



Analyse, traitement et qualification des données

Volet « Houle »

Lot 4 – référence T4.H

Auteur : Laurent LEBALLEUR, Guiomar Lopez

Référence : convention Shom/MTES [DGEC], SHOM n°79/2019, relative à la réalisation des études de reconnaissance environnementale de site en vue d'appels d'offres pour l'implantation d'éoliennes en mer au large des côtes françaises

Table des matières

1. Introduction.....	3
2. Observations des états de mer sur la zone EMR de Méditerranée	4
2.1 Campagne de mesures d'états de mer.....	4
2.2 Contrôle qualité.....	5
2.3 Comparaison des observations aux résultats du modèle.....	6
2.3.1 Comparaison de tous les événements	7
2.3.2 Comparaison des événements énergétiques.....	10
3. Analyse statistique.....	13
3.1 Analyse de l'état de mer total.....	13
3.2 Analyse de l'état de mer en fonction de la direction moyenne des vagues	20
3.3 Analyse de l'état de mer par systèmes de vagues	27
4. Estimation des hauteurs significatives extrêmes associées à des périodes de retour	34
4.1 Méthode de sélection des valeurs extrêmes.....	34
4.2 Estimation des hauteurs significatives extrêmes associées à des périodes de retour	34
4.2.1 Méthode stationnaire	35
4.2.2 Méthode non-stationnaire appliquée aux variations saisonnières	36
5. CONCLUSION.....	40
BIBLIOGRAPHIE.....	41
Annexe 1. Corrélogrammes pour les secteurs directionnels	43
Annexe 2. Sélection du seuil de tempête.....	51

1. Introduction

La présente étude vise à compléter les travaux menés dans le cadre du Lot 2 (Lopez et al., 2022). Ce dernier avait pour objectif de caractériser, à partir de résultats de modélisation numérique, la climatologie moyenne des états de mer en Méditerranée, sur la zone envisagée pour l'implantation d'un parc éolien en mer (zone B). Ce travail est complété ici par une étude statistique plus approfondie qui comporte notamment l'étude relative aux valeurs extrêmes d'états de mer sur la zone. De plus, la qualification du rejeté utilisé pour le Lot 2 est enrichie grâce aux données acquises en mer lors de la campagne relative au Lot 3.

Ce document s'articule autour de 3 chapitres. Le premier s'intéresse aux mesures d'états de mer acquises lors des travaux à la mer conduits lors du Lot 3 : les mesures y sont d'abord brièvement présentées, pour ensuite être exploitées pour la qualification des résultats du modèle d'états de mer utilisé lors des travaux des Lots 2 et 4. Cette comparaison permet d'évaluer la pertinence du modèle au niveau de la zone d'implantation considérée du parc éolien. Le second chapitre s'intéresse à l'analyse statistique du rejeté de vagues sur la zone d'implantation du parc éolien. Enfin, le rejeté d'états de mer est exploité pour caractériser le climat extrême d'état de mers sur la zone.

2. Observations des états de mer sur la zone EMR de Méditerranée

Dans le cadre de la présente étude, une campagne d'acquisition des mesures in situ a été réalisée entre Juin 2021 et Décembre 2022. Le présent paragraphe présente ces données et les résultats des comparaisons avec les résultats de modélisations numériques du jeu. Ce dernier, qui a été décrit et utilisé pour les analyses présentées dans le rapport du Lot 2 (Lopez et al., 2022), et couvre la période 2023 – 2022.

2.1 Campagne de mesures d'états de mer

Une campagne de mesures en mer a été réalisée sur le littoral méditerranéen du golfe du Lion, dans la zone EMR numérotée B, située au large de Gruissan. Un capteur de type Aanderaa Motus a été déployé du 10/06/2021 au 31/12/2022, à la position 43.01667°N – 3.55°E, par environ 90.5 m de profondeur. La Figure 1 présente la zone d'étude et la position du capteur Motus, ainsi que la grille de calcul du modèle numérique et les points sélectionnés pour l'analyse statistique.

Le capteur utilisé pour cette étude a été installé sur une plateforme flottante de 5 mètres de longueur pour 3 mètres de largeur. Les dimensions de cette plateforme sont très grandes pour suivre fidèlement les mouvements de l'eau, ce qui peut entraîner un filtrage des vagues les plus courtes. Les données de déplacements de la bouée MOTUS ont été traitées automatiquement par le logiciel constructeur installé avec le capteur. Les spectres de vagues ainsi que les paramètres ont été estimés à partir d'échantillons de 30 minutes suivant la procédure décrite dans le manuel du capteur (Aanderaa, 2020).

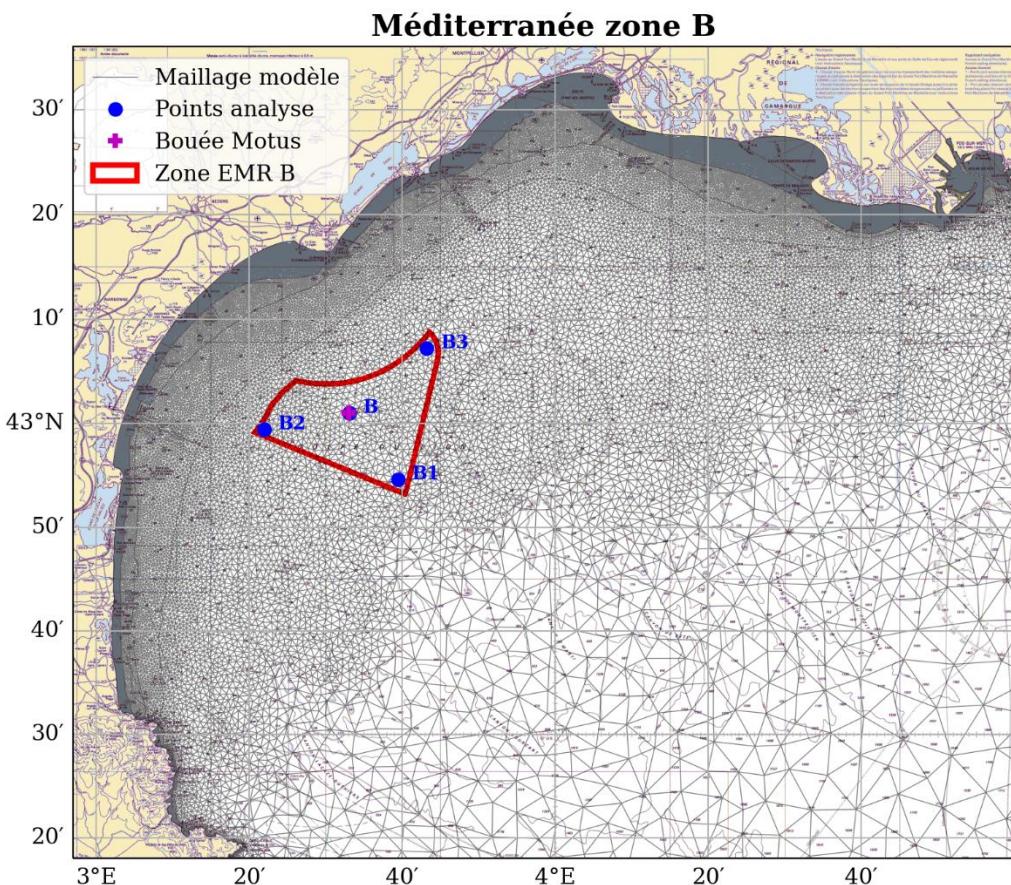


Figure 1: Représentation de la zone d'étude EMR (en rouge) superposée à la carte marine Shom 7008 (2023). Le capteur Aanderaa Motus déployé pour cette étude est représenté par une croix mauve, la grille de calcul du modèle numérique par des triangles noirs et les points sélectionnés pour l'analyse statistique par des points bleus.

2.2 Contrôle qualité

Un contrôle qualité de ces données a été réalisé afin de détecter les valeurs aberrantes. Cette analyse se base sur les trois critères suivants, dont les différents seuils sont répertoriés dans le Tableau 1.

- Le premier critère vise à supprimer automatiquement du jeu de données les valeurs en dehors des seuils d'acceptabilité des capteurs (min/max).
- Le second critère se base sur la répétition de valeurs identiques pendant les périodes de temps indiquées dans le Tableau 1.
- Le troisième critère cherche à détecter les pics anormaux, horsains (spike/outlier) en analysant les variations temporelles du signal sur trois enregistrements consécutifs.

Pour le second et le troisième critères, deux niveaux de seuil sont définis. Lorsque le premier niveau est dépassé, le point est identifié comme « suspect » et fera l'objet d'un contrôle manuel. En revanche, lorsque le second niveau est dépassé, le point est marqué comme mauvais et supprimé automatiquement de la série temporelle. Un exemple de contrôle qualité est présenté sur la Figure 2 avec les indices qualité correspondant au Tableau 2.

	H_s	T_{02}	T_p	D_p
Seuil d'acceptabilité (min – max)	0.02 m – 20 m	1.42 s – 25 s	1 s – 25 s	0° - 360°
Répétition (douteux - mauvais)	3 h – 6 h	3 h – 6 h	12 h – 14 h	3 h – 6 h
Différence entre 3 échéances consécutives (douteux - mauvais)	2 m – 2.5 m	2 s – 4 s	-	-

Tableau 1 : Récapitulatif des différents seuils utilisés pour les trois critères sur lesquels se base le contrôle qualité. Les unités des seuils pour les tests d'acceptabilité et de différence entre 3 échéances consécutives sont respectivement mètres, secondes et degrés, pour les paramètres H_s , T_{02} , T_p et D_p . Les seuils pour le test de répétition sont indiqués en heures.

QC flag (indice qualité)	Description
1	Valeur qui a passé le contrôle qualité avec succès
2	Valeur non évaluée
3	Valeur douteuse / suspectée mauvaise, à contrôler manuellement
4	Valeur qualifiée comme mauvaise, supprimée automatiquement
9	Valeur manquante

Tableau 2 : Code adopté pour qualifier la qualité de chacune des mesures.

Au final, il résulte de ce contrôle qualité que 14.7% des points des séries temporelles du capteur Motus ont été supprimés, essentiellement à cause de valeurs manquantes dans les acquisitions. La série temporelle finale de la hauteur significative mesurée est présentée sur la Figure 3, superposée aux résultats du modèle correspondants.

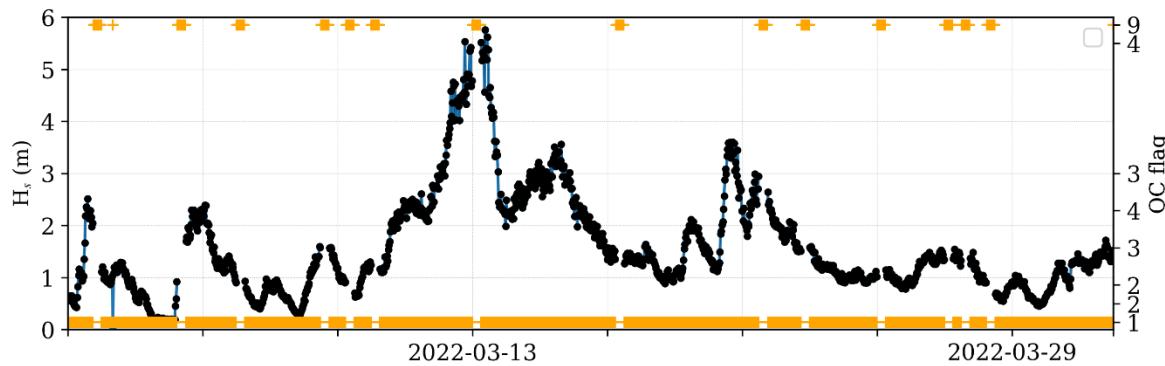


Figure 2 : Exemple des résultats du contrôle qualité (croix orange) pour la hauteur significative des vagues mesurée par le capteur Motus (ligne bleue). Illustration de valeurs manquantes dans le jeu de données au mois de mars 2022 (QC flag = 9). L'axe de droite correspond à l'indice qualité (voir Tableau 2).

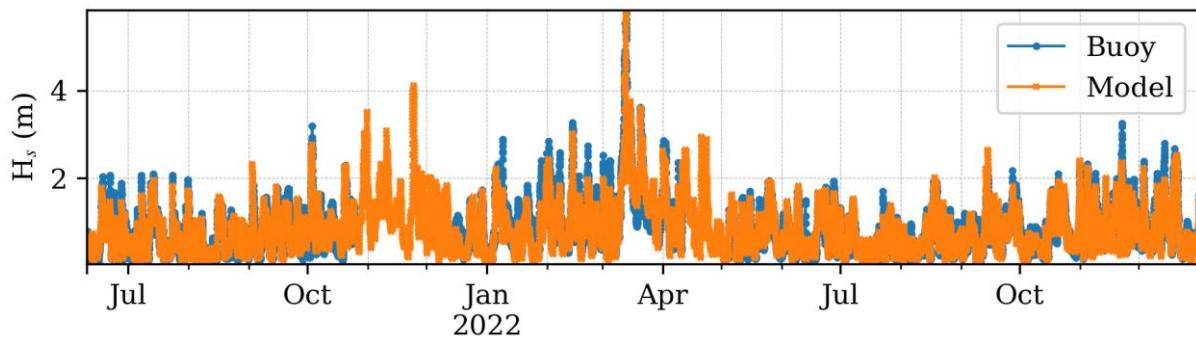


Figure 3 : Série temporelle de la hauteur significative mesurée par le capteur Motus après l'étape de contrôle qualité (en bleu), superposée aux résultats du modèle correspondants (en orange).

2.3 Comparaison des observations aux résultats du modèle

Ce paragraphe se concentre sur la validation des résultats du modèle numérique (rejet WW3-MEDUG) à partir des données de la campagne spécifique décrite aux paragraphes précédents. Nous rappelons également qu'une validation plus détaillée du rejet numériques a été réalisée dans le cadre des lots 2 et 5 (Lopez et al., 2022 et 2023). Elle se base notamment sur des observations de durée plus longue, disponibles notamment grâce à l'altimétrie satellitaire et aux bouées houlographes côtières du réseau CANDHIS¹.

La Figure 1 permet de visualiser la grille de calcul du modèle numérique ainsi que de localiser la bouée utilisée pour ce travail de validation. Le point de sortie de la grille du modèle le plus proche à l'emplacement de la bouée a été utilisé afin de comparer les valeurs mesurées avec celles modélisées (point B). Le recouvrement entre les observations et le rejet est d'environ 15 mois et demi, du 10 juin 2021 à fin décembre 2022 en tenant compte des données manquantes ou n'ayant pas passé l'étape du contrôle qualité.

¹ <https://candhis.cerema.fr/index.php>

2.3.1 Comparaison de tous les événements

Pour les paramètres spectraux H_s^2 , $T_{m(0,2)}$, T_p , D_m et D_p , les diagrammes de dispersion et les distributions entre les résultats du modèle et les observations sont présentés respectivement sur les Figure 4 et Figure 5, pour tous les événements de la période commune aux deux jeux de données. Le Tableau 3 synthétise les scores obtenus par comparaison des résultats du modèle numérique avec le capteur Motus, et pour chacun de ces paramètres.

En termes de hauteur significative, le biais du modèle est non significatif et l'erreur quadratique moyenne est faible (23 cm), soit une erreur quadratique moyenne normalisée de 24 %. Le coefficient de corrélation est très bon (0.91) et l'indice de dispersion est de 28.8%. On peut toutefois remarquer que les valeurs les plus fortes observées par la bouée sont sous-estimées par le modèle.

En ce qui concerne les périodes, l'analyse montre une bonne représentation des périodes moyennes malgré une légère tendance du modèle à la sous-estimation : le biais du modèle est inférieur à 0.3 s (en valeur absolue) et l'erreur quadratique moyenne est d'environ 0.69 s (soit une erreur quadratique moyenne normalisée autour de 18 %). Le coefficient de corrélation est de 0.74, pour un indice de dispersion de 18%. Pour les périodes pic, les scores sont moins bons : le biais du modèle est un peu plus élevé (+0.5 s), l'erreur quadratique moyenne est aussi plus forte (2 s, soit une erreur quadratique moyenne normalisée d'environ 39%). Le coefficient de corrélation est de 0.49, et l'indice de dispersion d'environ 41%. Cette période représentant la période associée au pic d'énergie du spectre d'états de mer, elle est fortement dépendante de la bonne représentation des différents systèmes et de leur niveau d'énergie relative. Or, cette zone est soumise à des houles en provenance du secteur sud et à des mers de vent de secteur ouest (voir paragraphes 3.2 et 3.3). Le diagramme de dispersion des périodes pic montrent que le modèle représente des événements de périodes pic comprises entre 6 et 14s alors que la bouée indique plutôt des périodes comprises entre 2 et 5s, ce qui résulte d'un mauvais positionnement du pic d'énergie dans le spectre modélisé pour des événements où la mer de vent devrait être dominante.

Enfin, les directions (pic et moyenne) sont bien représentées par le modèle. Le biais est inférieur à 5°, l'erreur quadratique moyenne est d'environ 60° (soit 33% en relatif). La corrélation est satisfaisante (0.75-0.78) et la dispersion un peu plus élevée (33-34%).

En général, les scores obtenus permettent d'avoir une bonne confiance dans les statistiques d'états de mer calculées à partir du rejet dans la suite de cette étude. Comme décrit dans ce paragraphe, l'incertitude la plus marquée est associée aux périodes pic et à la direction des vagues. Les résultats pour la hauteur des vagues et la période moyennes sont toutefois très satisfaisants.

Param.	Nobs	Bias	RMSE	NRMSE	SI (%)	a	b	R
H_s (m)	22 723	0.00	0.23	0.24	28.8	0.90	0.08	0.91
$T_{m(0,2)}$ (s)	22 723	-0.29	0.69	0.18	18.0	1.01	-0.33	0.74
T_p (s)	22 723	0.50	1.92	0.39	41.2	0.65	2.13	0.49
D_m (deg)	22 663	2.6	63.9	0.33	33.4	0.92	16.9	0.78
D_p (deg)	22 663	4.8	61.3	0.34	34.0	0.88	26.3	0.75

Tableau 3 : Scores du rejet numérique dérivés de la comparaison avec la bouée Motus sur la totalité de la période commune. Les résultats sont présentés pour la hauteur significative (H_s), période moyenne ($T_{m(0,2)}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p). Les colonnes sont respectivement, le paramètre, le nombre d'observations (Nobs), le biais (Bias), l'erreur quadratique moyenne (RMSE), l'erreur quadratique moyenne normalisée (NRMSE), le Scatter Index (SI), la pente (a) et l'ordonnée à l'origine (b) de la régression linéaire, et enfin le coefficient de corrélation R.

² La hauteur significative H_s est ici calculée à partir du moment d'ordre 0 du spectre d'état de mer $H_{m_0} = 4\sqrt{m_0}$.

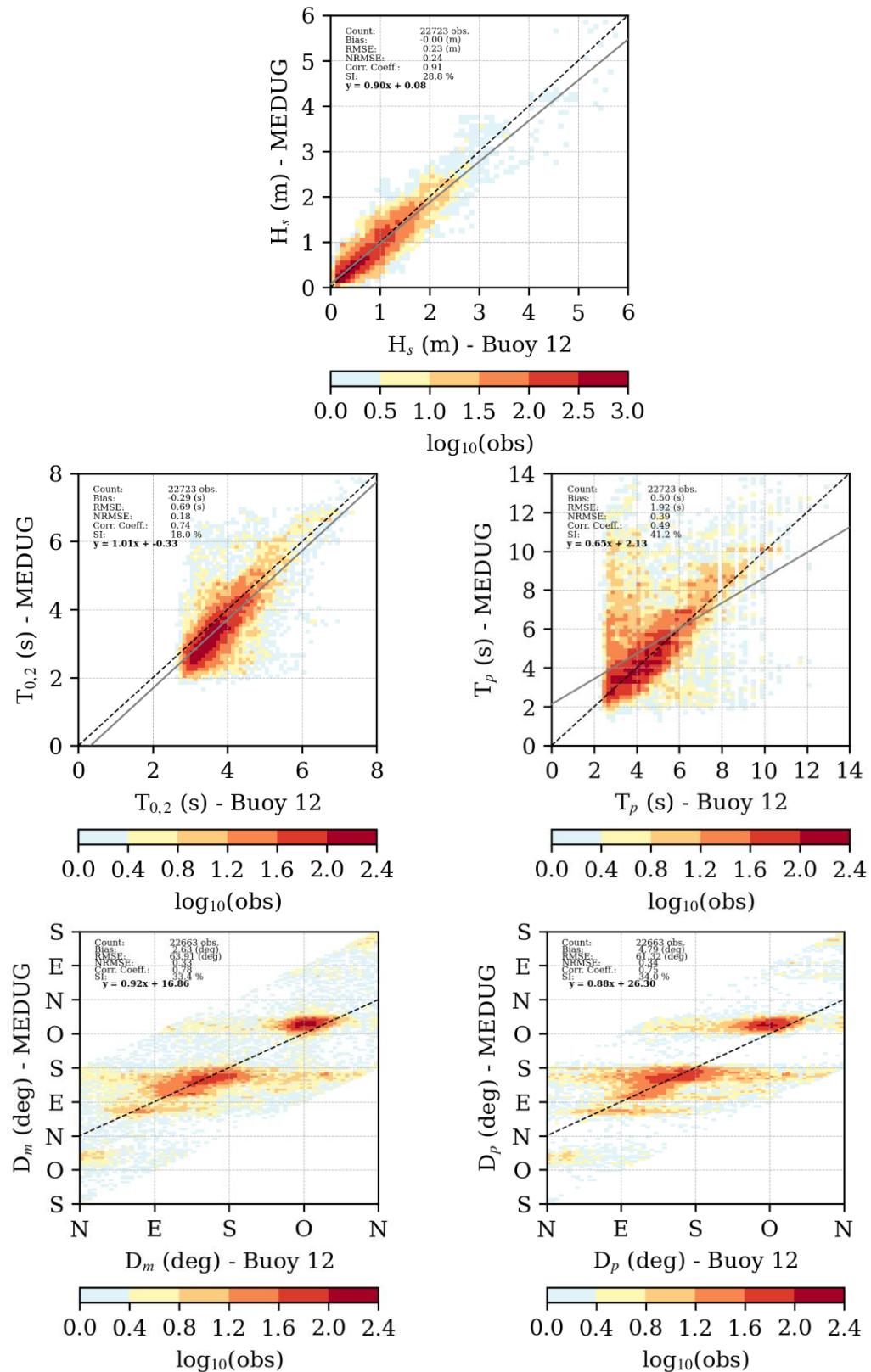


Figure 4 : Corrélation des valeurs de hauteur significative (H_s), période moyenne ($T_{m0,2}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p) modélisées et observées par la bouée Motus déployée dans la zone d'étude, pour tous les événements du jeu de données.

