramètres à inférer, ce qui peut être compliqué si l'on ne dispose pas suffisamment de données et il est difficile de gérer le point d'intersection des deux lois qui composent la loi hybride (mais il existe des études qui proposent des solutions quant au choix de ce point d'intersection).
- La variabilité interannuelle pourrait être intégrée dans le générateur stochastique d'événements de houle. Cependant, cela nécessitera de disposer des données sur une longue période.
- Une autre perspective est de coupler les séries temporelles de houles simulées avec celles des mers du vent simulées (voir Annexe A) afin de disposer d'une climatologie complète pour l'Afrique de l'Ouest.
- Il serait intéressant de réaliser un outil opérationnel, permettant de simuler des séries temporelles d'états de mer à partir de paramètres calés sur des spectres directionnels d'observations ou en sortie de modèle. On pourrait apporter certaines simplifications au générateur stochastique d'événements de houle pour l'automatiser dans un cadre opérationnel. Par exemple, étant donné que les lois marginales des paramètres sont désormais connues, il est possible d'estimer les paramètres de ces lois en fonction de la moyenne et de la variance des échantillons sans systématiquement passer par des ajustements. De plus, étant donné qu'on utilise dans le générateur la même loi pour tous les mois de l'année, les paramètres de cette dernière peuvent être obtenus en fonction de la moyenne annuelle modulée par une fonction sinusoïdale représentant la variabilité saisonnière. Ces simplifications permettraient de passer à un autre site en Afrique de l'Ouest sans reprendre toutes les étapes qui ont conduit à la conception du générateur.
- Il serait également intéressant de comparer, pour des applications sensibles à l'enchaînement des systèmes de vagues (érosion côtière par exemple), les résultats obtenus avec des climatologies reconstituées par utilisation de l'approche par événements et des résultats obtenus par utilisation des méthodes classiques (représentation par série temporelle), afin de quantifier la valeur ajoutée de l'approche par événements.
- Enfin, il est possible d'étendre cette approche par événements à d'autres zones géographiques ou encore de l'appliquer à d'autres types de variables océano-météo (courant, vent, etc.).



Annexes

Sommaire

A Climatologie des mers du vent	155
A.1 Climatologie du vent sur le site Girassol	156
A.2 Corrélation entre les caractéristiques du vent et celle de la mer du vent	157
A.3 Modèles stochastiques reliant les caractéristiques du vents à celle de la mer du vent	159
A.3.1 Modèle auto-régressif de la hauteur significative conditionné par la vi- tesse du vent	159
A.3.2 Modèle empirique basé sur la corrélation entre l'énergie adimensionnelle et l'inverse de l'âge des vagues	160
A.4 Conclusion	161
B Amélioration de l'estimation spectrale par utilisation de la structure temporelle des systèmes de vagues	163
C Étude comparative de fonctions paramétriques d'évolution de la puis- sance au sein d'un événement de houle	175

Climatologie des mers du vent

Cette étude présente une méthode de génération stochastique des données de mers du vent à partir des données du vent. Elle vise à compléter les séries temporelles de houles générées par le générateur stochastique, présenté dans le chapitre 6, par des séries temporelles simulées de mers du vent.

Les mers du vent sont des systèmes de vagues qui se forment sous l'action du vent local. De fait leurs caractéristiques (H_s , T_p , θ_m , etc.) dépendent de la vitesse et de la direction du vent ainsi que de son fetch et de sa durée d'action. En Afrique de l'Ouest, les mers du vent contribuent faiblement à l'énergie totale des états de mer (moins de 20% en général, voir chapitre 4). En effet, la vitesse moyenne du vent dans cette zone est sensiblement égale 5 m/s, limitant souvent la génération de vagues de fortes amplitudes. De plus, le régime des vents est très fluctuant, notamment en amplitude, conduisant à la génération des mers du vent dont la hauteur significative est aussi fluctuante. Pour ce type de climatologie, une représentation par événements n'est pas adaptée car la durée des événements ainsi que les périodes calmes (temps entre 2 événements successifs) sont généralement trop courtes.

Dans cette étude, l'approche adoptée (voir figure A.1) consiste à générer des séries temporelles de H_s , T_p et θ_m de la mer du vent à partir des caractéristiques (U_{10} et θ_w) du vent simulé. Cette approche nécessite de disposer préalablement d'un générateur de données de vent. De tels générateurs existent et reproduisent globalement bien la climatologie du vent (Monbet et al. [2007], Ailliot et al. [2015], Bessac et al. [2015]). Ensuite, à partir du vent simulé, des séries temporelles de H_s , T_p et θ_m peuvent être générées grâce à des modèles stochastiques reliant les caractéristiques du vent à celles de la mer du vent et dont la détermination fait l'objet de cette étude.

Ne disposant pas de suffisamment de données *in situ*, les données de hindcast du site Girassol (21 ans) seront essentiellement utilisées comme données de référence dans le cadre de cette étude.

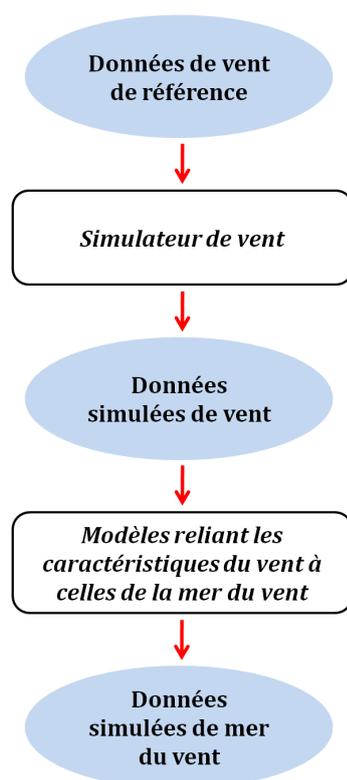


FIGURE A.1 – Approche adoptée pour la simulation des séries temporelles des caractéristiques des mers du vent

A.1 Climatologie du vent sur le site Girassol

On se propose d'analyser d'abord la climatologie du vent sur le site Girassol à partir de 21 ans de séries temporelles de l'intensité et de la direction du vent.

La figure A.2 montre l'histogramme (a) et la fonction de répartition empirique (b) des intensités du vent. On peut remarquer que les intensités du vent sont généralement faibles. Elles sont distribuées entre 0 et 10 m/s avec un mode situé autour de 5 m/s . Près de 70% des données affichent des intensités inférieures à 5 m/s et 90% des intensités du vent sont inférieures à 6 m/s . La rose des vents présentée dans le figure A.3 (a) montre que le vent est généralement des secteurs Sud et Sud-Ouest. La fonction d'autocorrélation du signal de l'intensité du vent A.3 (b) montre des corrélations relativement fortes aux pas de temps 24 h et 48 h signifiant ainsi une périodicité de 24 h du signal de l'intensité du vent. La situation générale correspond à un vent qui se lève à midi, atteint son pic à 18 h puis décroît jusqu'à minuit.

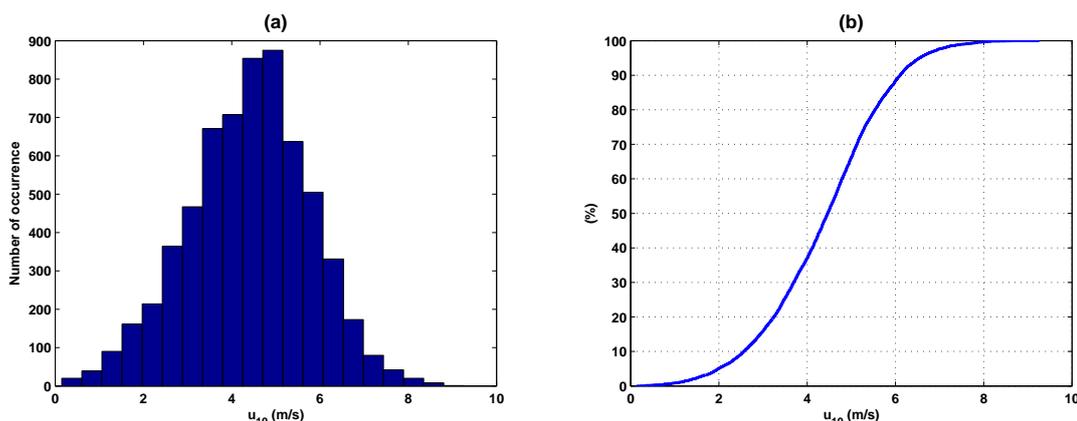


FIGURE A.2 – Histogramme (a) et la fonction de répartition empirique (b) des intensités du vent

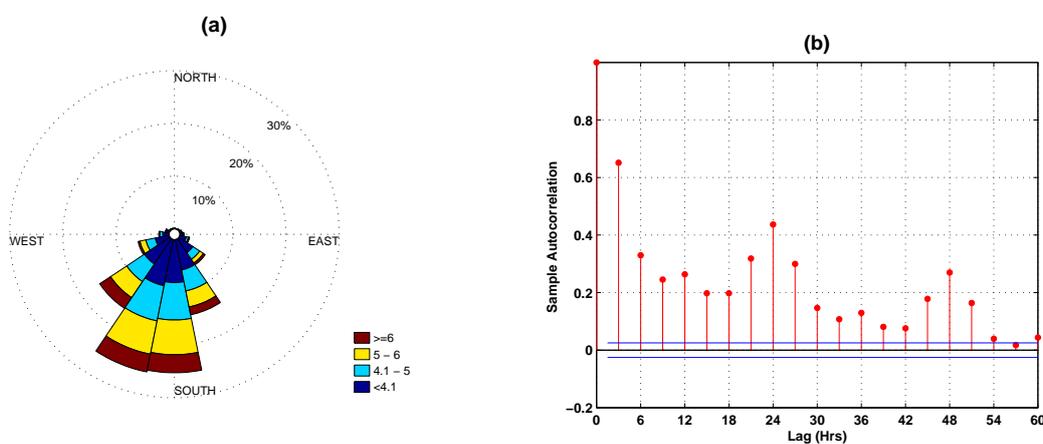


FIGURE A.3 – Rose des vents (a) et fonction d'autocorrélation de l'intensité du vent (b)

A.2 Corrélation entre les caractéristiques du vent et celle de la mer du vent

Les caractéristiques de mer du vent sont fortement dépendantes de celles du vent qui la génère. La hauteur significative et la période de pic de la mer du vent dépendent essentiellement de 3 facteurs : la distance sur laquelle le vent a soufflé (fetch), la durée d'action du vent et l'intensité du vent. Quant à la direction moyenne de la mer du vent, elle est aussi fortement dépendante de celle du vent.

La figure A.4 montre une séquence de série temporelle de la hauteur significative et de la direction moyenne de la mer du vent en comparaison avec le carré de la vitesse du vent et sa direction. Nous pouvons remarquer une corrélation entre H_s et U_{10}^2 . Cette corrélation est plus apparente en observant le nuage des points entre ces deux paramètres (figure A.5 (a)). Comme le montrent les séries temporelles, les variations rapides du U_{10}^2 et du H_s ne favorisent pas une segmentation par événements.

En ce qui concerne les directions, la différence entre la direction du vent et celle de

le mer du vent est assez faible. La figure A.5 (b) montre en effet l'histogramme de la différence entre la direction du vent et celle de la mer du vent ($\delta\theta = \theta_w - \theta_m$). On peut remarquer que la distribution de $\delta\theta$ admet un mode autour de zéro et la plupart des valeurs de $\delta\theta$ est contenue dans l'intervalle $[-45^\circ, 45^\circ]$ généralement autorisé pour la différence entre la direction du vent et celle de la mer du vent (Donelan et al. [1985]).

Notre objectif est d'utiliser ces corrélations pour déterminer des modèles reliant les caractéristiques de la mer du vent à celles du vent, à partir desquels des séries temporelles de H_s , T_p et θ_m peuvent être simulées, connaissant U_{10} et θ_w .

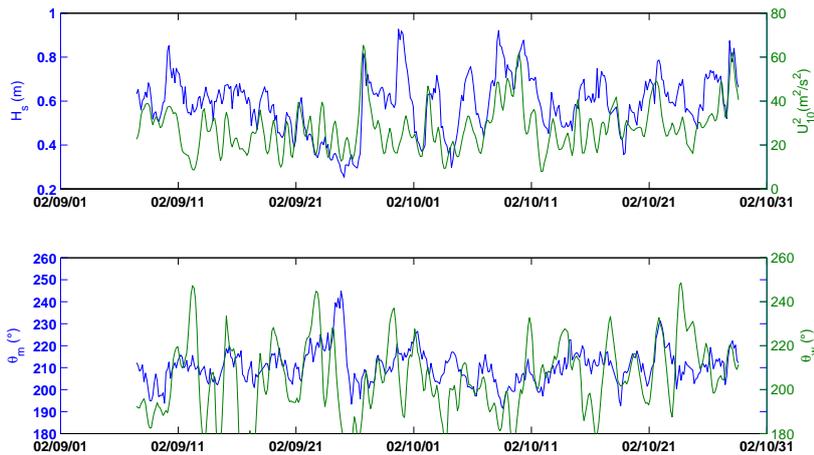


FIGURE A.4 – Série temporelle de la hauteur significative et de la direction moyenne de la mer du vent en comparaison avec le carré de la vitesse du vent et sa direction

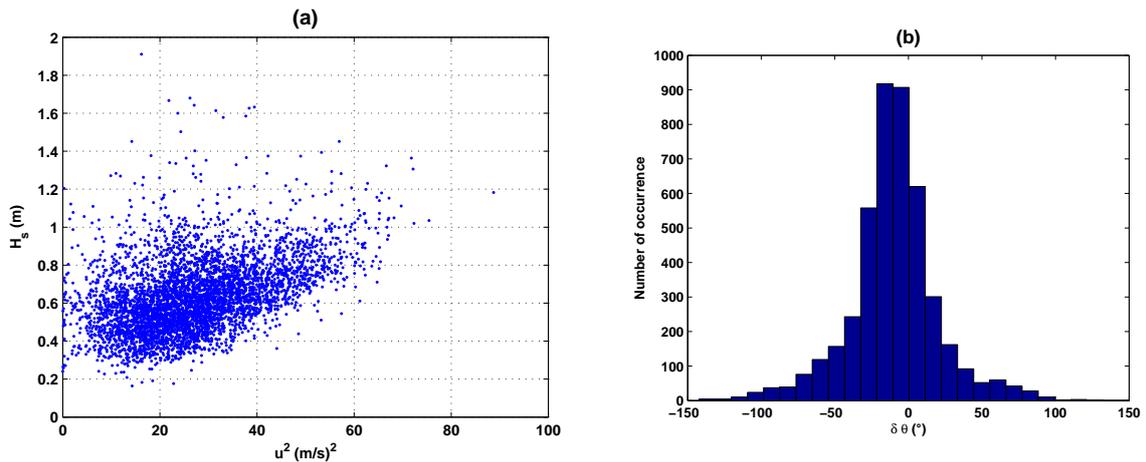


FIGURE A.5 – Nuage de points entre le H_s et le carré de la vitesse du vent (a) et histogramme des valeurs de la différence entre la direction du vent et celle de la mer du vent (b)

A.3 Modèles stochastiques reliant les caractéristiques du vents à celle de la mer du vent

Nous présentons dans cette partie, les modèles permettant de simuler des séries temporelles de H_s , T_p et θ_m à partir de l'intensité et de la direction du vent. Le premier est un modèle auto-régressif de la hauteur significative conditionné par la vitesse du vent, tandis que le deuxième est un modèle empirique basé sur la corrélation entre l'énergie adimensionnelle et l'inverse de l'âge des vagues.

A.3.1 Modèle auto-régressif de la hauteur significative conditionné par la vitesse du vent

Le premier modèle adopté dans cette étude est un modèle auto-régressif de la hauteur significative, conditionné par le carré de la vitesse du vent. En effet, comme nous l'avons souligné dans la Section A.2, il existe une corrélation entre le H_s et le carré de la vitesse du vent. Le modèle auto-régressif construit utilise comme covariable le carré de la vitesse du vent à l'instant $t - 1$ (pour prendre en compte le temps de réponse de la mer sous l'action du vent). La hauteur significative à l'instant t est donc obtenue sous forme de combinaison linéaire du H_s et du carré de la vitesse du vent à l'instant $t - 1$:

$$H_s(t) = a_0 + a_1 U_{10}^2(t-1) + a_2 H_s(t-1) \quad (\text{A.1})$$

Les 21 années de données ont été divisées en deux parties dans le but de réaliser l'apprentissage du modèle sur la première partie et d'utiliser l'autre partie pour la validation du modèle. L'apprentissage du modèle a conduit aux résultats suivants : $a_0 = 0.057$, $a_1 = 0.0145$, $a_3 = 0.873$.

Pour évaluer la pertinence du modèle, on calcule le coefficient de détermination (R^2) qui évalue à quel point une régression est adaptée pour décrire la distribution des points. Ce coefficient est défini par :

$$R^2 = 1 - \frac{\sum_{i=1}^n (H_{s_i} - \hat{H}_{s_i})^2}{(n-1)Var(H_s)} \quad (\text{A.2})$$

où n est le nombre d'échantillons, \hat{H}_s le hauteur significative prédite par le modèle et $Var(H_s)$ indique la variance des H_s . Pour le modèle A.1 nous avons obtenue $R^2 = 0.81$ signifiant ainsi que le modèle explique 81% de la variance en H_s .

La figure A.6 montre à gauche le diagramme quantile/quantile entre les H_s de référence et les H_s simulés. On peut remarquer globalement un bon accord entre les quantiles sauf pour les H_s inférieures à 0.4 m qui sont sur-estimés par le modèle et les $H_s > 1.4$ m qui sont en revanche sous-estimés par le modèle. Ces remarques peuvent également se faire sur la séquence des séries temporelles montrée à droite de la figure A.6. Toutefois cela n'a pas d'influence significative sur la qualité globale du modèle car la majeure partie de la distribution des H_s est bien reproduite par ce dernier.

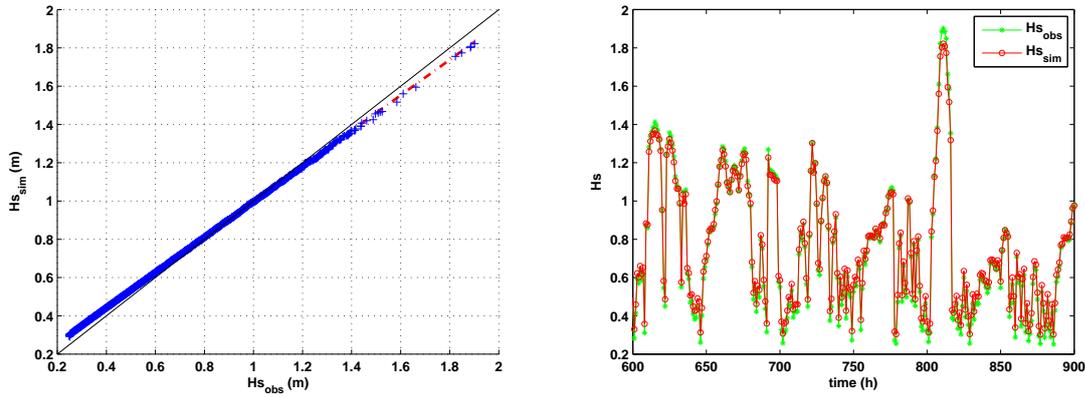


FIGURE A.6 – Diagnostic du modèle : diagramme quantile/quantile (à gauche) et série temporelle (à droite)

A.3.2 Modèle empirique basé sur la corrélation entre l'énergie adimensionnelle et l'inverse de l'âge des vagues

Les résultats obtenus lors du projet JONSWAP [Hasselmann et al. \[1973\]](#) et de nombreuses autres études ([Hwang et al. \[1998\]](#), [Donelan et al. \[1992\]](#), [Hanson, Phillips \[1999\]](#)) ont montré qu'il existe une corrélation entre l'énergie adimensionnelle (ϵ) et l'inverse de l'âge des vagues (U_{10}/C_p). La relation entre ces deux variables peut être écrite de la manière suivante :

$$\epsilon = \alpha \left(\frac{U_{10}}{C_p} \right)^\beta \quad (\text{A.3})$$

où α et β sont les paramètres du modèle à déterminer à partir des données U_{10} , vitesse du vent à 10m et C_p , vitesse de phase des vagues. L'énergie adimensionnelle est définie par :

$$\epsilon = \frac{g^2}{U_{10}^4} \int \int S(f, \theta) df d\theta \quad (\text{A.4})$$

avec g la constante de gravité. En eau profonde, la vitesse de phase est :

$$C_p = \frac{gT_p}{2\pi} \quad (\text{A.5})$$

L'énergie adimensionnelle peut être ré-écrite en fonction de la hauteur significative par :

$$\epsilon = \frac{g^2 H_s^2}{16U_{10}^4} \quad (\text{A.6})$$

et l'équation [A.3](#) peut se ré-écrire sous forme

$$\frac{g^2 H_s^2}{16U_{10}^4} = \alpha \left(\frac{2\pi U_{10}}{gT_p} \right)^\beta \quad (\text{A.7})$$

L'équation [\(A.7\)](#) relie la hauteur significative et la période de pic à l'intensité du vent. Les paramètres α et β peuvent être estimés à partir des données et nous avons obtenus

comme résultat $\alpha = 0.0008$ et $\beta = -3.73$, ce qui conduit à la relation :

$$\frac{g^2 H_s^2}{16U_{10}^4} = 0.0008 \left(\frac{2\pi U_{10}}{gT_p} \right)^{-3.73} \quad (\text{A.8})$$

La figure A.7 montre la régression entre l'énergie adimensionnelle et l'inverse de l'âge des vagues. Nous pouvons remarquer une forte corrélation entre les deux variables. Avec ce modèle, la proportion de la variance en ϵ expliquée est de 0.92%.

