

Consultez les discussions, les statistiques et les profils des auteurs de cette publication sur <https://www.researchgate.net/publication/356359712>

# Analyse de clustering du processus de transport des Sargasses : application à la prédiction des échouages aux Petites Antilles

Préimpression · Novembre 2021

DOI: 10.5194/os-2021-109

CITATIONS

0

LIT

26

5 auteurs, y compris:



Didier Bernard

Université des Antilles

56 PUBLICATIONS 444 CITATIONS

[VOIR LE PROFIL](#)



Emmanuel Biabiany

Université des Antilles

17 PUBLICATIONS 819 CITATIONS

[VOIR LE PROFIL](#)



Raphael Cece

Université des Antilles

37 PUBLICATIONS 118 CITATIONS

[VOIR LE PROFIL](#)



Romual Chéry

Université des Antilles

3 PUBLICATIONS 3 CITATIONS

[VOIR LE PROFIL](#)

Certains des auteurs de cette publication travaillent également sur ces projets connexes :



Changement climatique et conséquences sur les Antilles françaises [Voir le projet](#)



Groupe Régional d'Experts sur le Climat de la Guadeloupe (GREC Guadeloupe) [Voir le projet](#)



# Analyse de clustering du processus de transport des Sargasses : application à la prédiction des échouages aux Petites Antilles

Didier Bernard<sup>1</sup>, Emmanuel Biabiany<sup>2</sup>, Raphaël Cecé<sup>1</sup>, Romual Chéry<sup>1</sup>, et Naoufal Sekkat<sup>1</sup>

<sup>1</sup>LARGE, Université des Antilles françaises, 97157 Pointe-à-Pitre, Guadeloupe, France.

5 <sup>2</sup>LAMIA, Université des Antilles françaises, 97157 Pointe-à-Pitre, Guadeloupe, France.

Correspondance à : Didier Bernard ( didier.bernard@univ-antilles.fr )

**Abstract.** Les échouages massifs d'algues sargasses observés au cours de la dernière décennie constituent le nouveau risque naturel qui affecte actuellement les États insulaires de la région des Caraïbes (santé humaine, dommages environnementaux et pertes économiques). Cette étude vise à améliorer la prédiction de la dynamique des courants de surface conduisant aux échouages dans les Petites Antilles, en utilisant des méthodes d'analyse de clustering. Les courants de surface d'entrée, y compris l'effet de dérive, ont été dérivés du modèle Mercator et du modèle hybride de coordonnées océaniques (HYCOM). Les observations passées quotidiennes d'échouages de Sargasses sur les côtes guadeloupéennes ont également été intégrées. Quatre régimes de courant représentatifs ont été identifiés pour les données Mercator et HYCOM. L'analyse des séquences de courants rétrogrades conduisant à des échouages a montré que la récurrence de dix régimes de courant est liée aux pics d'échouage observés respectivement en mars et en août. Un classificateur d'arbre de décision a été construit et sa précision atteint 73,3 % avec des données HYCOM à l'échelle 0,04° et 50,8 % avec des données Mercator à l'échelle 0,08°. Cette différence de précision significative met en évidence le besoin de données de courant à très petite échelle (c'est-à-dire à une échelle inférieure à 5 km) pour évaluer le risque côtier de sargasses dans les Petites Antilles. Le présent système prédictif d'analyse de regroupement permettrait d'améliorer cette gestion des risques dans les îles de cette région. Cette différence de précision significative met en évidence le besoin de données de courant à très petite échelle (c'est-à-dire à une échelle inférieure à 5 km) pour évaluer le risque côtier de sargasses dans les Petites Antilles. Le présent système prédictif d'analyse de regroupement permettrait d'améliorer cette gestion des risques dans les îles de cette région. 15 région. Cette différence de précision significative met en évidence le besoin de données de courant à très petite échelle (c'est-à-dire à une échelle inférieure à 5 km) pour évaluer le risque côtier de sargasses dans les Petites Antilles. Le présent système prédictif d'analyse de regroupement permettrait d'améliorer cette gestion des risques dans les îles de cette région.

## 1. Introduction

Au cours des périodes 2011-2012, puis 2014-2019, des échouages massifs de Sargasses ont impacté la plupart des côtes des Petites Antilles (LA), principalement celles exposées à l'est et au sud-est. LA a reçu de grandes quantités d'algues sur la côte atlantique au vent, tandis que les zones côtières sous le vent des Caraïbes sont restées légèrement affectées (Franks et al., 2012, Gower et al., 2013, Johnson et al., 2014, Hu et al., 2016, Wang et Hu, 2016, Maréchal et al., 2017). Ces échouages en termes de fréquence et d'intensité peuvent désormais être considérés comme un nouveau risque naturel pour les îles Caraïbes et les côtes américaines. Des échouages ont également été observés en Afrique (Széchy et al., 2012). En effet, s'il a été démontré que les algues sargasses fournissent des services écosystémiques, un habitat et un abri pour divers organismes dans un écosystème océanique structurellement stérile (Witherington et al., 2012 ; Bertola et al., 2020), les échouages de la dernière décennie ont induit des risques sanitaires pour la population et ont eu des impacts socio-économiques considérables (Franks et al., 2012). Par exemple, si l'on regarde les Antilles françaises, l'archipel de la Guadeloupe et la Martinique, les résultats sont les suivants :

- (1) En dehors de 2013, l'afflux récent de radeaux de Sargasses sur les côtes de la Guadeloupe et de la Martinique, bien qu'irrégulier, n'a pas cessé depuis 2011, atteignant un paroxysme en 2015 (Florenne et al., 2016 ; Berline et al., 2020) . Les services de l'État ont estimé que les volumes échoués sur les rives étaient de l'ordre de 1,5 million de m<sup>3</sup>, d'octobre 2014 à octobre 2015 en Guadeloupe (Florenne et al., 2016). Seul un tiers d'entre eux a pu être collecté par les autorités et la priorité étant



accordées à des zones à enjeux telles que, des zones habitées, des rives avec des activités économiques ou touristiques et des écosystèmes ou d'autres niches environnementales. La particularité et la difficulté résidait dans le fait que 60% de ce littoral et/ou du volume échoué restaient inaccessibles aux techniques aujourd'hui éprouvées et à des coûts actuellement supportables.