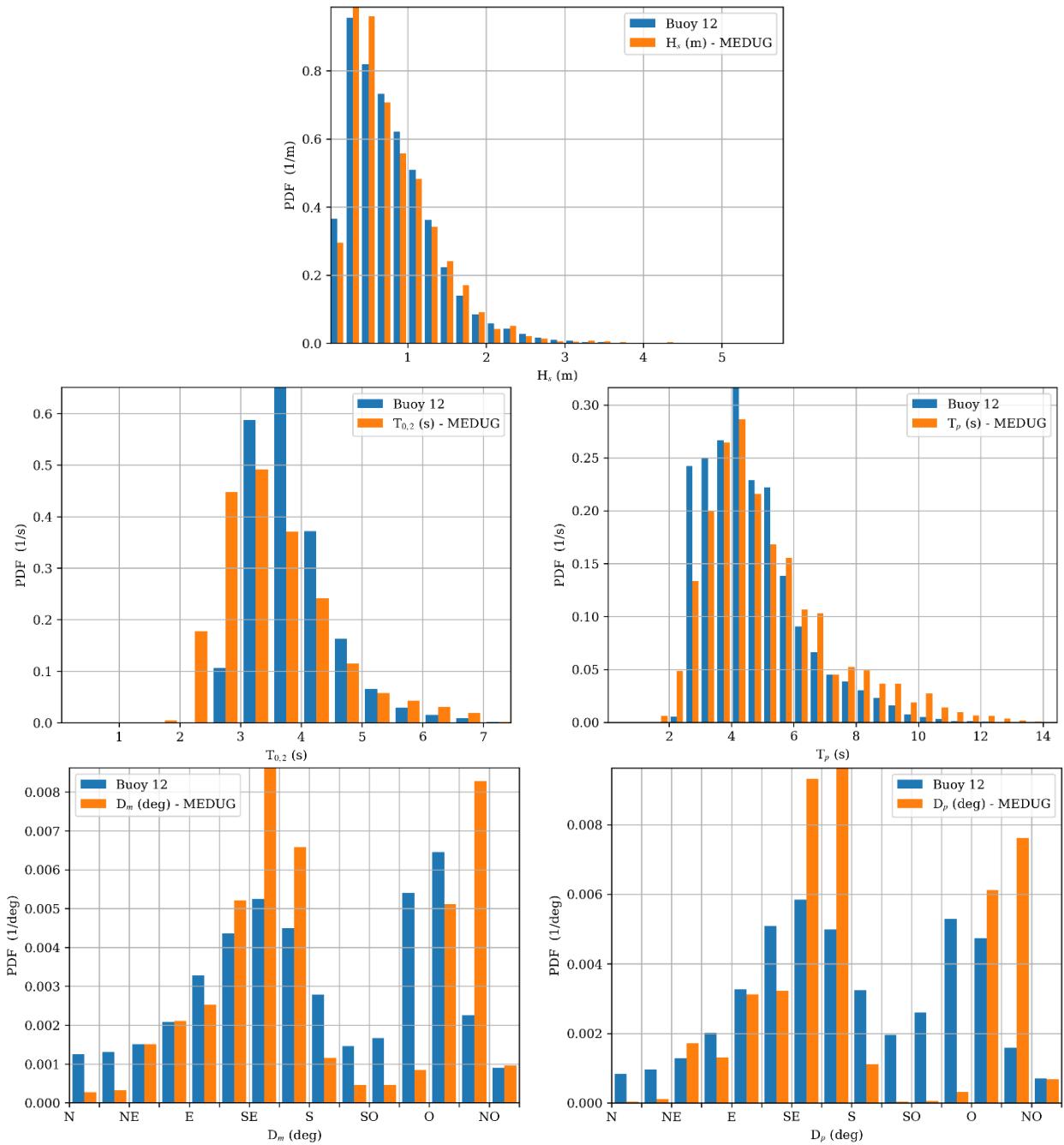


Figure 5 : Distributions des valeurs de hauteur significative (H_s), période moyenne ($T_{m0.2}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p) modélisées (en orange) et observées par la bouée Motus déployée dans la zone d'étude (en bleu), pour tous les événements du jeu de données.

2.3.2 Comparaison des événements énergétiques

Pour les paramètres spectraux H_s , $T_{m(0,2)}$, T_p , D_m et D_p , les diagrammes de dispersion et les distributions entre les résultats du modèle et les observations sont présentés respectivement sur les Figure 6 et Figure 7, pour les événements de la période commune dont la hauteur significative est supérieure au quantile 95 %. Le Tableau 4 synthétise les scores obtenus.

En termes de hauteur significative, le modèle ne montre pas de biais significatif (-4 cm), l'erreur quadratique moyenne est de 40 cm, soit une erreur quadratique moyenne normalisée de 14%. Le coefficient de corrélation est bon (0.84) et l'indice de dispersion est relativement faible 15%. Par rapport au paragraphe précédent qui considérait l'ensemble du jeu de données sur les périodes communes, cette analyse sur les événements plus énergétiques illustre une légère tendance à la sous-estimation des événements énergétiques, l'erreur RMS est un peu plus importante en valeur absolue mais plus faible en valeur relative. Bien que la corrélation diminue légèrement sur ces événements intenses, on note une amélioration de l'indice de dispersion.

En ce qui concerne les périodes, on peut remarquer une nette amélioration des scores pour ces événements intenses : le biais est non significatif, l'erreur quadratique normalisée est inférieure à 10%. La corrélation est aussi meilleure (0.90 et 0.89 contre 0.74 et 0.49) et la dispersion est réduite à moins de 10%.

Enfin, pour les directions, les scores sont moins bons : le biais est d'environ 20°, l'erreur quadratique moyenne est de près de 70 degrés (soit 50% en relatif). La corrélation diminue légèrement à 0.72 et la dispersion augmente à près de 50%. Ces scores sont induits par une partie des événements que la bouée enregistre en provenance d'ouest-nord-ouest, alors que le modèle les représente davantage en provenance du sud-sud-est. On peut aussi se demander si la structure métallique qui accueille le capteur Motus n'aurait pas une influence sur l'estimation des paramètres directionnels des vagues.

Param.	Nobs	Bias	RMSE	NRMSE	SI (%)	a	b	R
H_s (m)	601	-0.04	0.40	0.14	14.6	0.78	0.57	0.84
$T_{m(0,2)}$ (s)	601	0.05	0.36	0.06	6.4	1.03	-0.13	0.90
T_p (s)	601	0.16	0.70	0.09	9.5	0.98	0.27	0.89
D_m (deg)	601	-17.6	68.8	0.52	52.3	0.70	42.6	0.72
D_p (deg)	601	-21.2	67.6	0.50	49.7	0.64	51.1	0.72

Tableau 4 : Scores du rejet numérique dérivés de la comparaison avec la bouée Motus pour les événements de la période commune dont la hauteur significative est supérieure au quantile 95% de la hauteur significative. Les résultats sont présentés pour la hauteur significative (H_s), période moyenne ($T_{m0,2}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p). Les colonnes sont respectivement, le paramètre, le nombre d'observations (Nobs), le biais (Bias), l'erreur quadratique moyenne (RMSE), l'erreur quadratique moyenne normalisée (NRMSE), le Scatter Index (SI), la pente (a) et l'ordonnée à l'origine (b) de la régression linéaire, et enfin le coefficient de corrélation R.

Malgré le faible nombre d'observations disponibles pour évaluer les performances du modèle en conditions énergétiques, il ressort de cette analyse que les hauteurs significatives supérieures au quantile 95% ne présentent pas de biais significatif, bien qu'une tendance à la sous-estimation soit observée. L'erreur RMS associée est légèrement plus importante que pour l'ensemble des événements (environ 40 cm contre 25 cm) mais représente une erreur relative moindre (14% contre 24%). La corrélation est légèrement moins bonne, et reste très satisfaisante (0.84), avec un bon indice de dispersion (15%). Les périodes sont également très bien représentées. La plupart des directions de provenance sont bien représentées par le modèle, même si certains événements d'ouest-nord-ouest sont moins bien représentés par le modèle.

Les scores du rejet obtenus pour ces conditions énergétiques montrent sa bonne capacité à les représenter. Malgré certaines imperfections, le rejet est adapté pour l'estimation des valeurs extrêmes associées à des périodes de retour.

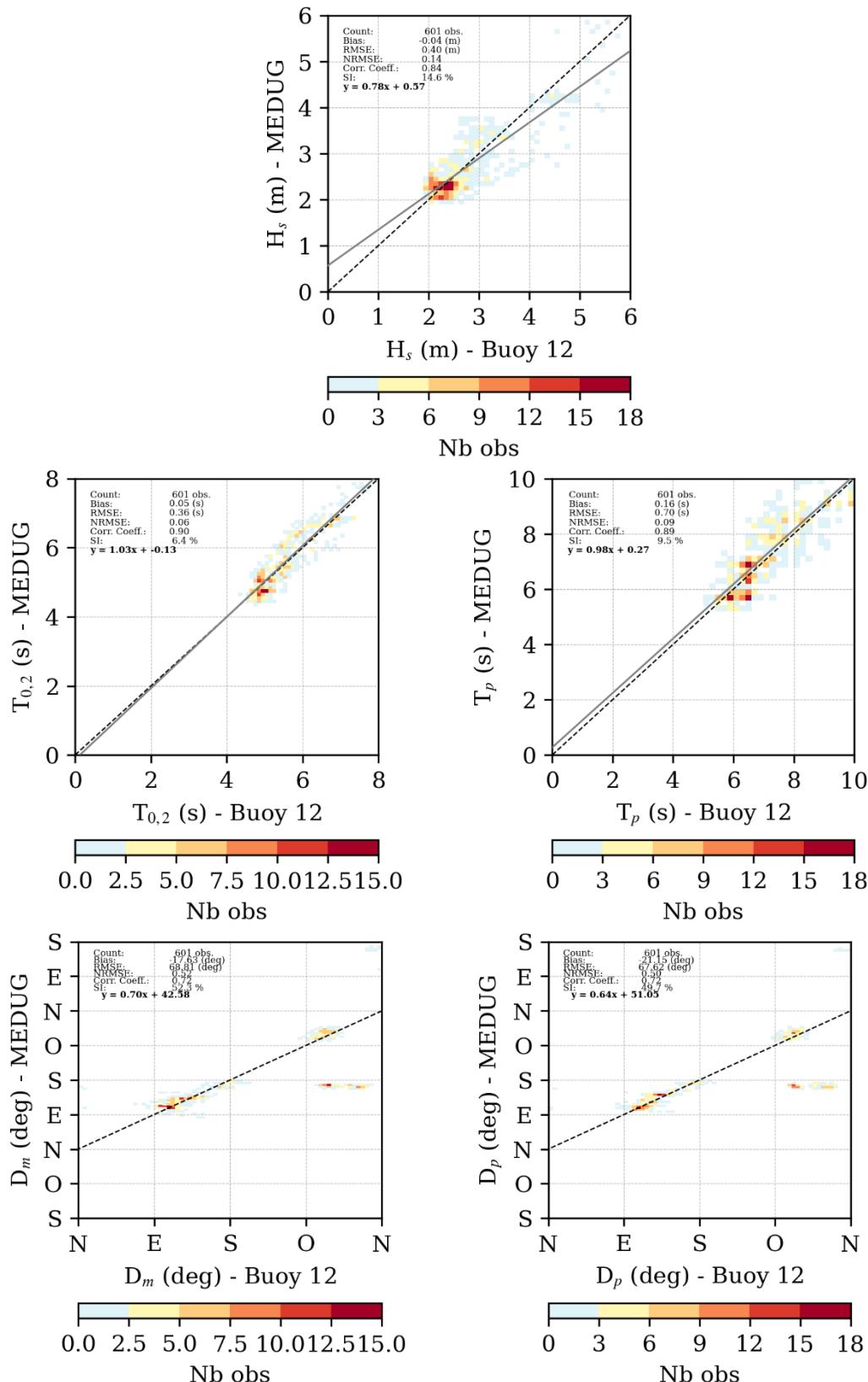


Figure 6 : Corrélation des valeurs de hauteur significative (H_s), période moyenne ($T_{m0,2}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p) modélisées et observées par la bouée Motus déployée dans la zone d'étude, pour les événements dont la hauteur significative est supérieure au quantile 95%.

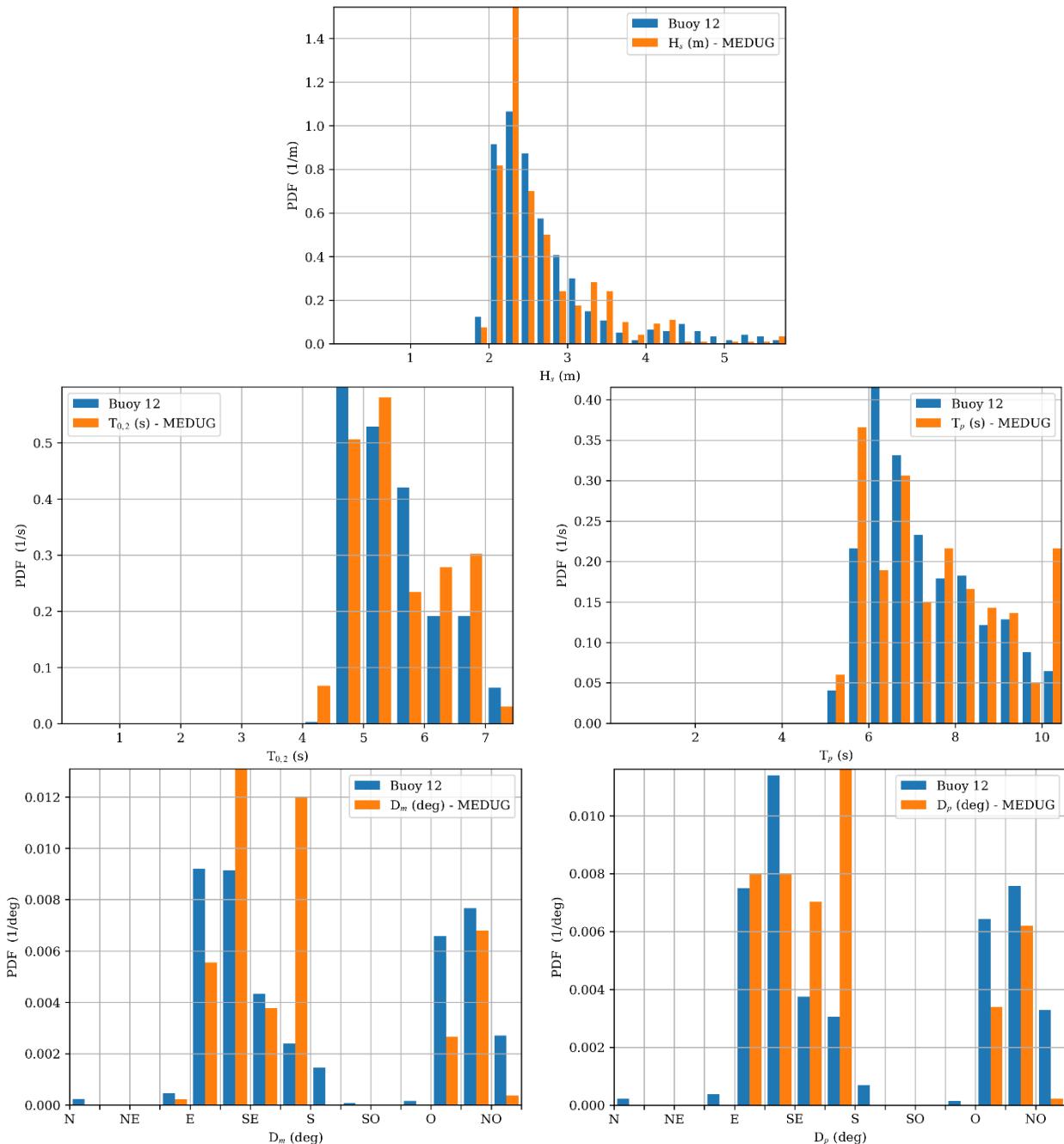


Figure 7 : Distributions des valeurs de hauteur significative (H_s), période moyenne ($T_{m0,2}$), période de pic (T_p), direction moyenne (D_m) et direction de pic (D_p) modélisées (en orange) et observées par la bouée Motus déployée dans la zone d'étude (en bleu), pour les événements dont la hauteur significative est supérieure au quantile 95%.

3. Analyse statistique

3.1 Analyse de l'état de mer total

L'étude réalisée dans le cadre du Lot 2 avait été produite à partir du rejet global de la Méditerranée à 6' de résolution sur la période 2003-2021. Elle avait montré que l'état de mer sur la zone EMR B est caractérisé par une hauteur significative moyenne des vagues comprises entre 0.85 m (B2) et 1.1 m (B1), ainsi qu'un gradient croissant du Nord-Ouest vers le Sud-Est, dans le même sens que la bathymétrie. Les hauteurs significatives les plus fortes (quasiment 9 m) sont atteintes au point B1, le plus au Sud-Est. Ces résultats sont confirmés à nouveau ici, avec l'utilisation du rejet à plus haute résolution et complété avec l'année 2022 (Figure 8, Figure 9 et Tableau 5), précisant les statistiques produites lors du Lot 2.

	H_s (m)					$T_{m(0,2)}$ (s)				
	mean	max	median	90 prc.	99 prc.	mean	max	median	90 prc.	99 prc.
B	0.96	8.92	0.80	1.80	3.42	3.42	9.31	3.25	4.74	6.54
B1	1.10	8.96	0.94	2.10	3.51	3.58	9.37	3.47	4.85	6.48
B2	0.84	8.57	0.69	1.56	3.23	3.34	9.34	3.12	4.80	6.65
B3	0.99	8.52	0.82	1.87	3.51	3.46	9.11	3.30	4.75	6.52

Tableau 5 : Statistiques des paramètres de hauteur significative H_s et de période moyenne $T_{m(0,2)}$ aux points B, B1, B2 et B3.

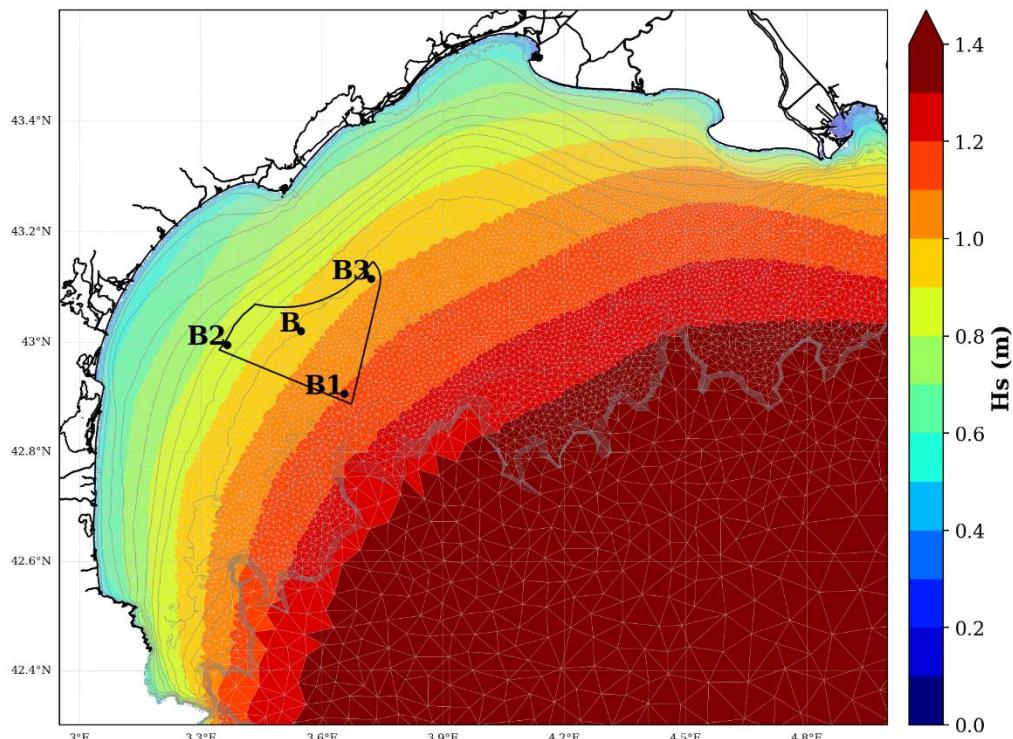


Figure 8 : Carte des hauteurs significatives moyennes (en mètres) modélisées sur la période 2003 – 2022 avec le rejet numérique. Le polygone noir représente la zone d'étude pour l'implantation du parc éolien. Les points B, B1, B2 et B3 utilisés pour les analyses qui suivent dans le rapport sont également représentés.

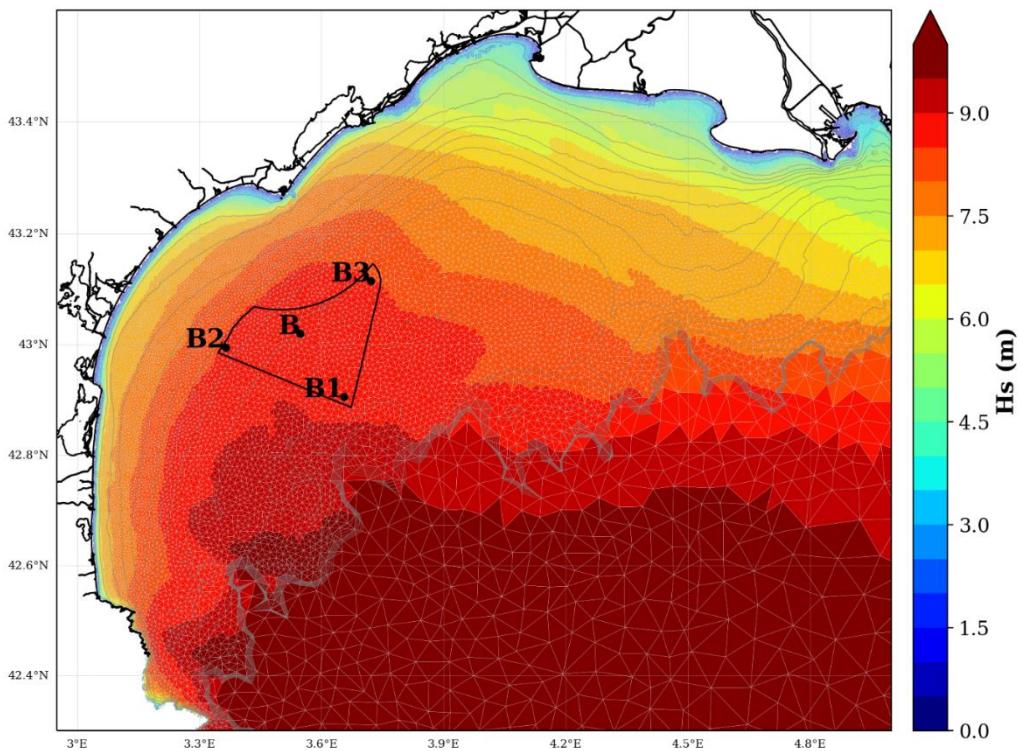


Figure 9 : Carte des hauteurs significatives maximales (en mètres) modélisées sur la période 2003 – 2022 avec le rejeu numérique. Le polygone noir représente la zone d'étude pour l'implantation du parc éolien. Les points B, B1, B2 et B3 utilisés pour les analyses qui suivent dans le rapport sont également représentés.

En complément de la description spatiale des hauteurs de vagues, une mise à jour de l'étude statistique présentée dans le lot 2 est présentée ici pour les quatre points localisés dans la zone d'implantation du parc éolien. Ces points sont notamment représentés sur les cartes des Figure 1, Figure 8 et Figure 9.

Pour ces quatre points, les Figure 10, Figure 11 et Figure 12 montrent respectivement la répartition de la hauteur significative H_{sig} , de la période moyenne $T_{m(0,2)}$ et de la direction moyenne de provenance des vagues θ_m . Le Tableau 5 regroupe les statistiques de valeurs moyennes, maximales, médianes ainsi que les percentiles 90 % et 99 % de la hauteur significative et de la période moyenne pour ces quatre mêmes points.

En ce qui concerne la hauteur significative, (Figure 10), on observe une certaine similitude entre les quatre points analysés, en particulier entre les points B et B3, ce qui montre une certaine homogénéité spatiale des hauteurs significatives sur la zone d'étude. Le point B1 montre une distribution légèrement plus forte alors qu'au point B2, la distribution est légèrement plus faible. La plupart des vagues ont des hauteurs significatives comprises entre 0 m et 4 m.

Les distributions de la période moyenne des vagues (Figure 11) sont également assez homogènes dans la zone d'étude et généralement comprise entre 2 et 6 secondes. Au point B1, la distribution est légèrement plus forte, et au point B2, elle est légèrement plus faible.

Pour les quatre points analysés, la Figure 12 montre que les vagues proviennent majoritairement de deux secteurs : le secteur d'ouest à nord-ouest et le secteur d'est à sud. Dans le paragraphe suivant, des roses de vagues fournissent des informations plus détaillées (Figure 15).