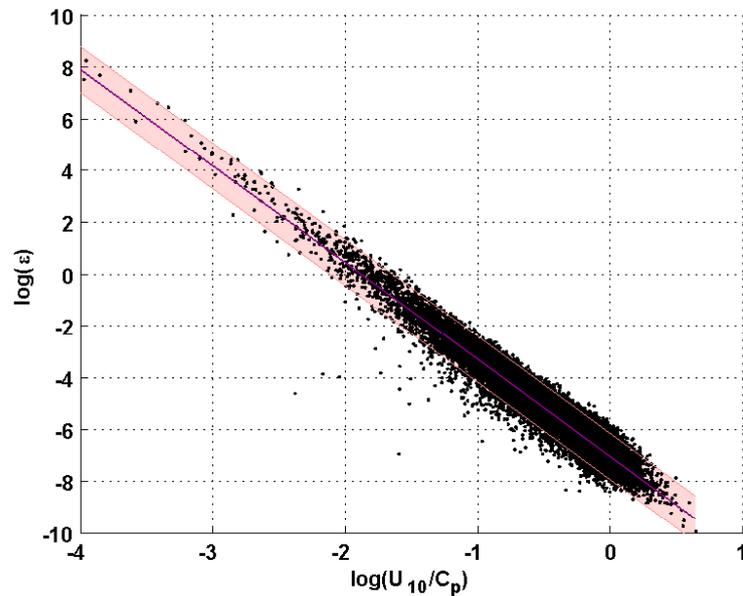


FIGURE A.7 – Régression entre l'énergie adimensionnelle et l'inverse de l'âge des vagues. Les intervalles de confiance sont aussi représentés

En résumé, disposant d'un générateur de vent et des modèles A.1 et A.7, il est possible de déterminer le H_s et le T_p de la mer du vent : le modèle A.1 sert d'abord à obtenir le H_s à partir de l'intensité du vent, puis on en déduit le T_p à partir du modèle A.7.

En ce qui concerne la direction de la mer du vent, au regard de l'histogramme de la figure A.5 (b), elle peut être obtenue à partir d'un modèle simple additionnant la direction du vent à un bruit blanc gaussien de moyenne 0 et de variance 45, pour des angles exprimés en degrés :

$$\theta_m(t) = \theta_w(t) + \Theta(t) \quad (\text{A.9})$$

avec $\Theta(t) \hookrightarrow N(\mu = 0, \sigma^2 = 45)$.

A.4 Conclusion

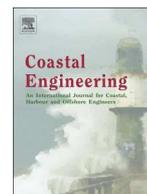
Nous avons proposé dans cette étude, des modèles permettant de simuler des séries temporelles des caractéristiques de la mer du vent (H_s , T_p et θ_m) à partir de l'intensité et de la direction du vent. L'utilisation de cette approche nécessite au préalable de disposer

d'un générateur de vent. Cependant, nous pouvons bénéficier des générateurs de vent qui existent déjà et dont les résultats sont globalement satisfaisants ([Monbet et al. \[2007\]](#), [Ailliot et al. \[2015\]](#), [Bessac et al. \[2015\]](#)).

Même si la climatologie des mers du vent en Afrique de l'Ouest n'en permet pas une représentation par événements, l'approche présentée dans cette étude peut servir à générer des séries temporelles de mers du vent, lesquelles peuvent être couplées avec les séries temporelles de houles générées pour reconstituer la climatologie complète des états de mer en Afrique de l'Ouest.

B

Amélioration de l'estimation spectrale par utilisation de la structure temporelle des systèmes de vagues



Improvement ocean wave spectra estimation using the temporal structure of wave systems



K. Agbéko Kpogo–Nuwoklo ^{a,*}, Pierre Ailliot ^b, Michel Olagnon ^a, Zakoua Guédé ^a, Sabine Arnault ^c

^a Laboratoire Comportement des Structures en Mer, IFREMER Centre de Brest, B.P. 70, F-29280 Plouzané, France

^b Laboratoire de Mathématiques de Bretagne Atlantique, Université de Bretagne Occidentale, Brest, France

^c LOCEAN/IPSL, Université Pierre et Marie Curie, Paris, France

ARTICLE INFO

Article history:

Received 3 June 2014

Received in revised form 11 November 2014

Accepted 14 November 2014

Available online xxxx

Keywords:

Directional wave spectrum

Spectral estimation

Periodogram smoothing

Kriging

ABSTRACT

Sea states are usually the combination of several time-evolving wave systems whereas the classical spectral estimation methods assume stationarity. A method that adapts to the dynamical evolution of the spectral components is proposed to improve both omnidirectional and directional sea wave spectral estimations. In this method, periodograms are computed for each sea state as in the conventional methods, and rather than only smoothing individual periodograms, the overall time-history of periodograms are simultaneously smoothed in frequency and time dimensions. Since a simple two dimensional averaging would not be appropriate because the temporal evolution of the wave systems reflects typical non-stationary behaviors, we use either kriging or adaptive 2D kernel density estimators that allow the taking care of the spectral component frequency–time evolutions. The method is successfully validated on sequences of spectra typical of sea-state conditions in West Africa. The comparison with the simple 2D averaging method and individual periodogram smoothing method shows that the proposed method gives higher effective numbers of degrees of freedom, better estimates of the spectral shape and reliable spectral moments. The method also provides a tool for sea wave spectra interpolation and may thus be used to fill in missing values and improve wave systems tracking for storm identification purposes.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

A description of random ocean waves is of vital importance for many activities and therefore a considerable effort is spent on the collection and analysis of wave data. The most frequent approach for analyzing the characteristics of random ocean waves is spectral analysis (Goda, 2010; Borgman, 1967). In that approach, a sea state is defined for a suitable short-term period of time (from half hour to several hours), and is considered fully characterized by its directional spectrum which provides the distribution of the wave energy as a function of frequency and direction. Reliable sea state spectral estimation is therefore important, especially in coastal engineering and other marine applications. Two types of methods are available for estimating the spectrum of a sea surface elevation time series: non-parametric and parametric methods. Non-parametric sea wave spectral estimation generally consists of periodogram smoothing with frequency windows (Schuster, 1898; Welch, 1967), or Fourier transformation of smoothed or

truncated auto-covariance function (Blackman and Tukey, 1959). In parametric spectral estimation, the task is to estimate the parameters of a model that describes the spectrum of the sea state (Phillips, 1985; Hasselmann et al., 1973). Parametric estimation is generally satisfactory for unimodal spectra, but much weaker when the spectrum exhibits many peaks (Olagnon et al., 2013).

In many cases, the sea state is a superimposition of coexisting wave systems (swells and wind sea) and the resulting spectra have many peaks. One often needs in such cases to partition the sea state into its wave systems for an accurate characterization (see (Aarnes and Krogstad, 2001; Gerling, 1992; Hanson and Phillips, 2001; Portilla et al., 2009)). Furthermore, some of the successive wave systems are created by the same meteorological event (storm, hurricane, pressure low, etc.) and thus, one can track in time the evolution of wave systems in order to link them to the same event (see (Hanson and Phillips, 2001; Aarnes and Krogstad, 2001)). Indeed, when looking at the time-history of periodograms, it is possible to observe temporal coherence in the spectral components. As an illustration, Fig. 1 shows on the left panel a periodogram for individual sea state (in dark) and on the top right panel the time-history of periodograms off Angola. Linear patterns can be observed in the low frequency ranges, corresponding to the time evolutions of swell systems. Some coherent patterns can also be observed in high frequency ranges, corresponding to local wind effects. It

* Corresponding author. Tel.: +33 298224144.

E-mail addresses: kkpogonu@ifremer.fr (K.A. Kpogo–Nuwoklo), pierre.ailliot@univ-brest.fr (P. Ailliot), michel.olagnon@ifremer.fr (M. Olagnon), zakoua.guede@ifremer.fr (Z. Guédé), sabine.arnault@locean-ipsl.upmc.fr (S. Arnault).

is legitimate to expect that taking into account the temporal coherence of the spectral components may be a good way to improve sea wave spectra estimation. That allows the overcoming of the limits imposed in the estimation quality by strict stationarity conditions.

This study proposes methods that adapt to the dynamical evolution of the spectral components to improve both omnidirectional and directional sea wave spectral estimations. Periodograms are first of all computed for each sea state as in the conventional methods. Yet, rather than only smoothing individual periodograms, we consider the overall time-history of periodograms as a random field that depends on frequency and time, and we smooth this field simultaneously in frequency and time dimensions. Since the temporal evolution of the wave systems reflects typical non-stationary behaviors, as for instance the linear increasing trend of the pic frequency of swell components, a simple two dimensional averaging would not be appropriate. Two techniques that allow the consideration of the dynamical evolution of wave systems are therefore proposed in this study to smooth the time-history of periodograms:

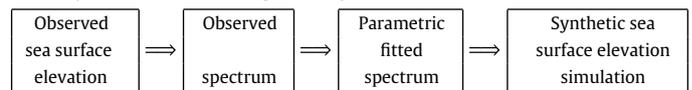
- Kriging that is a spatial optimal interpolation method based on the idea that the value at an unknown point should be a weighted average of the known values at its neighbors, where the optimal weights are calculated using the time and frequency correlations (Stein, 1999);
- Adaptive 2D kernel density estimator i.e. imposing a kernel that adapts to the dynamical evolution of wave systems (Silverman, 1986).

The addition and the use of the time dimension, taking into consideration the dynamical evolution of the wave systems, improve the sea state spectral estimation. Spectra estimates using the proposed approach dramatically enhance the temporal tracking of wave systems for sea storm identification purposes. Moreover, in situ wave measurements often suffer from missing data in consequence of instrument loss, malfunction, or delayed maintenance. For the many marine applications that need continuous measurements, the approach provides a straightforward way to interpolate spectra when faced with missing data (see Fig. 1 right panel). The extension of the method to directional spectra improves also the estimation of the directional spreading which is important in some marine applications (Table 1).

The paper is organized as follows. A brief description of the data used for this study is given in Section 2. In Section 3, we present the proposed method for omnidirectional spectral estimation. We then show some comparison results of the estimated spectra with those that are estimated by only smoothing individual periodograms, and with spectra that are estimated using a naive two dimensional averaging. The ability of the proposed method to interpolate the missing spectra is finally shown. Section 4 presents the extension of the

Table 1

Summary of the method used to generate synthetic time series of sea surface elevation.



methods to the directional spectral estimation. Conclusions are given in Section 5.

2. Data

The data used for this study are re-simulated times series of sea surface elevations derived from a time-history of parametric spectra in order to validate the proposed methods in an idealized situation where the “true” spectra are known. The measurement data, covering period from March 2001 to April 2004 are obtained from a directional waverider buoy, sampled at 1.28 Hz, and anchored by 1450 m water depth off the coast of Angola. Half hourly successive spectra derived from in situ measurements are first partitioned using a “dynamical partitioning algorithm” (Ailliot et al., 2013). Synthetic spectra are then obtained as linear combinations of parametric models fitted to the partitions. A wind sea component is fitted by a JONSWAP parametric model and swell components are fitted by log-normal parametric models as recommended in Forristal et al. (2013), Olagnon et al. (2013) for West Africa conditions (see Fig. 2). The resulting time-history of synthetic spectra provides a realistic sequence of spectra (reference spectra). From each synthetic spectrum, time series of sea surface elevation $z(\tau)$ are finally simulated, assuming that $z(\tau)$ is a Gaussian process, as described in Goda (1977) and Olagnon and Robin (1990). The goal is to estimate the spectra from the re-simulated sea surface elevation time series with the proposed methods and the conventional ones and to compare the results with the reference spectra.

3. Proposed method for omnidirectional spectral estimation

3.1. Method

The method consists of two steps:

1. block the data by short stationary periods (typically 30 min) and compute periodograms for recorded sea elevation time series using discrete Fourier transform on each block;
 - kriging, or
 - two dimensional kernel density estimator by imposing a kernel that follows the dynamical evolution of wave systems (adaptive kde2D).
2. smooth the time-history of log-periodogram using either.

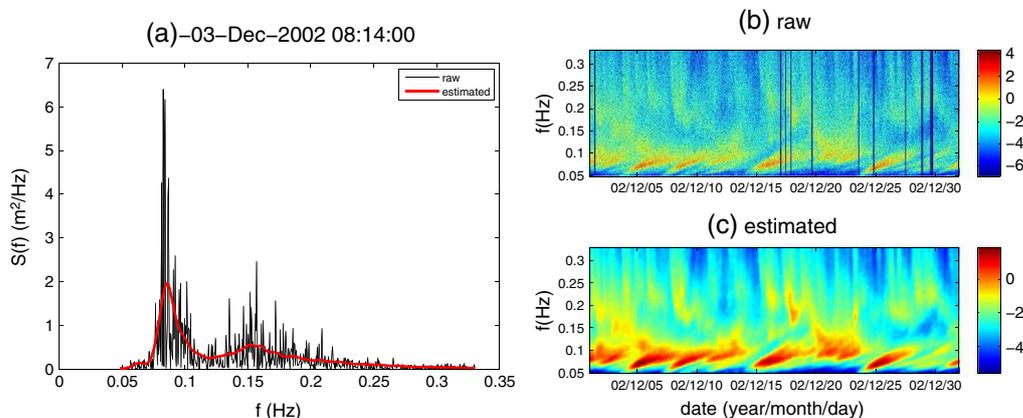


Fig. 1. Sea wave spectra estimation and interpolation applied to in situ data: (left) example of periodogram and estimated spectrum; (right) time-history of periodogram (up) and estimated spectra (below).

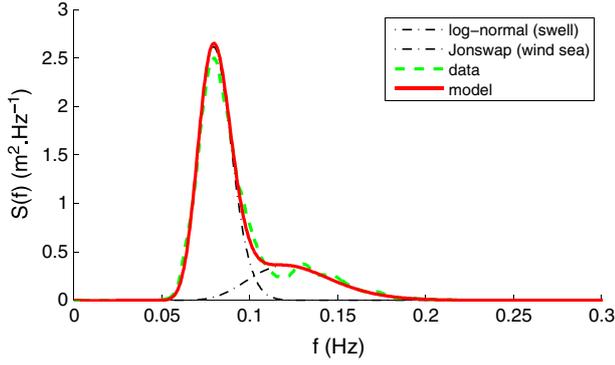


Fig. 2. Observed periodogram and fitted combined model (log-normal function for the swell component and JONSWAP for the wind sea).

3.1.1. Definitions

Consider a real-valued, zero mean process $\{z_\tau\}_{0 \leq \tau \leq T}$ over duration T and assume that $\{z_\tau\}_{0 \leq \tau \leq T}$ is divided into N zero mean stationary sub-processes $\{z_\tau^1\}, \{z_\tau^2\}, \dots, \{z_\tau^N\}$ of individual duration $\delta t = T/N$ and with finite size observed realization $(z_{0,\dots}^1, z_{2n-1}^1), (z_{0,\dots}^2, z_{2n-1}^2), \dots, (z_{0,\dots}^N, z_{2n-1}^N)$ at time $t \in \{\delta t, 2\delta t, \dots, N\delta t\}$.

At a given time $t = j\delta t$, the associated periodogram for the stationary sub-process $\{z_\tau^j\}$ is computed for a discrete set of frequencies $f \in [\delta f, 2\delta f, \dots, n\delta f]$ as

$$I(f, t) = \frac{1}{2n} \left| \sum_{\tau=0}^{2n-1} z_\tau^j e^{-i2\pi f \tau} \right|^2 \quad (1)$$

with $\delta f = 1/(2n\tau)$ as the sampling frequency.

Based on asymptotic properties, (Christensen, 1991; Fan and Kreutzberger, 1998) a frequently adopted model for $I(f, t)$ is

$$I(f, t) = S(f, t)\varphi(f, t)$$

with $\varphi(f, t)$ are $\sim^{iid} \exp(1)$ and $S(f, t)$ denotes the true spectral density. According to this model,

$$E[I(f, t)] = S(f, t)$$

and

$$Var[I(f, t)] = S^2(f, t).$$

Due to its very large variance, $I(f, t)$ is seldom used as an estimator of $S(f, t)$. A logarithmic transform is commonly applied to convert the above multiplicative noise model into an additive noise model and stabilize the variance (see, (Fan and Kreutzberger, 1998; Lee, 1997; Pawitan and O'sullivan, 1994))

$$L(f, t) = m(f, t) + \xi(f, t)$$

where $L(f, t) = \log(I(f, t))$, $m(f, t) = \log(S(f, t))$ is the true log-spectral density and $\xi(f, t) = \log(\varphi(f, t))$ are iid random variables.

The periodogram estimate is an asymptotically (in n) unbiased but inconsistent estimate of the underlying true spectrum (see (Stoica and Randolph, 1997)). In particular, the estimate suffers from a high variance which does not converge to zero as n increases. Consistent estimators of wave spectral density can be obtained by assuming the spectrum constant over some neighborhood and directly smoothing the periodogram or the log-periodogram (Wahba, 1978).

3.1.2. Smoothing the time-history of periodograms using kriging

Kriging is based on the idea that the value at an unknown point should be a weighted average of the known values at its neighbors. Kriging theory provides an optimal linear unbiased estimator, based

on assumptions on covariances and it is meant for estimation of a single realization of a random field (Cressie, 1993).

We assume that the true log-spectral density $m(f, t)$ is a second order time stationary process and denote

$$\begin{cases} \mu_m(f) & = E[m(f, t)] \\ cov[m(f, t), m(f', t')] & = C_m(f, f', t' - t). \end{cases}$$

As a consequence the log-periodogram $L(f, t)$ is also a second order stationary process such that

$$\begin{aligned} \mu_L(f) &= E[L(f, t)] \\ &= E[m(f, t) + \xi(f, t)] \\ &= \mu_m(f) + \alpha \end{aligned} \quad (2)$$

with $\alpha = E[\xi(f, t)]$ and

$$\begin{aligned} cov[L(f, t), L(f', t')] &= cov[m(f, t) + \xi(f, t), m(f', t') + \xi(f', t')] \\ &= \begin{cases} C_m(f, f', t - t') & \text{if } (f, t) \neq (f', t') \\ C_m(f, f', t - t') + \sigma^2 & \text{if } (f, t) = (f', t') \end{cases} \end{aligned} \quad (3)$$

with $\sigma^2 = var[\xi(f, t)]$.