35 (2) un impact sur la santé humaine et les écosystèmes car dans les baies peu profondes et exigües, les algues accumulées se dégradent par fermentation et émettent des composés chimiques tels que le sulfure d'hydrogène (H<sub>2</sub>S) et l'ammoniac (NH<sub>3</sub>) (Anses, 2017, Van Tussenbroek et al., 2017 ; Resière et al., 2018).

(3) L'enquête menée par les organismes chargés du développement socio-économique a estimé que la baisse du tourisme a entraîné une perte économique de 5,5 millions de dollars pour le premier semestre 2015 (<https://eos.org/features/sargassum-watchwarns- d'algues entrantes>).

40

Les volumes à collecter étant considérables par rapport à la taille de ces îles (< 1200 km<sup>2</sup>chacun) et la vulnérabilité de ces territoires. Ce phénomène nouveau a soulevé plusieurs questions scientifiques telles que leurs transports, leurs origines, les sources de nutriments favorisant leur croissance mais surtout les facteurs physiques qui ont conduit à l'apparition et au développement des radeaux de Sargasses dans l'Atlantique tropical et équatorial.

45 À l'aide d'observations à grande échelle avec la télédétection par satellite de la couleur de l'océan, d'observations hydrographiques historiques, de séries chronologiques du volume de sargasses collecté sur les navires, d'une réanalyse pluriannuelle du vent et du courant, des modèles numériques ont estimé à la fois le rôle de l'approvisionnement en nutriments souterrains et du transport des courants de surface. Plusieurs auteurs ont contribué à la compréhension des mécanismes et processus physicochimiques régissant le phénomène (Gower et al., 2006 ; Gower et King, 2011 ; Gower et al., 2013 ; Maréchal et al., 2017 ; Johns et al., 2020 ). Des systèmes opérationnels ont été développés tels que le Sargassum Watch System

50 SaWS (Hu, 2009 ; Hu et al., 2015) et le Sargassum Early Advisory System (SEAS) (Webster et Linton, 2013). Ils fournissent une évaluation temporelle et spatiale des augmentations et des diminutions saisonnières annuelles de la quantité d'algues sargasses sur de vastes zones de l'Atlantique tropical et des Caraïbes (Wang et Hu, 2016 ; Wang et Hu, 2017 ; Wang et al. 2019). Des séries temporelles issues de la télédétection ont été couplées à des modèles de distribution spatiale pour déterminer les mécanismes qui agrègent les algues sargasses le long d'une bande zonale dans l'Atlantique tropical en tenant compte des sources possibles de nutriments favorisant les efflorescences annuelles observées (Wang et al., 2018 ; Wang et al., 55 2019 ; Johns et al., 2020)).

Les courants et les vents de l'Atlantique tropical agrègent et transportent ces algues de façon saisonnière vers les Caraïbes (Franks et al., 2016 ; Brooks et al., 2018 ; Cuevas et al., 2018). Les études de modélisation ont principalement porté sur les propriétés de transport des radeaux de Sargasses par les courants du large (Wang et Hu, 2017 ; Brooks et al., 2018 ; Maréchal et al., 2017 ; Putman et al., 2018, 2020 ; Wang et al., 2019 ; Berline et al., 2020). Johns et al. (2020) ont étendu cette analyse pour mettre en évidence le transport anormal dû à l'anomalie NAO de 2009-2010 et 60 l'agrégation saisonnière par la zone de convergence intertropicale (ITCZ).

Une combinaison d'images satellite MODIS AFAI avec des données de prévision de courant de surface HYCOM a été utilisée par Maréchal et al. (2017) pour prédire à court terme les échouages de Sargasses pour la Guadeloupe et les Antilles françaises. Maréchal et al. (2017) ont montré que ce système de prévision à court terme (c'est-à-dire, la détection commençant à moins de 50-100 km des côtes) a fonctionné efficacement au cours de l'année 2015 avec un pourcentage de performance de 62 % et une incertitude de la date de prévision d'échouage inférieure à un jour.

65 Dans les travaux ci-dessus, la mise en œuvre de méthodes basées sur plusieurs jeux de données indépendants a conduit à la production de



connaissances et même au développement de systèmes de prévision à grande échelle. Aucun d'entre eux n'a utilisé la modélisation prédictive, y compris les classificateurs, pour déterminer la probabilité qu'un ensemble de données appartienne à un autre ensemble afin de découvrir des modèles répétables, permettant de produire une décision pour les responsables de la prévention des risques.

70 Dans cet article, nous proposons d'utiliser des méthodes de clustering et de classification par arbre de décision, combinant la réanalyse des courants de surface océaniques et des vents avec les échouages observés dans le passé pour obtenir un premier modèle prédictif de l'échouage des sargasses sur les côtes caribéennes. Ce modèle sera utilisé avec des données de prévision en entrée pour produire un système opérationnel d'aide à la décision.

Les données océaniques étant des champs spatio-temporels, des méthodes d'apprentissage automatique telles que K-Means (KMS) peuvent être utilisées pour obtenir un nombre fini de partitions k-cluster possibles des courants de surface. Ces méthodes ont été largement utilisées en prévision météorologique (Michelangeli et al., 1995 ; Cassou et al., 2004 ; Boé et Terray, 2008) mais sont beaucoup moins courantes en océanographie physique (Harms et Winant, 1998 ; 75 Hisaki, 2013). Pour optimiser le partitionnement final, une métrique supplémentaire basée sur la divergence de Kullback Leiber (Kulback et Leibler, 1951, Biabiany et al., 2020) sera incluse.