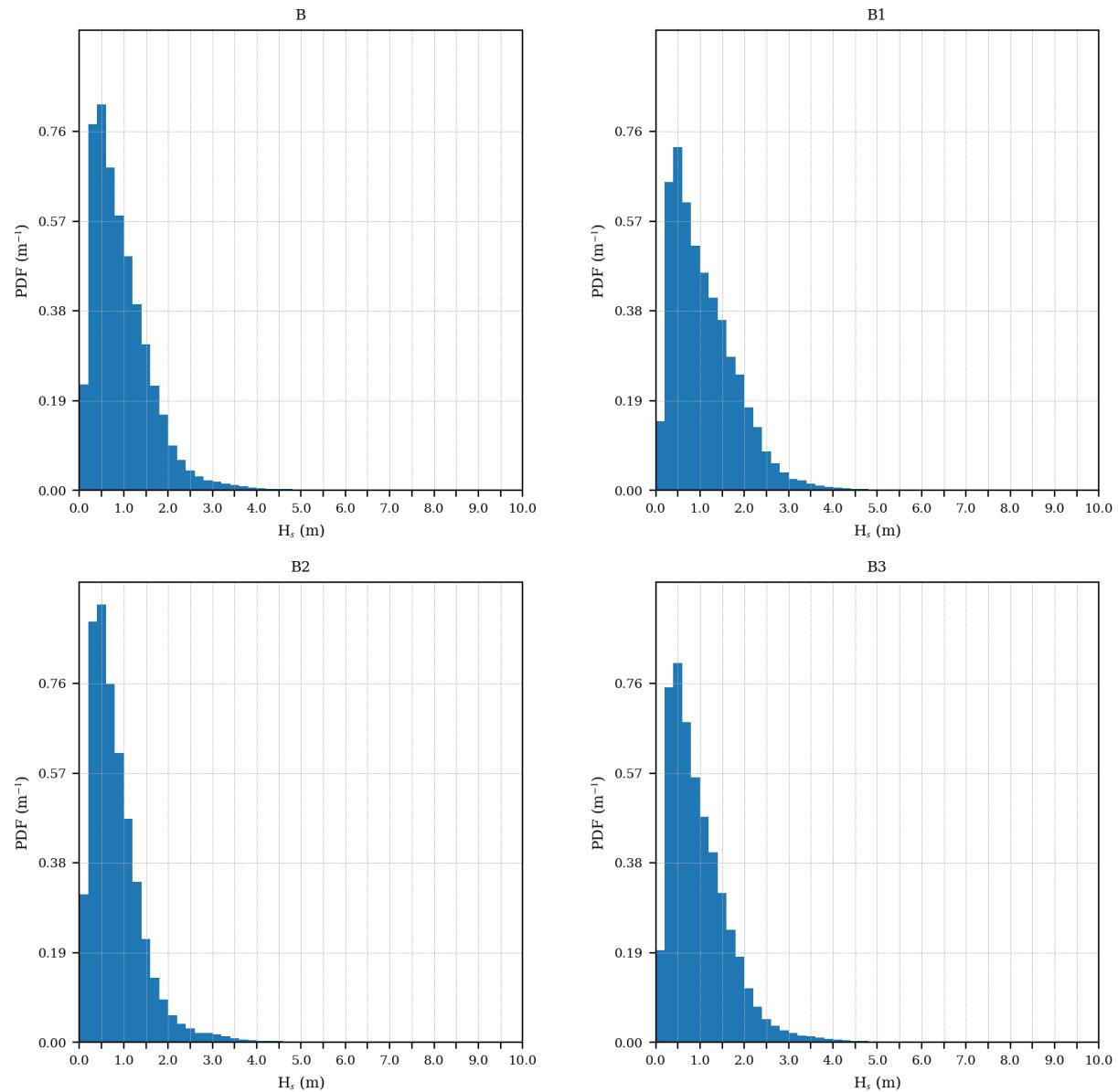


Figure 10 : Fonction de densité de probabilité des hauteurs significatives H_s aux points B, B1, B2 et B3, situés dans la zone d'étude pour l'implantation du parc éolien. Les intervalles utilisés pour le calcul de la fonction de densité de probabilité sont de 0.2 m, dont chacun inclut la borne inférieure et exclut la borne supérieure (ex: [0, 0.2[).

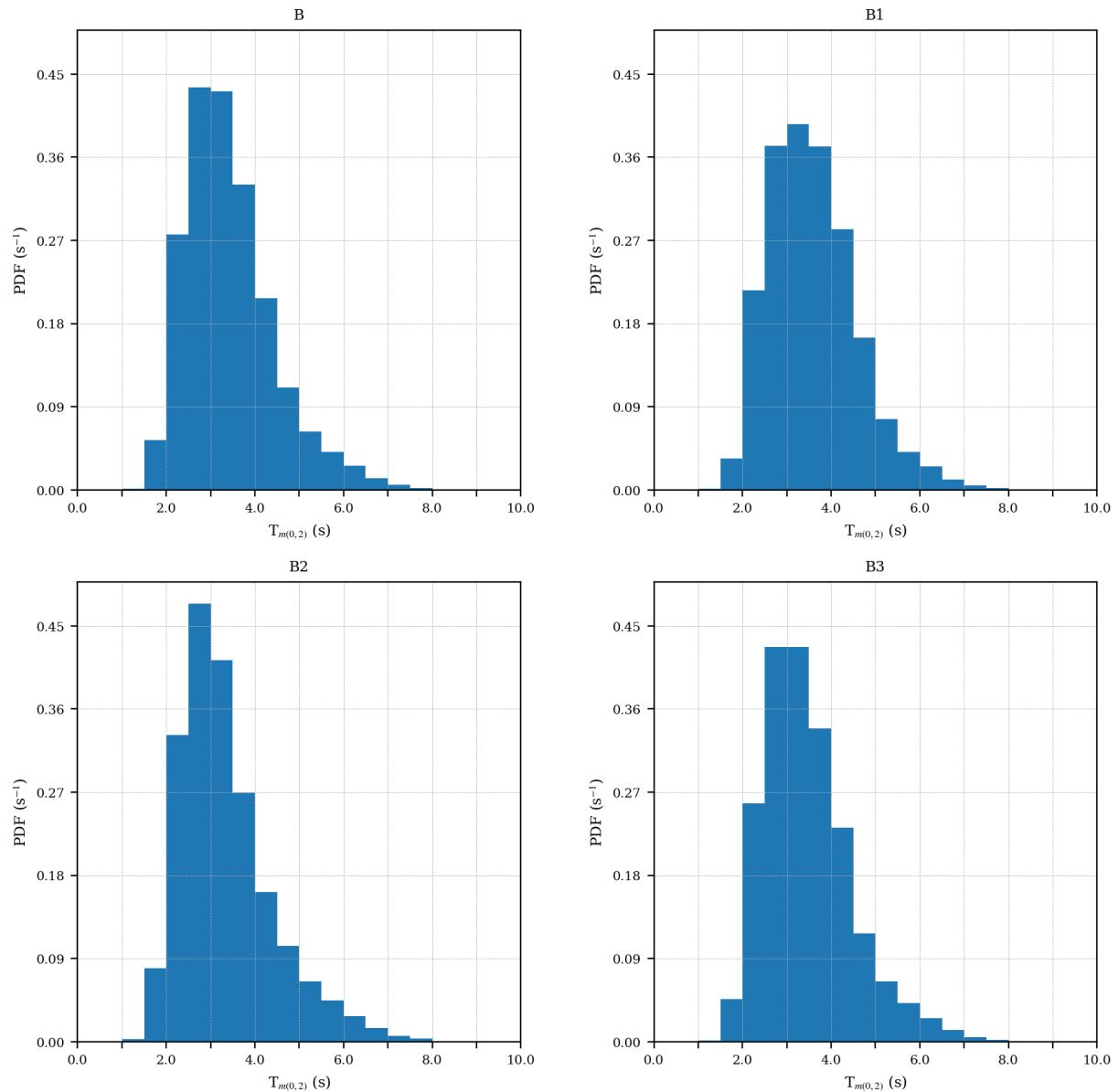


Figure 11 : Fonction de densité de probabilité de la période moyenne $T_{m(0,2)}$ aux points B, B1, B2 et B3, situés dans la zone d'étude pour l'implantation du parc éolien. Les intervalles utilisés pour le calcul de la fonction de densité de probabilité sont de 0.5 seconde, dont chacun inclut la borne inférieure et exclut la borne supérieure (ex : [5, 5.5]).

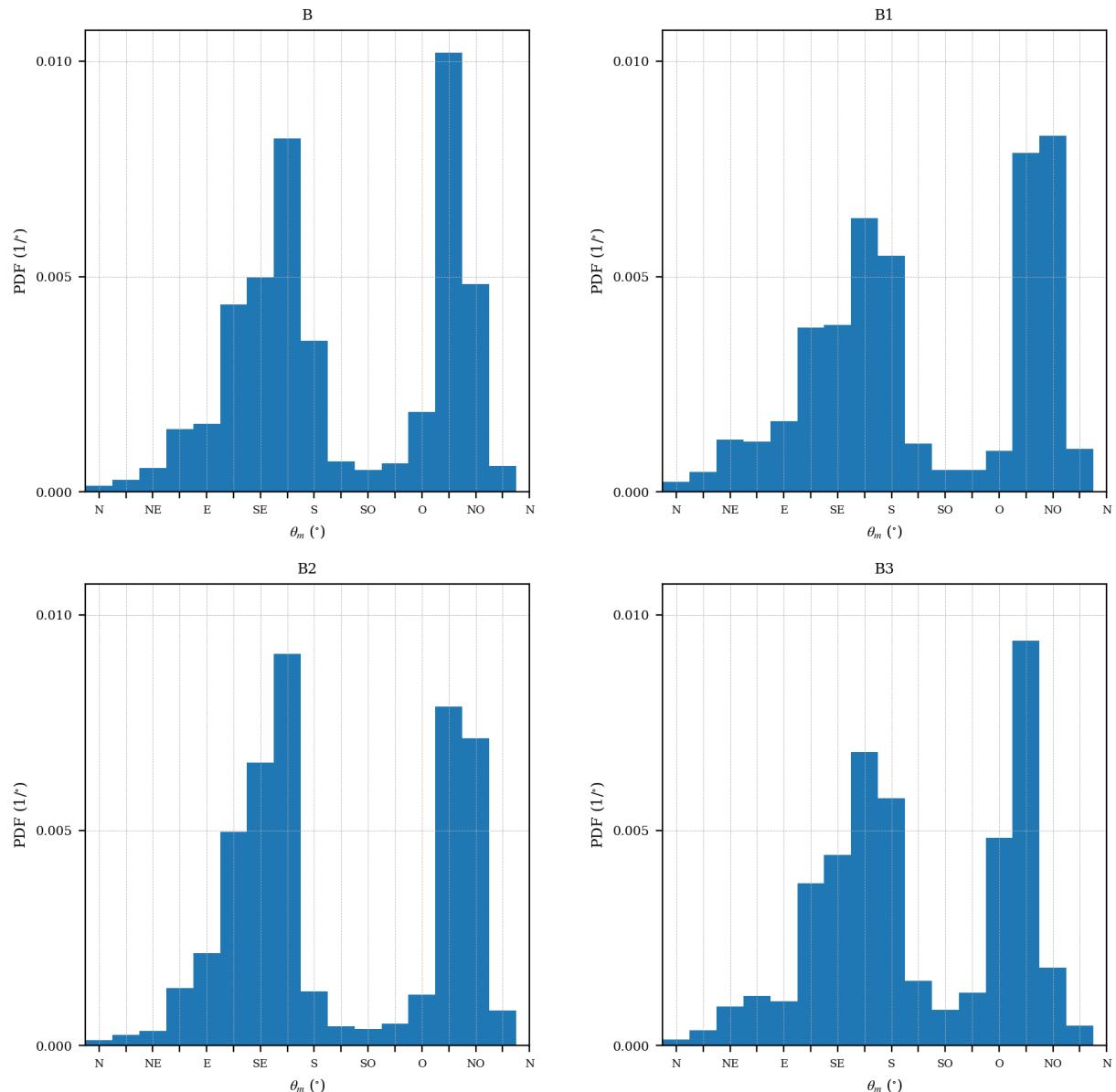


Figure 12 : Fonction de densité de probabilité des directions moyennes θ_m aux points B, B1, B2 et B3, situés dans la zone d'étude pour l'implantation du parc éolien. Les intervalles utilisés pour le calcul de la fonction de densité de probabilité sont de 22.5 degrés, dont chacun inclut la borne inférieure et exclut la borne supérieure (ex : [0, 22.5]).

La Figure 13 montre les corrélogrammes entre la hauteur significative des vagues et leur période moyenne, pour chacun des quatre points analysés. Globalement, ils confirment que le climat d'état de mer est assez homogène sur l'ensemble de la zone : il est principalement constitué de vagues avec des périodes moyennes généralement comprises entre 2 et 5 secondes, associées à des hauteurs significatives inférieures à 2.5 mètres. Les états de mer plus énergétiques sont représentés par des vagues pouvant dépasser les 7 m de hauteur significative associée à des périodes moyennes supérieures à 8 secondes. Ces événements restent cependant statistiquement assez rares.

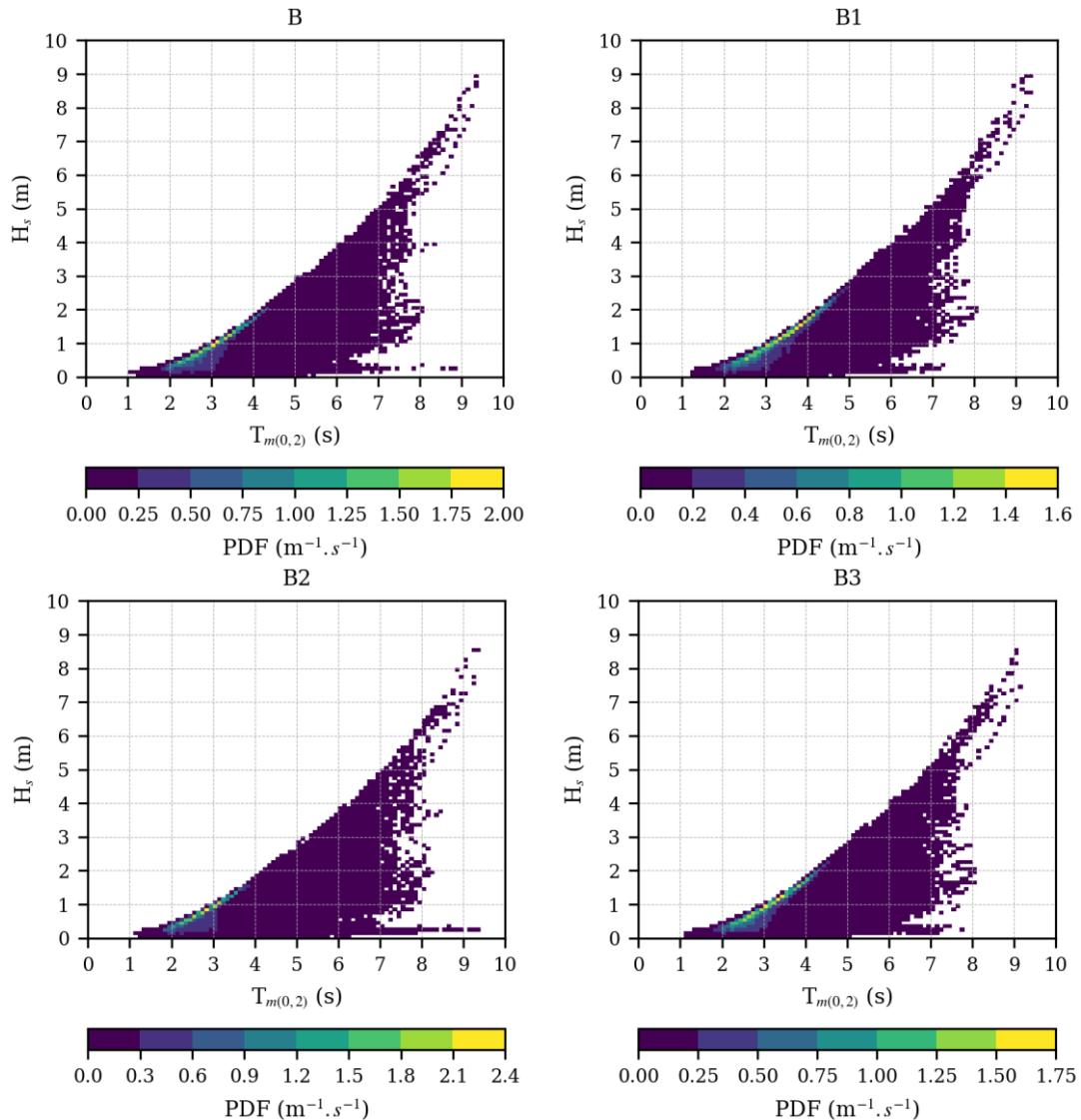


Figure 13 : Corrélogrammes hauteurs significatives (H_s) / périodes moyennes ($T_{m(0,2)}$) pour les quatre points du modèle analysés.

La Figure 14 présente les statistiques mensuelles de la hauteur significative pour les quatre points du modèle analysés. Elle confirme une faible variabilité spatiale sur la zone d'étude. On peut également y distinguer la variabilité saisonnière qui module la distribution de la hauteur significative au cours de l'année. Les moyennes et médianes mensuelles sont voisines de 0.5 m de juin à août; et supérieures à 1 m d'octobre à mars (excepté au point B2 où elles sont plus faibles). Les valeurs maximales ne dépassent pas 3 m en juin et juillet, et sont supérieures à 6 m de décembre à mars.

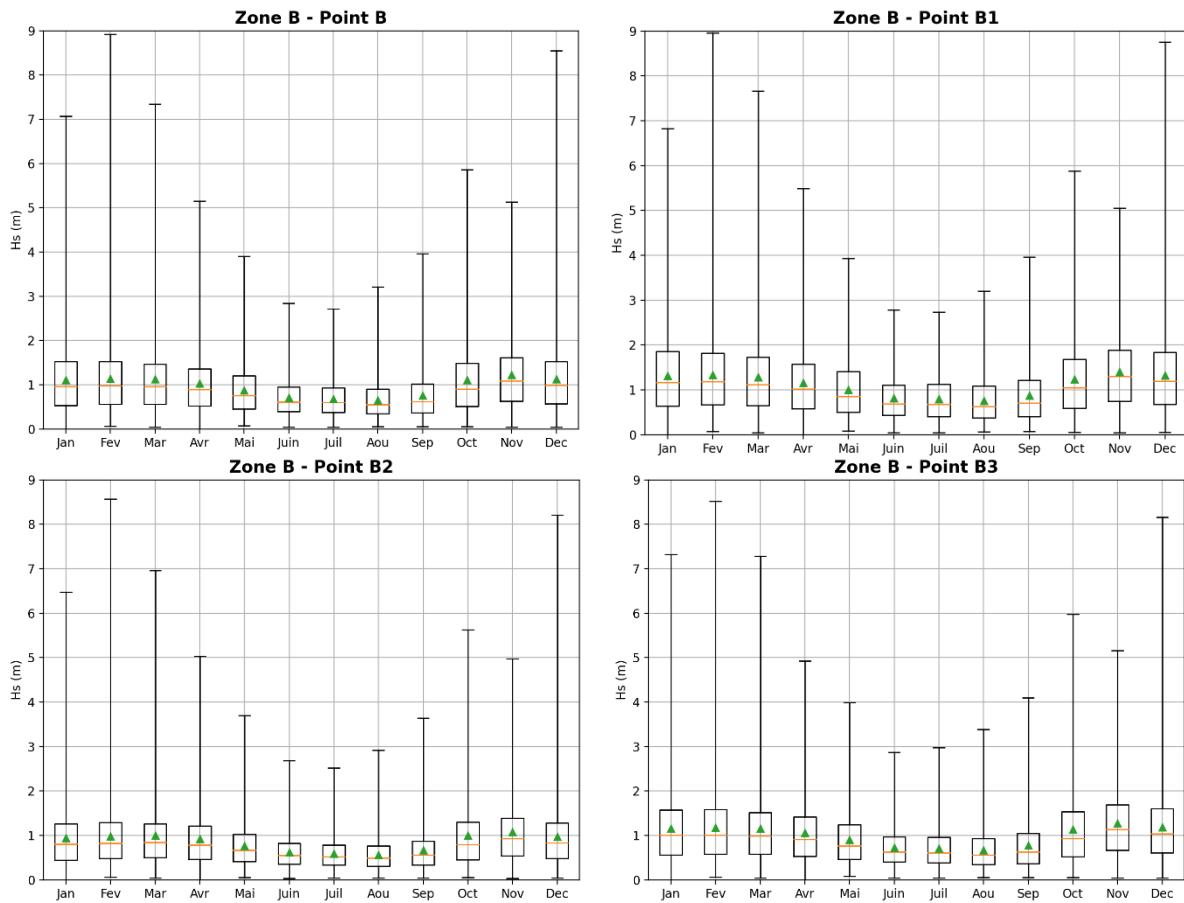


Figure 14 : Statistiques mensuelles de la hauteur significative (H_s) pour les quatre points du modèle analysés. Représentation sous forme de diagrammes en boîte : les extrémités des boîtes correspondent aux 1^{er} et 3^{ème} quartiles de la distribution mensuelle, les segments inférieurs et supérieurs représentent les valeurs minimales et maximales mensuelles, les traits orange sont les valeurs médianes mensuelles et les triangles verts sont les valeurs moyennes mensuelles.

3.2 Analyse de l'état de mer en fonction de la direction moyenne des vagues

Pour les quatre points analysés dans la zone, la Figure 15 montre la répartition des hauteurs significatives des vagues H_s en fonction de leur direction moyenne de provenance θ_m . Dans la colonne de gauche, ces roses de vagues sont calculées à partir de la totalité de la période disponible (2003-2022), tandis que dans la colonne de droite, elles sont calculées à partir des événements dont la hauteur significative dépasse 5 mètres.

Dans la colonne de gauche, on observe que les vagues proviennent majoritairement de deux secteurs : le secteur d'ouest à nord-ouest associé aux vents de terre avec des fetchs courts, et le secteur d'est à sud, associé à des vagues plus longues provenant du sud de la Méditerranée occidentale. En fonction du point considéré, on peut également distinguer quelques effets de réfraction sur ces directions d'incidence. La présence de la côte au nord-ouest de la zone entraîne un fetch court et limite en général le développement des vagues provenant de ces secteurs. Néanmoins, dans le Tableau 6, on peut observer que la hauteur significative maximale simulée au point B1 pour les secteurs ouest à nord-ouest dépasse 4 m.

Dans la colonne de droite qui considère uniquement les hauteurs significatives supérieures à 5 m, on peut constater que les conditions énergétiques proviennent quasi exclusivement du secteur est-sud-est pour les quatre points analysés. Cela est dû à un fetch bien plus important dans ce secteur, qui permet un développement plus important des vagues. Les vagues générées par des dépressions passant plus au sud peuvent également atteindre la zone.

En complément de ces roses de vagues, le Tableau 6 regroupe les statistiques de valeurs moyennes, maximales, médianes ainsi que les percentiles 90% et 95% de la hauteur significative H_s et de la période moyenne $T_{m(0,2)}$ pour les quatre points analysés et par direction de provenance des vagues.

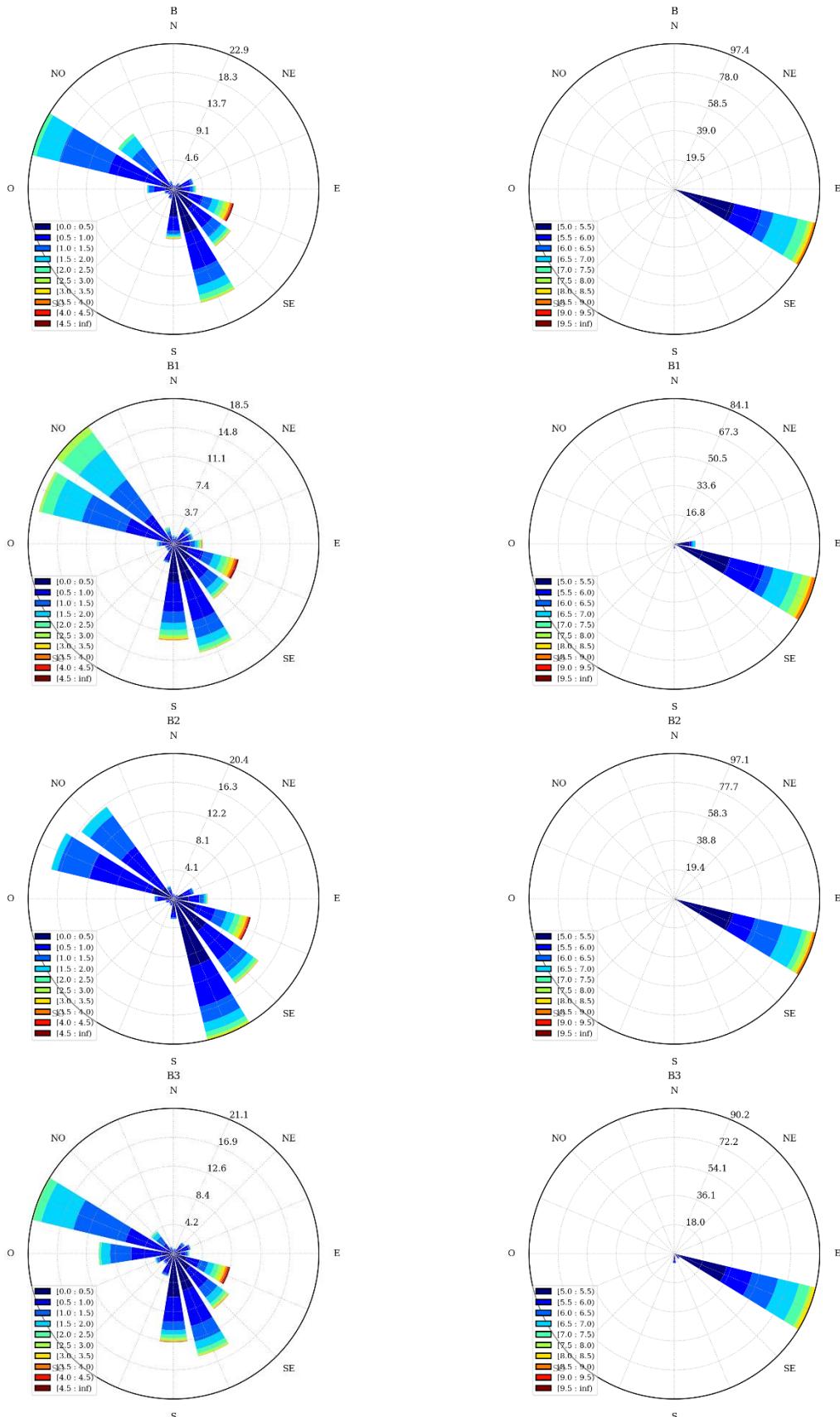


Figure 15 : Distribution de la hauteur significative de vagues H_s en fonction de la direction moyenne de provenance pour les points B, B1, B2 et B3 (de haut en bas). Les secteurs de directions sont de 22.5° , et l'unité radiale est %. La colonne de gauche considère l'ensemble des états de mer sur la période 2003-2022. La colonne de droite considère les états de mer dont la hauteur significative est supérieure à 5 m.