We construct an unbiased linear estimator of $m(f, t)$

$$\hat{m}(f, t) = \mu_m(f) + \sum_{\substack{i \\ |f_i - f| < Df \\ |t_i - t| < Dt}} \lambda_i (L(f_i, t_i) - \mu_L(f_i)) \quad (4)$$

with $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$ such that

$$E[(\hat{m}(f, t) - m(f, t))^2] = \lambda' K_L \lambda - 2\lambda' k_L + var[m(f, t)]$$

is minimum. λ is thus given by

$$\lambda = k_L K_L^{-1}$$

with

$$k_L = \begin{pmatrix} cov[m(f, t), L(f_1, t_1)] \\ cov[m(f, t), L(f_2, t_2)] \\ \vdots \\ cov[m(f, t), L(f_n, t_n)] \end{pmatrix} = \begin{pmatrix} C_m(f, f_1, t - t_1) \\ C_m(f, f_2, t - t_2) \\ \vdots \\ C_m(f, f_n, t - t_n) \end{pmatrix} \quad (5)$$

and

$$K_L = \begin{pmatrix} var[L(f_1, t_1)] & cov[L(f_1, t_1), L(f_2, t_2)] & \dots & cov[L(f_1, t_1), L(f_n, t_n)] \\ cov[L(f_2, t_2), L(f_1, t_1)] & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ cov[L(f_n, t_n), L(f_1, t_1)] & \dots & \dots & var[L(f_n, t_n)] \end{pmatrix}. \quad (6)$$

We estimate $cov[L(f, t), L(f', t + \Delta t)]$ by

$$\hat{C}_L(f, f', \Delta t) = \frac{1}{N} \sum_{t=1}^{N-\Delta t} L(f, t)L(f', t + \Delta t) - \frac{1}{N} \sum_{t=1}^{N-\Delta t} L(f, t) \sum_{t=1}^{N-\Delta t} L(f', t + \Delta t) \quad (7)$$

and $\mu_L(f)$ by

$$\hat{\mu}_L(f) = \frac{1}{N} \sum_{t=1}^N L(f, t)$$

and then deduce an estimate of μ_m and C_m using Eqs. (2) and (3) and then the weights λ by replacing C_L and μ_L by the estimates in Eqs. (5) and (6). Finally, the estimated spectral density at a given frequency and a given time is obtained by

$$\hat{S}(f, t) = \exp(\hat{m}(f, t)).$$

Note that it is also possible to compute the kriging variance

$$\text{var}[\hat{m}(f, t) - m(f, t)] = E\left[(\hat{m}(f, t) - m(f, t))^2\right] \quad (8)$$

which gives an estimation of the uncertainty and may be interesting for some applications.

In practice, the covariance is estimated by Eq. (7) using three year time-history of periodograms. In the case where only very short measurement durations would be available, covariance could still be estimated from long duration hindcast data. It is also possible to use a parametric model for the covariance function as it is often the case in geostatistics (see e.g. (Cressie, 1993)). The covariance should be constant at a regional scale and one should be in position to use the same covariance to estimate the spectra for any close enough site.

Fig. 3 shows the correlation matrices defined as

$$R(f, f', \Delta t) = \frac{\hat{C}_L(f, f', \Delta t)}{\hat{C}_L(f, f', 0)}.$$

A global yet expected observation is that the correlation decreases when Δt increases. Moreover, it is found that swell systems ($f < f_s = 0.11$ Hz) are still self-correlated up to two days time-lags whereas wind sea systems are only weakly correlated beyond 24 h. For swell systems the correlation structure moves to the right side meaning that the correlation is more important between $S(f, t)$ and $S(f', t')$ when $t > t'$ and $f > f'$. This reflects the increasing trend of frequencies along a swell event (Eq. (11)). The correlation structure of wind sea systems is also found to move slightly to the opposite side reflecting the decreasing trend of wind sea frequencies. The parameters D_f and D_t of Eq. (4) which describe the size of neighborhood used for kriging interpolation are $D_f = 10\delta f$ and $D_t = 12\delta t$, where δf and δt denote frequency and time increments respectively.

3.1.3. Smoothing the time-history of periodograms using 2D kernel density estimator

Kernel smoothing of the individual log-periodogram is a popular non-parametric method for spectral density estimation (Hannig and Lee, 2004). The usual kernel estimation (kde1D) for the log-spectrum is defined as

$$\hat{m}(f, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{f-f_i}{h}\right) L(f_i, t) \quad (9)$$

where kernel function K takes various forms in the literature, for instance a Hanning window, and is in fact a symmetric probability density function. The bandwidth h is a non-negative smoothing parameter that controls the amount of smoothing.

Since the time-history of log-periodogram is assumed to be a random field $L(f, t)$, we propose in the following to smooth it with a 2D kernel density estimator. The general form for the 2D density estimator is then

$$\hat{m}(f, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\det(\mathbf{H})} K\left\{\mathbf{H}^{-1}\left(\begin{bmatrix} f \\ t \end{bmatrix} - \begin{bmatrix} f_i \\ t_i \end{bmatrix}\right)\right\} L(f_i, t_i). \quad (10)$$

n is the number of observation points, K is the kernel function, and \mathbf{H} is the 2×2 bandwidth matrix. When \mathbf{H} is a diagonal matrix i.e. $\mathbf{H} = \begin{bmatrix} h_f & 0 \\ 0 & h_t \end{bmatrix}$ where the elements h_f and h_t control the smoothing in frequency and time ranges respectively, the corresponding estimator is equivalent to a simple 2D averaging (see Fig. 4(b)) and it is denoted hereafter by “kde2D”.

Yet we adapt in our proposed method a kernel that follows the frequency–time evolutions of wave systems. This is easily carried out by using a full bandwidth matrix i.e. $\mathbf{H} = \begin{bmatrix} h_f & \lambda \\ \lambda & h_t \end{bmatrix}$ where λ is used to control the frequency–time evolution of wave systems (“adaptive kde2D”). According to a linear wave theory in deep water, for a wind event that was bounded in time and space, the dominant wave frequency observed

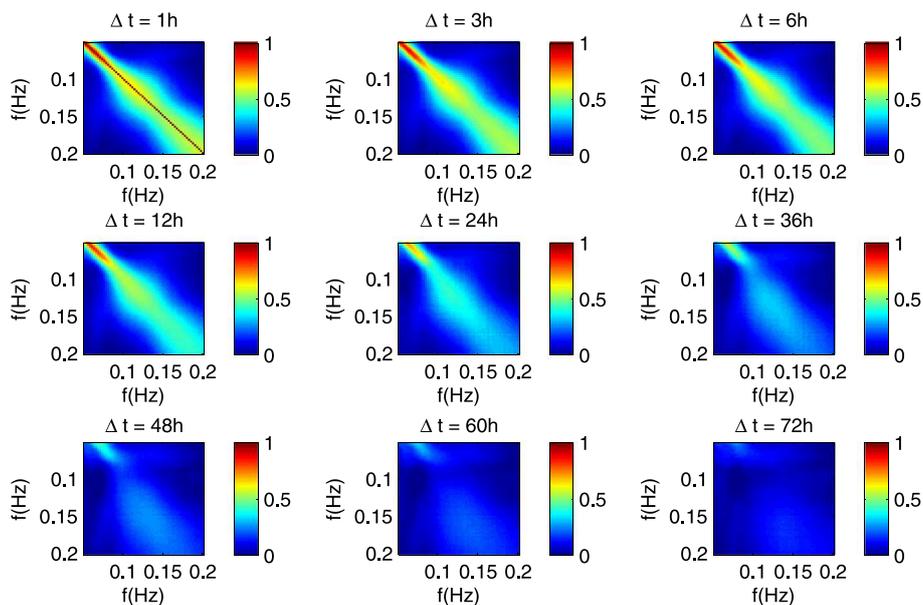


Fig. 3. Correlation function for different values of Δt .

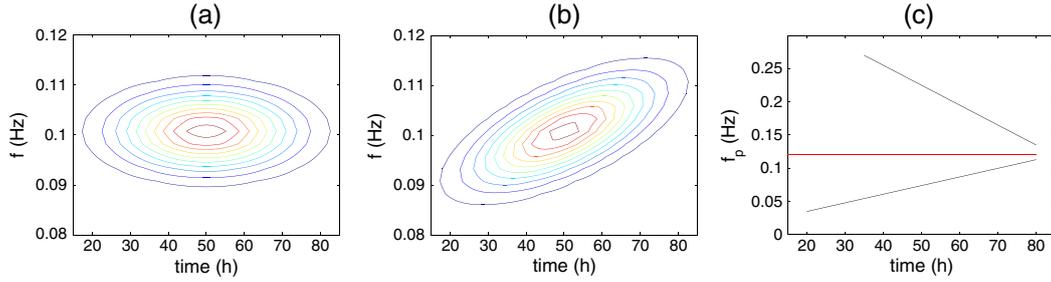


Fig. 4. Illustration of 2D kernel density estimator: (a) with diagonal bandwidth matrix (kde2D), (b) with full bandwidth matrix (adaptive kde2D) for the case $f < f_s$; and (c) wave systems peak frequency–time evolution.

at some distant location linearly increases over time at a rate inversely proportional to the distance traveled (see (Gjevik et al., 1988; Hanson and Phillips, 2001)):

$$\frac{df}{dt} = \frac{g}{4\pi d}. \quad (11)$$

The dynamical time evolution of wind sea systems however does not have any particular a priori pattern. It depends on the wind field characteristics (speed, direction, and time-displacement). Yet, according to Sverdrup and Munk (1947), at the beginning of wind wave development, the energy transferred from wind to waves helps to increase the period (decrease the frequency). For simplicity it is assumed that the frequency–time evolution of wind sea systems follows in practice a decreasing linear trend (Fig. 4(c)). Thus λ has a positive value for swell systems and negative value for wind sea systems. The use of a full bandwidth matrix allows the taking into account the dynamical evolution of the wave systems and therefore allows the overcoming of the limits imposed in the estimation by strict stationarity conditions.

It is well known that the choice of \mathbf{H} is much more crucial than the choice of K (e.g., see (Wand and Jones, 1993; Simonoff, 1996)). If the smoothing parameters are too small, the estimate is too noisy, exhibiting high variability. If they are too large, then the estimate may miss key features due to over-smoothing, washing out small details. In practice, to speed up computation and for its simplicity we use a triangular kernel defined by

$$K(u) = \begin{cases} 1-|u| & \text{if } |u| \leq 1. \\ 0 & \text{otherwise} \end{cases}$$

The bandwidth of the kernel is a free parameter that exhibits a strong influence on the resulting estimate. There are several methods to optimize the kernel bandwidth. In this study, the bandwidth is obtained by minimizing the mean integrated squared error (MISE) function. It is the most common optimality criterion used to select kernel bandwidth. The MISE is defined as

$$\text{MISE}(\mathbf{H}) = E \sum_f \sum_t (\hat{m}(f, t) - m(f, t))^2.$$

As the theoretical log-density $m(f, t)$ is unknown, the MISE cannot be estimated directly. Bowman (1984) has proposed the use of cross validation method to estimate the MISE. In this method, the observation sample i is removed for estimating $m(f_i, t_i)$.

For instance we find for West Africa the optimal bandwidth matrix $\mathbf{H} = \begin{bmatrix} h_f & \lambda \\ \lambda & h_t \end{bmatrix}$ for the adaptive 2D kernel density estimator which elements are:

$$h_f = 0.01 \text{ Hz} \quad h_t = 15 \text{ h} \quad \lambda = 0.45 \times \text{mHz} \cdot \text{h}^{-1}, \quad (\text{case of swell systems})$$

$$h_f = 0.04 \text{ Hz} \quad h_t = 9 \text{ h} \quad \lambda = -1.2 \times \text{mHz} \cdot \text{h}^{-1}, \quad (\text{case of wind sea systems}).$$

The swell system frequency domain smoothing bandwidth is smaller than the wind sea one justifying the fact that swells have narrower ranges of frequencies than wind sea; swell waves have dispersed from their generation area and have been dissipated. In contrast, the time domain smoothing bandwidth of swell systems is greater than the wind sea one because swells are generated by distant storms and the more remote is the storm, the greater is the duration between arrivals of successive frequencies according to the dispersion of sea waves in deep water (Eq. 11). Similar to the covariance in the case of kriging, the bandwidth matrix should be constant at a regional scale. One needs to estimate in practice only once the bandwidth matrix which can be therefore used for other neighboring sites. Knowledge of the local climate would allow one to choose without computation bandwidth matrix values close to these optimal ones.

3.2. Results

Re-simulated time series of sea surface elevation (see Section 2) are used to estimate the spectra by the proposed methods. Periodograms are first computed from the time series and the time-history of periodograms are smoothed using the kriging technique of Section 3.1.2 and adaptive 2D kernel density estimator technique (adaptive kde2D) of Section 3.1.3. Our goal is to compare the estimated spectra with the reference spectra (i.e. synthetic spectra that are used to simulate the time series of sea surface elevation). Furthermore, in order to check the relevance of our method, the spectra are also estimated, using a diagonal 2D kernel density estimator (kde2D) i.e. a naive 2D averaging, and using 1D kernel density estimator, that consists of periodogram smoothing in the frequency domain only (kde1D).

The first part of this section shows results related to the improvement in the spectral estimation using kriging and adaptive kde2D. The second part shows some results of the ability of the these techniques to interpolate the missing spectra.

3.2.1. Improvement in spectral estimation

Fig. 5 shows the time-history of the error $S(f, t) - \hat{S}(f, t)$, where $S(f, t)$ denotes the true spectra and $\hat{S}(f, t)$ the spectra estimated by the different methods (kde1D, kde2D, adaptive kde2D and kriging). The error is greater for the spectra estimated by kde1D and kde2D than those estimated using adaptive kde2D and kriging. In particular, kde1D and

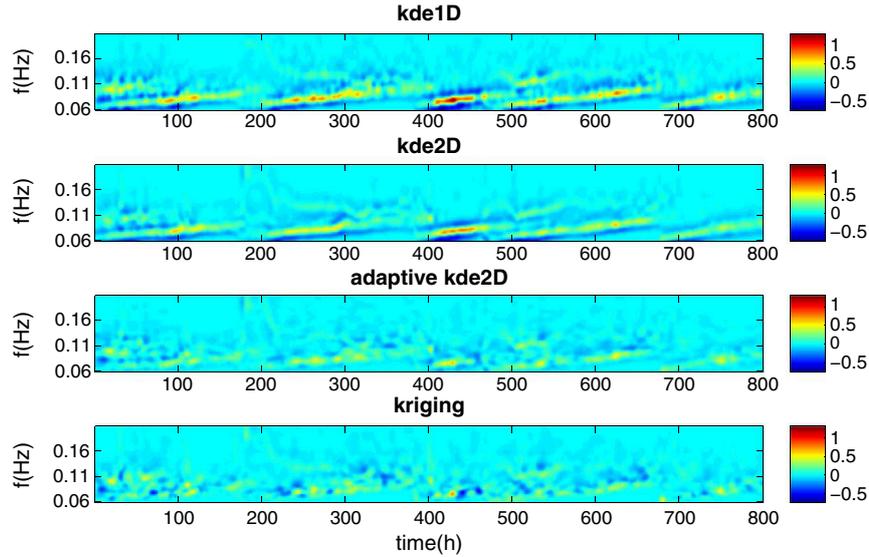


Fig. 5. Time-history of the difference between the true spectra and the estimated spectra.

kde2D often underestimate the peaks of swell systems in comparison with the proposed methods. Visual inspection shows that the spectra estimated using an adaptive kde2D or kriging are closer to the true spectra than those obtained by kde1D and kde2D. The same observation can be made on the individual spectra (see Fig. 6).

The results are also studied quantitatively on three statistical parameters: effective number of degrees of freedom of the spectral estimation, Goda's peakedness parameter, and the spectral moments.

- Effective number of degrees of freedom

Under the assumption that the periodograms are independent and exponentially distributed, the spectrum $\hat{S}(f, t)$ obtained by smoothing the periodograms, is shown (see for example (Elgar, 1987; Zhang et al., 2005)) to approximately obey a χ^2 distribution with $2\nu_{eff}$ degrees of freedom

$$2\nu_{eff} \frac{\hat{S}(f, t)}{S(f, t)} \sim \chi^2(2\nu_{eff}).$$

ν_{eff} is often used in spectral estimation as an indicator of goodness of the estimation. A classical assumption made to compute the number

of degrees of freedom is that the spectrum is constant in some neighborhood. Since this assumption is not fully valid, we compute an "effective number of degrees of freedom" defined by

$$\begin{aligned} \frac{1}{\nu_{eff}(f, t)} &= \text{var} \left[\frac{\hat{S}(f, t)}{S(f, t)} - 1 \right] \\ &= E \left[\left(\frac{\hat{S}(f, t)}{S(f, t)} - 1 \right)^2 \right], \end{aligned} \quad (12)$$

where $S(f, t)$ is the reference spectrum denoted by $S_{ref}(f, t)$ and $\hat{S}(f, t)$ is the estimated spectrum. The "effective number of degrees of freedom" integrated for the whole time-history can then be estimated by

$$\hat{\nu}_{eff}(f) = \frac{1}{\frac{1}{N} \sum_{t=1}^N \left(\frac{\hat{S}(f, t)}{S(f, t)} - 1 \right)^2} \quad (13)$$

where N is the total number of spectra. ν_{eff} can also be seen as the inverse of the normalized mean square error.

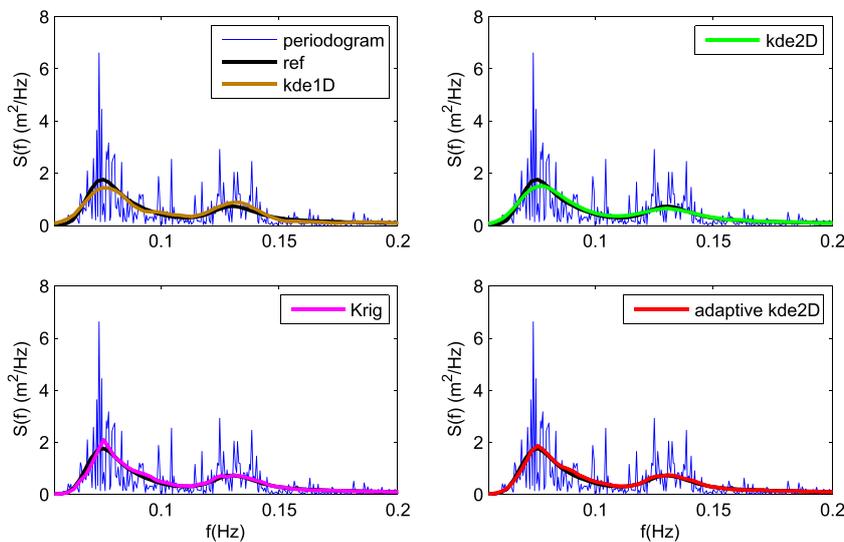


Fig. 6. Example of estimated spectrum in comparison with the reference one.

Fig. 7 shows the results of the “effective number of degrees of freedom” for the four estimation methods. Globally, $\nu_{eff}(f)$ has higher values for kriging and adaptive kde2D spectra than naive kde2D spectra which are slightly superior to the kde1D ones. $\nu_{eff}(f)$ varies proportionally with the spectral density and then reaches its peaks approximately at the same frequency as the peak frequency of the swell and the wind sea systems. The peaks of swell are better estimated by kriging and by adaptive kde2D than by kde2D and kde1D. For the wind sea systems the differences between adaptive kde2D, kde2D and kriging are not significant. The quality of the estimation is not good enough for the frequency interval between the swell and wind sea systems ([0.11 0.13 Hz]) and the tail of the spectrum $f > 0.18$ Hz. For those intervals of frequency, the spectral densities are low and no temporal structure can actually be identified.

• Goda’s peakedness parameter (Goda, 1978)

The objective is to investigate what improvements in the accuracy of the spectral bandwidth estimation one can have with our proposed methods. We used as spectral bandwidth the Goda’s peakedness parameter Q_p , and it is estimated for each component (wind sea and swell) of the spectrum as:

$$Q_p = \frac{2}{m_0^2} \int_{f_{min}}^{f_{max}} f S(f)^2 df$$

where $m_0 = \int_{f_{min}}^{f_{max}} S(f) df$, f_{min} and f_{max} are respectively the lower and the upper frequency bounds. We compute for the four estimation methods the normalized mean absolute error defined as

$$NMAE_{Q_p} = \frac{|Q_p - \hat{Q}_p|}{Q_p} \times 100$$

where Q_p is the reference Goda’s peakedness parameter and \hat{Q}_p is the one that is computed from the spectra estimated by kde1D, kde2D, adaptive kde2D or kriging. The results in Fig. 8 show the average of $NMAE_{Q_p}$ versus the peak frequency. The swell system spectral bandwidths are clearly better estimated using adaptive kde2D and kriging than the others. Taking into account the frequency–time evolution of the swell systems allows the reduction of approximately by 10% the error in the estimation of the spectral bandwidth. The improvement in the wind sea system bandwidth parameter estimation is not significant. The wind sea has in fact a broad spectral bandwidth and short stationary durations and it is difficult to have a precision in its estimation.

• Spectral moments

Spectral moments are used to compute sea wave parameters (significant wave height, mean period, etc.). The ability of the proposed methods to give good estimates of spectral moments is therefore another important criterion for choosing spectral estimates for ocean waves. The objective is to verify that the reduction of the variance in the spectral estimation using our method does not

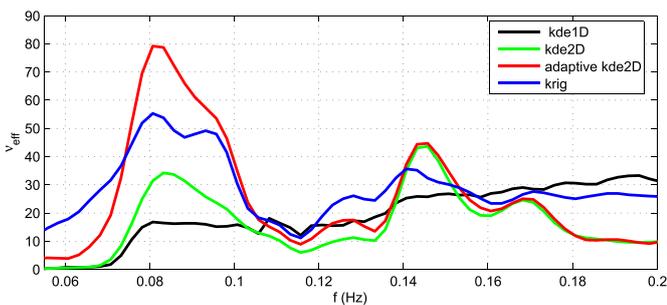


Fig. 7. Effective number of degrees of freedom as function of frequency for the three estimation methods.