Nous nous sommes concentrés sur la zone offshore couvrant de part et d'autre les Petites Antilles, entre 55-66°W et 8-17°N (Fig. 1a). L'analyse visuelle des cartes mensuelles du SaWS indique que cette région reste la principale voie d'accès des radeaux de sargasses de l'océan Atlantique à la mer des Caraïbes. Le Courant Nord Equatorial (NEC), le Courant de Guyane (GC), les tourbillons et le front de 80 rétroflexion du Courant Nord Brésil (NBC) sont les principaux contributeurs de ce transport. La figure 1b décrit la zone focalisée divisée en un premier sous-ensemble "LA1" pour la mer des Caraïbes, un deuxième, "LA2" entre 18°N et 14,5°N (Guadeloupe, Dominique, Martinique, Sainte-Lucie) et un troisième "LA3" au sud de 14,5°N (Saint-Vincent, Barbade, Trinité-et-Tobago). Cette région océanique correspond aux cases CA et TA1 de Johns (2020).

Les questions sont les suivantes. Les schémas dynamiques des courants de surface dans les Petites Antilles peuvent-ils être résumés en un ensemble discret de 85 cas ? Quelle est leur récurrence temporelle ? Quelles combinaisons de courants favorisent l'arrivée et l'échouage des radeaux de Sargasses sur les côtes des Petites Antilles ? Quel est l'apport de ce type de modélisation prédictive à la prévention de ce nouveau risque naturel ? La base de données, les méthodes de regroupement et l'arbre de décision utilisés dans cette étude sont décrits dans la Sect. 2. Les régimes de courant obtenus, leur relation avec le risque Sargassum et les performances du système d'aide à la décision sont présentés dans la Sect. 3. Ces résultats sont discutés dans la Sect. 4.

## 90 2 Jeux de données et méthodes

### 2.1 Jeu de données de courant de surface HYCOM

Données de courant de surface à échelle fine du modèle d'analyse 1/25 degré HYCOM + NCODA du golfe du Mexique (version GOMu0.04/ expt\_90.1m000, Hogan et al, 2014 ; Helber et al., 2013 ; Cummings et Smedstad, 2013 ; Cummings , 2005) entre le 1<sup>er</sup> janvier 2019 (c'est-à-dire la date de début des données disponibles) et le 31 décembre 2020 ont été analysés. Des champs journaliers 12Z donnant les composantes u et v du courant 95 à 50 cm de profondeur ont été utilisés. Ces données de courant à résolution fine n'ont pas été utilisées dans les études précédentes traitant du danger des sargasses (Putman et al., 2018 ; Johns et al., 2020).

### 2.2 Jeu de données Mercator sur les courants de surface

Les composantes quotidiennes de courant à 50 cm de profondeur du système d'analyse 3D PSY4V3R1 Mercator 1/12 degré, y compris la version 3.1 du modèle océanique NEMO (Lellouche et al., 2018 ; Gasparin et al., 2019) ont également été analysés sur la même période que HYCOM



100 données un. Ces données ont une résolution qui est la moitié de celle des données HYCOM (échelle 1/25 degré). Par conséquent, le modèle Mercator  
donne une représentation moins précise des formes des îles. La comparaison entre les résultats HYCOM et Mercator aiderait à mieux comprendre  
les effets de la résolution spatiale sur les modèles de courants de surface dans la région ciblée.

### 2.3 Jeu de données ERA-5 : vents de surface

Le vent de surface influence le transport des radeaux d'algues flottants, avec un facteur optimal de  $= 0,01$ , qui correspond au coefficient de  
105 traînée ou fardage, d'après Johns et al. (2020). Un premier clustering (KMS-L2) sur analyse Mercator sans dérive avait été proposé par  
Bernard et al. (2019). Berline et al. (2017), Putman et al. (2018) et Johns et al. (2020) ont montré que la dérive améliore les simulations  
lagrangiennes du transport des radeaux de sargasses dans la région des Caraïbes. La dérive a été incluse dans le regroupement actuel des  
courants de surface. Les données de vent de surface (à 1000 hPa) du modèle ERA-5 pour la période 2019 à 2020 ont été intégrées aux  
courants Mercator selon cette formule :

$$110 \quad (u, v) = (u_s, v_s) + (u_w, v_w) \quad (1)$$

où  $(u, v)$  représente les courants océaniques de surface avec dérive,  $(u_s, v_s)$  la vitesse des courants océaniques de surface,  $(u_w, v_w)$  la dérive et la vitesse  
des vents de surface. Cette approche est conforme à Putman et al. (2016) et Johns et al. (2020) études. La région analysée dans le  
présent travail correspond à la région CA - TA1 définie dans Johns et al. (2020).

### 2.4 Données d'observation des échouages (Guadeloupe)

115 Une base de données de référencement incluant les échouages observés sur les côtes guadeloupéennes a été utilisée dans la présente étude. La période de  
temps sélectionnée est la même que celle des données de courant de surface : de 1<sup>er</sup> janvier 2019 au 31 décembre 2020. Cette période comprend 730 jours  
d'observation avec 110 jours d'échouages observés. Ces données d'observation basées sur la télédétection et les données in situ sont archivées en ligne par la  
Direction Régionale de l'Environnement, de l'Aménagement et du Logement de la Guadeloupe (<http://www.guadeloupe.developpementdurable.gouv.fr/sargasses-r999.html>).

### 120 2.5 Analyse de clustering avec déviation d'expert

Pour traiter ces données, nous utiliserons une nouvelle approche appelée Expert Deviation (ED) qui intègre l'analyse d'images dans des méthodes  
d'apprentissage non supervisé (Clustering) pour apporter des connaissances de terrain dans le processus d'analyse (Biabiany et al., 2020). Cette méthode a  
permis une amélioration significative de l'analyse de clustering traitant des données climatiques caractérisées par une forte variabilité spatio-temporelle,  
comme les précipitations (Biabiany et al., 2020).