Secteur	Point	H_s (m)					$T_{m(0,2)}$ (s)				
		Mean	Max	Median	q90	q95	Mean	Max	Median	q90	q95
N	B	0.85	2.77	0.72	1.58	1.90	2.99	5.26	2.86	3.97	4.34
	B1	1.02	3.36	0.87	1.98	2.32	3.25	5.64	3.12	4.42	4.80
	B2	0.74	2.43	0.65	1.39	1.66	2.83	5.05	2.72	3.81	4.13
	B3	0.91	2.81	0.78	1.63	1.97	3.05	5.21	2.94	4.00	4.33
NE	B	0.72	2.79	0.62	1.32	1.58	3.05	5.29	2.98	3.91	4.20
	B1	0.81	3.50	0.69	1.51	1.83	3.17	5.83	3.08	4.11	4.46
	B2	0.67	2.45	0.57	1.27	1.51	3.02	5.07	2.94	3.94	4.24
	B3	0.73	2.98	0.64	1.28	1.53	2.94	5.49	2.86	3.72	4.03
E	B	1.16	8.55	0.74	2.74	3.58	4.04	9.27	3.74	6.12	6.67
	B1	1.37	8.96	0.97	3.09	3.87	4.30	9.37	4.17	6.23	6.74
	B2	1.05	8.57	0.68	2.44	3.23	4.06	9.31	3.76	6.08	6.66
	B3	1.07	7.63	0.70	2.53	3.34	3.82	8.91	3.47	5.89	6.48
SE	B	0.92	8.92	0.70	1.92	2.41	3.77	9.31	3.64	5.28	5.85
	B1	0.96	6.52	0.75	1.96	2.46	3.78	8.20	3.66	5.25	5.81
	B2	0.83	6.04	0.63	1.75	2.18	3.71	9.34	3.59	5.19	5.77
	B3	1.07	8.52	0.79	2.23	2.86	3.94	9.11	3.80	5.55	6.08
S	B	0.77	5.51	0.52	1.74	2.26	3.37	8.78	3.12	4.89	5.36
	B1	0.85	5.88	0.61	1.85	2.34	3.44	7.82	3.23	4.88	5.32
	B2	0.64	4.84	0.43	1.41	1.99	3.32	7.75	3.03	4.92	5.44
	B3	0.79	5.98	0.54	1.74	2.29	3.43	7.96	3.20	4.94	5.42
SO	B	0.58	3.07	0.47	1.11	1.40	2.64	5.70	2.47	3.60	3.95
	B1	0.64	3.39	0.52	1.18	1.49	2.77	5.92	2.58	3.79	4.15
	B2	0.52	2.61	0.44	0.99	1.24	2.55	5.38	2.40	3.40	3.71
	B3	0.62	3.56	0.50	1.20	1.50	2.73	5.94	2.57	3.72	4.13
O	B	0.90	3.58	0.84	1.52	1.72	2.95	5.73	2.91	3.82	4.05
	B1	0.95	4.41	0.83	1.73	1.99	3.07	6.34	2.97	4.12	4.40
	B2	0.70	2.85	0.64	1.19	1.35	2.63	5.10	2.57	3.40	3.61
	B3	1.08	4.12	1.03	1.78	1.97	3.23	6.13	3.21	4.12	4.31
NO	B	1.15	3.08	1.11	1.80	1.97	3.28	5.35	3.28	4.10	4.29
	B1	1.38	4.20	1.34	2.19	2.41	3.59	6.20	3.61	4.53	4.74
	B2	0.92	2.92	0.88	1.44	1.60	2.94	5.16	2.93	3.68	3.87
	B3	1.21	3.30	1.18	1.90	2.10	3.37	5.43	3.38	4.22	4.42

Tableau 6 : Statistiques (valeurs moyenne, maximale, médiane, quantiles 90% et 95%) des paramètres de hauteur significative H_s et de période moyenne $T_{m(0,2)}$ par direction de provenance des vagues, pour les points B, B1, B2 et B3.

Les Figure 16 et Figure 17 présentent les fonctions de densité de probabilité (PDF : Probability Density Function) associées à chaque secteur directionnel pour les paramètres H_s et $T_{m(0,2)}$. Sur cette figure, la distribution de chaque secteur directionnel est superposée avec une couleur différente.

Pour la majorité des secteurs directionnels, on observe une prédominance des hauteurs significatives autour d'un mètre et une densité très faible pour les vagues de plus de quatre mètres. On retrouve des distributions légèrement plus fortes pour les secteurs est et sud-est.

La période moyenne est généralement comprise entre 2 et 6 s. Seuls les secteurs est, sud-est et sud présentent une densité significative de vagues de plus de 6 s.

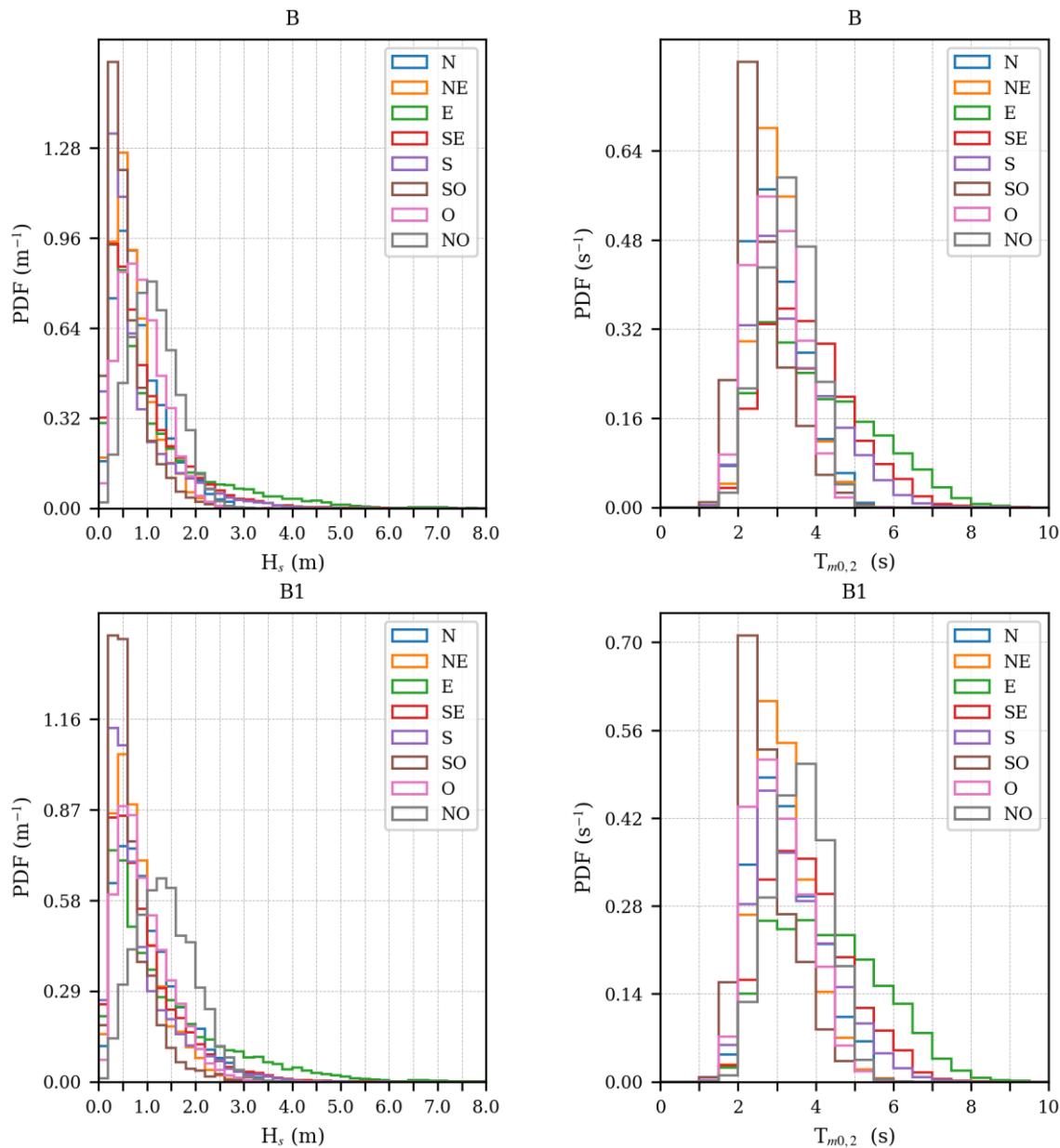


Figure 16 : Densité de probabilité des hauteurs significatives H_s (à gauche) et des périodes moyennes $T_{m0,2}$ (à droite) par direction de provenance des vagues au point B (en haut) et B1 (en bas), pour la période modélisée (2003-2022). Chaque couleur représente la distribution d'un secteur directionnel. Les intervalles utilisés pour le calcul de la fonction de densité de probabilité sont 0.2 m pour H_s et de 0.5 seconde pour $T_{m0,2}$, dont chacun inclut la borne inférieure et exclut la borne supérieure (ex : [1.5, 2.5[).

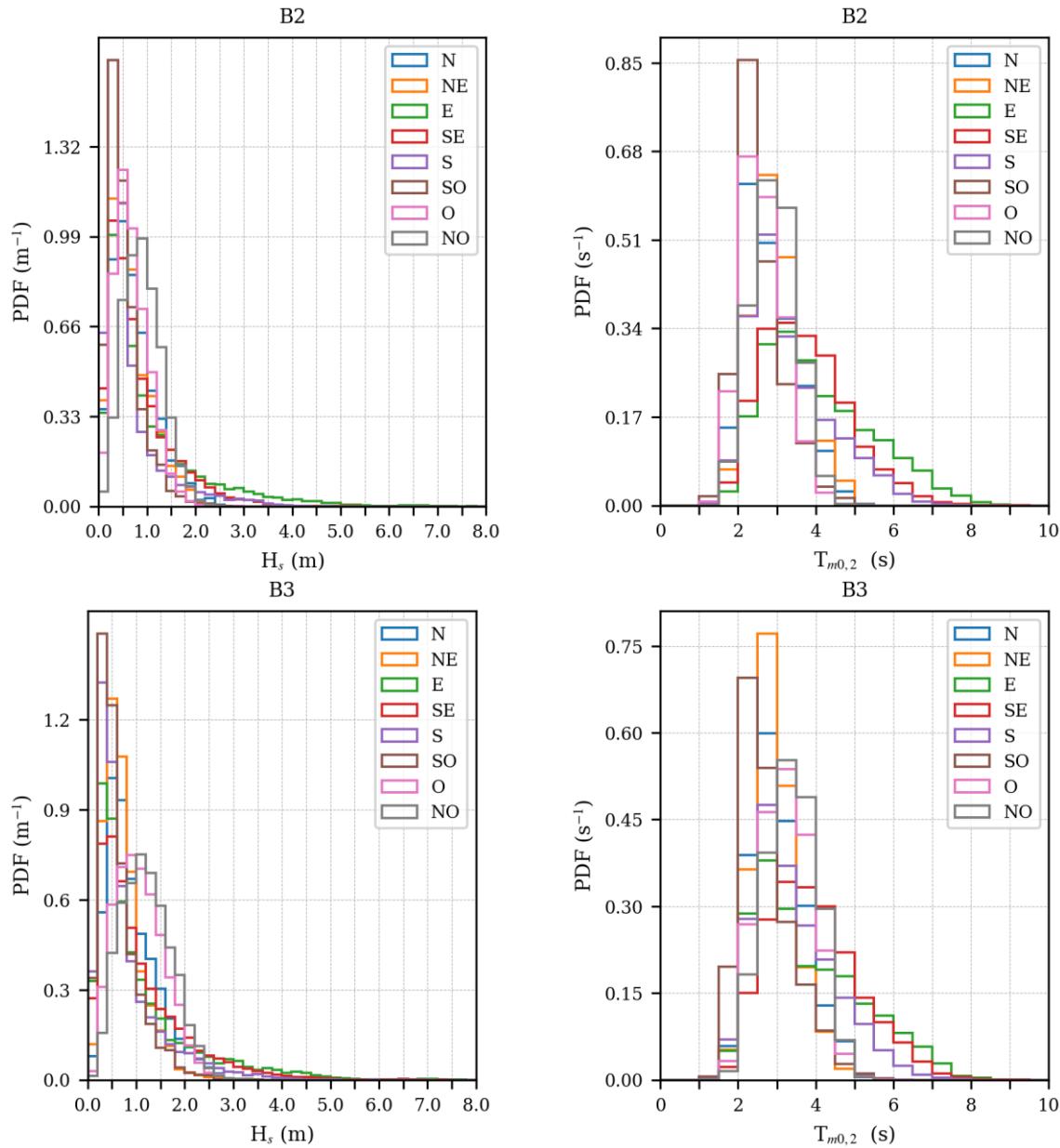


Figure 17 : Densité de probabilité des hauteurs significatives H_s (à gauche) et des périodes moyennes $T_{m0,2}$ (à droite) par direction de provenance des vagues au point B2 (en haut) et B3 (en bas), pour la période modélisée (2003-2022). Chaque couleur représente la distribution d'un secteur directionnel. Les intervalles utilisés pour le calcul de la fonction de densité de probabilité sont 0.2 m pour H_s et de 0.5 seconde pour $T_{m0,2}$, dont chacun inclut la borne inférieure et exclut la borne supérieure (ex : [1.5, 2.5[).

Les Figure 18 et Figure 19 montrent les corrélogrammes entre la hauteur significative des vagues et leur période moyenne, pour les vagues provenant respectivement de l'est et du nord-ouest. Les figures pour les autres secteurs directionnels se trouvent en Annexe 1.

Le climat d'état de mer en provenance du secteur Est est principalement constitué de vagues relativement courtes avec des périodes moyennes comprises entre 2 et 5 secondes et des hauteurs significatives inférieures à 2 m. Cependant, dans ce secteur, on retrouve les états de mer les plus énergétiques qui se caractérisent par des vagues de hauteur significative pouvant dépasser 8 m, pour des périodes moyennes de 9 à 10 s.

Pour les états de mer en provenance du nord-ouest, le fetch relativement court empêche le développement de vagues longues. Ainsi, les périodes moyennes restent inférieures à 5 – 6 secondes, et les hauteurs significatives inférieures à 3 – 4 mètres.

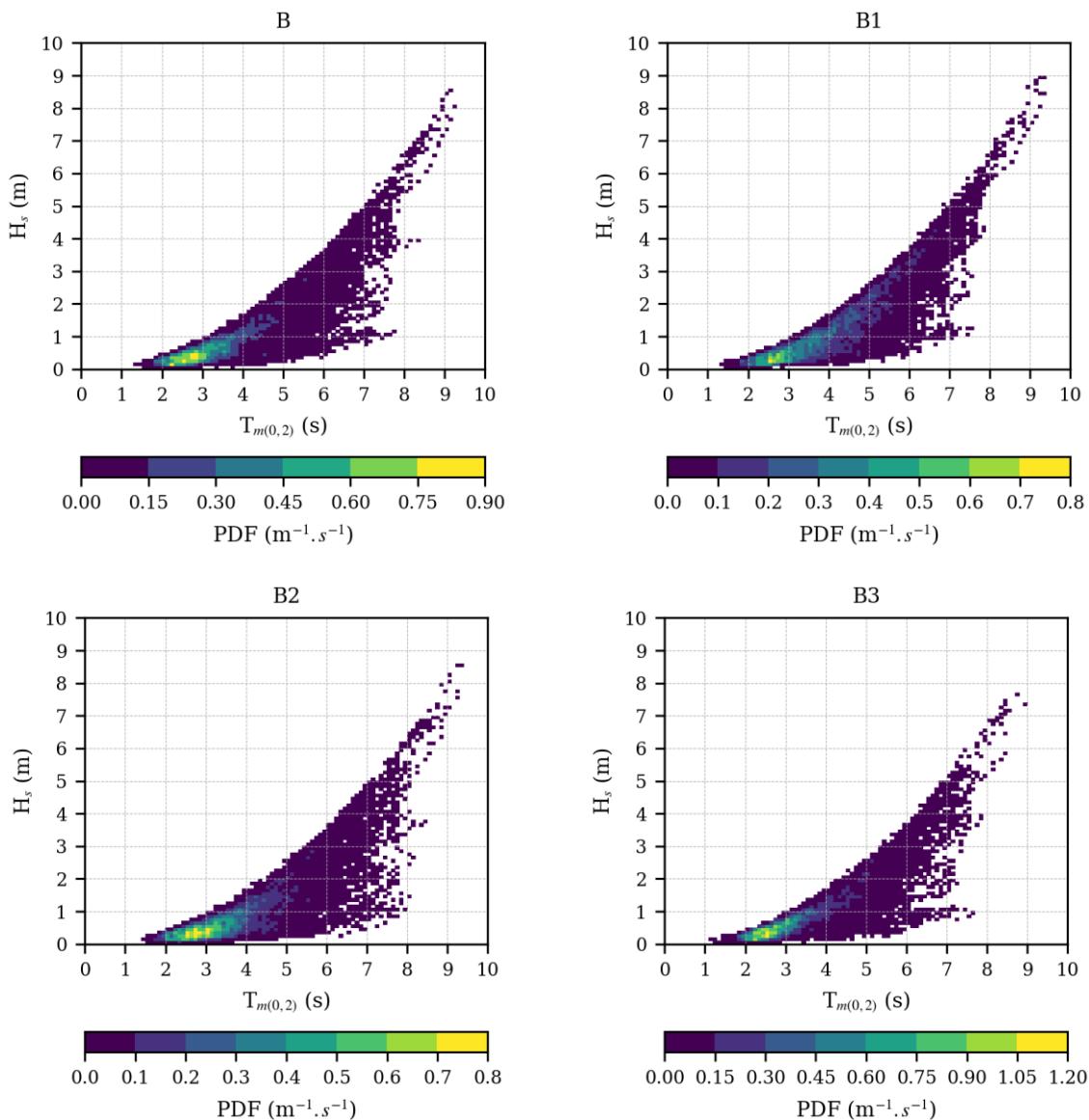


Figure 18 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant de l'Est ($67.5^\circ < \theta_m \leq 112.5^\circ$).

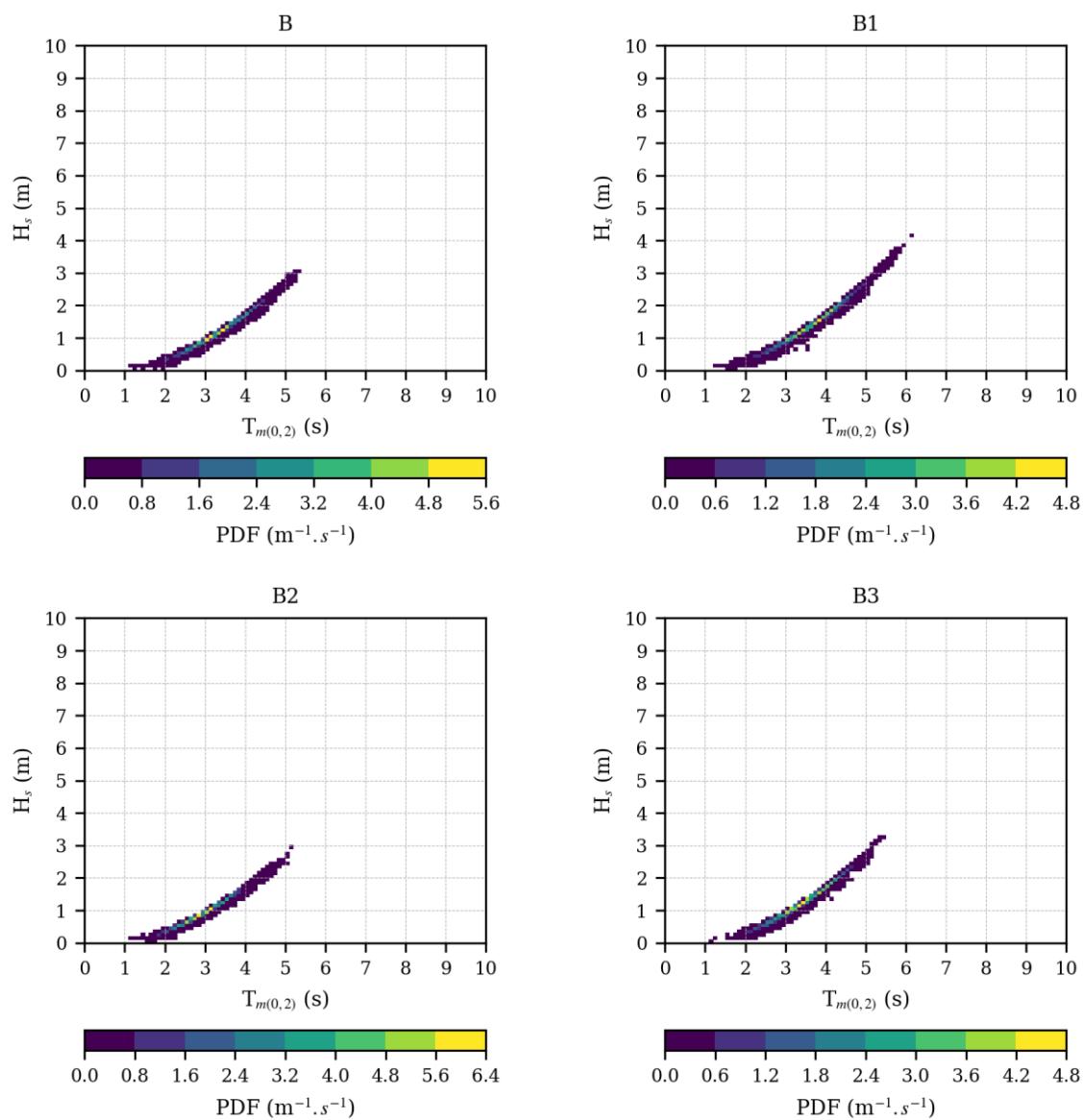


Figure 19 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Nord-Ouest ($292.5^\circ < \theta_m \leq 337.5^\circ$).

3.3 Analyse de l'état de mer par systèmes de vagues

L'état de mer peut être potentiellement séparé en différents systèmes de vagues (ou partitions). Le modèle WAVEWATCH III® permet une séparation automatique des différents systèmes de vagues. Dans la suite du paragraphe, une analyse par partition de vagues est proposée.