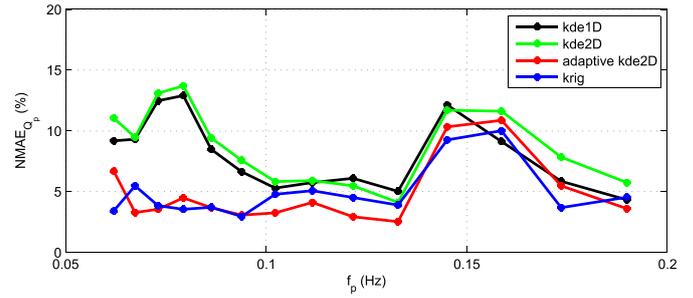


Fig. 8. Normalized mean absolute error for the Goda’s peakedness parameter as function of the peak frequency.

decrease the accuracy of the spectral moments. We thus investigate the relative difference between the reference spectral moments and the estimated ones defined for the n th moment as:

$$Dm_n(t) = \frac{\left| \int f^n S_{ref}(f, t) df - \int f^n \hat{S}(f, t) df \right|}{\int f^n S_{ref}(f, t) df}$$

Fig. 9 shows the boxplot of Dm_n distribution for the $-1, 0, 1, 2$ orders of spectral moments. The difference between the reference spectral moments and the estimated ones are lower for kriging and adaptive kde2D. It can be concluded that the spectral moments, and thus the sea wave environmental parameters, may be accurately estimated by the two proposed methods. The reduction of the spectral variance by taking into account the temporal evolution of wave systems is thus very useful and does not affect the accuracy of the sea state parameter estimation.

3.2.2. Interpolation ability

In situ deployed instruments, such as buoys, are sources for continuous long-term oceanographic data acquisition. Given the rough and unfavorable conditions in the ocean, data gaps are a common consequence of instrument loss, malfunctioning, or delayed maintenance or data collection, and are a major problem for data analysis. The previously proposed methods, namely kriging and adaptive kde2D, can be effective tools for missing data interpolation. The ability of the methods to perform spectral interpolation is shown in the present part.

We take a time-history of periodograms $I(f, t)_{1 \leq t \leq N}$ split at half-hour intervals. For each time step t , gaps with durations $dt \in \{1h, 3h, 6h, 12h, 24h\}$ are arbitrarily generated by removing the corresponding number of periodograms $\{2, 6, 12, 24, 48\}$ centered on t . The spectra in the gap extent are then interpolated with kriging and adaptive kde2D methods. We then compare the interpolated spectra with the reference ones.

Fig. 10 shows an interpolated spectrum for different gap lengths. From the visual inspection of this figure we can see a good agreement between the reference spectrum and the two interpolated ones onto 12 h. The overall accuracy of the interpolated spectra as a function of frequency is evaluated by the “effective number of degrees of freedom” (Eq. (13)). The “effective number of degrees of freedom” curves are presented in Fig. 11 for the two methods and for different gap lengths. Globally, as could have been expected ν_{eff} decreases when the length of the gaps increases, reflecting that the quality of the interpolation decreases when the length of the gaps increases. The first one (0 h) correspond to ν_{eff} when there is no gap; it is the same as in Fig. 7. Up to 6 h, the kriging ν_{eff} curves are most of the time under the adaptive kde2D ones, reflecting that adaptive kde2D is slightly better than kriging for those gap lengths. Regarding then 12 h gap length, kriging interpolation is slightly better than the adaptive kde2D. The available information for

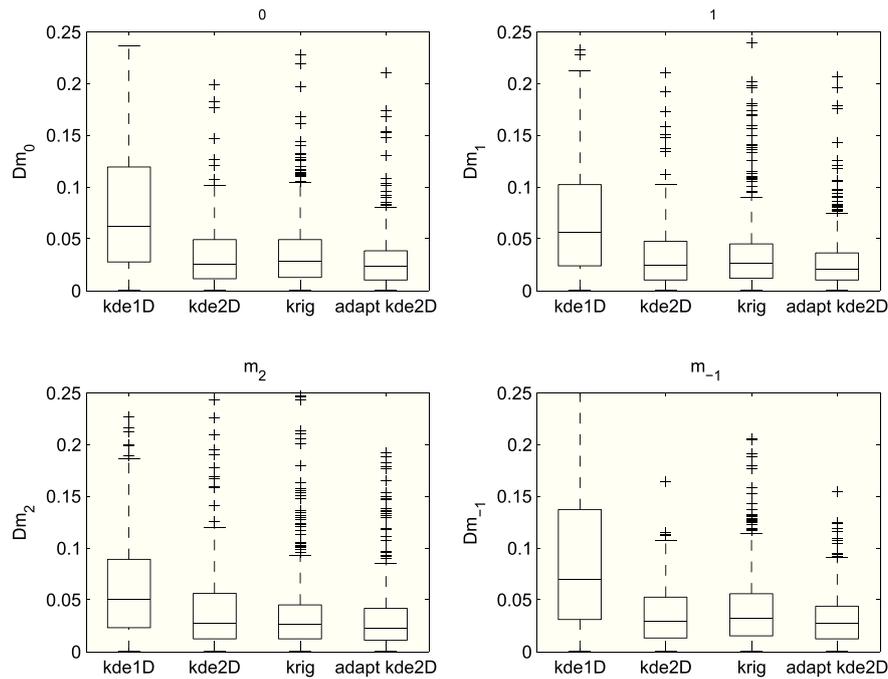


Fig. 9. Boxplot of relative error between the reference spectral moments and the estimated ones.

this gap length is not sufficiently rich for the accurate estimation of the peaks that correspond to the highest density values. Above a gap length of 24 h, the spectra cannot be interpolated by adaptive kde2D since the optimal time domain bandwidth of the wind sea systems h_{DWS} is equivalent to 9 h. The kriging interpolation for those gap lengths also shows poor quality. It is clear that longer periods of missing spectra cannot be properly interpolated.

3.2.3. Discussion

The results show that the spectral components of temporal correlation structure can be used to improve sea wave spectrum estimation. Both the kriging method and the adaptive kde2D method give better estimates for spectral shape and for overall spectral moments than the kde1D method or naive kde2D method. The dependence between successive sea states inside the same meteorological event cannot be taken into account when smoothing an individual sea wave periodogram. This dependence needs to be taken into account by also considering additional information provided by neighboring periodograms. However,

as it is shown by the results, a simple 2D averaging (kde2D) of the periodograms is not sufficient. Introducing the time dimension in the spectral estimation requires the taking care of the spectral components evolutions in order to have reliable spectra that can be used with confidence in marine applications.

Moreover, the results also show that between the adaptive kde2D and kriging, the first has a slight superiority. They are roughly equivalent in the sense that we are looking in the two cases for a linear estimator of the spectral density $S(f, t)$ by using the neighboring spectral densities weighted by the neighbors' distance to $S(f, t)$. The difference between the two methods is probably due to the size of data used to estimate the covariance function which is a key input in the kriging method. The latest results are obtained with covariance estimated from only three years of data. Yet, even with three years of data, the quality of the spectra estimates using kriging is better than those obtained with kde1D or a simple 2D averaging. The accuracy of the spectra estimates using kriging might be improved by using more than three years. Moreover, in contrast to adaptive kde2D, the kriging method has the

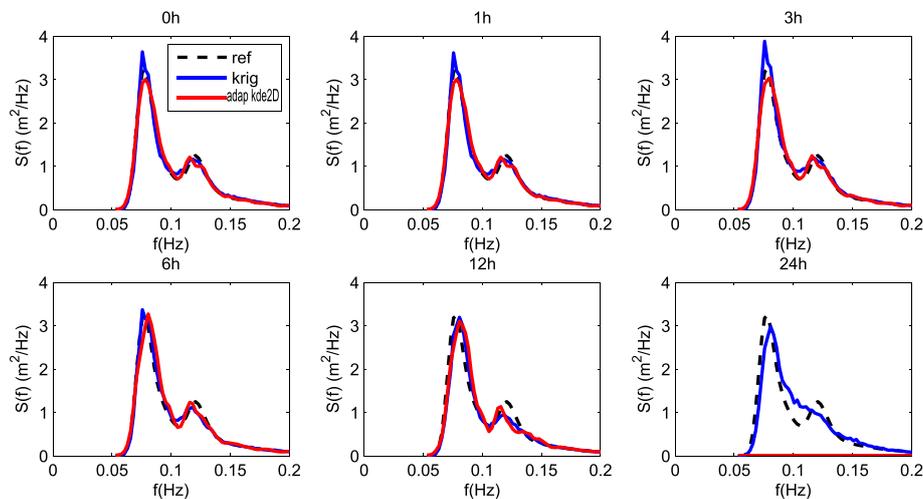


Fig. 10. Example of interpolated spectrum with the reference one for different gap lengths.

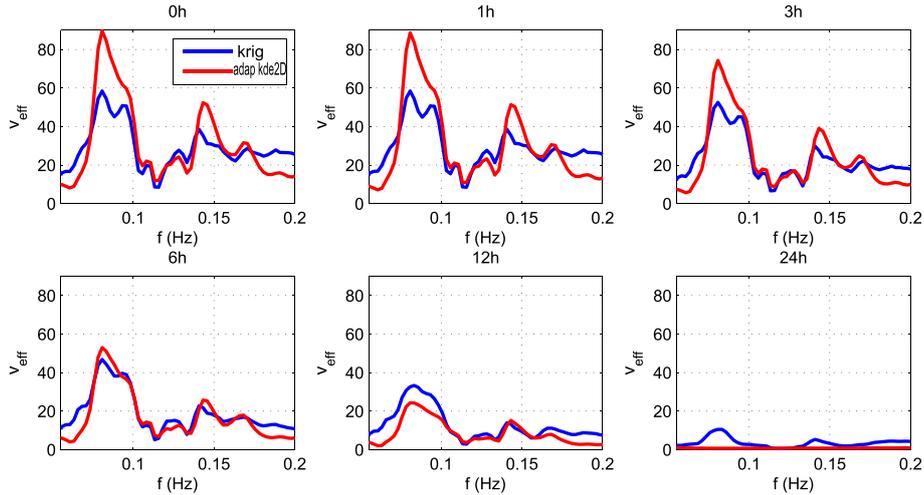


Fig. 11. “Effective number of degrees of freedom” v_{eff} for the two methods and for different gap lengths.

advantage of not requiring a prior model for the time evolution of spectral components. The time evolution of the peak frequencies of swell systems is usually easy to model, but that of the wind sea one lacks specific repeatable models. Choosing a prior model for wind sea peak frequency–time evolution may slightly affect spectral estimation using adaptive kde2D.

4. Extension of the directional spectral estimation

Directional information is crucial in many ocean engineering applications. This information is condensed in a directional wave spectrum described by

$$S(f, \theta) = S(f)D(f, \theta), \quad (14)$$

where $S(f)$ is the energy density spectrum and $D(f, \theta)$ is the directional spreading function. The directional spreading function can be expanded in an angular Fourier series,

$$D(f, \theta) = \frac{1}{\pi} \left[\frac{a_0(f)}{2} + \sum_{n=1}^{\infty} a_n(f) \cdot \cos(n\theta) + b_n(f) \cdot \sin(n\theta) \right], \quad (15)$$

where $a_n(f)$ and $b_n(f)$ are the Fourier coefficients. Based on the three simultaneous wave measurements recorded at the same location, such as for instance heave, pitch and roll of a buoy, it is known that only the first and second Fourier order coefficients can be obtained based on the cross-spectra (Longuet-Higgins et al., 1963). It is common to use a parametric form for $D(f, \theta)$, one of the most used directional spreading function model is *cosine – 2s* defined by

$$D(f, \theta) = K(f) \cos^{2s(f)} \left(\frac{\theta - \theta_m(f)}{2} \right), \quad (16)$$

where $K(f)$ is a normalization factor, $\theta_m(f)$ the mean wave direction, and $s(f)$ the directional spreading parameter.

In the case of the measurements recorded by a pitch/roll buoy the first Fourier coefficients may be obtained by

$$a_1(f) = \frac{Q_{zx}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, b_1(f) = \frac{Q_{zy}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, \quad (17)$$

where subscripts x , y , and z denote East, North, and vertical particle displacements respectively; and the real and imaginary parts C_{ij} and Q_{ij} of a

cross-spectrum between wave records i and j . The mean wave direction, and the directional width are related to the first harmonic through

$$\theta_{m1}(f) = \arg(a_1(f) + ib_1(f))$$

$$\sigma(f) = \sqrt{2(1 - (a_1^2(f) + b_1^2(f)))}.$$

The directional spreading parameter $s(f)$ is then given by

$$s(f) = \frac{2}{\sigma^2(f)} - 1.$$

4.1. Method

The generalization of the methods to the 2D spectra estimation can be described by three steps.

1. Cross-periodogram computation

The periodogram of sea surface elevation $z(\tau)$ (Eq. (1)) can be seen as a cross-periodogram between z and itself. Generally the cross-periodogram between discrete-time signals $y(\tau)$ and $x(\tau)$ of length n is computed as

$$I_{xy}(f, t) = \frac{1}{N} X^*(f, t) Y(f, t) = C_{xy}(f, t) - iQ_{xy}(f, t) \quad (18)$$

with

$$X^*(f, t) = \frac{1}{2n} \sum_{\tau=0}^{2n-1} x_{\tau}^j e^{i2\pi f\tau}$$

and

$$Y(f, t) = \frac{1}{2n} \sum_{\tau=0}^{2n-1} y_{\tau}^j e^{-i2\pi f\tau}.$$

Following Eq. (18) the periodogram of the co-spectra ($C_{zz}(f, t)$, $C_{xx}(f, t)$, $C_{yy}(f, t)$) and the quadrature spectra ($Q_{zx}(f, t)$, $Q_{zy}(f, t)$, $C_{xy}(f, t)$) can be computed.

2. Smoothing the time-history of the logarithmic periodograms of the co-spectra and the quadrature spectra
The time-history of the logarithmic periodograms of the co-spectra

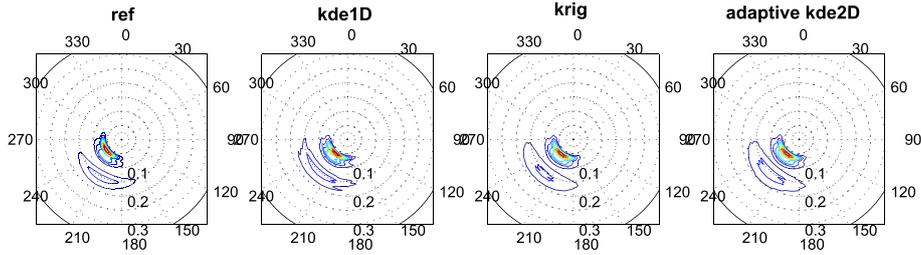


Fig. 12. Example of directional spectrum estimates.

and the quadrature spectra are treated as a random field and are separately smoothed using the same smoothing method (Kriging (Section 3.1.2) or kde2D (Section 3.1.3)) as in the 1D case in order to get estimates of co-spectral densities ($\hat{C}_{zz}(f, t)$, $\hat{C}_{xx}(f, t)$, $\hat{C}_{yy}(f, t)$ and $\hat{C}_{xy}(f, t)$), and estimates of quadrature spectral densities ($\hat{Q}_{zx}(f, t)$ and $\hat{Q}_{zy}(f, t)$).

3. Directional spreading function

From those estimated cross-densities, the last step is to calculate the Fourier coefficients (Eq. (17)), then the directional spreading function $D(f, \theta)$ using Eq. (16) and finally the directional spectrum $S(f, \theta)$ using Eq. (14).

4.2. Results

As for the 1D case the validation is carried out on a simulated time series of sea surface horizontal ($x(\tau)$ and $y(\tau)$) and vertical ($z(\tau)$) motions. The simulated time series are derived from synthetic directional spectra obtained by the fitting of parametric models on partitions derived from West Africa in situ data. The wind sea component is fitted by a JONSWAP parametric model; swell components are fitted by a log-normal shape, and the directional spreading function is modeled by a $\cos^{2\sigma}$ function. To avoid imposing ill-estimated values of σ for partitions of low energy, we use at times the empirical directional width parameter σ derived in Forristal et al. (2013) for West Africa locations. σ is a function of frequency and can be expressed as:

$$\sigma(f) = 0.87f^{0.25}, \quad (19)$$

where $\sigma(f)$ is in radians.

Directional spectra are then estimated from the simulated time series using kriging, adaptive kde2D and kde1D methods. Fig. 12 shows an example of estimated directional spectrum in comparison with the reference one. Since the directional spectrum is the product of $S_{zz}(f)$ by $D(f, \theta)$ the same results on spectral moments may be obtained as in the 1D case. The directional width $\sigma(f)$ seems that the natural parameter can be investigated for checking the quality of directional spectra estimates. Fig. 13 presents the result of directional width estimates with the three methods for 500 consecutive time steps in comparison with the reference directional width (Eq. (19)). The directional width

obtained with the kde1D method exhibits the highest dispersion compared with the other methods. The mean difference between quantile 5% and quantile 95% is 5° against 3.5° for kriging and 3° for adaptive kde2D. These differences do not look significant, because the chosen reference directional width (West Africa climate) corresponds to the narrow spectra; wide spectra might show greater differences since they are expected to be proportional to the directional width. The results however show that it is more likely to get accurate directional width estimates using kriging or (slightly better) adaptive kde2D.

5. Conclusions

Two methods based on kriging and adaptive kernel density estimator are proposed for both 1D and 2D ocean wave spectral estimations and interpolation. The main idea behind those two methods is the use of the temporal correlation structure of wave systems (wind sea and swell) to improve the wave spectra estimation. Regarding 1D spectrum estimation, the first stage common to both methods is to calculate the periodograms (raw spectra) of recorded sea surface elevation time series. The spectral density is then considered as a function of frequency and time and the next step is to treat the time-history of the log-periodogram as a random field and smooth it using either kriging or adaptive kde2D method. The methods are extended to 2D spectra estimation by using the same technique as for the 1D spectra to estimate separately auto-densities and cross-densities, and then computing the directional distribution function with Fourier coefficients obtained from them.

The estimation methods were tested on synthetic yet field-similar data and the results were compared with the conventional kde1D method and a naive 2D averaging method. We have found that both kriging and adaptive kde2D-based proposed methods provide better spectral estimates, i.e. a better representation of the spectra shape and better values of spectral moments compared with the kde1D method and a naive 2D averaging method. Regarding directional spectra, it is found that the proposed methods allow the better estimation of the directional spreading parameter. It is also found that the adaptive kde2D-based method gives slightly better results than kriging. However, the adaptive kde2D method requires a prior model of time evolution for peak frequencies of wind sea and swell systems. The kriging method has the

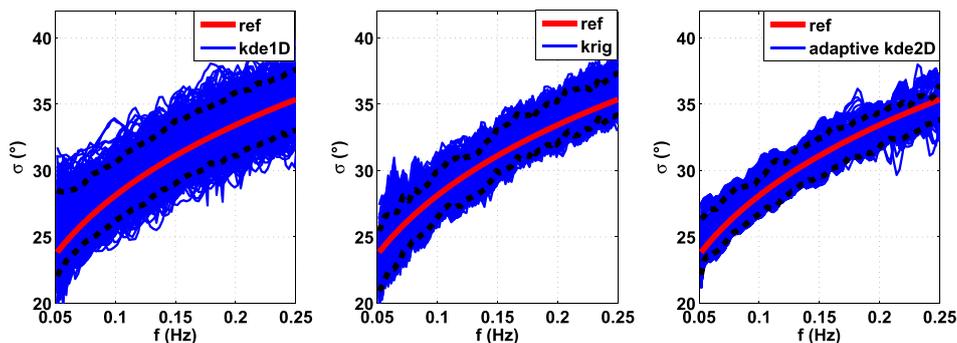


Fig. 13. Directional width obtained with the three methods; 5% and 95% fluctuation intervals are plotted in dashed black line. The red curve corresponds to the true value given by Eq. (19).

advantage of not considering any prior model for the time evolution of the peak frequencies of wave systems. It is based on the estimation of covariance matrices, the quality of which depends on the size of the data available.