#### 125 2.5.1 Processus de regroupement des méthodes

Des méthodes d'apprentissage non supervisées telles que Hierarchical Agglomerative Clustering (HAC) et les algorithmes K-means sont utilisées dans la  
présente étude. Outre les mesures et les classes de distance entre objets telles que la distance euclidienne pour les K-means et la méthode de Ward pour le  
HAC, une nouvelle métrique a également été ajoutée (Biabiany et al. 2020). Cette métrique intègre un ensemble de connaissances sur la dynamique des  
données à partitionner ainsi que leurs propriétés spatio-temporelles. Le résultat est une analyse automatisée avec sa propre expertise sur les données  
130 d'entrée.

#### 2.5.2 Utilisation de la déviation experte dans les algorithmes de clustering

La métrique ED, qui semble plus adaptée à cette étude, a été utilisée. Les méthodes de clustering L2 peuvent conduire, au sein d'un même cluster, à  
rassemblements de situations physiques différentes (Biabiany et al., 2020). Pour supprimer ces biais liés au clustering L2, la première étape de



la méthode utilisée ici consiste à considérer la variabilité spatiale de la dynamique des courants de surface quotidiens analysés à partir de  
135 L2. La zone d'étude LA a été séparée en trois parties (Fig. 1b) sur la base des centres d'action de transport des radeaux de Sargassum  
rapportés dans la littérature (Franks et al., 2016 ; Berline et al., 2020). A l'ouest de LA, la première zone, LA1, est centrée sur la mer des  
Caraïbes. A l'est, la zone atlantique a été scindée en deux zones vers 13,5°N, juste au-dessus de l'île de la Barbade. Au sud-est se trouve la  
zone LA3 sous l'influence de la région de recirculation nord équatoriale (NERR) et de ses anneaux de réflexion, tandis qu'au nord-est se  
trouve la zone LA2, plus représentative du courant nord équatorial. Les champs journaliers analysés comprennent au total 14 279 mailles (4  
140 282 mailles en LA1, 3 407 mailles en LA2 et 4 536 mailles en LA3).

La deuxième étape a consisté à regrouper les informations portées par les champs de vitesse de courant journaliers  
conditionnellement aux trois zones données dans des histogrammes. La similarité des champs les plus similaires est estimée par  
paire et par zone à partir de la divergence symétrisée de Kullback-Leibler (KL) calculée à partir des histogrammes (Kullback et  
Leibler, 1951). Cela permet d'exprimer l'entropie entre deux distributions sans avoir de raisonnement a priori sur la distribution de  
145 probabilité. La similarité entre deux histogrammes a été quantifiée de cette façon. La dernière étape a consisté à calculer la  
moyenne des valeurs de divergence pour chaque zone. Cela permet d'avoir une valeur unique, nommée Expert Distance (ED)  
quantifiant la similarité entre les individus de la base de données lors du clustering.

L'indice SaMk défini dans Biabiany et al. (2020) a été utilisé. Ceci permet d'exprimer la qualité d'un clustering, par la moyenne de la qualité  
de chaque cluster, qui est elle-même la moyenne des indices de silhouettes sur les éléments du cluster. Cet indice est défini comme suit :

$$= 1 \times \sum_{i=1}^n \left( \frac{1}{|C_i|} \sum_{j \in C_i} (s_{ij}) \right) \quad (2)$$

## 2.6 Analyse de clustering sur les séquences inversées échouées

Afin de mieux comprendre la dynamique du régime actuel qui peut conduire à des échouages de Sargasses sur les côtes de la Guadeloupe, les  
rétroséquences d'échouage de courant de 30 jours passées ont été analysées. Alors que 110 jours d'échouage observés ont été enregistrés entre  
155 janvier 2019 et décembre 2020, seules 107 rétro-séquences ont été étudiées ici. Cela s'explique par le fait que les jours d'échouage enregistrés en  
janvier 2020 ont été supprimés pour éviter les rétro-séquences manquantes des données de la période de décembre 2018. Ces séquences arrière de  
107 brins ont été examinées avec le modèle de courant de surface à plus haute résolution, c'est-à-dire les champs HYCOM. Les dissemblances entre  
ces séquences rétrogrades ont été calculées avec des méthodes d'appariement optimal avant de diviser la population en plusieurs groupes en  
utilisant une classification hiérarchique (Larmarange et al., 2015). La méthode de la plus longue sous-séquence commune (LCS) a été utilisée pour  
160 calculer les distances entre les séquences inversées (Elzinga et Struder, 2015 ; Studer et Ritschard, 2016). Un dendrogramme a été calculé en utilisant  
l'algorithme de Wald. Le critère de perte d'inertie relative le plus élevé a permis de déterminer le nombre optimal de partitions (package TraMiner  
(Gabadinho et al., 2011)).

## 2.7 Système d'aide à la décision

Pour déterminer la probabilité d'échouement de Sargasses à un endroit donné, un arbre de décision a été construit à l'aide d'éléments complémentaires  
165 appelés « modules » (Fig. 2). Ils génèrent chacun des informations basées sur les données d'entrée, y compris les courants de surface avec des effets de vent



(Mercator, HYCOM et ERA-5) et des observations passées d'échouages en Guadeloupe. Ainsi, pour une journée donnée, le système proposé fonctionne comme suit :

● Le module A prend en entrée le mois du jour sélectionné et renvoie la probabilité mensuelle associée (fréquence) de échouage ;

170 ● Module B qui attribue un numéro de cluster au jour focalisé après le clustering ED des courants de surface quotidiens. Puis il construit à partir de ce jour des séquences rétrogrades empiriques de nombres entre 1 et 4 (type de cluster) sur une période de 30 jours passés ;