Parmi les différents systèmes de vagues qui peuvent constituer un état de mer, on distingue la mer de vent et des systèmes de houle. La mer de vent présente des vagues, en cours de développement sous l'effet du vent local, tandis qu'un système de houle correspond à la propagation de vagues formées précédemment à des distances plus ou moins éloignées et qui se sont propagées. La mer de vent est associée à des périodes / longueurs d'onde faibles, et présente des étalements fréquentiels et directionnels importants. Du fait de la dispersion des vagues au cours de leur propagation, les systèmes de houle sont associés à des périodes / longueurs d'onde plus importantes et présentent des spectres d'énergie plus étroits en fréquence et direction.

Trois systèmes ont été considérés : mer de vent, houles primaire et secondaire Pour chacun de ces trois systèmes de vagues :

- les Figure 20 et Figure 21 montrent les densités de probabilité des hauteurs significatives H_s , des périodes de pic T_p et des directions de provenance,
- les Figure 22 et Figure 23 montrent les corrélogrammes $H_s - T_p$,
- les Figure 24 et Figure 25 présentent la répartition des hauteurs significatives H_s en fonction de la direction au pic d'énergie Θ_p sous forme de roses de vagues.

Il ressort de cette analyse que la mer de vent est principalement constituée d'états de mer provenant du secteur ouest à nord-ouest, associés aux vents de terre. Elle est caractérisée par des hauteurs significatives d'énergie moyenne, généralement inférieures à 2.5-3 m et des périodes de pic comprises entre 2 et 6 s. Ce système est aussi associé aux états de mer les plus énergétiques qui se caractérisent par des vagues de hauteur significative pouvant dépasser 8 m, pour des périodes de pic pouvant dépasser 10-12 s, mais en provenance du secteur est-sud-est désormais.

La houle primaire se caractérise par des états de mer provenant des secteurs est à sud (principalement sud-sud-est et sud), généralement peu énergétiques, avec des hauteurs significatives inférieures à 1-2 m, et des périodes de pic majoritairement comprises entre 6 et 10 s, et pouvant exceptionnellement dépasser les 16 s.

La houle secondaire a des caractéristiques similaires à la houle primaire, mais se distingue par un niveau d'énergie plus faible, avec des hauteurs significatives généralement inférieures à 0.5-1 m, des périodes de pic un peu plus élevées (entre 8 et 12 s), et pouvant exceptionnellement dépasser les 16 s. L'incidence de la houle secondaire est également similaire à celle de la houle primaire : en provenance des secteurs est à sud, avec une dominance des secteurs sud-sud-est et sud, auxquels se rajoute le secteur est.

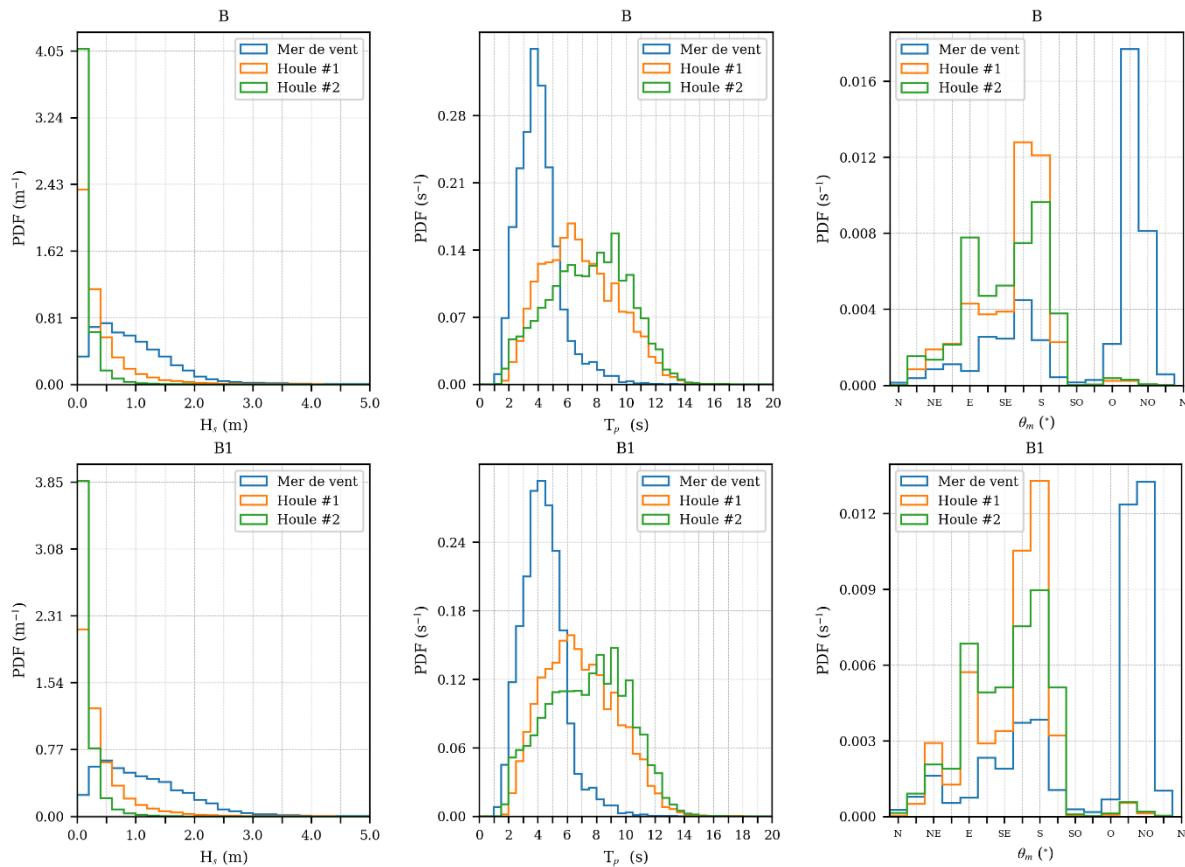


Figure 20 : Densité de probabilité des hauteurs significatives H_s (à gauche), des périodes de pic T_p (au milieu) et des directions de provenance (à droite), pour les différents systèmes de vagues, calculées aux points B (en haut) et B1 (en bas), pour l'ensemble de la période modélisée (2003-2022). La courbe bleue représente la densité de probabilité de la mer de vent, les courbes orange et verte celles des deux systèmes de houles les plus énergétiques (houle primaire et secondaire).

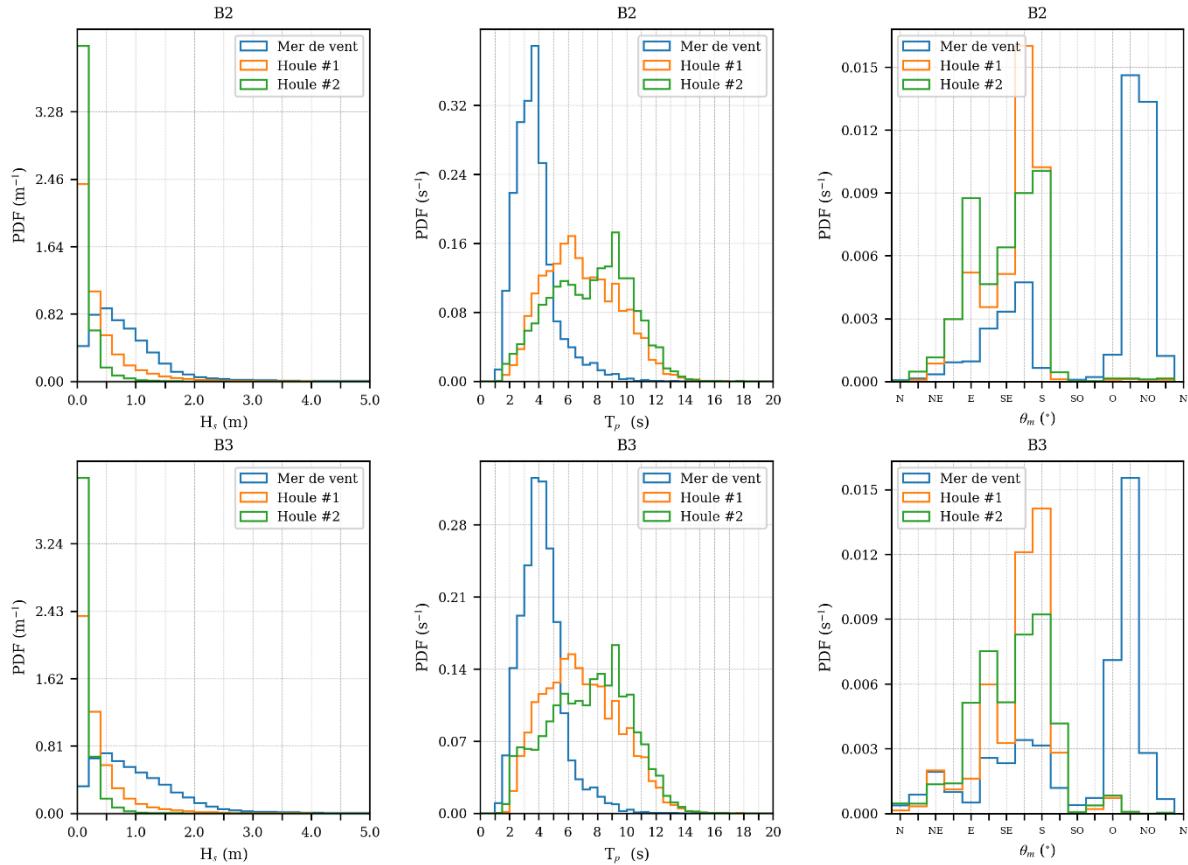


Figure 21 : Densité de probabilité des hauteurs significatives H_s (à gauche), des périodes de pic T_p (au milieu) et des directions de provenance (à droite), pour les différents systèmes de vagues, calculées aux points B2 (en haut) et B3 (en bas), pour l'ensemble de la période modélisée (2003-2022). La courbe bleue représente la densité de probabilité de la mer de vent, les courbes orange et verte celles des deux systèmes de houles les plus énergétiques (houle primaire et secondaire).

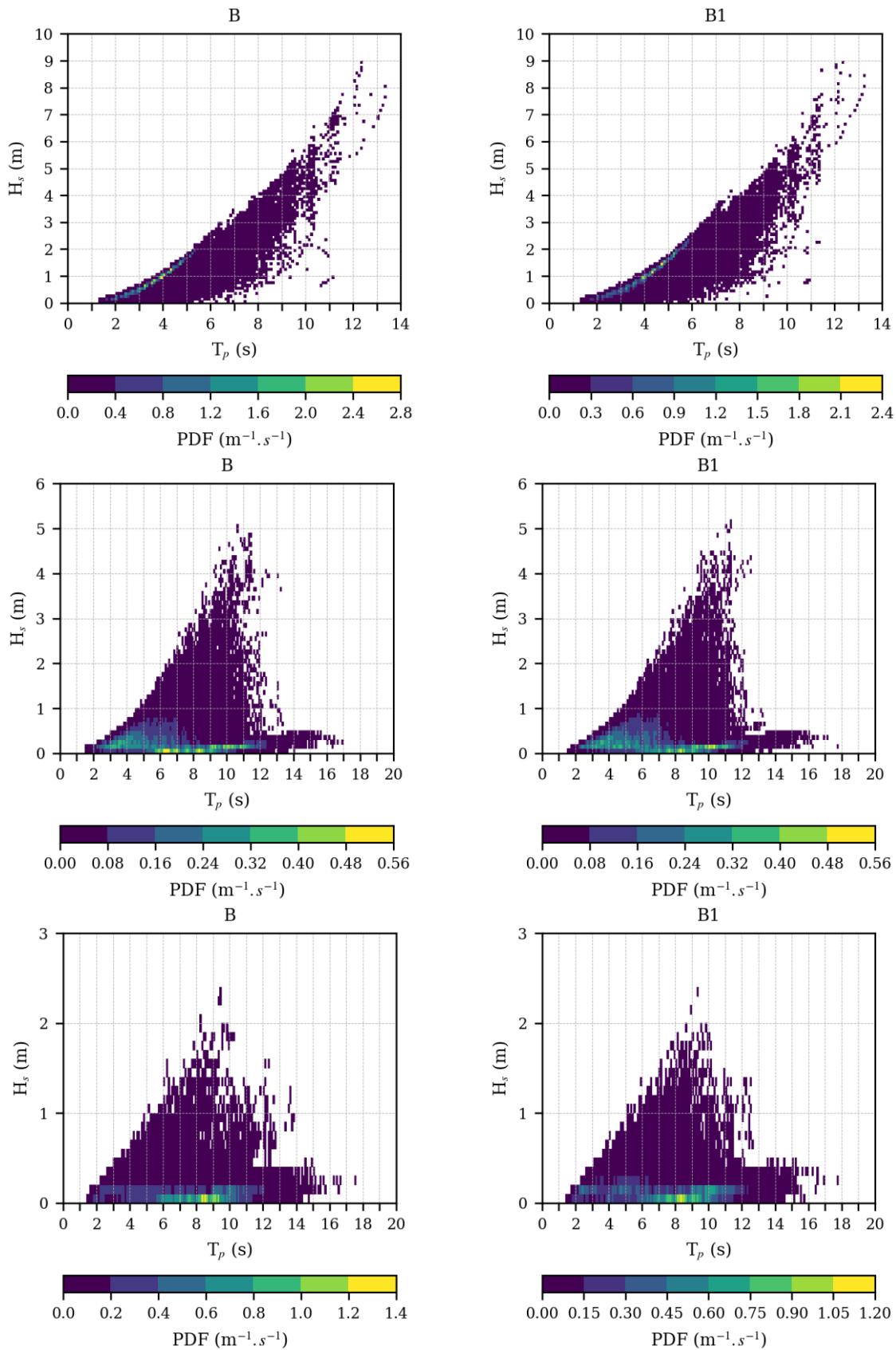


Figure 22 : Corrélogrammes H_s - T_p pour les différents systèmes de vagues : mer de vent (en haut) houle primaire (au milieu), et houle secondaire (en bas). Les colonnes de gauche et droite correspondent aux points B et B1, respectivement.

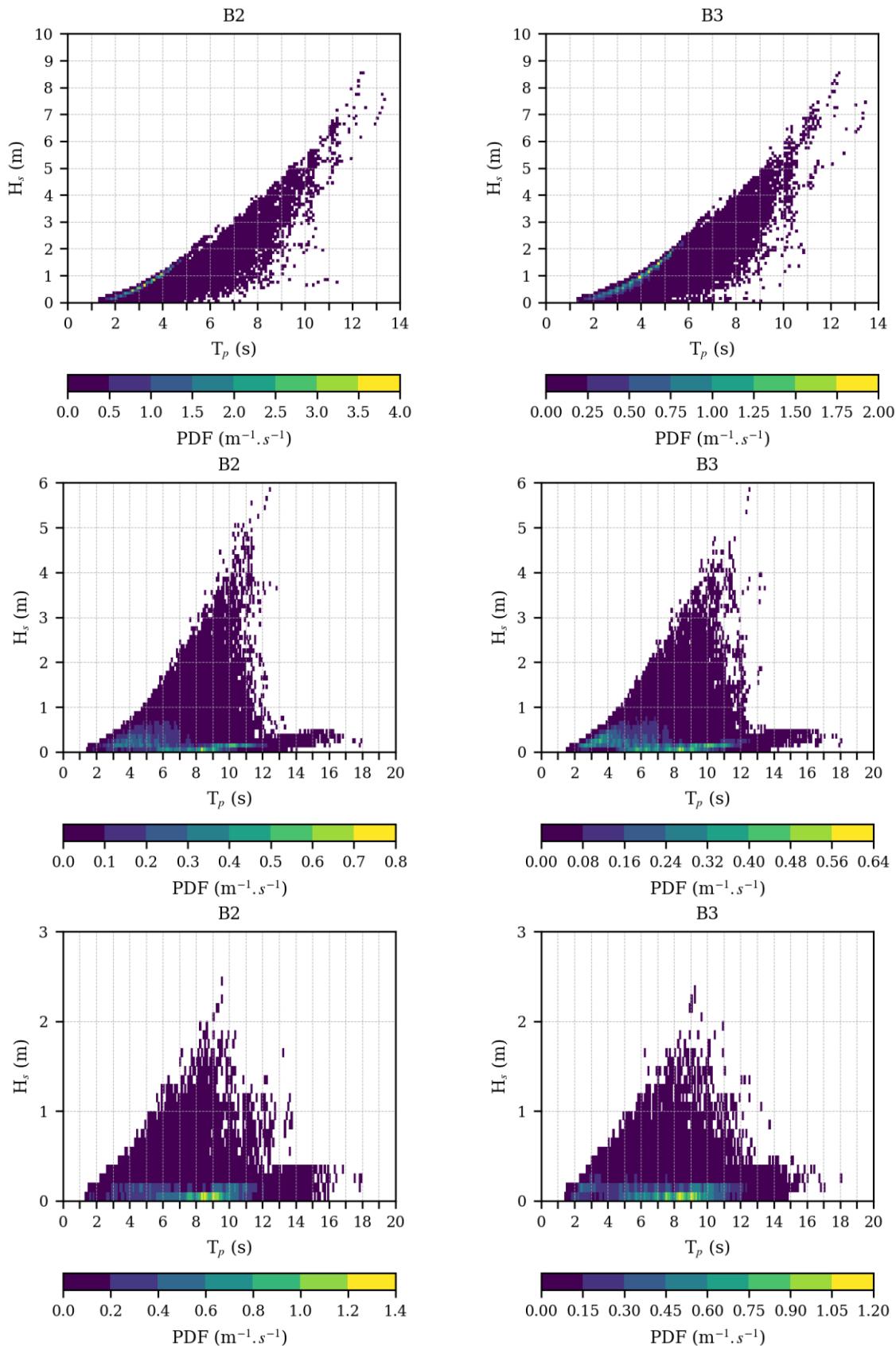


Figure 23 : Corrélogrammes H_s - T_p pour les différents systèmes de vagues : mer de vent (en haut) houle primaire (au milieu), et houle secondaire (en bas). Les colonnes de gauche et droite correspondent aux points B2 et B3, respectivement.

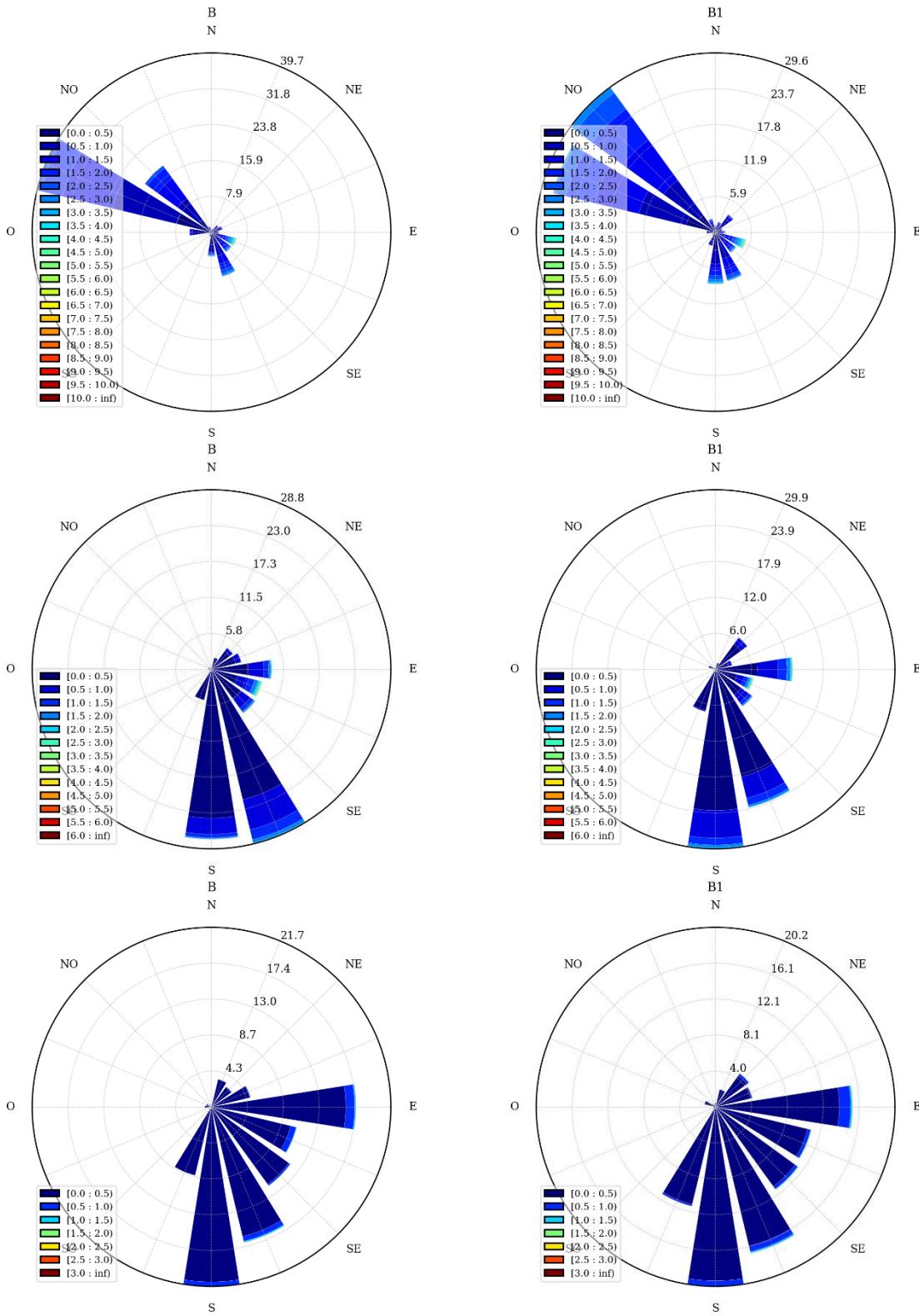


Figure 24: Répartition des hauteurs significatives H_s en fonction de la direction au pic d'énergie θ_p pour les différents systèmes de vagues : mer de vent (en haut) houle primaire (au milieu), et houle secondaire (en bas). Les colonnes de gauche et droite correspondent aux points B et B1, respectivement. L'unité radiale est %.

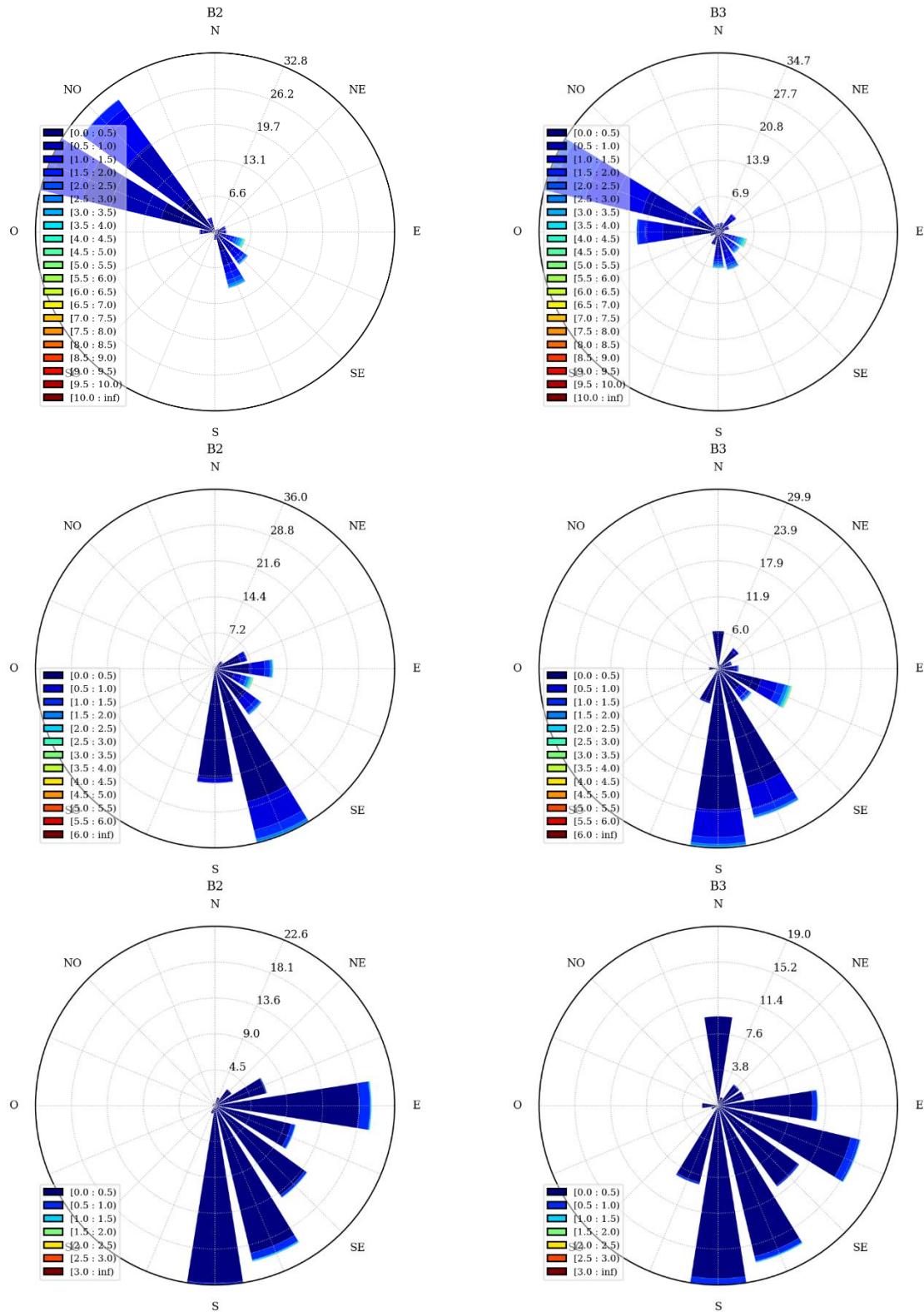


Figure 25 : Répartition des hauteurs significatives H_s en fonction de la direction au pic d'énergie θ_p pour les différents systèmes de vagues : mer de vent (en haut) houle primaire (au milieu), et houle secondaire (en bas). Les colonnes de gauche et droite correspondent aux points B2 et B3, respectivement. L'unité radiale est %.