The exploitation of the time dimension correlation structure in spectral estimation is also demonstrated to be a good tool for spectra interpolation over gaps when dealing with in situ data, which often suffer from such missing data. With the proposed methods, we found that spectra can be properly interpolated up to 12 h gap lengths, with the quality of interpolated spectra decreasing when gap length increases.

Some improvements can still be brought to the proposed estimation methods. Especially, for adaptive kde2D, the use of wind data may be used for the improvement of the time evolution model of peak frequencies of the wind sea. Regarding kriging, a choice of an appropriate covariance structure model rather than using the empirical one may improve the spectral estimates.

The proposed estimation methods can improve the tracking of wind sea and swell systems as much as the wave system time evolution is taken into account in spectra estimation. Kriging might also be used when one accesses data from an array of buoys. It can similarly be applied to swell time-space tracking.

Acknowledgment

The data used in the study is the joint property of TOTAL E&P Angola and Esso Exploration Angola (Block 17) Ltd./BP Exploration (Angola) Ltd./Statoil Angola Block 17 A.S./Norsk Hydro Dezassete A.S./SONANGOL, Sociedade Nacional de Combustíveis de Angola, EP. The authors of the papers are grateful to them for their authorization to use the data for this paper.

References

- Aarnes, J., Krogstad, H., 2001. Partitioning sequences for the dissection of directional ocean wave spectra: a review. Part of ENVIWAVE Project.
- Ailliot, P., Maisondieu, C., Monbet, V., 2013. Dynamical partitioning of directional ocean wave spectra. *Probab. Eng. Mech.* 33, 95–102.
- Blackman, R.B., Tukey, J.W., 1959. *The measurement of power spectra: from the point of view of communications engineering*. vol. 190. Dover, New York.
- Borgman, L.E., 1967. *Ocean wave simulation for engineering design*. Technical Report, DTIC Document.
- Bowman, A.W., 1984. An alternative method of cross-validation for the smoothing of density estimates. *Biometrika* 71, 353–360.
- Brockwell, P.J., Davis, R.A., 1991. *Time series: theory and methods*. Springer-Verlag, New York.
- Christensen, R., 1991. *Linear models for multivariate, time series and spatial data*. Springer-Verlag, New York.
- Cliff, A.D., Ord, J.K., 1974. Space-time modeling with an application to regional forecasting. *Trans. Pap. Inst. Br. Geogr.* 66, 119–128 (1991).
- Cressie, N., 1993. *Statistics for Spatial Data: Wiley Series in Probability and Statistics*.
- Elgar, S., 1987. Bias of effective degrees of freedom of a spectrum. *J. Waterw. Port Coast. Ocean Eng.* 113 (1), 77–82.
- Fan, J., Kreutzberger, E., 1998. Automatic local smoothing for spectral density estimation. *Scand. J. Stat.* 25 (2), 359–369.
- Forristal, G.Z., Ewans, K., Olagnon, M., Prevosto, M., 2013. *The West Africa Swell Project (WASP). Proceedings of the 32nd Int. Ocean, Offshore and Arctic Eng. Conf., ASME. Paper Number 11264*.
- Gerling, T.W., 1992. Partitioning sequences and arrays of directional ocean wave spectra into component wave systems. *J. Atmos. Ocean. Technol.* 9 (4), 444–458.
- Gjevik, B., Krogstad, H., Lygre, A., Rygg, O., 1988. Long period swell wave events on the Norwegian Shelf. *J. Phys. Oceanogr.* 18 (5), 724–737.
- Godá, Y., 1977. Numerical experiments on statistical variability of ocean waves. *Rep. Port Harb. Res. Inst.* 16 (2), 3–26.
- Godá, Y., 1978. The observed joint distribution of periods and heights of sea waves. *Coast. Eng. Proc.* 1 (16).
- Godá, Y., 2010. *Random Seas and Design of Maritime Structures*. World Scientific.
- Hannig, J., Lee, T., 2004. Kernel smoothing of periodograms under Kullback–Leibler discrepancy. *Signal Process.* 84 (7), 1255–1266.
- Hanson, J.L., Phillips, O.M., 2001. Automated analysis of ocean surface directional wave spectra. *J. Atmos. Ocean. Technol.* 18 (2), 277–293.
- Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Ewing, J.A., Gienapp, H., Hasselmann, D.E., Kruseman, P., et al., 1973. *Measurements of Wind-Wave Growth and Swell Decay During the Joint North Sea Wave Project (JONSWAP)*.
- Lee, T.C.M., 1997. A simple span selector for periodogram smoothing. *Biometrika* 84 (4), 965–969.
- Longuet-Higgins, M.S., Cartwright, D.E., Smith, N.D., 1963. *Observations of the Directional Spectrum of Sea Waves Using the Motions of a Floating Buoy*.
- Olagnon, M., Robin, A., 1990. Numerical Simulation of Actual Irregular Wave Properties. Olagnon, M., Ewans, K., Forristal, G.Z., Prevosto, M., 2013. *West Africa swell spectral shapes*. Proc. 32nd Int. Ocean, Offshore and Arctic Eng. Conf. OMAE 2013-11228.
- Pawitan, Y., O’Sullivan, F., 1994. Nonparametric spectral density estimation using penalized whittle likelihood. *J. Am. Stat. Assoc.* 89 (426), 600–610.
- Phillips, O.M., 1985. Spectral and statistical properties of the equilibrium range in wind-generated gravity waves. *J. Fluid Mech.* 156, 505–531.
- Portilla, J., Ocampo-Torres, F.J., Monbaliu, J., 2009. Spectral partitioning and identification of wind sea and swell. *J. Atmos. Ocean. Technol.* 26 (1), 107–122.
- Schuster, A., 1898. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terr. Magn.* 3 (1), 13–41.
- Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis*. vol. 26. CRC press.
- Simonoff, J.S., 1996. *Smoothing Methods in Statistics*. Springer.
- Stein, M.L., 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer.
- Stoica, P., Randolph, L., 1997. *Introduction to Spectral Analysis*. vol. 1. Prentice hall, New Jersey.
- Sverdrup, H.U., Munk, W.H., 1947. *Wind Sea, and Swell: Theory of Relations for Forecasting*.
- Wahba, G., 1978. *Automatic Smoothing of the Log Periodogram*. University of Wisconsin, Department of Statistics.
- Wand, M.P., Jones, M.C., 1993. Comparison of smoothing parameterizations in bivariate kernel density estimation. *J. Am. Stat. Assoc.* 88 (422), 520–528.
- Welch, P., 1967. The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 15 (2), 70–73.
- Zhang, Y., Baggeroer, A.B., Bellingham, J.G., 2005. The total variance of a periodogram-based spectral estimate of a stochastic process with spectral uncertainty and its application to classifier design. *IEEE Trans. Signal Process.* 53 (12), 4556–4567.

Étude comparative de fonctions paramétriques d'évolution de la puissance au sein d'un événement de houle

Comparison of parametric shapes for swell events

K. Agbéko Kpogo-Nuwoklo^{#1}, Michel Olagnon^{#2}, Christophe Maisondieu^{#3}, Sabine Arnault^{*4}

[#]Ifremer Centre de Brest, CS10070 - 29280 Plouzané, France

¹kkpogonu@ifremer.fr

²michel.olagnon@ifremer.fr

³christophe.maisondieu@ifremer.fr

^{*}Université Pierre et Marie Curie, Paris, France

⁴sa@locean-ipsl.upmc.fr

Abstract—Knowledge of the time-evolution of sea state parameters within a storm is crucial for optimizing the energetic performance as well as for conducting a reliability assessment of a wave energy device. Because sea states are often composed of two or more superimposed wave systems (swells and wind sea), with different directions and of characteristics that evolve over time, their refined description should distinguish the time-evolution of wind sea systems from time-evolution of swell systems. This study is focused on the time-evolution of swell systems (swell events) with the objective of identifying a parametric model providing a good statistical distribution of the most energetic waves in a swell event which can be used to model the typical time-evolution of wave power within a swell event. Based on the analysis of wave power time series, three parametric shapes for power evolution within a swell event are proposed (Triangular, Log-normal and Gamma). Analysis of statistical results shows that the Gamma-model gives the best results and therefore may be considered as a good model for the time-evolution of the wave power within the swell event.

Index Terms—wave energy; swell event modelling; swell energy; time-evolution of wave power

I. INTRODUCTION

Estimation of wave energy potential as well as design, deployment and maintenance of wave energy harvesters require a refined description of sea states. The common description of sea states generally considers the existence of one single consistent wave system, in order to reduce the number of representing parameters. Though this method makes data management easier and provides sufficient elements for many analyses and applications, it does not reflect the complexity of actual encountered sea states. Sea states are often composed of two or more superimposed wave systems (swells and wind sea) with different energetic contributions to the total energy, and with characteristics that evolve over time. Previous studies [1], [2] have shown that the description of the sea state refined at the level of wave systems dramatically enhances the precision of wave energy potential assessment.

Furthermore, because installation and maintenance of wave energy devices can only be done during favourable sea states conditions, the way wave systems parameters evolve over time is crucial. Knowledge of the time-evolution of sea state parameters within an event is also important for optimizing the energetic performance of wave energy harvesting systems. The concept of event is related here to the extension along the time dimension of the wave systems defined for individual sea

states. Indeed, an event is defined as a coherent succession of wave systems, obtained after partitioning of consecutive sea states and that can be linked to the same meteorological origin (storm, low pressure system, etc). Previous studies [3], [4], [5] that have addressed this issue consider events identified from the global sea state parameters without distinction between wind sea systems and swell systems. However, since wind sea and swell systems are generated by different weather systems and evolve according to different dynamics, one should distinguish swell events from wind sea events. Swell and wind sea events can be identified from time-series of directional spectra and many studies have been dedicated to this topic [6], [7], [8].

This study focuses on swell systems with the aim of analysing and modelling the typical time-evolution of wave power within swell events. Swells are often characterized by longer wave periods and can be considered highly energetic despite usually moderate significant wave heights. Additionally, the energy of a swell is focused within a rather narrow spectral band, hence is probably more beneficial for resonant wave energy converters. More specifically, our objective is to identify a parametric shape that can be used to model the typical time-evolution of wave power within a swell event. Parametric shape has the advantage that it requires a small yet sufficient number of parameters from which the whole event evolution can be accurately reconstructed with an analytical function. When the parametric shape is defined, swell climate can be derived from the joint distributions of the parameters. This approach to obtain swell climate also has the advantage of keeping the time coherence in swell systems.

In this study we use the same set of sea surface displacement data measured by a buoy located off the coast of West Africa as in [9]. Swell events are first identified within sea states using a partitioning method. Visual inspection of the behaviour of the time-evolution of the wave power within swell events then suggests to study three parametric shapes (Triangular, Log-normal and Gamma) to model such events. These models are then fitted individually to the actual swell events using weighted least square methods. An important criterion to take into account when choosing the best parametric shape is the capacity of this to accurately reproduce the statistical distribution of the most energetic waves, required in the design of wave energy harvesters against extreme loads and fatigue.

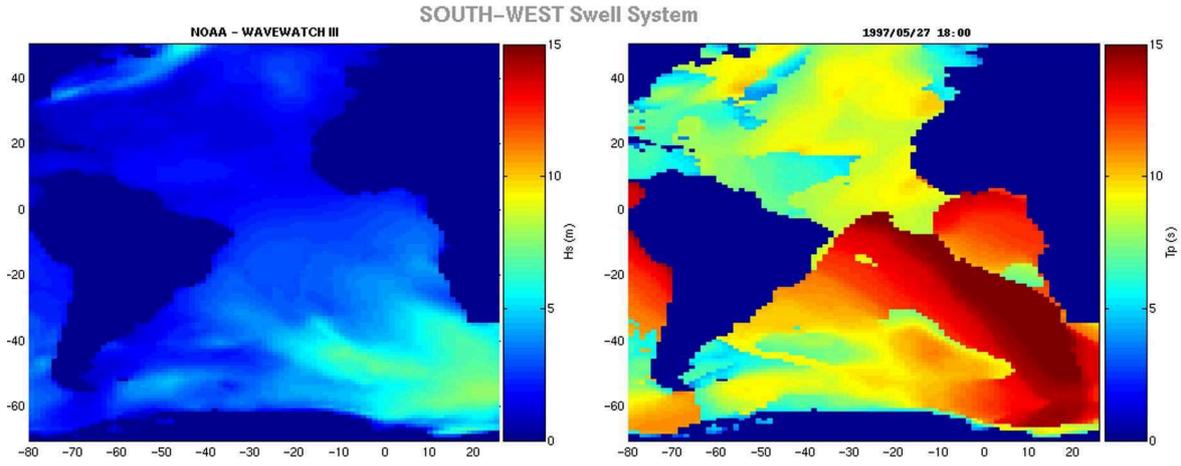


Fig. 1. West Africa sea state conditions: (left) significant wave height , (right) peak period . The sea states are dominated by swells coming mostly from the South Atlantic. (from [10])

Thus, the statistical distributions of extreme wave crests are first computed from the actual measured swell events, and then from the fitted parametric shapes. The results are compared so as to identify the shape providing the closest distributions of extreme wave crests to that obtained from the actual events.

II. SWELL EVENTS IDENTIFICATION FROM TIME-SERIES OF WAVE BUOY DATA

The data used for this study was obtained from in situ measurement conducted off the coast of West Africa [9]. Sea states in that area are dominated by swells; the energy provided by the swells accounts for over 75% on average of the total energy of the sea states. Additionally, swells propagating towards West African locations generally originate from a succession of very repetitive storms (every three days on average) travelling from Cape Horn or the coast of Argentina to Good Hope [10] (see Figure 1). Two reasons make West Africa potentially interesting in terms of wave energy harvesting. The first one is related to the focusing of swell energy within a rather narrow spectral band resulting from the dispersion process over the long distances travelled by the waves. This may be of interest for an optimal exploitation of resonant wave energy converters. The second reason is related to the relatively low seasonal variability of these swells providing a rather constant distribution of the convertible energy all year long in spite of slightly weaker summers in the southern hemisphere [11].

The primary data are the displacements $\{z(t), x(t), y(t)\}$ recorded by a Heave/Surge/Sway buoy anchored off the coast of Angola in 1450 m water depth from which wave spectra are derived. Measurements recorded with a 1.28 Hz sampling frequency extend from March 2001 to March 2003, with some gaps in between due to sensor problems. For the purpose of this study it was necessary to extract from the original data set a new data base consisting in swell events only. The successive steps to achieve this are summarised as a block diagram in Figure 2. The process can be divided in two parts.

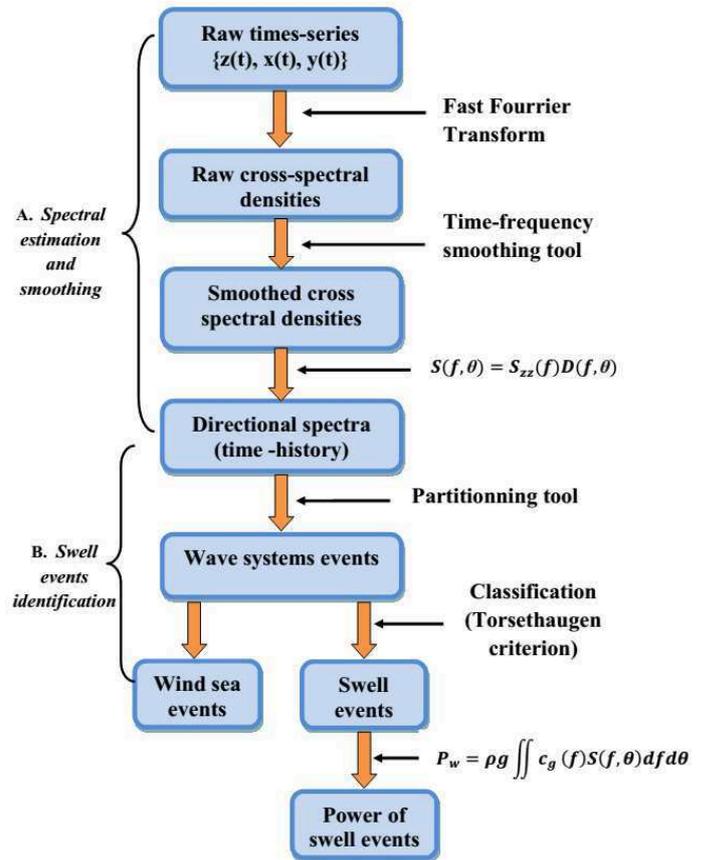


Fig. 2. Block diagram of swell events identification starting from time-series of sea surface displacements

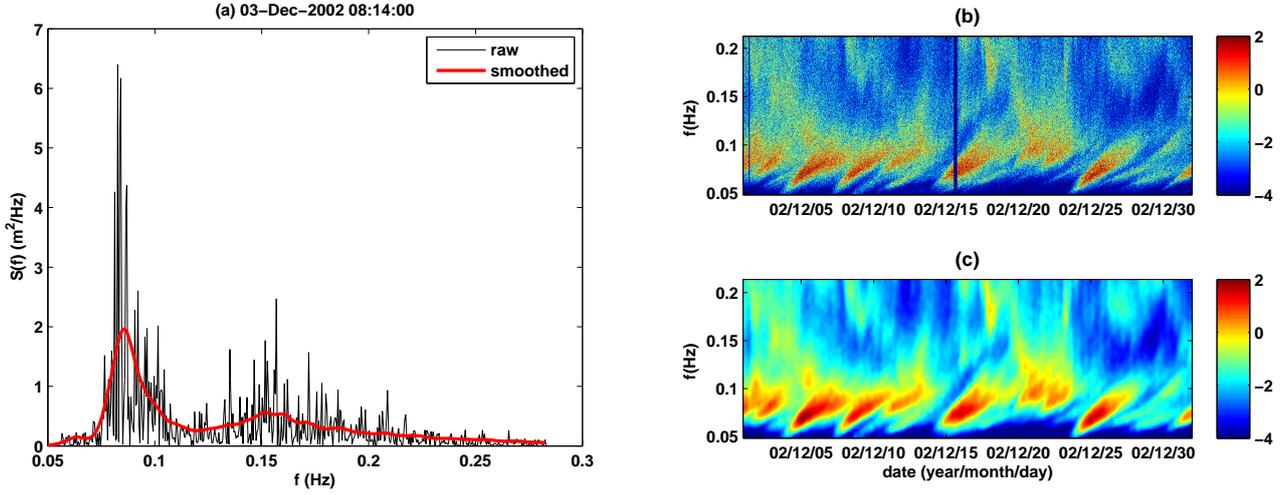


Fig. 3. Sea wave spectra estimation and interpolation applied to in-situ data: (left) example of estimated spectrum, (right) time-history of spectra. The linear energetic patterns observable in low frequencies ranges are typically swell events signatures. (Log-scale)

Directional spectra are first estimated from the records of the buoy displacements; then they are partitioned so as to identify time-series of swell events.

A. Spectral estimation and smoothing

Sea surface displacement data were split into short subsets (30 min) and raw auto- and cross-spectral densities were computed using discrete Fourier transform on each block. Figure 3 shows (a) an example of a raw auto spectral density (in black) and (b) a one month time-history of raw auto spectral densities for the sea surface elevation $z(t)$. Temporal structures can be observed in low frequency ranges, corresponding to the time-evolution of swell events, and also in high frequency ranges related to the local wind effect. Classically, non-parametric wave spectrum estimation consists of periodogram (raw spectral density) smoothing with frequency windowing [12], [13]. Here, rather than only smoothing individual raw spectral densities, we considered the overall time-history of raw spectral densities as a random field depending on both frequency and time, and we smoothed this field simultaneously in frequency and time. Since the time-evolution of the wave systems reflects typical non-stationary behaviours, such as the linear increasing trend of the peak frequency of swell components with time, a simple two dimensional averaging would not have been appropriate. Hence, we used for time-frequency smoothing an adaptive 2D kernel density estimator i.e. imposing a kernel that adapts to the dynamical evolution of wave systems. The method is fully described in [14] where it is shown to improve spectral estimation and to be an efficient tool for the interpolation requested for filling gaps in the case of missing spectra.