● Module C qui prend en entrée le nombre de grappes quotidiennes produites par le module B et renvoie la probabilité (fréquence) de échouage associé au type de cluster. Cette probabilité est calculée, par type de cluster, à partir des échouages observés sur les côtes de l'archipel guadeloupéen. Le système a 107 séquences d'échouage en arrière de 30 jours. Ces séquences à rebours démarrent la journée de standing sur les côtes de la Guadeloupe. Cet ensemble de séquences inversées de brin référencées est appelé BASE (Fig. 2b);

175 ● Module D, qui compare la séquence arrière du jour donné aux séquences arrière d'échouage avec la distance de Jaccard. Le module D est interconnecté à BASE et au module B. Il renvoie le pourcentage de correspondance entre eux. Dans la littérature, la moyenne des différents modules est souvent utilisée comme opérateur de décision (Bo. et al., 2020 ; Swain et Hauska, 1977). Dans le présent travail, le pourcentage d'échouage pour un jour donné a été déterminé à l'aide des pourcentages fournis par les modules A, C et D, selon la

180 formule suivante :

$$P(i) = \frac{P_A(i) + P_C(i) + P_D(i)}{3} \quad (3)$$

où  $P(i)$  est la quantité utilisée dans la conception de la règle de décision. Cette règle est simplement la combinaison linéaire des pourcentages des modules A, C et D, calculée selon :

$$P(i) = \frac{1}{3} \times \sum_{j \in \{A, C, D\}} P_j(i) \quad (4)$$

185 où  $i \in \mathcal{J}$ , l'ensemble des jours passés (2019-2020) et  $DECISION(i)$  est une réponse (logique) de l'arbre de décision pour un jour donné c'est-à-dire exprimée sous forme binaire. L'arbre proposé dans la Fig. 2 a été expérimenté sur les 120 premiers jours de l'année 2021, à partir du 1<sup>er</sup> janvier 2021 au 30 avril 2021, soit 120 tests.

### 3. Résultats

#### 3.1 Courants de surface dans la zone focalisée

190 Les déciles des vitesses des courants de surface y compris la dérive, selon l'équation (1), sont présentés dans le tableau 1. Pour les deux modèles HYCOM et Mercator, les intensités de vitesse ne dépassent pas 2,57 ms<sup>-1</sup> et 90% d'entre eux restent en dessous de 0,65 ms<sup>-1</sup>. Les données Mercator ont une médiane de 0,28 ms<sup>-1</sup>, la moyenne de 0,33 ms<sup>-1</sup>, alors que pour HYCOM ces valeurs sont respectivement égales à 0,32 ms<sup>-1</sup> et 0,36 ms<sup>-1</sup>. Le rapport entre le premier et le dernier décile est proche de 6. La figure 3 montre des distributions asymétriques avec une asymétrie égale à 1,31 et 1,21. La masse de distribution est concentrée à gauche. Il existe des valeurs extrêmes indiquant des vitesses de courant de surface avec des écarts 5  
195 fois supérieurs à l'écart type.

Pour évaluer la contribution de chacune des trois régions (c'est-à-dire, LA1, LA2, LA3) aux déciles, la fréquence relative par rapport aux seuils de déciles indiqués dans le tableau 1 est illustrée à la Fig. 4. Trois formes différentes peuvent être observées. Dans la mer des Caraïbes, les distributions de fréquences relatives LA1 de HYCOM et Mercator sont presque horizontales, indiquant une contribution assez constante (~ 3%) sur toute la vitesse



Des classes. Dans l'océan Atlantique (c'est-à-dire, y compris LA2 et LA3), les distributions de vitesse des courants HYCOM et Mercator sont assez similaires. Les distributions de fréquence montrent deux comportements opposés respectivement pour LA2 et LA3. Dans la partie nord Atlantique LA, zone LA2, la fréquence diminue avec la vitesse du courant. Les vitesses actuelles supérieures à 0,65 ms<sup>-1</sup> sont très rares. Au contraire, dans la partie Atlantique sud LA, zone LA3, l'augmentation de la fréquence est observée avec une fréquence maximale liée à des vitesses de courant supérieures à 0,65 ms<sup>-1</sup>. Ces trois distributions de vitesse de courant spécifiques significatives associées à LA1, LA2 et LA3 confirment la nécessité de séparer ces trois zones dans le processus de regroupement des métriques ED.

Les différences entre les vecteurs de courant HYCOM et Mercator ont également été examinées pour chaque cellule de la grille (Fig. 5). Globalement, en mer, les écarts de vitesse des courants sont faibles et restent inférieurs à 0,15 ms<sup>-1</sup>. Ces différences entre HYCOM et Mercator augmentent près des îles avec une valeur moyenne de 0,3 ms<sup>-1</sup>. Les plus grandes différences, supérieures à 0,5 ms<sup>-1</sup> sont observées dans la partie sud de l'arc de Los Angeles, autour de Trinidad et Tobago.

A chaque point de la grille, les écarts angulaires trouvés entre les médianes des directions des vecteurs vitesse du courant de surface peuvent être divisés en trois groupes de magnitude de 45°. Les écarts de direction des courants entre 0 et 45° sont le groupe le plus fréquent dans la région, tandis que ceux entre 45 et 90° restent localisés en aval des îles. Enfin, ceux au-dessus de 90° se produisent exclusivement autour de Trinidad.

### 3.2 Analyse de regroupement

Pour identifier les modèles de courants de surface dans la région, puis ceux qui conduisent au transport des radeaux de Sargassum vers les côtes des îles LA, le regroupement des données maillées selon l'équation (1) a été effectué.