4. Estimation des hauteurs significatives extrêmes associées à des périodes de retour

Le présent paragraphe traite de l'estimation statistique des hauteurs significatives extrêmes pour différentes périodes de retour.

4.1 Méthode de sélection des valeurs extrêmes

Une des approches de sélection des valeurs extrêmes souvent utilisée est celle nommée « *block maxima* ». Cette méthode utilise uniquement la valeur maximale rencontrée sur une période de temps donnée : un an ou un mois typiquement. Puis, ces valeurs maximales sont utilisées pour ajuster une distribution des valeurs extrêmes, typiquement selon la distribution GEV (Generalized Extreme Value). Le principal problème de la méthode des maxima annuels (mensuels) est la perte d'information engendrée par la sélection d'une seule valeur par an (mois), alors que plusieurs événements extrêmes peuvent se produire dans une même année (mois).

Une deuxième approche consiste à retenir toutes les valeurs qui dépassent un seuil préétabli. En utilisant cette méthode, nommée d'après son acronyme anglais POT (Peaks Over Threshold), tous les événements considérés comme « extrêmes » sont donc retenus pour l'analyse statistique. Par conséquent, cette approche est souvent préférable à la méthode *block maxima*. Néanmoins, il y a des difficultés associées à l'utilisation de la méthode POT, notamment liées au choix du temps minimal entre 2 pics de tempête, nécessaire pour assurer l'indépendance des événements, ainsi qu'au choix du seuil, qui reste une décision subjective malgré l'existence d'indicateurs basés sur les propriétés théoriques des distributions statistiques de valeurs extrêmes.

Pour cette étude, nous avons utilisé la méthode POT afin de sélectionner les pics de tempête qui servent ensuite à ajuster une distribution statistique spécifique aux valeurs extrêmes. Deux pics consécutifs de tempête sont considérés comme appartenant à la même tempête s'ils sont distants de moins de 24 heures. Cette valeur a été choisie suite à des essais avec différentes valeurs : 12 h, 24 h, 48 h et 72 h. Des analyses d'autocorrélation réalisées sur la série temporelle de H_s ont aussi confirmé l'adéquation de cette fenêtre temporelle. Les techniques utilisées pour la sélection du seuil sont décrites dans l'annexe 2.

4.2 Estimation des hauteurs significatives extrêmes associées à des périodes de retour

Dans ce paragraphe, nous allons estimer les hauteurs significatives associées à des périodes de retour en utilisant les 20 ans du rejet numérique au point B1 (Figure 1), situé le plus au large et pour lequel les conditions d'états de mer sont les plus sévères (Figure 9).

Après avoir testé l'adéquation de plusieurs distributions statistiques, la distribution de Pareto Généralisée (GPD : Generalized Pareto Distribution) telle que décrite par l'équation (Eq. 1) a été sélectionnée pour l'ajustement statistique aux valeurs extrêmes.

$$G(x) = \begin{cases} \left[1 - \left(1 + \frac{\xi(x-u)}{\sigma_u} \right)^{-1/\xi} \right] & \text{si } \xi \neq 0 \\ \left[1 - \exp \left(\frac{-(x-u)}{\sigma} \right) \right] & \text{si } \xi = 0 \end{cases} \quad (\text{Eq. 1})$$

où ξ et σ sont respectivement les paramètres de forme et d'échelle de la GPD, et u correspond au seuil pour la sélection des valeurs extrêmes de $H_s(x)$. Dans l'équation (Eq. 1), la ligne du bas correspond au cas particulier où le paramètre de forme ξ est égal à zéro. Dans ce cas, la GPD est équivalente à la distribution exponentielle.

Pour estimer les valeurs extrêmes de la hauteur significative, nous allons utiliser deux méthodes :

- Une méthode dite stationnaire, présentée au paragraphe 4.2.1, qui consiste à appliquer la méthode POT associée à la GPD de manière usuelle. Elle est qualifiée de stationnaire car les dates des valeurs retenues pour l'analyse n'influent pas dans les calculs.
- Cette méthode est ensuite adaptée aux paragraphes 4.2.2 pour l'utiliser sous forme non stationnaire, de façon à prendre en compte les variations saisonnières des valeurs de H_s .

L'influence des variations climatiques interannuelles sur les niveaux de retour a également été étudiée en essayant de prendre en compte différents indices climatiques (WeMO³ : Western Mediterranean Oscillation Index; MOI⁴ : Mediterranean Oscillation Indices; SaOI : Saharan Oscillation Index). Par analogie avec l'article de Izaguirre et al. (2010) et le travail de Leckler et al. (2018), une covariable linéaire supplémentaire a été introduite dans le paramètre de position du modèle statistique non-stationnaire présenté au paragraphe 4.2.2. Cependant, cette analyse n'a pas montré de corrélation claire entre ces indices climatiques et l'intensité des états de mer sur la période étudiée (2003-2021). L'influence des variations climatiques interannuelles sur les niveaux de retour extrêmes n'a donc pas été jugée suffisamment significative pour que les résultats soient présentés dans ce rapport.

4.2.1 Méthode stationnaire

Nous avons premièrement fait le choix du seuil de tempête à l'aide des tests de vérification des propriétés statistiques des distributions, notamment la variation linéaire des excès de H_s au seuil en fonction du seuil, et la propriété asymptotique de la loi GPD. Ces tests sont présentés en Annexe 2. En utilisant ces méthodes et après avoir réalisé des tests de sensibilité, la valeur de 4.02 mètres a été retenue pour le seuil de tempête. 58 pics de tempête ont ainsi été sélectionnés, soit 2.9 événements tempétueux par an en moyenne.

Avec le seuil choisi et en ajustant les distributions GPD et exponentielle (Eq. 1) aux pics de tempête à l'aide de la méthode d'estimation du maximum de vraisemblance (MLE : Maximum Likelihood Estimation), nous obtenons les résultats présentés dans le Tableau 7 et sur la Figure 26. Les résultats obtenus avec les ajustements aux deux lois paramétriques sont assez proches et commencent à différer pour des périodes de retour supérieures à 10 ans (Figure 26). À partir de cette période de retour, la loi GPD est plus conservative que la loi exponentielle.

Période retour (année)	H_s (m)		Intervalle de confiance (m)	
	GPD	Exponentielle	GPD	Exponentielle
1	5.16	5.18	[4.9 ; 5.5]	[4.8 ; 5.5]
5	6.97	6.94	[6.1 ; 7.7]	[6.1 ; 7.7]
10	7.77	7.69	[6.7 ; 8.7]	[6.7 ; 8.7]
50	9.70	9.45	[7.8 ; 11.5]	[8.0 ; 10.9]
100	10.57	10.20	[8.2 ; 12.8]	[8.6 ; 11.8]

Tableau 7 : Hauteurs significatives associées aux périodes de retour de 1, 5, 10, 50 et 100 ans. Les valeurs ont été obtenues à partir de l'ajustement des distributions GPD et exponentielle aux données du jeu numérique pour un seuil de 4.02 m. La dernière colonne montre les intervalles de confiance 95% des estimations, obtenus par la méthode Monte Carlo.

³ <http://www.ub.edu/gc/wemo/>

⁴ <https://crudata.uea.ac.uk/cru/data/moi/>

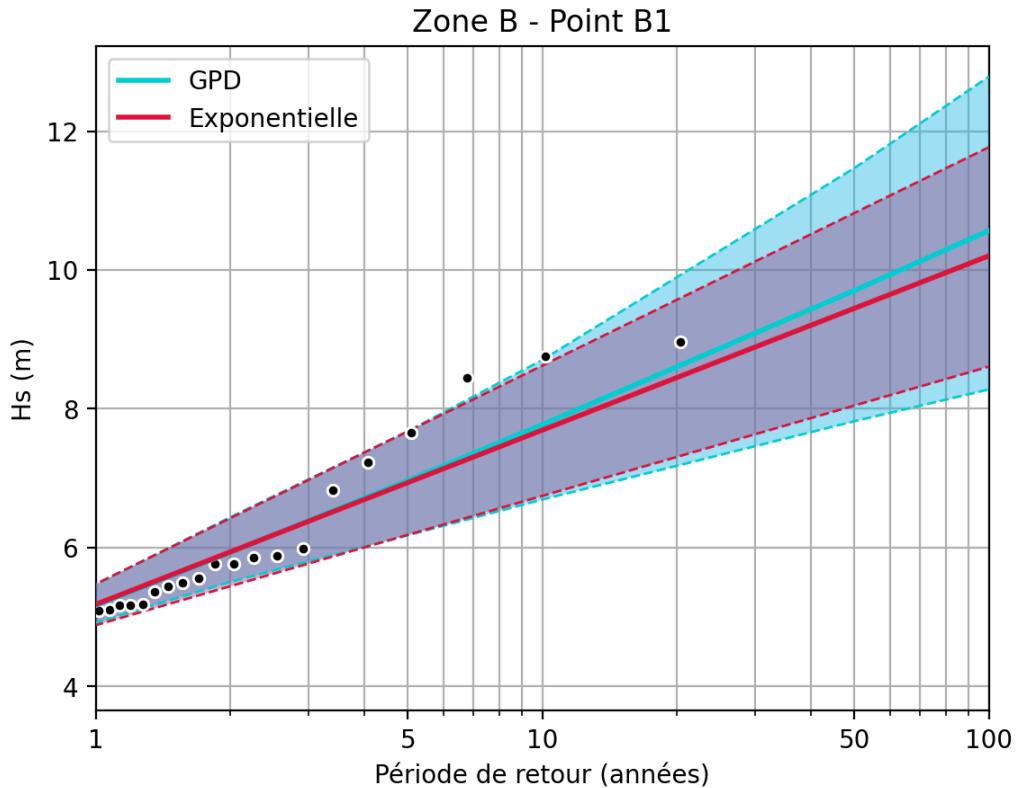


Figure 26 : H_s pour les périodes de retour de 1 à 100 ans. Les points noirs représentent les données du rejet. La ligne bleue est l'ajustement de la loi GPD (avec $\xi = 0.031$ et $\sigma = 1.057$), et la ligne rouge correspond à l'ajustement de la loi exponentielle (avec $\sigma = 1.091$). Les lignes pointillées correspondent aux intervalles de confiance 95% calculés par la méthode Monte Carlo.

4.2.2 Méthode non-stationnaire appliquée aux variations saisonnières

De manière à étudier des valeurs extrêmes d'une façon plus exhaustive, l'étude de l'intensité des tempêtes peut être accompagnée d'une analyse de leur fréquence d'occurrence. Pour cela, nous pouvons utiliser le modèle « *Point Process* » (PP), introduit par Smith (1989), et Davison et Smith (1990).

Cette méthode est équivalente au modèle Poisson-GPD, qui consiste à ajuster d'une part les pics de tempête au-dessus d'un seuil à une GPD, et de l'autre la fréquence d'occurrence des dépassements du seuil à une distribution de Poisson. Toutefois, le modèle PP permet également d'ajuster de façon simultanée la fréquence et l'intensité des dépassements du seuil. Pour cela, la paramétrisation du modèle est faite selon la distribution GEV, et donc les paramètres obtenus sont équivalents à ceux de la distribution GEV. De plus, le paramètre de forme est aussi équivalent à celui de la GPD, et le paramètre d'échelle peut être transformé pour obtenir celui de la GPD (voir Coles, 2001 - Chapitre 7).

Pour un nombre N de pics de tempête x_1, \dots, x_N , observés sur une période de n_y (ans), la fonction de vraisemblance du modèle PP non-stationnaire est exprimée de la façon suivante (Gilleland et al., 2016) :

$$\begin{aligned}
 l(\mu(t), \sigma(t), \xi; x_1, \dots, x_n) &= -\ln \sigma(t_i) - (1/\xi + 1) \sum_{i=1}^n \left[1 + \frac{\xi}{\sigma(t_i)} (x_i - u) \right]_{x_i > u} \\
 &\quad - n_y \left(1 + \frac{\xi}{\sigma(t_i)} (u - \mu(t_i)) \right)^{-1/\xi}
 \end{aligned} \tag{Eq. 2}$$

où μ, σ, ξ correspondent respectivement aux paramètres de position, échelle et forme, et sont équivalents à ceux de la distribution GEV. n_y désigne le nombre d'années des données, u correspond au seuil de tempête, et t le temps des observations.

Une fois que les paramètres μ, σ, ξ ont été estimés par l'équation (Eq. 2), les hauteurs significatives associées à des périodes de retour sont calculées par :

$$Hs_m(t) = \mu(t) - \frac{\sigma(t)}{\xi} [1 - [-\log(1 - q)]^{\xi}], \quad \text{pour } \xi(t) \neq 0 \quad (\text{Eq. 3})$$

où $Hs_m(t)$ est la valeur associée à la période de retour $m = 1/q$ (avec $0 < q < 1$, égal à la probabilité de dépassement) en fonction du temps.

Afin de prendre en compte les variations saisonnières de H_s , nous considérerons les paramètres μ et σ comme dépendants du temps. Nous notons que d'autres auteurs ont appliqué cette méthode non-stationnaire de manière comparable pour les vagues ou les précipitations (Mendez et al., 2006; Rueda et al., 2016; Menéndez et al., 2009; Mentaschi et al., 2016; Niroomandi et al., 2018; Leckler et al. 2018; Rahimpour et al., 2016).

Une série de Fourier (Eq. 4) peut être utilisée pour décrire n'importe quelle fonction périodique.

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cdot \cos\left(nx \frac{2\pi}{T}\right) + b_n \cdot \sin\left(nx \frac{2\pi}{T}\right) \quad (\text{Eq. 4})$$

Le cycle annuel de hauteurs significatives peut être correctement décrit par une fonction harmonique. L'équation (Eq. 5) montre la forme prise à l'ordre 1 par les paramètres ξ, μ , et σ du modèle PP en fonction du temps :

$$\begin{aligned} \xi &= \xi_0 \\ \mu(t) &= \mu_0 + \mu_1 \cdot \cos(2\pi t) + \mu_2 \cdot \sin(2\pi t) \\ \sigma(t) &= \sigma_0 + \sigma_1 \cdot \cos(2\pi t) + \sigma_2 \cdot \sin(2\pi t) \end{aligned} \quad (\text{Eq. 5})$$

Les résultats présentés par la suite seront issus du modèle utilisant la méthode PP dans une forme non-stationnaire pour représenter le cycle annuel. Les paramètres $\xi, \sigma(t)$ et $\mu(t)$ tels que décrits par l'équation (Eq. 5), sont ajustés aux pics de tempête modélisés en utilisant la méthode d'estimation du maximum de vraisemblance (Eq. 2).

Afin d'évaluer l'aptitude du modèle à estimer l'intensité des pics de tempête, nous avons tracé les quantiles des dépassements du seuil de H_s simulés, par rapport aux quantiles empiriques de la distribution exponentielle. Pour cela, les paramètres non-stationnaires obtenus avec l'équation (Eq. 2) ont été d'abord transformés pour pouvoir être comparés à ceux de la distribution exponentielle. Si une variable aléatoire Y_t suit une GPD non-stationnaire (nous rappelons ici que les paramètres du modèle PP peuvent être transformés pour obtenir ceux de la GPD) avec paramètres ξ et $\sigma(u, t)$, la variable transformée est alors :

$$\epsilon_t = [1/\xi] \ln\{1 + \xi [Y_t/\sigma(u, t)]\} \quad (\text{Eq. 6})$$

La variable transformée, ϵ_t , suit une distribution exponentielle. Or, un QQ-plot peut être généré et comparé avec les résultats théoriques pour la distribution exponentielle (Katz, 2012). La Figure 27 (à gauche) montre un bon accord entre les quantiles théoriques de la distribution exponentielle et nos résultats transformés.

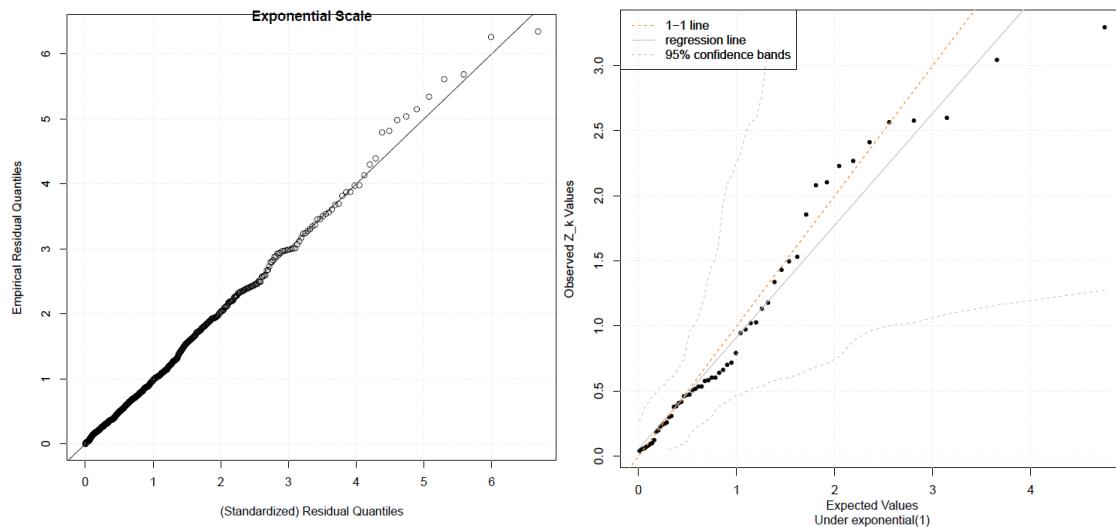


Figure 27 : Résultats des diagnostics du modèle PP non-stationnaire : QQ plot (à gauche) et Z plot (à droite).

Afin d'évaluer la capacité de l'ajustement à estimer la fréquence des événements extrêmes, nous avons utilisé la « statistique Z » (Smith et Shively, 1995). Cette statistique correspond au taux cumulatif des dépassements consécutifs du seuil (Z_k), qui doit être distribué de façon exponentielle avec moyenne égale à 1, si le modèle est correct. De cette manière, en traçant les quantiles des valeurs Z_k par rapport aux quantiles attendus pour une distribution exponentielle de moyenne égale à 1, nous pouvons évaluer dans quelle mesure notre modèle est capable d'estimer la fréquence d'occurrence des dépassements du seuil. Sur la Figure 27 (à droite), nous pouvons constater un bon ajustement du modèle aux valeurs attendues.

Après ajustement, nous obtenons les paramètres qui décrivent la saisonnalité. Ainsi, l'équation (Eq. 5) devient :

$$\begin{aligned} \xi &= -0.111 \\ \mu(t) &= 4.486 + 1.962 \cdot \cos(2\pi t) + 0.494 \cdot \sin(2\pi t) \\ \sigma(t) &= 0.669 + 0.489 \cdot \cos(2\pi t) + 0.085 \cdot \sin(2\pi t) \end{aligned} \quad (\text{Eq. 7})$$

Les valeurs extrêmes de hauteur significative pour les périodes de retour 1, 5, 10, 50 et 100 ans sont représentées sur la Figure 28. Les valeurs maximales atteintes et les intervalles de confiance à 95% associés sont regroupés dans le Tableau 8. Elles valent respectivement 5.13 m, 8.12 m, 8.83 m, 10.19 m et 10.70 m. Ces valeurs correspondent à la période de retour si tous les jours de l'année se comportaient en termes de tempêtes comme le pire jour de janvier. On peut également remarquer que la fonction harmonique sinusoïdale introduite ici pour représenter la variabilité du cycle annuel des hauteurs significatives ne permet pas de reproduire fidèlement cette variabilité : les valeurs maximales étant modélisées au cours des mois de février et décembre (Figure 14). Elle permet tout de même d'avoir une bonne approximation de ce cycle en première approche.

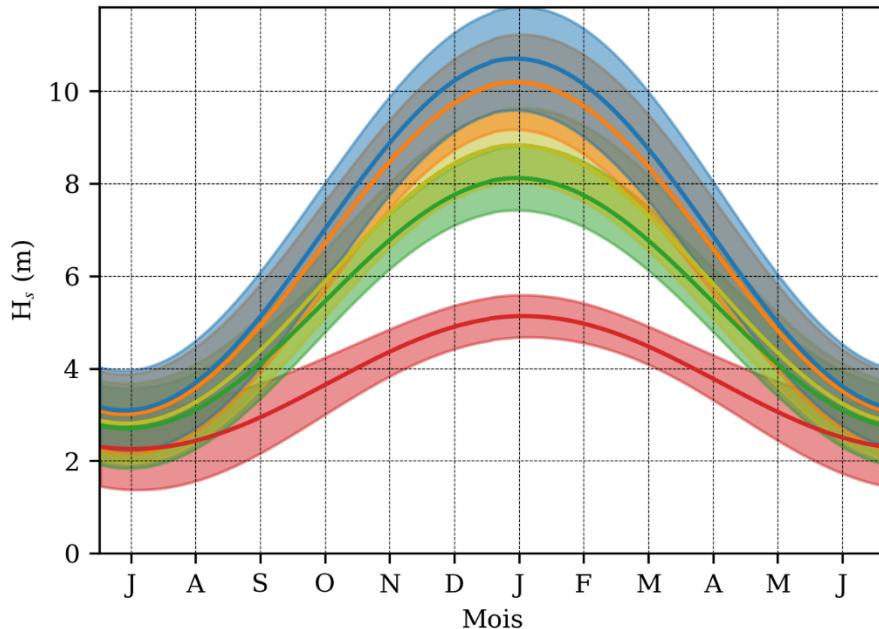


Figure 28 : Variation annuelle de la hauteur significative et ses intervalles de confiance calculés par la méthode delta (Coles, 2001), pour les périodes de retour de 1 an (rouge), 5 ans (vert), 10 ans (jaune), 50 ans (orange), et 100 ans (bleu). L'axe des abscisses montre les initiales de chaque mois, en commençant par juillet.

Période retour (année)	$H_{s,\max}$ (m)	Intervalle de confiance (m)
1	5.13	[4.6 ; 5.6]
5	8.12	[7.4 ; 8.9]
10	8.83	[8.0 ; 9.7]
50	10.19	[9.1 ; 11.3]
100	10.70	[9.5 ; 11.9]

Tableau 8 : Hauteurs significatives associées aux périodes de retour de 1, 5, 10, 50 et 100 ans. $H_{s,\max}$ corresponds à la valeur maximale des courbes annuelles. La dernière colonne montre l'intervalle de confiance 95% des estimations, obtenu par la méthode delta (Coles, 2001).