Following this method, the auto-spectral density of $z(t)$ at a given frequency and time, denoted by $\hat{S}_{zz}(f, t)$ is obtained

as

$$\hat{S}_{zz}(f, t) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\det(\mathbf{H})} K \left\{ \mathbf{H}^{-1} \left(\begin{bmatrix} f \\ t \end{bmatrix} - \begin{bmatrix} f_i \\ t_i \end{bmatrix} \right) \right\} I_{zz}(f_i, t_i) \quad (1)$$

$I_{zz}(f_i, t_i)$ are the raw spectral densities, n is the number of observation points, K is the Kernel function, and \mathbf{H} is a full 2×2 bandwidth matrix. $\mathbf{H} = \begin{bmatrix} h_f & \lambda \\ \lambda & h_t \end{bmatrix}$ where the elements h_f and h_t control the smoothing in frequency and time ranges respectively and λ is used to control the time-frequency evolution of wave systems. It is well known that the choice of \mathbf{H} is much more crucial than the choice of K (e.g., see [15]). If the smoothing parameters are too small, the estimate are too noisy, and exhibits high variability. If the smoothing parameter is too large, then the estimate may miss key features due to over-smoothing washing out small details. The bandwidth of the kernel is obtained by minimizing the mean integrated squared error function which is the most frequently optimality criterion used to select kernel bandwidth (see [14]). Figure 3 (c) shows the estimated smoothed time-history of spectra obtained after applying the equation 1 to the raw time-history of spectra plotted in Figure 3 (b). In Figure 3 (a) an example of individual raw spectrum (black) is plotted with its corresponding estimated smoothed spectrum (red).

The auto- and cross-spectral densities \hat{S}_{xx} , \hat{S}_{yy} , \hat{S}_{zx} , \hat{S}_{zy} and \hat{S}_{xy} may also be estimated separately in the same manner. Directional spectra are computed as

$$S(f, \theta) = S_{zz}(f) D(f, \theta), \quad (2)$$

where $\cosine - 2s$ is chosen as the directional spreading function:

$$D(f, \theta) = N(f) \cos^{2s(f)} \left(\frac{\theta - \theta_m(f)}{2} \right), \quad (3)$$

where $N(f)$ is a normalisation factor, $\theta_m(f)$ the mean wave direction, and $s(f)$ the directional spreading parameter. $\theta_m(f)$

and $s(f)$ are obtained respectively from the first Fourier coefficients as

$$\theta_m(f) = \arg(a_1(f) + ib_1(f)) \quad (4)$$

$$s(f) = \frac{2}{\sqrt{2(1 - (a_1^2(f) + b_1^2(f)))}} - 1, \quad (5)$$

where

$$a_1(f) = \frac{Q_{zx}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, \quad (6)$$

$$b_1(f) = \frac{Q_{zy}(f)}{[C_{zz}(f)(C_{xx}(f) + C_{yy}(f))]^{1/2}}, \quad (7)$$

with C_{ij} and Q_{ij} respectively the real and imaginary parts of the cross-spectrum between wave records i and j .

B. Identification of the swell events from the time-history of spectra

Sea states generally result from the combination of local wind waves and swell radiated from distant storms. The resulting spectra in such case exhibit two or more peaks. It is possible to separate the significant peaks in the sea state spectrum and this operation is often referred to as partitioning. Although partitioning is a purely formal procedure, the partitions can still be physically interpreted as separate independent wave systems (wind sea or swell) [6]. Furthermore, partitions obtained at the same point at different consecutive times and dates can be associated in order to link them to the same originating meteorological event. A swell event in particular is a sequence of successive swell partitions that can be traced back to a unique distant storm.

The existing methods for the identification of wave systems events generally comprise two steps:

- The first step is the partitioning of individual sea states. Some authors use parametric optimization algorithms where the separation is implicit in the procedure [16], [17]. Others use image segmentation algorithms of which the most widely used is the watershed algorithm [18], [6], [7], [19].
- Once wave systems are identified in each sea state, the temporal tracking of their parameters is carried out in order to link them to the same event. Many authors use empirical criteria often based on simple distance criteria in the parameter space (significant wave height H_s , peak frequency f_p , and peak direction θ_p) [7], [6].

The main issue with partitioning algorithms is that they usually find more partitions than can be expected. It is thus common to use somewhat arbitrary parametrizations to assess whether partitions are significant or spurious. Moreover, empirical criteria are also used in the tracking step and these make the implementation of automated procedures difficult, if not impossible, to calibrate [19].

In [8] a wave system events identification method that combines both the partitioning and tracking steps was proposed. This method is also based on the watershed, but differs from other methods in that the watershed is applied in 3D i.e.

directly to the whole time-history of directional wave spectra. Thus, it allows direct extraction of the wave systems events and does not require the tracking phase. The time-domain expansion of this spectra partitioning technique dramatically improved the identification of the wave systems. Hence this method is adopted in this study to identify the wave systems events.

Waves systems may be classified into two types: wind sea or swell. When wind data is available, the most commonly used method for wind sea selection is the wave age criterion [6], [7]. When wind data is not available, a default constant separation frequency is often set, tuned to the local dynamics of winds and waves. However, the separation frequency normally evolves in time, and arbitrarily fixing a constant value could introduce errors in the classification. As a consequence in this study, an H_s dependent separation frequency is adopted based on the work of Torsethaugen and Haver [20]. For the locally fully developed sea, the spectral peak frequency f_{pf} is a function of the significant wave height,

$$f_{pf} = \frac{1}{a_f H_s^{1/3}} \quad (8)$$

where a_f is a constant derived from the JONSWAP experiment, and is slightly dependent on the fetch (see [20]). For instance $a_f = 6.6sm^{-1/3}$ for 370 Km of fetch and $a_f = 5.3sm^{-1/3}$ for 100 Km of fetch. This classifies sea states as wind sea for $f_p > f_{pf}$ and as swell for $f_p < f_{pf}$. For the data used in this study, after testing different values of a_f , we obtain a good result with $a_f = 6.6sm^{-1/3}$. Finally, an event is classified as a swell event if its peak frequency is below the separation frequency at least 75% of the time, otherwise as a wind sea event. Figure 4 shows an example of extracted swell events for the month of December 2002. The top panel shows the time history of the omni-directional spectra and the watershed lines. The other panels show for each event, the corresponding evolutions of the significant wave height, the mean energy frequency and the wave power. Swell events are easily recognizable through their peak frequencies that fall below the separation frequency and whose time-evolution are almost linear.

Finally, a total of 201 swell events are extracted from the available dataset. Even though events are usually characterized in terms of significant wave height, in the following we will consider wave power (energy flux) as it is best suited to the purpose of this study. The wave power quantifies the interdependence of significant wave height and mean period. It also presents a smoother time-evolution than H_s within a swell event which is suitable for modelling. The wave power P_w is expressed as

$$P_w = \rho g \int \int c_g(f, h) S(f, \theta) df d\theta \quad (9)$$

where $c_g(f, h)$ denotes the group celerity of wave frequency f , which depends on the local water depth h (m). Constants ρ and g stand for sea water mass density ($\rho \approx 1025kgm^{-3}$) and the gravitational constant respectively. It may be shown

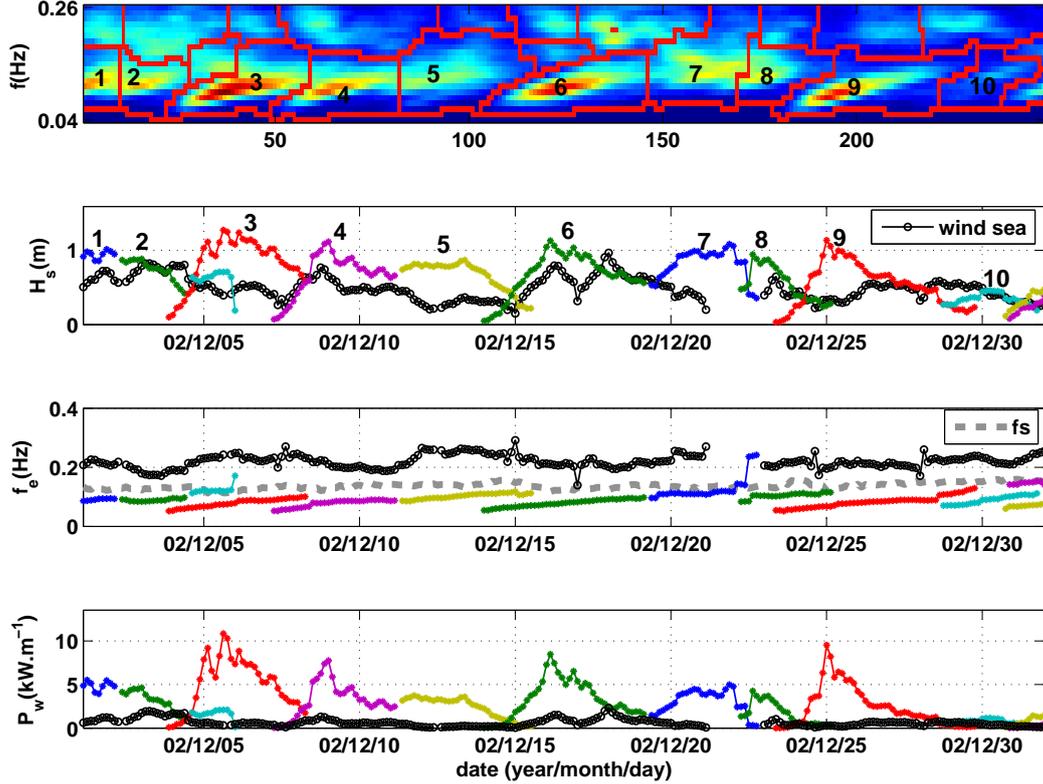


Fig. 4. Example of wave systems events identification after partitionning. The top panel shows the time history of omni-directional spectra and the numbers indicate identified swell events. For each swell event, the corresponding wave parameters significant wave height (H_s), mean frequency (f_e) and wave power (P_w) are plotted with the same colour. The wind sea is plotted in black line with circles

that the group celerity in deep water ($h \rightarrow \infty$) is given by $c_g(f, \infty) = g/(4\pi f)$. Hence, the approached expression for the wave power (in kWm^{-1}) in deep water

$$P_w = \frac{\rho g^2}{64\pi} H_s^2 T_e \approx 0.49 H_s^2 T_e \quad (10)$$

with T_e the mean energy period.

III. PARAMETRIC SHAPES FOR THE TIME-EVOLUTION OF WAVE POWER WITHIN A SWELL EVENT

Swells are regular surface gravity waves, generated by distant weather systems and it can be expected that the time-evolution of sea state parameters within swell events exhibit similar behaviours. For instance it is well known that the peak frequency increases linearly in time along a swell event according to the wave dispersion relationship in deep water, as can be observed in Figure 4 (middle panel). Similarly, previous studies [21], [22], [6], [9] indicate that, within a swell event, the significant wave height increases with the arrival of the swell front, reaches a maximum and then decreases.

Our objective here is to identify a unique parametric shape that can be used to model the time-evolution of wave power

within a swell event. This is similar to finding a parametric spectral shape (Pierson-Moskowitz, Jonswap, etc) for individual spectra. The parametric shape for a swell event has the advantage that it requires a small yet sufficient number of parameters from which the whole event evolution can be accurately reconstructed with an analytical function.

Figure 5 shows the time-evolution of the power of all the extracted swell events shifted in time so that the maximum power is at zero lag. It can be observed that P_w increases at first, reaching a maximum before decreasing and that the rising phase is steeper than the descending one. Such a behaviour could not be modelled by a symmetric shape. Taking this into account, three different parametric shapes (Triangular, Log-normal and Gamma) are selected to model the time-evolution of wave power within a swell event.

- **Triangular shape**

$$P_w(t) = \begin{cases} E\alpha_1(t - t_0) + P_{w_{max}} & \text{if } t_i \leq t < t_0 \\ E\alpha_2(t - t_0) + P_{w_{max}} & \text{if } t_0 \leq t \leq t_f \end{cases}$$

where α_1 and α_2 characterise, the ascending and de-

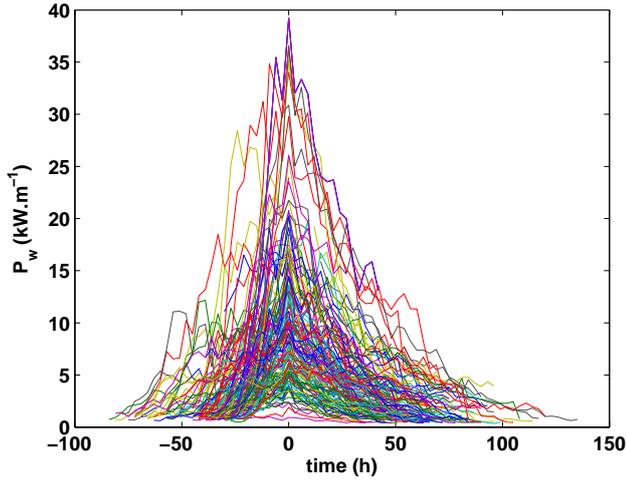


Fig. 5. Plot of wave power of 201 extracted events; with the time of their maxima is set at zero lag

ascending slopes respectively, t_i and t_f are respectively the starting and the ending times of the event, t_0 the time of the wave power maximum, $P_{w_{max}}$ the maximum wave power, and

$$E = \int_{t_i}^{t_f} P_w(t) dt \quad (11)$$

in $kWh.m^{-1}$ is the swell event total energy per meter of wave crest.

- **Log-normal shape**

$$P_w(t) = \frac{E}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(t) - \mu)^2}{2\sigma^2}\right), t_i \leq t \leq t_f \quad (12)$$

where the parameter E is the same as in (11) and σ and μ are the classical Log-normal parameters.

- **Gamma shape**

$$P_w(t) = \frac{E}{\tau^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\tau}\right), t_i \leq t \leq t_f \quad (13)$$

where $\Gamma(\cdot)$ is the Gamma function, α is the shape parameter, τ the characteristic duration of the event and E is the same as in (11).

In the process, power is at first normalized by the total energy (E) of the event. Parametric shapes are then fitted to the actual events using a weighted least squares method. As our goal is also to obtain a good distribution of the most energetic waves in the parametric shapes, we choose the weights to be proportional to the values of the wave power.

Figure 6 shows an example of a swell event and the three fitted parametric shapes. Visual inspection reveals that the Log-normal and the Gamma shape fit the swell event much better compared to the Triangular shape. The Triangular shape is less flexible in its ability to properly fit a wave power time-evolution within a swell event. Indeed, using a triangle with a

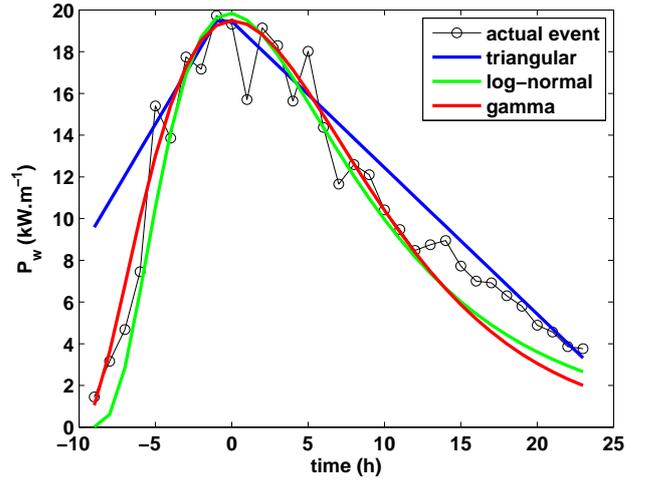


Fig. 6. Triangular, Log-normal and Gamma shapes fitted to a swell event

shorter base could provide a better fit of the medium and low values of P_w but would then misrepresent the highest values which are to be accurately taken into account as it is also important to correctly model the most energetic components of the event. Additional criteria are necessary to assess the actual efficiency of Log-normal and Gamma shapes. The chosen evaluation method consists in comparing the statistical distribution of the most energetic wave crest obtained using the parametric shapes with the distribution obtained from the actual swell events.

IV. SECOND-ORDER DISTRIBUTION OF WAVE CRESTS IN A SWELL EVENTS

The design of wave energy harvesters against extreme loads and fatigue requires accurate knowledge of the most energetic waves. To evaluate how the parametric shapes give good statistical distributions of extreme waves compared to the ones given by the actual swell event, we use the second-order distribution of the wave crest in the swell event as an indicator. Two different models are used for the second-order distribution of the wave crest: the Forristall model [23] and the Prevosto model [24]. For this assessment, we compare the probability of exceedance of the crest height in a swell event given by both the Triangular, the Log-normal and the Gamma shapes in order to derive the parametric shape that gives the closest results to the ones assessed using the actual swell events.

A. Probability of exceedance of the crest height in swell event

To the first order, the water surface can be represented as a Gaussian noise with a reasonably narrow frequency band. The crest heights then have the same distribution as the envelope of the noise, which has the Rayleigh distribution. In fact, because of nonlinearity, the Rayleigh law tends to under-predict the crests amplitude and to overpredict the troughs amplitude. More realistic wave crest distributions may be obtained using the second-order theory.

The probability of exceedance of the crest height varies with the sea states during the event. If we assume that crest heights are stochastically independent of each other, the probability of exceedance of the maximum wave crest during a swell event whose duration is D may be written as:

$$P(\eta_{max} > \eta|[0, D]) = 1 - \exp \left\{ \sum_{i=1}^N \frac{\Delta t_i}{T_{z_i}} \ln[1 - P_i(\eta_c > \eta)] \right\} \quad (14)$$

where, for the i th sea state of the event, $P_i(\eta_c > \eta)$ is the non-linear crest exceeding probability, Δt_i , T_{z_i} and $\Delta t_i/T_{z_i}$ are respectively the duration, the mean zero-crossing period and the number of zero-crossing waves. Note that N defines the number of sea states in the event so that $D \equiv \sum_{i=1}^N \Delta t_i$. Following Borgman [25], the limit of (14) as Δt_i tends to zero, is:

$$P(\eta_{max} > \eta|[0, D]) = 1 - \exp \left\{ \int_0^D \frac{\ln[1 - P(\eta_c > \eta)]}{\bar{T}_z} dt \right\} \quad (15)$$

In equation 15 the only unknown term is $P(\eta_c > \eta)$, the non-linear crest exceedance probability. Different models of $P(\eta_c > \eta)$ can be found in the literature. In this study we will focus on two of them: the Forristall model and the Prevosto model.

One can also calculate the maximum expected wave crest in a swell event as the integral over $(0, \infty)$ of the probability of exceedance of the wave crest,

$$\bar{\eta}_{max} = \int_0^\infty \left(1 - \exp \left\{ \int_0^D \frac{\ln[1 - P(\eta_c > \eta)]}{\bar{T}_z} dt \right\} \right) d\eta. \quad (16)$$

1) *Forristall model* [23]: This model is based on a Weibull law with the two parameters written as polynomials in the steepness and Ursell number. In this model, the probability of exceedance of the crest height is defined as the probability that a crest height is greater than η in a sea state with significant wave height H_s :

$$P(\eta_c > \eta) = \exp \left[- \left(\frac{\eta}{\alpha H_s} \right)^\beta \right]. \quad (17)$$

$$S_1 = \frac{2\pi H_s}{g T_e^2} \quad (18)$$

is the steepness and

$$U_r = \frac{H_s}{k_1^2 d^3} \quad (19)$$

the Ursell number with water depth d . Starting from simulations based on a synthetic directional spectrum data base and different water depths, the coefficients α and β of the polynomials were identified respectively as

$$\alpha = \frac{1}{\sqrt{8}} + 0.2568S_1 + 0.0800U_r \quad (20)$$

$$\beta = 2 - 1.7912S_1 - 0.5302U_r + 0.2824U_r^2 \quad (21)$$

2) *Prevosto model* [24]: The model is based on the narrowband and infinite crested wave equation:

$$P(\eta_c > \eta|H_s; f_m; d) = \exp \left(-8 \frac{\left(\frac{-1 + \sqrt{1 + 4\alpha(\eta - \beta)}}{2\alpha} \right)^2}{H_s^2} \right) \quad (22)$$

This model uses the exact asymptotic narrowband transfer coefficients given in [24]. In order to take into account the spectral bandwidth and directional spreading, the same model is considered but with a modified significant wave height \widetilde{H}_s and mean frequency \widetilde{f}_m .

$$\widetilde{H}_s = \alpha_{H_s} H_s, \quad \widetilde{f}_m = \alpha_{f_m} f_m$$

In looking at different directional spectral climatologies, including the WACSIS [26] data base, and different water depths, the best α_{H_s} and α_{f_m} formulations have been determined from simulations and theoretical considerations to be:

$$\alpha_{H_s} = 1 - \frac{1}{2} (\tanh(k_m d) - 0.9) \sqrt{\frac{2}{1 + s}} \quad (23)$$

$$\alpha_{f_m} = \frac{1}{1.23} \quad (24)$$

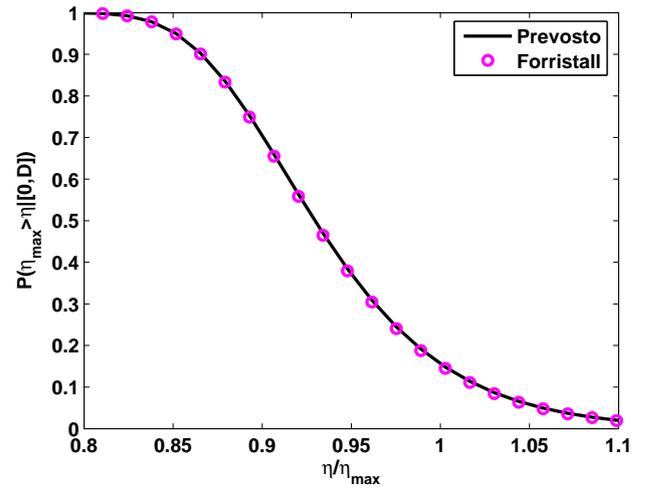


Fig. 7. Comparison of the probability of exceedance of crests height for the swell event plotted in Figure 6

A comparison of the probability of exceedance of the wave crest during a swell event obtained by the Prevosto model and the Forristall model, shows that both models give the same results (Figure 7). Such agreement is also mentioned in [27]. Indeed the main difference between these two models is the consideration of directional spreading of the wave systems in the Prevosto model. The directional spreading is very small for swell and is not considered in this study. Hence, only the Forristall model is used in the following to compute the probability of exceedance of the wave crests during a swell event.