#### 3.2.1 Évaluation du regroupement

L'une des incertitudes connues dans la méthode des k-moyennes est induite par le nombre de clusters sélectionné. Pour trouver un nombre optimal de clusters et identifier la meilleure partition (Biabiany et al. 2020), l'évolution de l'indice de silhouette (SaMk) en fonction du nombre de clusters, k, est représenté sur la Fig. 6. Les indices de silhouette obtenus par la méthode KMS-ED, sont en général supérieurs à 0,2 pour tout k < 15, et restent supérieurs à ceux des méthodes KMS-L2, HAC-L2 et HAC-ED. Ces valeurs indiquent que la qualité des clusters est bien meilleure avec la méthode KMS-ED. Le point d'inflexion de la courbe KMS-ED se produit pour le même nombre de clusters, k = 4, pour les données Mercator et HYCOM. Cela met en évidence quatre régimes de courant représentatifs dans la région étudiée, nommés respectivement MC1, MC2, MC3, MC4 pour Mercator et HC1, HC2, HC3 et HC4 pour HYCOM.

#### 3.2.2 Analyse visuelle des régimes actuels

Les quatre types de circulation de courant de surface, obtenus en intensité et en direction, sont représentés sur les Fig. 7 et 8, respectivement pour l'analyse Mercator et HYCOM. Le parangon qui est le jour le plus proche du centroïde a été choisi pour représenter chaque type de cluster. Les quatre groupes peuvent être distingués par l'expansion NBC et par les emplacements des anneaux de réflexion induits. Les vitesses des courants de surface et leurs lignes de courant associées sont régies par les structures suivantes :

- ceux qui pénètrent par la mer des Caraïbes par le sud, en restant presque parallèles au plateau continental. Ils se produisent dans les régions LA3 et LA1 ;





- Celles dues à la propagation des caractéristiques dynamiques tourbillonnaires liées aux anneaux de réflexion du NBC. Ils viennent du sud de la région LA3, le long du versant atlantique de l'arc des Petites Antilles, avant de traverser la mer des Caraïbes vers 12-14° N ;
- Ceux venant généralement du nord-est des régions LA1 et LA2, représentant la limite sud du gyre subtropical qui coupe les Petites Antilles à environ 15° N. Ils gardent leur direction initiale et sont cisailés par les courants du Sud-Est.

Le nombre de jours correspondant à chaque cluster est donné dans le tableau 2. MC1, HC2 et HC3 sont les plus fréquents sur la période étudiée. Chacun d'eux représente près de 30 % de la production journalière. Cependant, aucun des quatre clusters ne se démarque vraiment. Pour les deux analyses, les différences entre les occurrences de cluster restent inférieures à 10 %.

### 3.2.3 Jours correspondants entre clusters

Les clusters trouvés sont également liés par un ensemble de jours en commun. Les pourcentages de correspondance ont été calculés à l'aide de la formule suivante.

$$p(m,h) = \frac{N(m,h)}{N(m) + N(h) - N(m,h)} \quad (5)$$

où  $p(m,h)$  est le pourcentage de correspondance entre le cluster  $C_m$  et groupe  $C_h$  dérivés respectivement des ensembles de données Mercator et HYCOM.  $N(m,h)$  est le nombre de jours partagés par ces deux clusters. Le tableau 3 montre les résultats.

MC4 - HC2 est la paire de clusters avec le score de correspondance le plus important (69,8%). Il est suivi du couple MC2 - HC1 (60,4 %), puis MC3 - HC3 (56,7 %) et MC1 - HC4 (50,6 %).

### 3.2.4 Distribution et comparaison des intensités

Des déciles ont été utilisés pour étudier et analyser les distributions de vitesse caractérisant chaque cluster. Les évolutions de la fréquence relative de  $U_s(x,y,t)$  en fonction des déciles (Tableau 1) sont présentées dans les Figs. 9 et 10. Pour l'ensemble de l'analyse, les valeurs des déciles restent fixes et constantes, et les courbes sont tracées pour les trois régions décrites dans la Fig.1.

Pour les deux modèles, globalement, trois schémas principaux sont identifiés. Le premier motif comprend les clusters suivants MC1, MC3, HC1 et HC3. Ce schéma est caractérisé par l'augmentation de la courbe de fréquence relative dans les régions LA1 et LA3 et sa diminution dans la région LA2. Les éléments de ces clusters incluent de fortes vitesses de courant supérieures à la médiane de 0,28 ms<sup>-1</sup>. Le deuxième modèle comprend les clusters MC2 et HC2 qui sont caractérisés par la diminution ou la fréquence relative pour les trois régions (c'est-à-dire, LA1, LA2, LA3). Le dernier motif comprend des clusters MC4 et HC4 et correspond à trois courbes concaves avec des maximums situés à des seuils de vitesse différents selon la région étudiée.

Pour examiner les relations possibles, pour une région donnée, entre les deux variables, les seuils de vitesse décile et les grappes identifiées, des tableaux de contingence ont été construits (non présentés) et le test du chi carré a été effectué. Pour les trois domaines, la valeur de  $p$  était bien inférieure à 0,01. Les résultats du test du chi carré ont indiqué que pour les LA1, LA2 et LA3, la distribution de la vitesse dépend du cluster identifié.

### 3.2.5 Saisonnalité

La distribution mensuelle de chaque grappe est tracée (Figs. 11 et 12). Les différences sont relativement nettes pour les deux analyses de modèles avec une variation saisonnière marquée. Les régimes MC3 et HC3 sont observés durant la première moitié de l'année avec un maximum en mars,



265 suivi de MC2 et HC1 d'avril à juillet. Les deux derniers régimes sont observés d'août à décembre. Le couple MC4 HC2, atteint un maximum en septembre tandis que MC1 et HC4 persistent jusqu'en février de l'année suivante.

### 3.3 Liens avec les échouages de Sargasses

Comme pour de nombreux objets flottants, avant de débarquer sur les côtes de LA, les algues sargasses s'accumulent à la surface de l'océan en grande quantité et forment des nappes, ou structures filamenteuses, entrecoupées de zones vides, sous l'influence des courants. Ces structures dynamiques régulièrement observées depuis les satellites, les avions et les navires, ont une certaine inertie (Maximenko et al., 2012).