Par rapport aux résultats obtenus au paragraphe précédent (4.2.1) avec la méthode stationnaire, il ressort que les valeurs obtenues pour les périodes de retour de 1 et 100 ans sont très similaires. En revanche, il y a entre 50 cm et 1.10 m de différence pour les périodes de retour de 5, 10 et 50 ans. En effet, l'ajustement obtenu avec la méthode stationnaire (Figure 26) montre que les valeurs du rejet correspondant à ces périodes de retour sont à la limite ou sortent de l'intervalle de confiance associé aux deux lois utilisées. Le modèle non-stationnaire utilisé ici permet d'introduire davantage de degrés de liberté pour permettre de mieux représenter ces valeurs. Cependant, il faut garder à l'esprit que la durée du rejet numérique utilisé (20 ans) est tout juste acceptable pour estimer des niveaux de retour centennaux et qu'un rejet de plus longue durée permettrait de diminuer les incertitudes et les intervalles de confiance associés aux grandes périodes de retour, et d'augmenter ainsi la confiance dans ces résultats.

5. CONCLUSION

Un rejet d'états de mer à haute résolution (≈ 200 m à la côte), réalisé à l'aide du modèle numérique spectral WAVEWATCH III® sur la période 2003-2022, a été utilisé pour caractériser les conditions d'états de mer sur la zone d'implantation envisagée pour le parc éolien en Méditerranée. La validation des résultats de modélisation numérique, réalisée précédemment en utilisant des données mesurées par des satellites altimétriques et des bouées houlographes (Lopez et al., 2022 et 2023), a été étendue ici pour inclure la validation des résultats récents (2021-2022) par rapport aux mesures du capteur MOTUS déployé pour cette étude, dans la zone EMR. Tel qu'il avait été observé lors de la validation par rapport aux bouées du réseau CANDHIS, les résultats de cette nouvelle validation démontrent à nouveau une bonne capacité du modèle à reproduire les propriétés statistiques des états de mer sur la zone d'étude. Notamment, les résultats pour la hauteur des vagues sont très satisfaisants même pour les valeurs les plus fortes, et sont donc bien adaptés pour estimer les hauteurs significatives extrêmes associées à des périodes de retour.

Une analyse statistique des états de mer sur la zone EMR a ensuite été réalisée à partir du rejet. La zone présente des états de mer moyens peu énergétiques, avec des hauteurs significatives moyennes d'environ 1 m. Les deux directions prédominantes de provenance des vagues sur la zone sont le secteur Ouest à Nord-Ouest, pour lesquels les fetchs sont relativement courts) et le secteur Est à Sud, qui présentent des fetchs plus importants. L'essentiel des événements très énergétiques modélisés sur la zone montre des vagues provenant d'Est-Sud-Est avec des périodes d'environ 9 s et des hauteurs significatives inférieures à 9 m. La hauteur significative maximale modélisée dans la zone EMR durant la période 2003-2022 est de 8.96 m.

Enfin, la théorie des valeurs extrêmes a été appliquée pour estimer les hauteurs significatives extrêmes associées à des périodes de retour. Dans un premier temps, l'approche choisie est stationnaire et basée sur l'ajustement des lois GPD et exponentielle aux pics de tempête sélectionnés selon la méthode POT. Avec cette méthodologie, la hauteur significative estimée pour un événement centennal sur la zone EMR est comprise entre 10.2 et 10.6 m. Ensuite, nous avons utilisé une méthode non stationnaire pour considérer l'influence des variations saisonnières. De manière similaire à la méthode stationnaire, les pics de tempête ont été sélectionnés avec la méthode POT. Ensuite, une méthodologie « *point process* », qui permet d'ajuster de façon simultanée la fréquence et l'intensité des dépassements du seuil, a été utilisée. La paramétrisation de la méthode « *point process* » est faite selon la loi statistique GEV. Ainsi, la saisonnalité a été considérée de manière paramétrique, en introduisant les paramètres de position et d'échelle de la loi statistique GEV sous forme d'une série de Fourier au premier ordre. Avec cette méthodologie, la hauteur significative maximale estimée pour un événement centennal est de 10.7 m.

Remerciements

Cette étude a utilisé de nombreuses données d'observation ; les auteurs tiennent à remercier les différents organismes qui les ont collectées et mises à disposition : le CEREMA pour les données du réseau d'observation de la houle CANDHIS, Météo-France pour les données des bouées Lion et Nice, le CERSAT pour les données altimétriques, et le CEPMMT pour les données atmosphériques.

BIBLIOGRAPHIE

Aanderaa, 2020. TD 316 Operating Manual – Motus Wave Sensor 5729. 117p. Disponible sur : <https://www.aanderaa.com/media/pdfs/td316-manual-motus-5729.pdf>

Alday, M, Accensi, M., Arduin, F, Dodet, G., 2021: A global wave parameter database for geophysical applications. Part 3: Improved forcing and spectral resolution, Ocean Modelling, Volume 166, <https://doi.org/10.1016/j.ocemod.2021.101848>.

Barnston, A. G. and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, 115, 1083–1126.

Charles, E., D. Idier, J. Thiébot, G. Le Cozannet, R. Pedreros, F. Arduin, and S. Planton, 2012 : Present Wave Climate in the Bay of Biscay : Spatiotemporal Variability and Trends from 1958 to 2001. *Journal of Climate*, 25, 2020–2039.

Coles, S., 2001: An Introduction to Statistical Modelling of Extreme Values. Springer-Verlag London.

Datawell BV, 2020 : Datawell Waverider MKIII Manual. Disponible sur: https://datawell.nl/wp-content/uploads/2022/09/datawell_manual_dwr-mk3_dwr-g_wr-sg.pdf

Davison A, Smith R., 1990 : Models for Exceedances over High Thresholds. *Journal of the Royal Statistical Society B*, 52, 237–254.

Gilleland, E., & Katz, R. W. (2016). extRemes 2.0: An Extreme Value Analysis Package in R. *Journal of Statistical Software*, 72(8), 1–39. <https://doi.org/10.18637/jss.v072.i08>

Hurrell et al., 2003 : The North Atlantic Oscillation: Climate Significance and Environmental Impact, Geophysical Monograph 134, American Geophysical Union.

Izaguirre, C., F. J. Mendez, M. Menendez, A. Luceño, and I. J. Losada, 2010 : Extreme wave climate variability in southern europe using satellite data. *Journal of Geophysical Research : Oceans*, 115.

Jones, P. D., T. Jonsson, and D. Wheeler, 1997 : Extension to the north Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *International Journal of Climatology*, 17, 1433–1450.

Katz, R. , 2012 : Statistical methods for non-stationary extremes. In : AghaKouchak A, Easterling D., Hsu K., Schubert S., Sorooshian S. (eds) *Extremes in a changing climate*, Springer.

Leckler, F., Osinski, R., Michaud, H., Lathuilière, C., Pasquet, A., Le Bec, N., Tychensky, A., 2018 : Modélisation et étude des états de mer sur la zone EMR d'Oléron. Convention SHOM/DGEC n°2201103478, SHOM n°52/2017. Rapport interne Shom.

Lopez, G., Leballeur, L., Michaud, H., 2022 : Dossier de l'état de connaissance initiale de la zone d'implantation du parc éolien en mer Méditerranée – Zone B – Lot 2 - Volet Houle. Convention Shom/MTES [DGEC], SHOM n°79/2019.

Lopez, G., Leballeur, L., Michaud, H., 2023 : Connaissance statistique des régimes de houle pour la zone Méditerranée – Lot 5. Convention Shom/MTES [DGEC], SHOM n°79/2019.

Mazas, F., Hamm, L., 2011: A multi-distribution approach to POT methods for determining extreme wave heights. *Coastal Engineering*, Volume 58, Issue 5, 385-394, <https://doi.org/10.1016/j.coastaleng.2010.12.003>.

Méndez, F.J., Menéndez, M., Luceño, A. and Losada, I.J., 2006: Estimation of the long-term variability of extreme significant wave height using a time-dependent peak over threshold (pot) model. *Journal of Geophysical Research: Oceans*, 111(C7).

Menéndez, M., F. J. Méndez, C. Izaguirre, A. L. no, and I. J. Losada, 2009 : The influence of seasonality on estimating return values of significant wave height. *Coastal Engineering*, 56, 211 – 219.

Mentaschi, L., M. Vousdoukas, E. Voukouvalas, L. Sartini, L. Feyen, G. Besio, and L. Alfieri, 2016 : The transformed-stationary approach : a generic and simplified methodology for non-stationary extreme value analysis. *Hydrology and Earth System Sciences*, 20, 3527–3547.

Morales-Márquez, V., Orfila, A., Simarro, G., and Marcos, M. 2020 : Extreme waves and climatic patterns of variability in the eastern North Atlantic and Mediterranean basins, *Ocean Sci.*, 16, 1385–1398, <https://doi.org/10.5194/os-16-1385-2020>.

Niroomandi, A., G. Ma, X. Ye, S. Lou, and P. Xue, 2018 : Extreme value analysis of wave climate in Chesapeake Bay. *Ocean Engineering*, 159, 22 – 36.

Pinto, J.G., Zacharias, S., Fink, A.H. et al., 2009 : Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Clim Dyn* **32**, 711–737. <https://doi.org/10.1007/s00382-008-0396-4>

Rahimpour, V., Y. Zeng, C. M. Mannaerts, and Z. B. Su, 2016 : Attributing seasonal variation of daily extreme precipitation events across The Netherlands. *Weather and Climate Extremes*, 14, 56 – 66.

Rueda, A., Camus, P., Mendez, F., Tomas, A., Luceño, A. 2016: An extreme value model for maximum wave heights based on weather types. *Journal of Geophysical Research: Oceans*, 121, 1262–1273. <https://doi.org/10.1002/2015JC010952>

Smith R., 1989 : Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone. *Statistical Science*, 4, 367–393. doi:10.1214/ss/1177012400

Smith R., Shively T., 1995 : Point Process Approach to Modeling Trends in Tropospheric Ozone. *Atmospheric Environment*, 29, 3489–3499. doi:10.1016/1352-2310(95)00030-3

Annexe 1. Corrélogrammes pour les secteurs directionnels

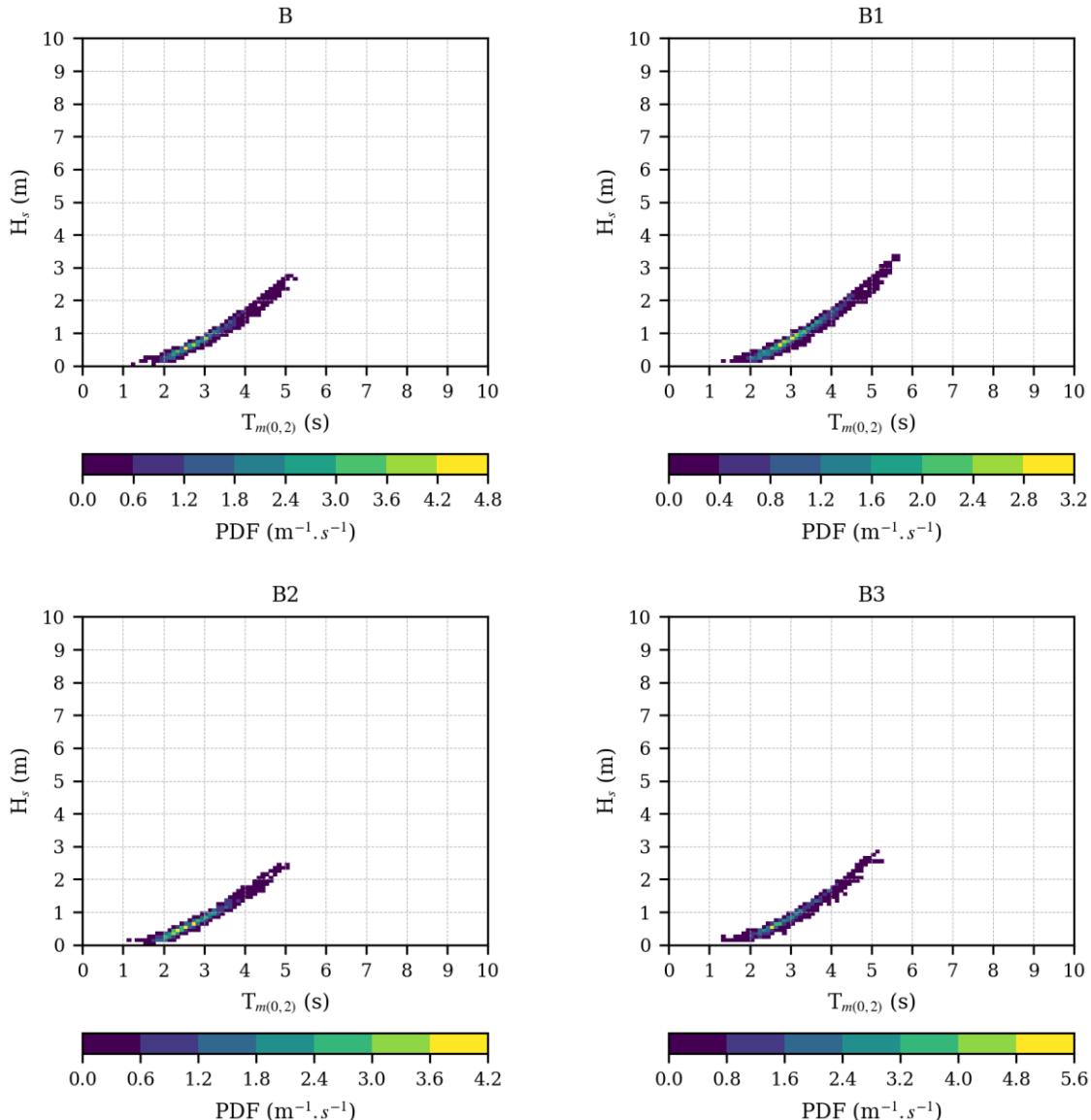


Figure 29 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Nord ($337.5^\circ < \theta_m \leq 22.5^\circ$).

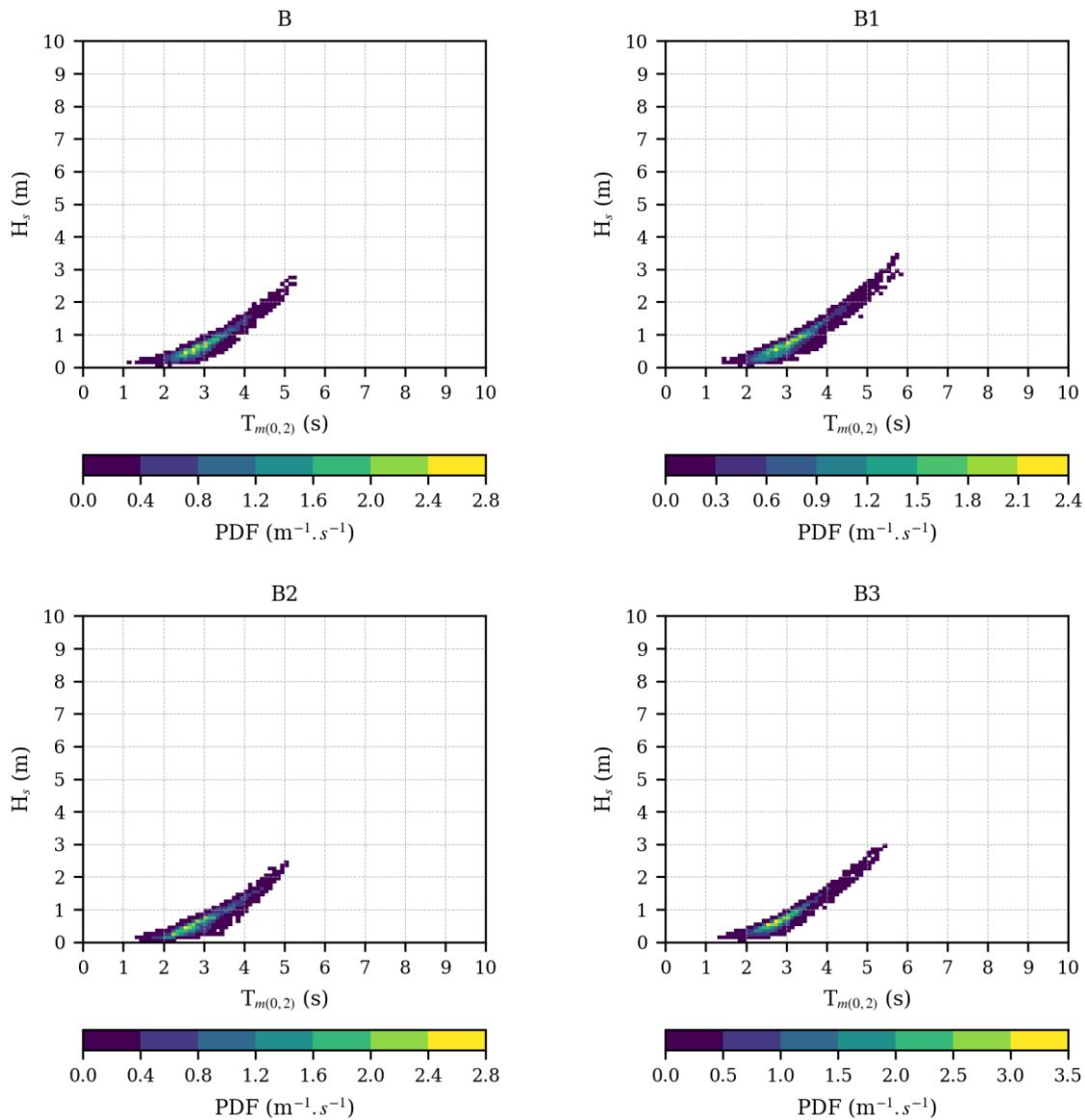


Figure 30 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Nord-Est ($22.5^\circ < \theta_m \leq 67.5^\circ$).

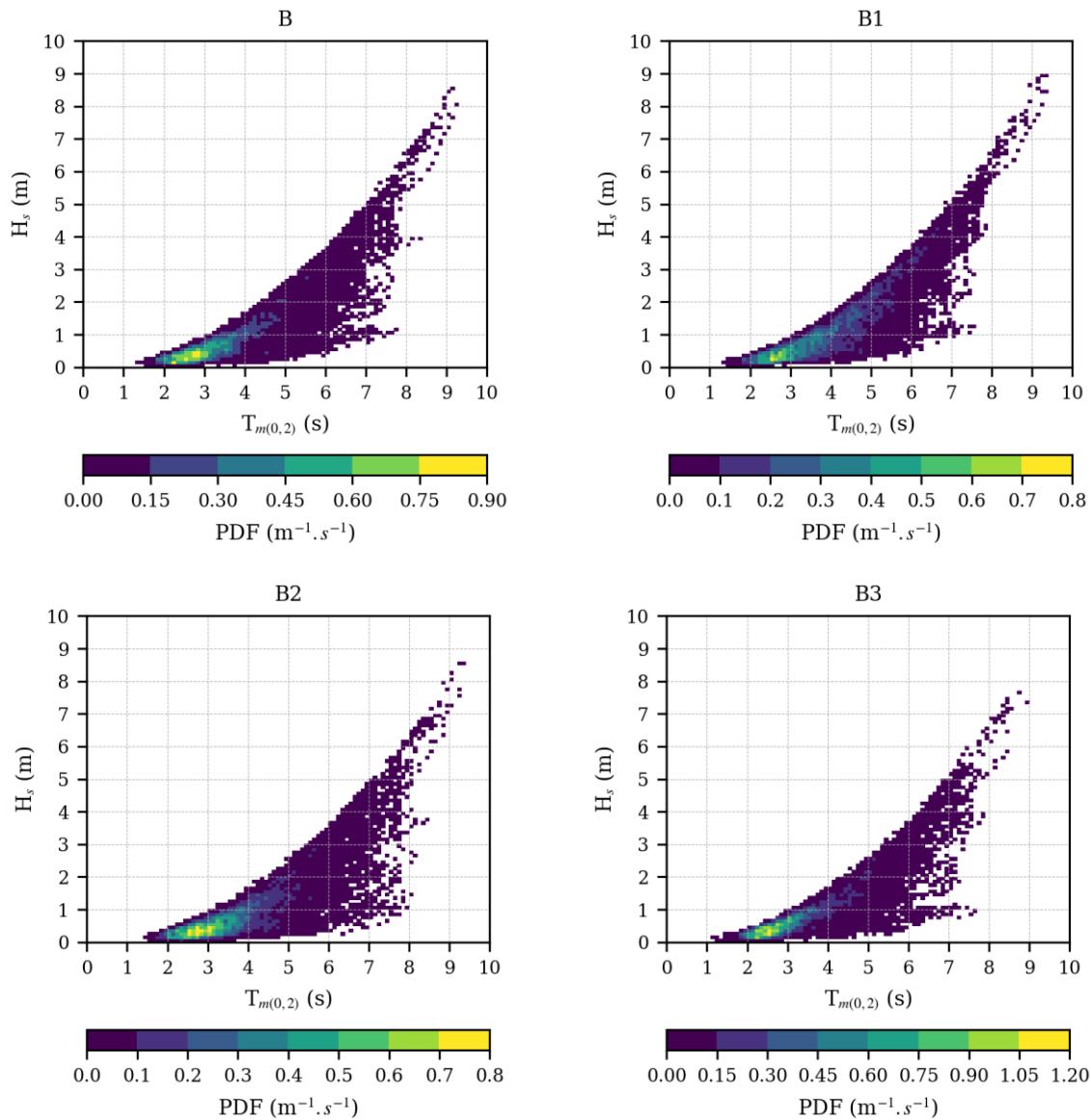


Figure 31 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant de l'Est ($67.5^\circ < \theta_m \leq 112.5^\circ$).

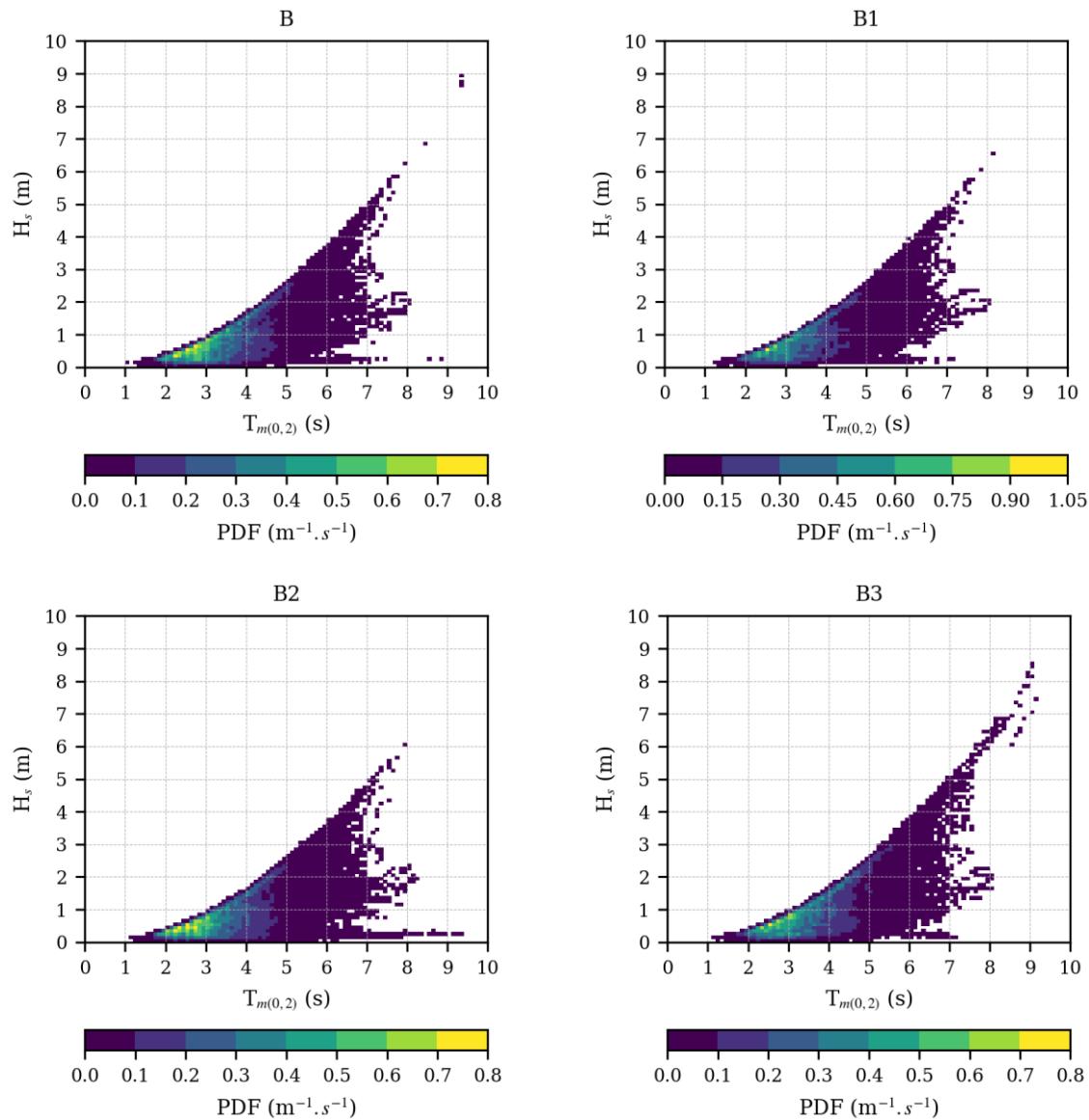


Figure 32 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Sud-Est ($112.5^\circ < \theta_m \leq 157.5^\circ$).