B. Comparison of the probability of exceedance of the wave crests during a swell event

1) *Typical case (swell event plotted in Figure 6):* The probability of exceedance of the wave crest during a swell event is estimated using equation (15). The actual swell event is now replaced by its three parametric shapes (Triangular, Log-normal and Gamma) and the probability is estimated for each of them.

Figure 8 shows the probability of exceedance of the wave crest during the swell event. The probability given by the Gamma shape is the closest to the one given considering the actual swell event.

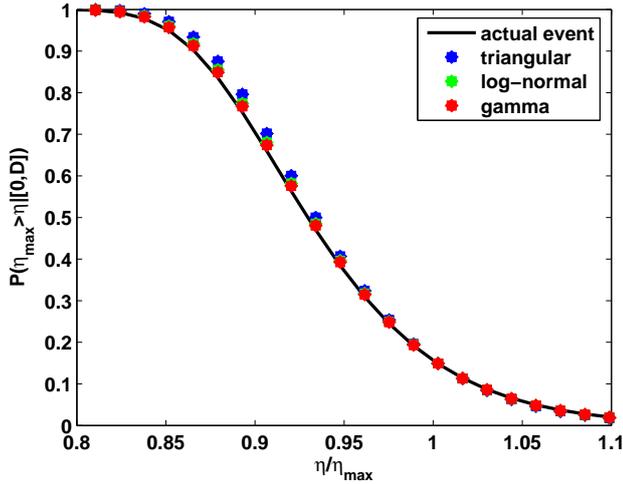


Fig. 8. Probability of exceedance of crests height for the swell event plotted in Figure 6 using Forristall shape.

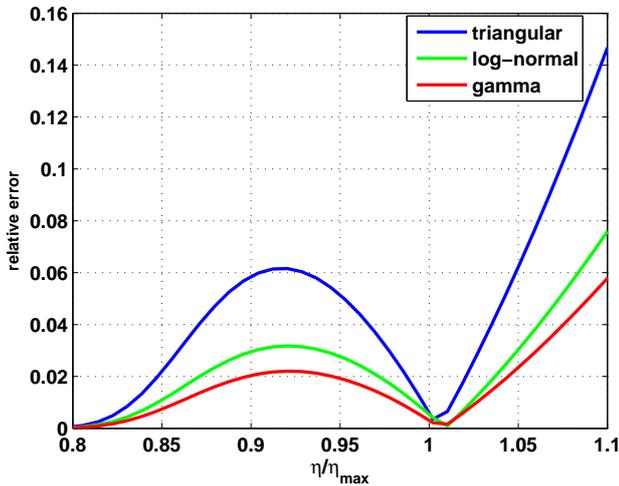


Fig. 9. Relative error of the probability of exceedance of the wave crests given by the actual swell event and the one given by the parametric shapes (example of Figure 6)

To quantify how $P(\eta_{max} > \eta | [0, D])$ given by the parametric shapes are close to the one obtained with the actual swell

events, we compute a relative error defined as

$$er = \left| \frac{P_{act}(\eta_{max} > \eta | [0, D]) - P_{mod}(\eta_{max} > \eta | [0, D])}{P_{act}(\eta_{max} > \eta | [0, D])} \right| \quad (25)$$

where $P_{act}(\eta_{max} > \eta | [0, D])$ is the probability of exceedance of the wave crest given by the actual swell event and $P_{mod}(\eta_{max} > \eta | [0, D])$ the one given by the parametric shape. Figure 9 shows the absolute relative error for the same event plotted in Figure 6. A general observation is that the relative error increases with crest height, reaches a maximum, and then decreases to reach a minimum value. This minimum corresponds to the maximum expected crest height in the event. This is well in line with the chosen fitting method i.e. allotting the highest weights to the maximum of the wave power. The Gamma and Log-normal shapes displayed lower errors compared to the Triangular shape, and between the Gamma and the Log-normal, the first is slightly better.

2) *All the swell events:* Figure 10 shows the distribution of the relative error computed for some values of normalized wave crest heights, for the 201 identified swell events from the available data. The distribution of the relative error appears to be lower for the Gamma shape than for the others and between the Log-normal shape and the Triangular shape the former gives better results in this case. In light of these results, the Gamma shape can be considered as the best choice for modelling the time-evolution of the wave power within a swell event.

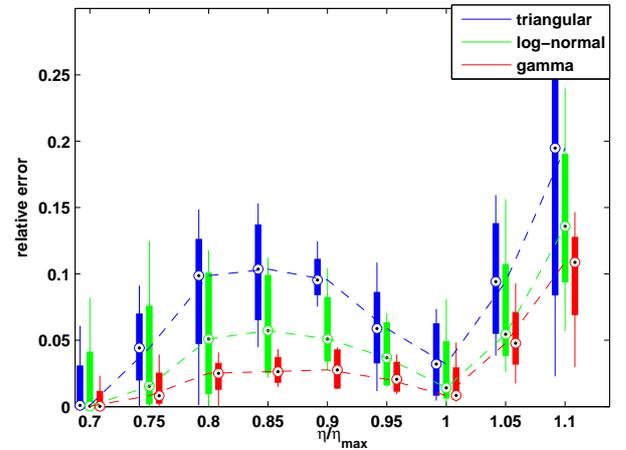


Fig. 10. Boxplot of the difference between the relative error of the probability of exceedance of the wave crest

V. CONCLUSION

We studied the ability of different parametric shapes to model the time-evolution of wave power within a swell event. Swell events were first extracted from the time-history of directional spectra using a new method that consists in applying the watershed algorithm in 3D i.e. on the time-history of directional spectra. This method has the advantage that it

does not require the tracking step often used in other existing methods, which can make automated processing difficult to implement.

Three parametric shapes (Triangular, Log-normal and Gamma) were used to fit the power evolution of the identified swell events. For each event and its parametric shapes, we calculated the probability of exceedance of the non-linear wave crest using the Borgman expression. We have used either the Forristall model or the Prevosto model for the second-order distribution of the wave crest. Both models gave the same results as pointed out in earlier studies. Among the three parametric shapes, the probabilities of exceedance of the non-linear wave crest obtained by the Gamma shape are the closest to the ones given by the actual swell events. The Log-normal shape also gives much better results than the Triangular shape. The Gamma shape appears to be the best choice for modelling the time-evolution of the wave power within a swell event according to the criterion of the statistical distribution of wave crests. This parametric shape reproduces well both extremes and other wave power values.

Through statistics on the parameters of the Gamma shape, synthetic swell events can be simulated and swell climates can be fully reconstructed. This approach to obtain swell climates has the advantage of keeping the time dependence in swell systems. Finally, it should be noted that wind sea systems can also be studied in the same manner. Ongoing research is dedicated to this topic.

ACKNOWLEDGMENT

The data used in the study is the joint property of TOTAL E&P Angola and Esso Exploration Angola (Block 17) Ltd. / BP Exploration (Angola) Ltd. / Statoil Angola Block 17 A.S. / Norsk Hydro Dezassete A.S. / SONANGOL, Sociedade Nacional de Combustiveis de Angola, EP. The authors of the paper are grateful to them for the authorization of using the data for this study. The authors are also grateful to the Natural Ocean Engineering Laboratory (Reggio Calabria, Italia) where a part of this work was done.

REFERENCES

- [1] M.-A. Kerbiriou, M. Prevosto, C. Maisondieu, A. Babarit, and A. Clément, "influence of an improved sea state description on a wave energy converter production," *Proceedings of OMAE 2007 - 29254*, San Diego, 2007.
- [2] J.-B. Saulnier, M. Prevosto, and C. Maisondieu, "Refinements of sea state statistics for marine renewables: A case study from simultaneous buoy measurements in portugal," *Renewable Energy*, vol. 36, no. 11, pp. 2853–2865, 2011.
- [3] B. Drew, A. Plummer, and M. N. Sahinkaya, "A review of wave energy converter technology," *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 223, no. 8, pp. 887–902, 2009.
- [4] M. O'Connor, T. Lewis, and G. Dalton, "Weather window analysis of irish west coast wave data with relevance to operations & maintenance of marine renewables," *Renewable energy*, vol. 52, pp. 57–66, 2013.
- [5] F. Arena, V. Laface, G. Malara, A. Romolo, A. Viviano, V. Fiamma, G. Sannino, and A. Carillo, "Wave climate analysis for the design of wave energy harvesters in the mediterranean sea," *Renewable Energy*, vol. 77, pp. 125–141, 2015.
- [6] J. L. Hanson and O. M. Phillips., "Automated analysis of ocean surface directional wave spectra," *Journal of atmospheric and oceanic technology*, vol. 18, no. 2, pp. 277–293, 2001.
- [7] J. Aarnes and H. Krogstad, "Partitioning sequences for the dissection of directional ocean wave spectra: A review," *Part of ENVIWAVE Project*, 2001.
- [8] K. A. Kpogo Nuwoklo, M. Olagnon, and Z. Guédé, "Wave spectra partitioning and identification of wind sea and swell events," in *Proc. 33rd Int. Ocean, Offshore and Arctic Eng. Conf., OMAE*, 2014.
- [9] M. Olagnon, K. A. Kpogo-Nuwoklo, and Z. Guédé, "Statistical processing of west africa wave directional spectra time-series into a climatology of swell events," *Journal of Marine Systems*, vol. 130, pp. 101–108, 2014.
- [10] G. Forristal, K. Ewans, M. Olagnon, and M. Prevosto, "The west africa swell project (wasp)," in *Proceedings of the 32nd Int. Ocean, Offshore and Arctic Eng. Conf., ASME. paper number 11264*, 2013.
- [11] R. Nerzic, C. Frelin, M. Prevosto, V. Quiniou-Ramus *et al.*, "Joint distributions of wind/waves/current in west africa and derivation of multivariate extreme i-form contours," in *The Seventeenth International Offshore and Polar Engineering Conference*. International Society of Offshore and Polar Engineers, 2007.
- [12] A. Schuster, "On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena," *Terrestrial Magnetism*, vol. 3, no. 1, pp. 13–41, 1898.
- [13] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *Audio and Electroacoustics, IEEE Transactions on*, vol. 15, no. 2, pp. 70–73, 1967.
- [14] K. A. Kpogo-Nuwoklo, P. Ailliot, M. Olagnon, Z. Gud, and S. Arnault, "Improvement ocean wave spectra estimation using the temporal structure of wave systems," *Coastal Engineering*, vol. 96, no. 0, pp. 81 – 91, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378383914002051>
- [15] M. Wand and M. Jones, "Kernel smoothing. vol. 60 of monographs on statistics and applied probability," 1995.
- [16] P. Ailliot, C. Maisondieu, and V. Monbet, "Dynamical partitioning of directional ocean wave spectra," *Probabilistic Engineering Mechanics*, vol. 33, pp. 95–102, 2013.
- [17] K. Ewans, E. Bitner-Gregersen, and C. G. Soares, "Estimation of wind-sea and swell components in a bimodal sea state," *Journal of offshore mechanics and Arctic engineering*, vol. 128, no. 4, pp. 265–270, 2006.
- [18] T. W. Gerling, "Partitioning sequences and arrays of directional ocean wave spectra into component wave systems," *Journal of atmospheric and Oceanic Technology*, vol. 9, no. 4, pp. 444–458, 1992.
- [19] J. Portilla, F. J. Ocampo-Torres, and J. Monbaliu, "Spectral partitioning and identification of wind sea and swell," *Journal of atmospheric and oceanic technology*, vol. 26, no. 1, pp. 107–122, 2009.
- [20] K. Torsethaugen and S. Haver, "Simplified double peak spectral model for ocean waves," in *Proceedings of 14th international offshore and polar engineering conference. ISOPE*, vol. 3, 2004, pp. 76–84.
- [21] F. E. Snodgrass, G. W. Groves, K. Hasselmann, G. Miller, W. Munk, and W. Powers, "Propagation of ocean swell across the pacific," *Philosophical Transactions for the Royal Society of London. Series A, Mathematical and Physical Sciences*, pp. 431–497, 1966.
- [22] B. Gjevik, H. Krogstad, A. Lygre, and O. Rygg, "Long period swell wave events on the norwegian shelf," *Journal of physical oceanography*, vol. 18, no. 5, pp. 724–737, 1988.
- [23] G. Z. Forristall, "Wave crest distributions: Observations and second-order theory," *Journal of Physical Oceanography*, vol. 30, no. 8, pp. 1931–1943, 2000.
- [24] M. Prevosto, H. E. Krogstad, and A. Robin, "Probability distributions for maximum wave and crest heights," *Coastal Engineering*, vol. 40, no. 4, pp. 329–360, 2000.
- [25] L. E. Borgman, "Maximum wave height probabilities for a random number of random intensity storms," *Coastal Engineering Proceedings*, vol. 1, no. 12, 1970.
- [26] G. Z. Forristall, S. F. Barstow, H. E. Krogstad, M. Prevosto, P. H. Taylor, and P. S. Tromans, "Wave crest sensor intercomparison study: an overview of wacsis," *TRANSACTIONS-AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF OFFSHORE MECHANICS AND ARCTIC ENGINEERING*, vol. 126, no. 1, pp. 26–34, 2004.
- [27] M. Prevosto and G. Z. Forristall, "Statistics of wave crests from models vs. measurements," in *ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering*. American Society of Mechanical Engineers, 2002, pp. 463–473.



Références

- Aarnes J., Krogstad H.* Partitioning sequences for the dissection of directional ocean wave spectra : A review // Part of ENVIWAVE Project. 2001.
- Agnus V, Ronse C, Heitz F.* Segmentation spatio-temporelle de séquences d'images // 12ème Congrès Francophone Reconnaissance des Formes et Intelligence Artificielle. 1. 2000. 619–627.
- Ailliot Pierre, Bessac Julie, Monbet Valérie, Pene Françoise.* Non-homogeneous hidden Markov-switching models for wind time series // Journal of Statistical Planning and Inference. 2015. 160. 75–88.
- Akaike Hirotogu.* Information theory and an extension of the maximum likelihood principle // Selected Papers of Hirotugu Akaike. 1998. 199–213.
- Anderson Theodore W, Darling Donald A.* A test of goodness of fit // Journal of the American Statistical Association. 1954. 49, 268. 765–769.
- Ardhuin Fabrice.* Les vagues : un compartiment important du système terre. 2012.
- Ardhuin Fabrice, Chapron Bertrand, Collard Fabrice.* Observation of swell dissipation across oceans // Geophysical Research Letters. 2009. 36, 6.
- Ardhuin Fabrice, Hanafin Jenny, Quilfen Yves, Chapron Bertrand, Queffelec Pierre, Obrebski Mathias, Sienkiewicz Joseph, Vandemark Doug.* calibration of the IOWAGA global wave hindcast (1991–2011) using ECMWF and CFSR winds // Proc. 12th International Workshop of Wave Hindcasting and Forecasting, Hawaii. 2011.
- Ardhuin Fabrice, Jenkins Alastair D.* On the interaction of surface waves and upper ocean turbulence // Journal of Physical Oceanography. 2006. 36, 3. 551–557.
- Ardhuin Fabrice, Roland Aron.* The development of spectral wave models : coastal and coupled aspects // Proceedings of Coastal Dynamics. 2013. 2013. 7th.
- Arena F., Pavone D.* Return period of nonlinear high wave crests // Journal of Geophysical Research : Oceans (1978–2012). 2006. 111, C8.

- Arena F., Pavone D.* A generalized approach for long-term modelling of extreme crest-to-trough wave heights // *Ocean Modelling*. 2009. 26, 3^U4. 217 – 225.
- Barber Norman Frederick, Ursell F.* The generation and propagation of ocean waves and swell. I. Wave periods and velocities // *Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*. 1948. 240, 824. 527–560.
- Barnett TP, Sutherland AJ.* A note on an overshoot effect in wind-generated waves // *Journal of Geophysical Research*. 1968. 73, 22. 6879–6885.
- Analyzing multidirectional wave spectra : a tentative classification of available methods. // . 1997.
- Benoit Michel, Marcos Frederic, Becq Françoise.* Development of a third generation shallow-water wave model with unstructured spatial meshing // *Coastal Engineering Proceedings*. 1996. 1, 25.
- Berg Daniel.* Copula goodness-of-fit testing : an overview and power comparison // *The European Journal of Finance*. 2009. 15, 7-8. 675–701.
- Bessac Julie, Ailliot Pierre, Monbet Valérie.* Gaussian linear state-space model for wind fields in the North-East Atlantic // *Environmetrics*. 2015. 26, 1. 29–38.
- Bitner-Gregersen EM, Hagen O.* Extreme value analysis of wave steepness and crest using joint environmental description // *Proceedings of the I 8th International Conference on Offshore Mechanics and Arctic Engineering, ASME, Paper OMAE99-6033*. 1999.
- Bitner-Gregersen Elzbieta M.* Joint probabilistic description for combined seas // *ASME 2005 24th International Conference on Offshore Mechanics and Arctic Engineering*. 2005. 169–180.
- Bitner-Gregersen Elzbieta M, Haver Sverre, others .* Joint environmental model for reliability calculations // *The First International Offshore and Polar Engineering Conference*. 1991.
- Boccotti P.* On coastal and offshore structure risk analysis // *Excerpta of the Italian Contributions to the Field of Hydraulic Engineering*. 1986. 1. 19–36.
- Boccotti Paolo.* Wave mechanics for ocean engineering. 64. 2000.
- Booij N, Ris RC, Holthuijsen Leo H.* A third-generation wave model for coastal regions : 1. Model description and validation // *Journal of Geophysical Research : Oceans* (1978–2012). 1999. 104, C4. 7649–7666.
- Boudière E., Maisondieu C., Ardhuin B., Accensi M., Pineau-Guillou L., Lepesqueur J.* A suitable metocean hindcast database for the design of Marine energy converters // *International Journal of Marine Energy*. 2013. 3-4, 7. e40–e52.
- Braun Aviva J.* A comparison of subtropical storms in the South Atlantic basin with Australian east-coast cyclones // *28th Conference on Hurricanes and Tropical Meteorology*. 2008.