270 Au-delà de la production biologique, ce sont donc les conditions dynamiques particulières des courants de surface et des vents de surface qui peuvent conduire à des échouages massifs de Sargasses sur les zones côtières caribéennes.

L'évolution mensuelle des jours d'échouage observés sur les côtes guadeloupéennes, l'évolution mensuelle de l'abondance des Sargasses sur la région Centre Atlantique (SaWS, <https://optics.marine.usf.edu/projects/SaWS.html>) ont également été analysées sur le période 2019-2020 (Fig. 11 et 12). Au cours de ces deux années, la quantité de sargasses sur la région de l'Atlantique central a augmenté de manière

275 significative de février à juillet, puis a diminué de juillet à novembre.

Deux valeurs maximales d'échouage sont trouvées : une en mars et la seconde en août. Les dates d'échouages et les dates d'occurrence des clusters ont également été comparées dans le tableau 4. Le couple MC3 - HC3 rassemble le plus grand nombre de similitudes, suivi des clusters MC1 et HC2.

Les couples (MC1, HC2) et (MC3, HC3) regroupent le plus grand nombre de jours d'échouage observés en Guadeloupe (Tableau 4).

280 Ces paires d'amas seraient favorables au transport de ces algues vers les côtes des îles des Petites Antilles. MC2 et HC1 sont les deux clusters avec le plus petit nombre de jours d'échouage.

### 3.4 Séquences rétrogrades du régime actuel conduisant à des échouages

L'analyse de clustering HAC sur les séquences rétrogrades du régime actuel conduisant aux jours d'échouage observés a permis de répartir les 107 séquences rétrogrades en quatre classes, respectivement appelées Seq1, Seq2, Seq3 et Seq4. Cette analyse n'a intégré que les données de courant de surface HYCOM qui ont une plus grande résolution que Mercator. Au cours de la période ciblée (c'est-à-dire 2019-2020), Seq4 (39,3 %) et Seq2 (37,4 %) ont la plus grande occurrence (tableau 5). Seq1 et Seq3 ont une occurrence respective de 16,8% et 6,5%. La figure 13 montre que Seq2, Seq3 et Seq4 sont caractérisés par les régimes de courant modaux respectifs HC3, HC1 et HC2. Pour les séquences arrière Seq1, il n'y a pas de régime de courant dominant clair. La distribution mensuelle des principales classes de séquences rétrogrades Seq2 et Seq4 met en évidence un clivage saisonnier important (Fig. 14). Les séquences inversées Seq2 se sont produites de décembre à juin tandis que celles de Seq4 se sont produites de

290 juillet à novembre. Ces deux distributions semblent également significativement corrélées avec les occurrences mensuelles d'échouages observés. Alors que le premier pic d'échouage survenant en mars est lié aux occurrences maximales de Seq2, le deuxième pic d'échouage survenu en août est lié aux occurrences maximales de Seq4.

### 3.5 Résultats du système d'aide à la décision

Le tableau 6 présente les résultats obtenus pour Mercator et HYCOM. Pour HYCOM, les résultats montrent que le modèle (clustering + arbre de

295 décision) a correctement classé 41,7% (37,5% pour Mercator) des observations en classe échouage, et 81,3% (54,2% pour Mercator) des observations parmi les non-échouages (Tableau 6). Globalement, la performance de l'arbre de décision atteint 50,8% pour la base de données Mercator et 73,3 % pour HYCOM. Le comportement de chaque module est présenté sur la Fig. 15. En général, les modules A et C restent avec des probabilités



inférieur à 0,2 et 0,5 respectivement. Ces deux valeurs sont attendues. Ils sont induits par la forte variabilité spatio-temporelle du phénomène.

300 Les pourcentages d'échouage par cluster associés au module C montrent des probabilités empiriques proches de 0,3 indiquant qu'un tiers des jours dans les clusters concernés sont des jours d'échouage. Le module D produit des probabilités empiriques liées aux liens entre les séquences observées passées et les séquences correspondant au jour de prévision. Dans notre cas, ils peuvent atteindre 0,95 (Fig. 15a) indiquant de fortes similitudes entre les séquences.

#### 4. Discussion

##### 305 4.1. Indices de performance et qualité de clustering

La performance du clustering et la qualité des clusters ont été évaluées à l'aide du coefficient de silhouette. L'évolution de ce coefficient (Fig. 6) montre clairement que d'une part, les méthodes basées sur l'algorithme HAC produisent des valeurs inférieures à celles obtenues par les algorithmes KMS. En revanche, pour ED, les indices de silhouette sont largement supérieurs à ceux trouvés par la distance L2 comme écrit par Biabiany et al. (2020). Cette évolution du coefficient de silhouette permet de retenir quatre types représentatifs de régimes actuels dans cette partie de la région Caraïbe.

310 Cependant, en raison du manque de travaux pour cette région, les comparaisons entre les résultats actuels et d'autres études étaient très limitées. Dans d'autres études, des auteurs ont proposé un nombre similaire de régimes dominants à grande échelle, dans le Pacifique tropical (Fereday et al., 2008),  
315 Dans notre cas, les distributions de vitesse montrent quatre profils singuliers confirmant la bonne performance du clustering. Chaque grappe avait également des distributions mensuelles distinctes. Cette analyse a permis de mieux comprendre la variabilité des circulations des courants de surface dans cette région.

##### 4.2. Analyse des courants de surface appliquée à l'aléa Sargassum

En termes de distribution spatiale, les clusters présentent des différences notables pour les deux types d'analyse de modèle et trois facteurs de variabilité peuvent être identifiés.