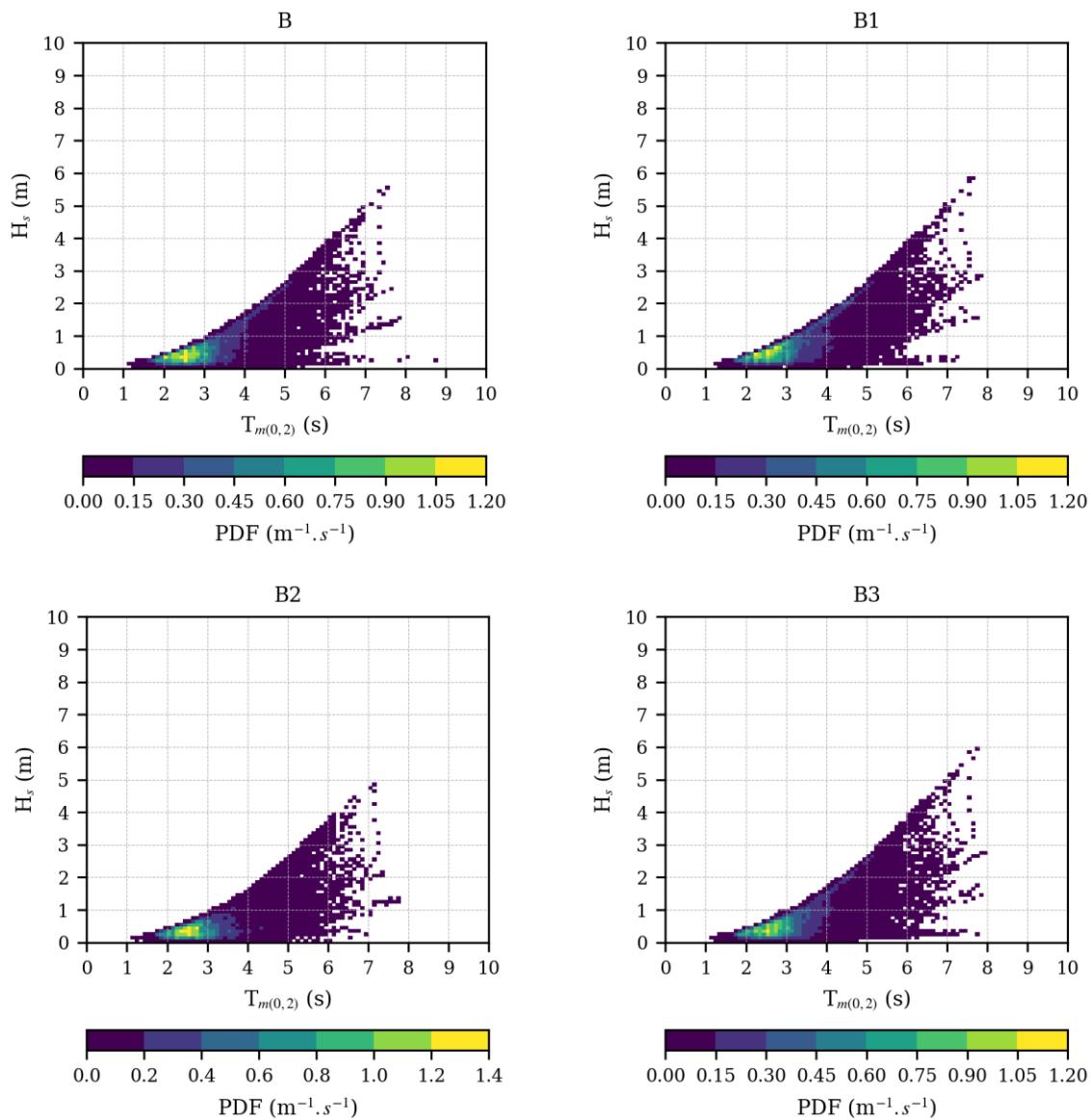


Figure 33 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Sud ($157.5^\circ < \theta_m \leq 202.5^\circ$).

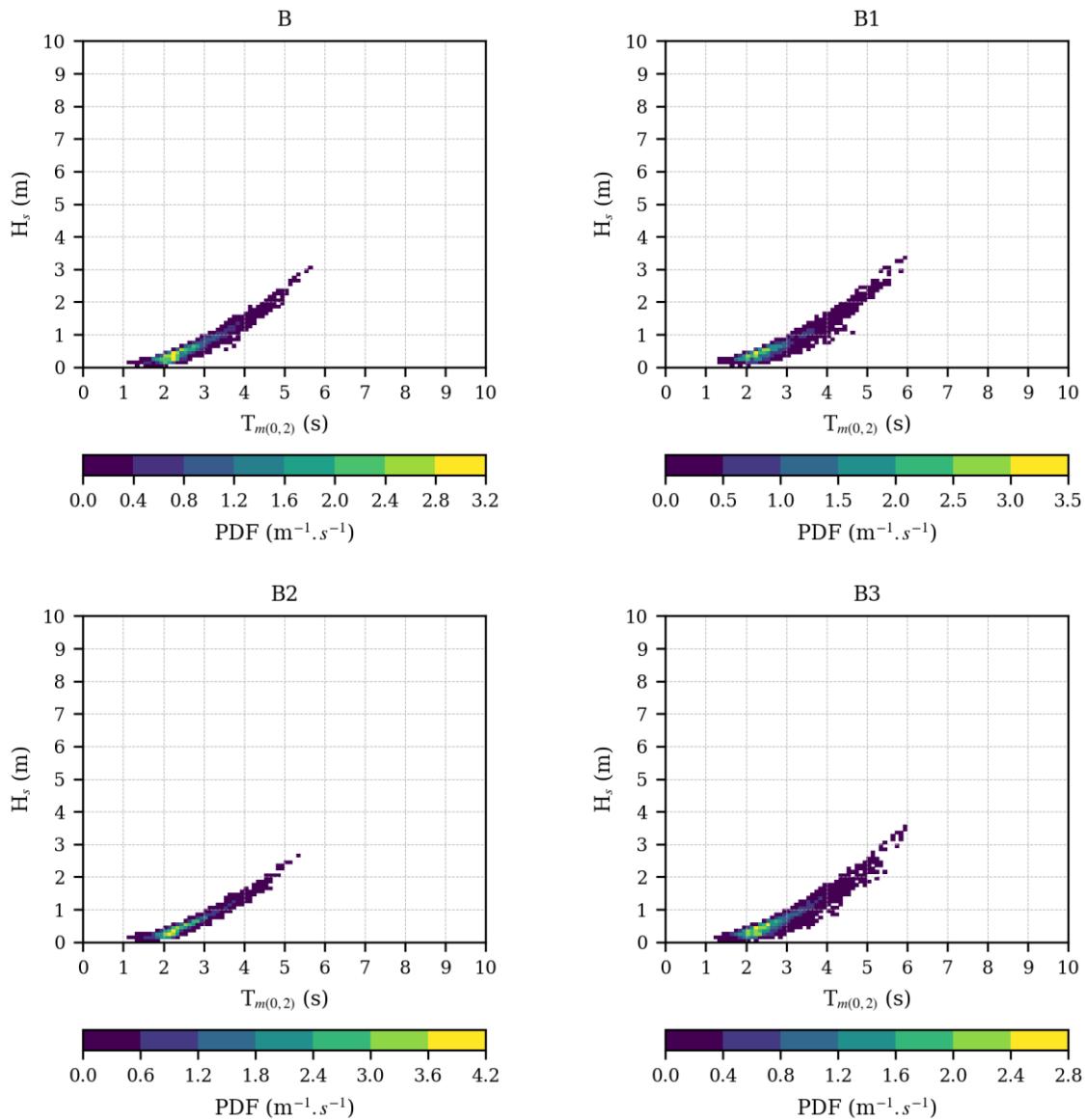


Figure 34 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Sud-Ouest ($202.5^\circ < \theta_m \leq 247.5^\circ$).

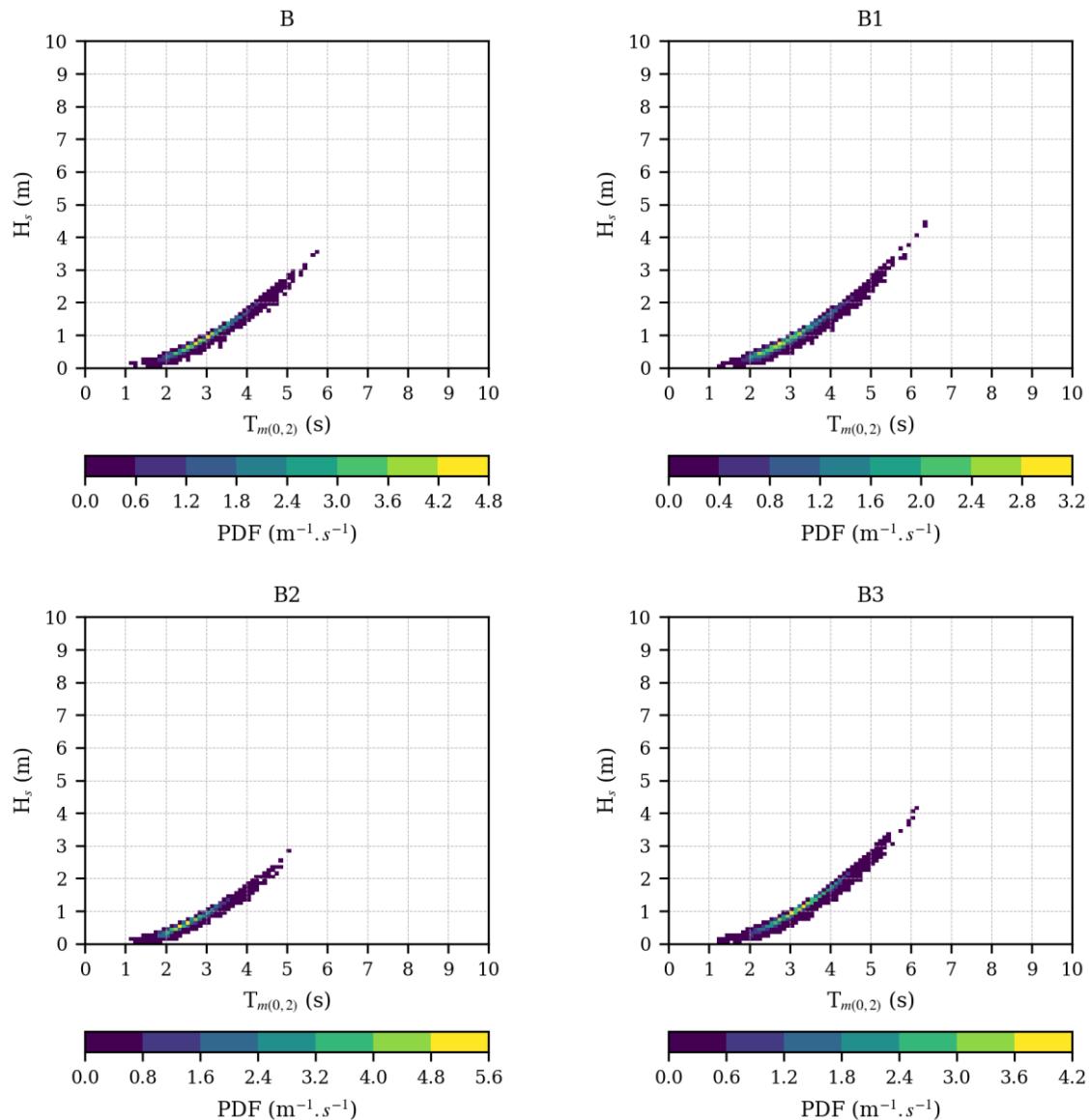


Figure 35 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant de l'Ouest ($247.5^\circ < \theta_m \leq 292.5^\circ$).

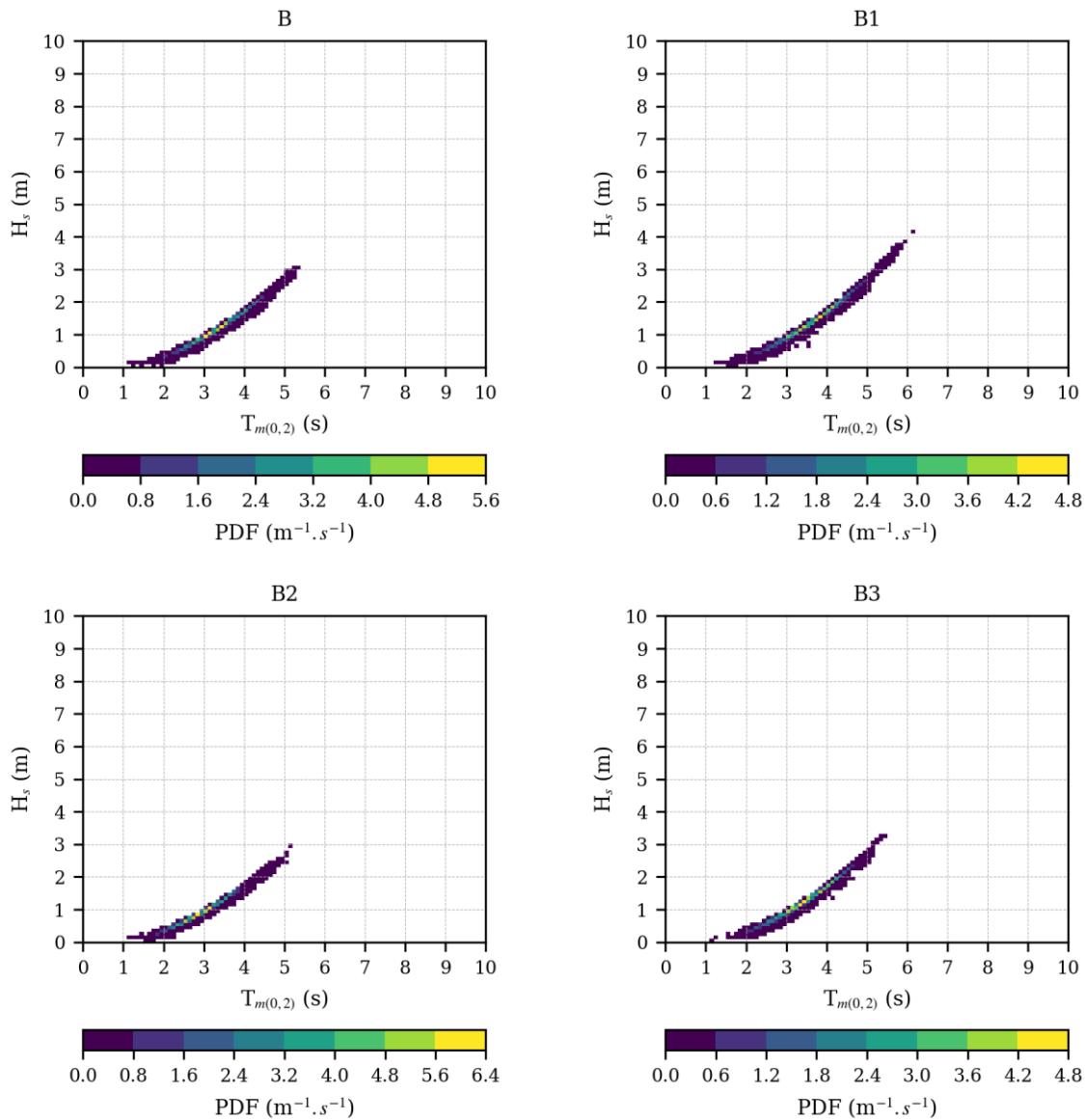


Figure 36 : Corrélogrammes $H_s - T_{m(0,2)}$ pour les points B, B1, B2 et B3 de la zone d'étude et pour les vagues provenant du Nord-Ouest ($292.5^\circ < \theta_m \leq 337.5^\circ$).

Annexe 2. Sélection du seuil de tempête

Des tests ont été réalisés pour déterminer la valeur du seuil de tempête pour la méthode POT. Notamment, nous avons vérifié que la moyenne des excès de H_s au-delà du seuil varie linéairement avec la valeur de ce dernier. Ce test est lié aux propriétés de la distribution GPD. Si la fonction de distribution GPD est valide pour les excès de H_s au seuil u_0 , la distribution est aussi valide pour les seuils supérieurs ($u > u_0$), après modification du paramètre d'échelle. Puis, la valeur moyenne de la distribution GPD est :

$$E(X - u | X > u) = \frac{\sigma(u)}{1 - \xi} = \frac{\sigma(u_0 + \xi u)}{1 - \xi} \quad (\text{Eq. 8})$$

Puisque le résultat de l'équation ci-dessus est une fonction linéaire du seuil, les excès de H_s doivent varier linéairement avec le seuil choisi.

Un deuxième test a été réalisé pour vérifier la propriété asymptotique de la loi GPD. Pour cela, nous avons examiné la stabilité des paramètres d'échelle modifié et de forme, qui devraient rester relativement constants lorsque la valeur du seuil augmente.

L'ensemble des résultats montrés sur la Figure 37 et Figure 38 indiquent qu'une valeur entre 3 et 4.5 mètres peut être un bon choix pour le seuil de tempête. Après avoir réalisé ces tests, nous avons choisi la valeur de 4.02 mètres pour le seuil. Avec cette valeur, le nombre moyen de tempêtes par an est égal à 2.9 (58 pics sur 20 ans). Ce seuil choisi a été utilisé pour les analyses présentées dans le paragraphe 4.2.

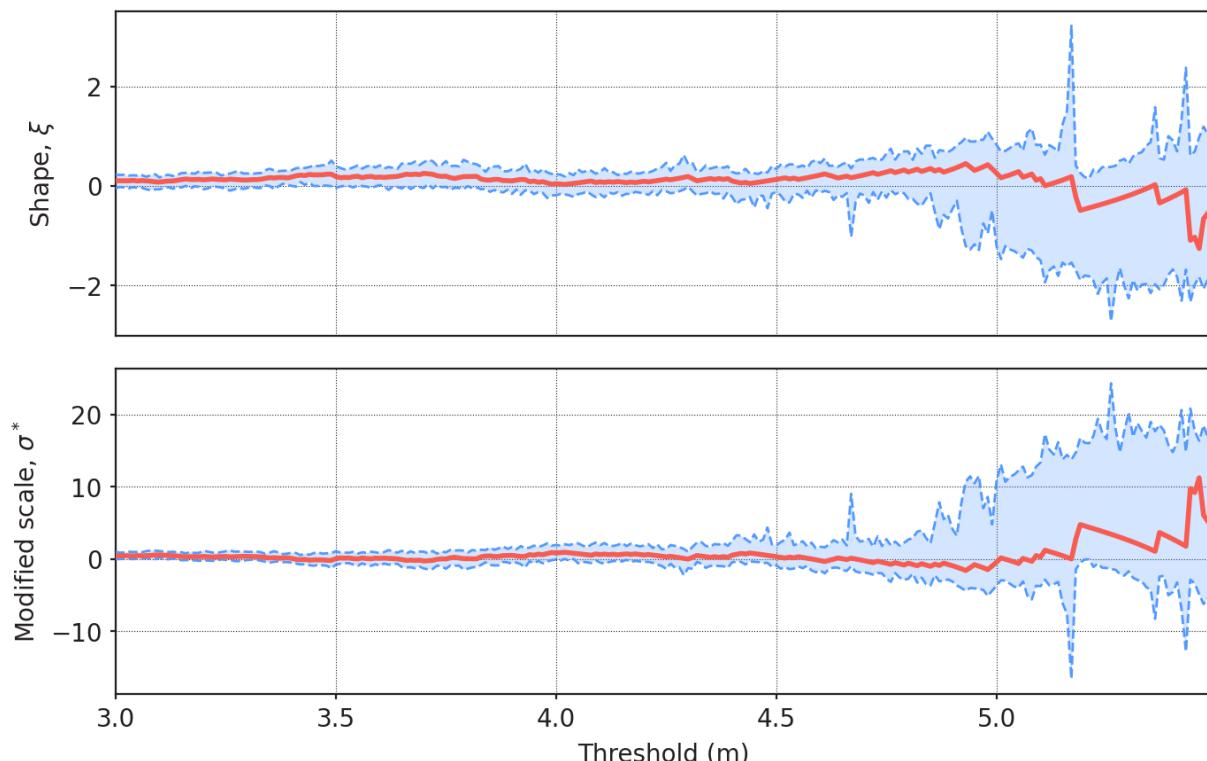


Figure 37 : Paramètre de forme (shape, en haut) et d'échelle (scale, en bas), en fonction du seuil de tempête (threshold). Sur les deux figures, l'axe horizontal montre les valeurs du seuil en mètres. Les courbes pointillées bleues forment l'enveloppe de l'intervalle de confiance à 95%.

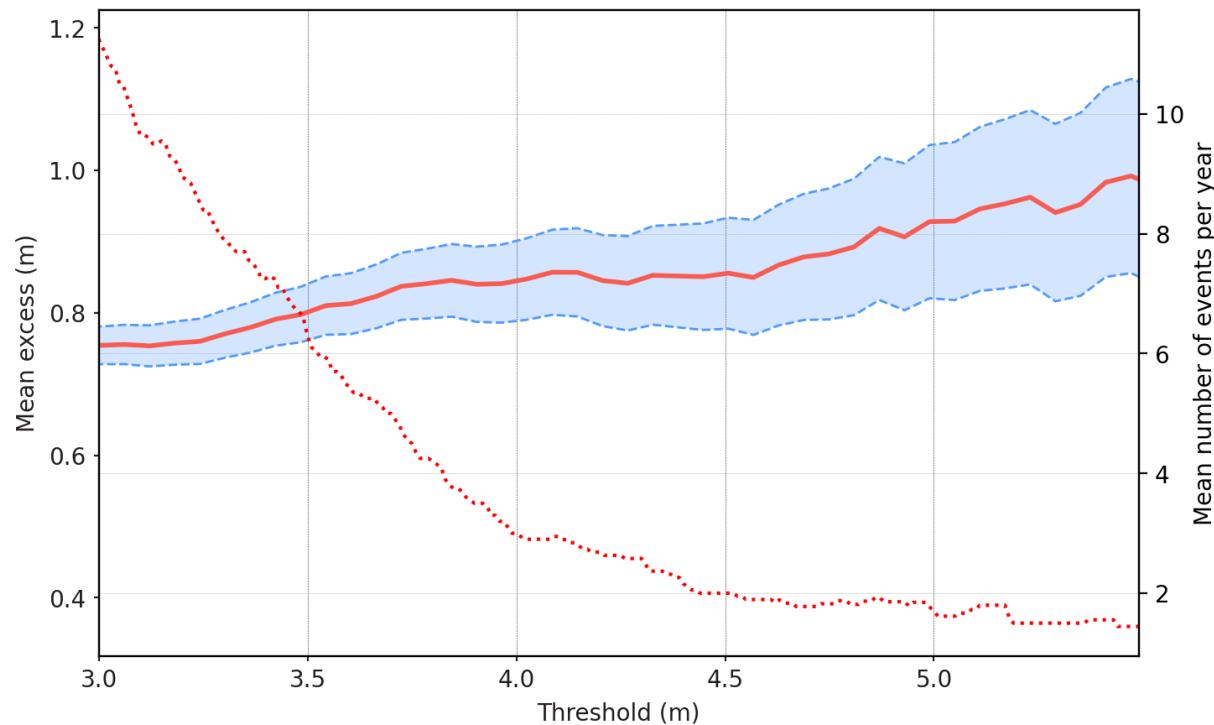


Figure 38 : « Mean excess plot » pour H_s . La ligne rouge correspond au résultat de l'équation (Eq. 8) et les lignes pointillées bleues forment l'intervalle de confiance à 95%. La ligne pointillée rouge indique le nombre moyen de pics de tempête par an selon le seuil choisi. L'axe horizontal montre les valeurs du seuil en mètres.