-
- Callaghan DP, Nielsen P, Short A, Ranasinghe R.* Statistical simulation of wave climate and extreme beach erosion // *Coastal Engineering*. 2008. 55, 5. 375–390.
- Castillo Enrique.* Extreme value theory in engineering. 2012.
- Cavaleri L, Curiotto S, Mazzoldi A, Pavanati M.* Long term directional wave recording in the Northern Adriatic Sea // *Nuovo cimento della Società italiana di fisica*. C. 1997. 20, 1. 103–110.
- Chandler Richard E.* On the use of generalized linear models for interpreting climate variability // *Environmetrics*. 2005. 16, 7. 699–715.
- Cherubini Umberto, Luciano Elisa, Vecchiato Walter.* Copula methods in finance. 2004.
- Coe R, Stern RD.* Fitting models to daily rainfall data // *Journal of Applied Meteorology*. 1982. 21, 7. 1024–1031.
- Coles Stuart, Bawa Joanna, Trenner Lesley, Dorazio Pat.* An introduction to statistical modeling of extreme values. 208. 2001.
- Collard Fabrice, Ardhuin Fabrice, Chapron Bertrand.* Monitoring and analysis of ocean swell fields from space : New methods for routine observations // *Journal of Geophysical Research : Oceans (1978–2012)*. 2009. 114, C7.
- Corbella S., Stretch D. D.* Simulating a multivariate sea storm using Archimedean copulas // *Coastal Engineering*. 2013. 76. 68–78.
- Corbella S, Stretch DD.* Multivariate return periods of sea storms for coastal erosion risk assessment // *Natural Hazards and Earth System Science*. 2012. 12, 8. 2699–2708.
- Coster Michel, Chermant Jean-Louis.* Précis d'analyse d'images // CNRS. 1985.
- Cox Andrew T, Cardone Vincent J, Swail Val R.* On the use of in situ and satellite wave measurements for evaluation of wave hindcasts // *Proc. WMO Workshop on Advances in Marine Climatology (CLIMAR99)*. 1999. 8–15.
- De Michele C., Salvadori G., Passoni G., Vezzoli R.* A multivariate model of sea storms using copulas // *Coastal Engineering*. 2007. 54, 10. 734 – 751.
- De Waal DJ, Gelder PHAJM van.* Modelling of extreme wave heights and periods through copulas // *Extremes*. 2005. 8, 4. 345–356.
- Deheuvels Paul.* La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance // *Acad. Roy. Belg. Bull. Cl. Sci.(5)*. 1979. 65, 6. 274–292.
- Delpy Matthias T, Ardhuin Fabrice, Collard Fabrice, Chapron Bertrand.* Space-time structure of long ocean swell fields // *Journal of Geophysical Research : Oceans (1978–2012)*. 2010. 115, C12.
- Der Kiureghian A., Pei-Ling L.* Structural reliability under incomplete probability information // *Journal of Engineering Mechanics*. 1986. 112, 1. 85–104.

- Digabel H, Lantuéjoul Christian.* Iterative algorithms // Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine. 19, 7. 1978. 8.
- Donelan M, Skafel M, Graber H, Liu P, Schwab D, Venkatesh S.* On the growth rate of wind-generated waves // Atmosphere-Ocean. 1992. 30, 3. 457–478.
- Donelan Mark A, Hamilton J, Hui W_H.* Directional spectra of wind-generated waves // Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences. 1985. 315, 1534. 509–562.
- Embrechts Paul, McNeil Alexander, Straumann Daniel.* Correlation : pitfalls and alternatives // Risk-London Magazine Limited. 1999. 12. 69–71.
- Fang Z S, Hogben N.* Analysis and prediction of long term probability distributions of wave height and periods // NMI Report R146. 1982.
- Favre Anne-Catherine, El Adlouni Salaheddine, Perreault Luc, Thiémonge Nathalie, Bobée Bernard.* Multivariate hydrological frequency analysis using copulas // Water resources research. 2004. 40, 1.
- Ferreira J .A., Guedes Soares C.* Modelling distributions of significant wave height // Coastal Engineering. 2000. 40. 361–374.
- Forristall G., Ewans K., Olagnon M., Prevosto M.* The West Africa Swell Project (WASP) // Proceedings of the 32nd Int. Ocean, Offshore and Artic Eng. Conf., ASME. paper number 11264. 2013.
- Forristall G. Z.* Wave crest distributions : Observations and second-order theory // Journal of Physical Oceanography. 2000. 30, 8. 1931–1943.
- Galiatsatou Panagiota, Prinios Panagiotis.* Bivariate Analysis of Extreme Wave and Storm Surge Events. Determining the Failure Area of Structures // Open Ocean Engineering Journal. 2011. 4. 3–14.
- Genest Christian, Rémillard Bruno.* Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models // Annales de l'IHP Probabilités et statistiques. 44, 6. 2008. 1096–1127.
- Genest Christian, Rémillard Bruno, Beaudoin David.* Goodness-of-fit tests for copulas : A review and a power study // Insurance : Mathematics and economics. 2009. 44, 2. 199–213.
- Gerling T. W.* Partitioning sequences and arrays of directional ocean wave spectra into component wave systems // Journal of atmospheric and Oceanic Technology. 1992. 9, 4. 444–458.
- Gjevik B., Krogstad H., Lygre A., Rygg O.* Long period swell wave events on the Norwegian shelf // Journal of physical oceanography. 1988. 18, 5. 724–737.
- Goda Y.* Numerical experiments on statistical variability of ocean waves // Report of the Port and Harbour Research Institute. 1977. 16, 2. 3–26.

-
- Gran Sverre*. A course in ocean engineering. 8. 1992.
- Group The Wamdi*. The WAM model—a third generation ocean wave prediction model // Journal of Physical Oceanography. 1988. 18, 12. 1775–1810.
- Guédé Z, Olagnon M, Pineau H, Francois M, Quiniou V*. Fast spectral rainflow fatigue damage assessment under wideband multipeak loading // 11th International Symposium on Practical Design of Ships and Other Floating Structures Rio de Janeiro, RJ, Brazil. 2010.
- Guishard Mark P, Evans Jenni L, Hart Robert E*. Atlantic subtropical storms. Part II : Climatology // Journal of Climate. 2009. 22, 13. 3574–3594.
- Hanson J. L., Phillips. O. M.* Automated analysis of ocean surface directional wave spectra // Journal of atmospheric and oceanic technology. 2001. 18, 2. 277–293.
- Hanson Jeffrey L, Phillips Owen M*. Wind sea growth and dissipation in the open ocean // Journal of physical oceanography. 1999. 29, 8. 1633–1648.
- Hasselmann K., Barnett T. P., Bouws E., Carlson H., Cartwright D. E., Enke K., Ewing J. A., Gienapp H., Hasselmann D. E., Kruseman P., others* . Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP) // report, Deutches Hydrographisches Institut. 1973.
- Hasselmann S., Brüning C., Hasselmann K., Heimbach P*. An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra // Journal of Geophysical Research. 1996. 101, C7. 16615–16.
- Measuring and analysing the directional spectra of ocean waves. // . 2005.
- Holthuijsen Leo H*. Waves in oceanic and coastal waters. 2007.
- Husson Romain*. Development and validation of global observation-based swell model using wave mode operating Synthetic Aperture Radar. 2012.
- Hwang Paul A, Teague William J, Jacobs Gregg A, Wang David W*. A statistical comparison of wind speed, wave height, and wave period derived from satellite altimeters and ocean buoys in the Gulf of Mexico region // Journal of Geophysical Research : Oceans (1978–2012). 1998. 103, C5. 10451–10468.
- IAHR , PIANC* . List Of Sea State Parameters. 1986.
- Jasper H. Norman*. Statistical Distribution Patterns of Ocean Waves and of Wave Induced Ship Stresses and Motions with Engineering Applications. 921. 1956.
- Jaworski Piotr, Durante Fabrizio, Hardle Wolfgang Karl, Rychlik Tomasz*. Copula theory and its applications. 2010.
- Jobe J Marcus*. Lognormal Distributions : Theory and Applications // Technometrics. 1989. 31, 3. 392–392.
- Joe Harry*. Multivariate models and multivariate dependence concepts. 1997.

- Johnson Norman Lloyd, Kotz Samuel.* Distributions in statistics. 1969.
- Kendall Maurice G.* A new measure of rank correlation // *Biometrika*. 1938. 81–93.
- Kerbaol Vincent, Chapron Bertrand, Vachon Paris W.* Analysis of ERS-1/2 synthetic aperture radar wave mode imageries // *Journal of Geophysical Research : Oceans* (1978–2012). 1998. 103, C4. 7833–7846.
- Kerbiriou M-A., Prevosto M., Maisondieu C., Babarit A., Clément A.* influence of an Improved Sea State Description on a Wave Energy Converter Production // *Proceedings of OMAE 2007 - 29254*, San Diego. 2007.
- Kerbiriou Marie-Aurélie.* Energie des Vagues : estimation du potentiel récupérable en regard du potentiel naturel du site. 2007.
- Kpogo-Nuwoklo K. A., Ailliot P., Olagnon M., Guédé Z., Arnault S.* Improvement ocean wave spectra estimation using the temporal structure of wave systems // *Coastal Engineering*. 2015a. 96, 0. 81 – 91.
- Kpogo-Nuwoklo K. A., Olagnon Michel, Maisondieu Christophe, Arnault Sabine.* Comparison of parametric shapes for swell events // *proceedings of the European Wave and Tidal Energy Conference*, 6 au 11 Septembre 2015, Nantes, France. 2015b.
- Krogstad Harald E, Barstow Stephen F, others .* Directional distributions in ocean wave spectra // *Proceedings of the 9th ISOPE Conference*. 3. 1999. 79–86.
- Kuik AJ, Vledder Gerbrant P. van, Holthuijsen LH.* A method for the routine analysis of pitch-and-roll buoy wave data // *Journal of physical oceanography*. 1988. 18, 7. 1020–1034.
- Laugel Amélie.* Climatologie des états de mer en Atlantique nord-est : analyse du climat actuel et des évolutions futures sous scénarios de changement climatique par descente d'échelle dynamique et statistique. 2013. Thèse de doctorat dirigée par Benoit, Michel Génie Côtier Paris Est 2013.
- Li F., Gelder P.H.A.J.M. van, Ranasinghe R., Callaghan D.P., Jongejan R.B.* Probabilistic modelling of extreme storms along the Dutch coast // *Coastal Engineering*. 2014. 86, 0. 1 – 13.
- Li Ye, Yu Yi-Hsiang.* A synthesis of numerical methods for modeling wave energy converter-point absorbers // *Renewable and Sustainable Energy Reviews*. 2012. 16, 6. 4352–4364.
- Longuet-Higgins MS, Cartwright DE, Smith ND.* Observations of the directional spectrum of sea waves using the motions of a floating buoy // *Prentice-Hall*. 1963.
- Longuet-Higgins Michael S.* On the statistical distributions of sea waves // *J. mar. Res.* 1952. 11, 3. 245–265.
- Longuet-Higgins Michael S.* The statistical analysis of a random, moving surface // *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*. 1957. 249, 966. 321–387.

-
- Maisondieu Christophe, Johanning Lars, Weller Samuel.* Best practice report–Operation and Maintenance requirements. Deliverable 3.6. 3 from the MERiFIC Project // MERiFIC Project. 2014.
- Martín-Hidalgo Melva, Martín-Soldevilla M^a Jesús, Negro Vicente, Aberturas Paloma, López-Gutiérrez JS.* Storm evolution characterization for analysing stone armour damage progression // Coastal Engineering. 2014. 85. 1–11.
- Masina Marinella, Lamberti Alberto, Archetti Renata.* Coastal flooding : A copula based approach for estimating the joint probability of water levels and waves // Coastal Engineering. 2015. 97. 37–52.
- Massey Jr Frank J.* The Kolmogorov-Smirnov test for goodness of fit // Journal of the American statistical Association. 1951. 46, 253. 68–78.
- Mathiesen Martin, Goda Yoshimi, Hawkes Peter J, Mansard Etienne, Martín María Jesús, Peltier Eric, Thompson Edward F, Van Vledder Gerbrant.* Recommended practice for extreme wave analysis // Journal of hydraulic Research. 1994. 32, 6. 803–814.
- Mathisen J., Bitner-Gregersen E.* Joint Distribution for Significant Wave Height and Wave Zero-Crossing Period // Applied Ocean Research. 1990. 1. 59–75.
- McCullagh Peter, Nelder John A.* Generalized linear models. 37. 1989.
- McShane Blake, Adrian Moshe, Bradlow Eric T, Fader Peter S.* Count models based on weibull interarrival times // Journal of Business & Economic Statistics. 2008. 26, 3.
- Meyer Fernand.* Topographic distance and watershed lines // Signal processing. 1994. 38, 1. 113–125.
- Miller Gregory K, Bhat U Narayan.* Estimation for renewal processes with unobservable gamma or Erlang interarrival times // Journal of statistical planning and inference. 1997. 61, 2. 355–372.
- Mitsuyasu H., Tasai Fukuzo, Suhara Toshiko, Mizuno Shinjiro, Ohkusu Makoto, Honda Tadao, Rikiishi Kunio.* Observations of the Directional Spectrum of Ocean Waves Using a Cloverleaf Buoy // Journal of Physical Oceanography. 1975. 5, 4. 750–760.
- Monbet Valérie, Ailliot Pierre, Prevosto Marc.* Survey of stochastic models for wind and sea state time series // Probabilistic engineering mechanics. 2007. 22, 2. 113–126.
- Moore David S.* Chi-Square Tests. 1976.
- Morton ID, Bowers J.* Extreme value analysis in a multivariate offshore environment // Applied Ocean Research. 1996. 18, 6. 303–317.
- Munk Walter Heinrich, Miller GR, Snodgrass FE, Barber NF.* Directional recording of swell from distant storms // Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences. 1963. 255, 1062. 505–584.
- Nataf Andre.* Statistique mathématique-détermination des distributions de probabilités dont les marges sont les données // Comptes rendus hebdomadaires des séances de l'académie des sciences. 1962. 255, 1. 42.

- Nelsen Roger B.* An Introduction to Copulas. 1999.
- Nerzic R., Frelin C., Prevosto M., Quiniou-Ramus V., others .* Joint distributions of wind/waves/current in West Africa and derivation of multivariate extreme I-FORM contours // The Seventeenth International Offshore and Polar Engineering Conference. 2007.
- Nestegård Arne, Ronæss Marit, Hagen Øistein, Ronold Knut O, Bitner-Gregersen Elzbieta, others .* New DNV Recommended Practice DNV-RP-C205 on Environmental Conditions and Environmental Loads // Proceedings. 2006.
- Nordenström W.* A method to predict long-term distributions of waves and wave-induced motions and loads on ships and other floating structures. 1973. (Doktorsavhandlingar vid Chalmers Tekniska Högskola).
- O. Falcão António F. de.* Wave energy utilization : A review of the technologies // Renewable and Sustainable Energy Reviews. 2010. 14, 3. 899 – 918.
- Ochi M. K.* Wave statistics for the design of ships and ocean structures. 1978.
- Olagnon M., Ewans , K. G.Z. Forristall, Prevosto M.* West Africa Swell Spectral Shapes // Proc. 32nd Int. Ocean, Offshore and Arctic Eng. Conf., OMAE 2013-11228. 2013.
- Olagnon M., Kpogo-Nuwoklo K. A., Guédé Z.* Statistical processing of West Africa wave directional spectra time-series into a climatology of swell events // Journal of Marine Systems. 2014. 130. 101–108.
- Phillips Oliver M.* The equilibrium range in the spectrum of wind-generated waves // Journal of Fluid Mechanics. 1958. 4, 04. 426–434.
- Pierson Jr Willard J, Neumann Gerhard, James Richard W.* Practical methods for observing and forecasting ocean waves by means of wave spectra and statistics. 1971.
- Pierson Willard J, Moskowitz Lionel.* A proposed spectral form for fully developed wind seas based on the similarity theory of SA Kitaigorodskii // Journal of geophysical research. 1964. 69, 24. 5181–5190.
- Portilla J., Ocampo-Torres F. J., Monbaliu J.* Spectral partitioning and identification of wind sea and swell // Journal of atmospheric and oceanic technology. 2009. 26, 1. 107–122.
- Prevosto M.* Statistics of wave crests from second order irregular wave 3D models // Rogue Waves 2000 : Proceedings of a Workshop Organized by Ifremer and Held in Brest, France, 29-30 November 2000, Within de Brest SeaTechWeek 2000. 2001. 59.
- Prevosto M., Forristall G. Z.* Statistics of wave crests from models vs. measurements // ASME 2002 21st International Conference on Offshore Mechanics and Arctic Engineering. 2002. 463–473.
- Prevosto Marc, Ewans Kevin, Forristall George Z, Olagnon Michel.* Swell genesis, modeling and measurements in West Africa // ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering. 2013. V02BT02A028–V02BT02A028.

-
- Price William Gerraint, Bishop Richard Evelyn Donohue.* Probabilistic theory of ship dynamics. 1974.
- Rice Stephen O.* Mathematical analysis of random noise // Bell System Technical Journal. 1944. 23, 3. 282–332.
- Rosenblatt Murray.* Remarks on a multivariate transformation // The annals of mathematical statistics. 1952. 470–472.
- Ross Sheldon M, others .* Stochastic processes. 2. 1996.
- Sagrilo L. V. S., Lima E. C. Prates de, Papaleo Arnaldo.* A joint probability model for environmental parameters // Transactions of the ASME-O-Journ of Offsh Mechan and Arctic Engin. 2011. 133, 3. 031605.
- Salvadori G, De Michele C.* On the use of copulas in hydrology : theory and practice // Journal of Hydrologic Engineering. 2007. 12, 4. 369–380.
- Santos Silva Ralph dos, Lopes Hedibert Freitas.* Copula, marginal distributions and model selection : a Bayesian note // Statistics and Computing. 2008. 18, 3. 313–320.
- Scherrer B.* Biostatistique, Vol. 1 // Gaëtan Morin éditeur (816 pp.). 2007.
- Schwarz Gideon, others .* Estimating the dimension of a model // The annals of statistics. 1978. 6, 2. 461–464.
- Sim CH.* Point processes with correlated gamma interarrival times // Statistics & probability letters. 1992. 15, 2. 135–141.
- Sklar A.* Fonction de répartition dont les marges sont données, Inst // Stat. Univ. Paris Publ. 1959. 8. 229–231.
- Smith Murray D.* Modelling sample selection using Archimedean copulas // The Econometrics Journal. 2003. 6, 1. 99–123.
- Snodgrass F. E., Groves G. W., Hasselmann K.F., Miller G.R., Munk W.H., Powers W.H.* Propagation of ocean swell across the Pacific // Philosophical Transactions for the Royal Society of London. Series A, Mathematical and Physical Sciences. 1966. 431–497.
- Soares C Guedes, Lopes LC, Costa MD.* Wave climate modelling for engineering purposes // Anais do Instituto Hidrográfico. 1988.
- Stern RD, Coe R.* A model fitting analysis of daily rainfall data // Journal of the Royal Statistical Society. Series A (General). 1984. 1–34.
- Sverdrup Harald Ulrik, Munk Walter Heinrich.* Wind, sea, and swell : theory of relations for forecasting. 1947.
- Swail VR, Ceccacci EA, Cox AT.* The AES40 North Atlantic wave reanalysis : validation and climate assessment // 6th International Workshop on Wave Hindcasting and Forecasting November. 2000. 6–10.

- Teng J., Palao I.* wave height and period distributions from long-term measurement // Coastal Engineering Proceedings. 2001. 1, 25.
- Thompson Paul, Cai Yuzhi, Reeve Dominic, Stander Julian.* Automated threshold selection methods for extreme wave analysis // Coastal Engineering. 2009. 56, 10. 1013–1021.
- Tolman HL.* Validation of WAVEWATCH III version 1.15 for a global domain // Technical Note. 2002. 213. 33.
- Tolman Hendrik L, others .* User manual and system documentation of WAVEWATCH III TM version 3.14 // Technical note, MMAB Contribution. 2009. 276.
- Torsethaugen K., Haver S.* Simplified double peak spectral model for ocean waves // Proceedings of 14th international offshore and polar engineering conference. ISOPE. 3. 2004. 76–84.
- Tucker Malcolm John, Pitt Edward G.* Waves in ocean engineering. 2001.
- Vladder Gerbrant van, Goda Yoshimi, Hawkes Peter, Mansard Etienne, Martin Maria Jesus, Mathiesen Martin, Peltier Eric, Thompson Edward.* Case studies of extreme wave analysis : a comparative analysis // Ocean Wave Measurement and Analysis (1993). 1993. 978–992.
- Voorrips AC, Makin VK, Hasselmann S.* Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model // Journal of Geophysical research. 1997. 102, C3. 5829–5849.
- Vrijling JK, Bruinsma J.* Developments in hydraulic engineering related to the design of the Oosterschelde storm surge barrier in the Netherlands // Proceedings on hydraulic aspects of coastal structures. 1980.
- Wahl T, Mudersbach C, Jensen J.* Assessing the hydrodynamic boundary conditions for risk analyses in coastal areas : a multivariate statistical approach based on Copula functions // Natural Hazards and Earth System Science. 2012. 12, 2. 495–510.
- Walker Rich T, Nieuwkoop-McCall Joana van, Johanning Lars, Parkinson Richard J.* Calculating weather windows : Application to transit, installation and the implications on deployment success // Ocean Engineering. 2013. 68. 88–101.
- Weisstein Eric W.* Log normal distribution // MathWorld—A Wolfram Web Resource. 2010.
- Yan Zhongwei, Bate Steven, Chandler Richard E, Isham Valerie, Wheeler Howard.* An analysis of daily maximum wind speed in northwestern Europe using generalized linear models // Journal of Climate. 2002. 15, 15. 2073–2088.
- Young Ian R.* Wind generated ocean waves. 2. 1999.