320 Le premier est l'évolution saisonnière du front de réflexion NBC (Baklouti et al. 2007). Le NBC alimente le courant de Guyane (GC) mais se sépare aussi nettement, vers 6°–8°N, de la côte sud-américaine et se rétroréfléchit pour alimenter, cette fois, le NECC vers l'est. De grands anneaux isolés se déplacent vers le nord-ouest vers la mer des Caraïbes, sur un parcours parallèle au littoral sud-américain, puis interagissent avec les Petites Antilles (Fratantoni et al., 2002, 2006). Ces deux structures dynamiques, les anneaux GC et NBC, contribuent de manière significative au transfert des eaux de surface de l'Atlantique Sud vers les Caraïbes. Ces structures dynamiques se retrouvent sur les quatre clusters identifiés et semblent fonctionner  
325 toute l'année avec des variations d'intensité.

Une autre partie de cette variabilité est causée par les anneaux du NBC qui se déplacent vers le nord-ouest depuis l'Atlantique équatorial et interagissent avec la topographie escarpée de l'arc des Petites Antilles. MC2 et HC1 sont deux cas typiques. Les interactions avec le chapelet insulaire provoquent des perturbations importantes de l'afflux par les passages sud avec un blocage. Cela permet un transport méridien des eaux de surface vers le nord, le long de l'arc LA (Fratantoni et Richardson, 2006 ; Huang et al., 2021). L'arc des Petites Antilles a clairement détourné la dérive  
330 initialement vers le nord-ouest des anneaux NBC vers un parcours plus au nord parallèle à l'arc insulaire. Johns et al. (2002) ont montré que la traversée de l'afflux de l'Atlantique vers la mer des Caraïbes par les passages des îles du Vent (c'est-à-dire les Petites



Antilles sud de Trinidad à la Martinique) a un cycle saisonnier très asymétrique, avec un maximum en juin et un minimum en septembre-octobre. La distribution annuelle des clusters MC2 et HC3 est proche de celle trouvée par Johns et al. (2002).

Le dernier facteur identifié est lié aux courants de surface présents dans la région de l'Atlantique Nord en raison du courant Nord et de la circulation de gyre associée. Dans cette partie de la zone d'étude, plusieurs clusters présentent des vitesses de courant plus faibles et des zones avec de grands écarts angulaires de direction ont également été identifiées. Dans la zone LA2 (c'est-à-dire la zone Atlantique entre 14,5°N et 18°N), les fréquences relatives des vitesses supérieures à la moyenne sont les plus faibles avec les alizés du nord-est. Les zones de cisaillement vent-courant sont également les plus étendues.

L'écoulement éolien se produit de l'emplacement du gyre subtropical jusqu'à 15°N, près de l'île de la Martinique (Johns et al., 2002). Les passages par les îles Sous-le-Vent ont un apport maximum en septembre et un minimum en juin.

La comparaison entre les situations météorologiques à grande échelle correspondant aux parangons a montré que les principales différences entre les clusters de régime actuel sont liées à la localisation et à l'extension des centres anticycloniques, au positionnement de l'ITCZ, à l'intensité de la basse Caraïbe Jet.

Tous les clusters contiennent des jours d'échouage en abondance relative, 12 à 36 % des jours d'échouage pour les deux années 2019, 2020. La répartition mensuelle des clusters et la répartition des échouages observés en Guadeloupe sont décalées. Le premier pic d'échouages, en mars et semble lié à la fréquence maximale des clusters MC3 et HC3. Le deuxième pic d'échouages observés se produit en août et semble associé aux clusters MC1, HC2 et HC4. Johns et al (2020) ont constaté que le forçage au vent induit par la convergence du vent accumule des radeaux de sargasses dans l'ITCZ entre avril et septembre. Cette accumulation contribuerait au pic d'échouage observé en août. L'analyse de clustering sur les rétroséquences du courant d'échouage a confirmé que la récurrence de HC3 (entre décembre et juin) et HC2 (entre juillet et novembre) induirait des échouages importants sur les côtes guadeloupéennes durant ces périodes respectives.

Le régime de courant HC2 est caractérisé par la prédominance du gyre de l'Atlantique Nord avec de faibles vitesses dans l'Atlantique Centre-Ouest et les lignes de courant zonales. Quant au régime de courant HC3, il est caractérisé par un fort Courant de Guyane avec des vitesses élevées dans la région LA3 et des lignes de courant méridiennes presque parallèles à l'Arc des Petites Antilles.

#### 4.3 Performances du modèle prédictif

Une méthode basée sur l'apprentissage automatique pour prédire l'échouage des sargasses a été proposée et a été construite à partir d'un arbre de décision. Cette méthode a déjà été utilisée pour d'autres paramètres et elle permet d'améliorer à la fois la précision de la prédiction et l'effet de boîte noire du réseau de neurones. Comparé aux méthodes statistiques paramétriques habituelles, il peut efficacement surmonter la multicolinéarité des variables indépendantes (par exemple, le courant océanique et le vent de surface). La précision de l'arbre de décision atteint 73,3% pour HYCOM contre 50,8% pour Mercator. Des scores de performance similaires ont été trouvés pour les arbres de décision prédisant les précipitations estivales à Chongqing (Chine) (Bo et al., 2020) ou les risques de glissement de terrain dans la province de Yen Bai (Vietnam) (Pham et al., 2020). Cependant, des performances asymétriques ont été mises en évidence avec de meilleurs résultats pour les vrais négatifs que pour les vrais positifs (tableau 6). Ceux-ci peuvent être attribués à l'algorithme et à la faible capacité du modèle à gérer différents ensembles de données. Ces erreurs de prédiction sont plus importantes pour Mercator.

Plusieurs pistes d'amélioration du modèle prédictif ont été identifiées. Le manque de données d'observation dans le temps (c'est-à-dire seulement deux ans) peut affaiblir la décision finale et induire un surajustement. L'arborescence pourrait également être améliorée en pondérant et hiérarchisant les différents modules pour augmenter leur pertinence. L'amélioration des résultats peut être trouvée en optimisant la règle de calcul de décision proposée (3) pour mieux intégrer les caractéristiques du phénomène observé.



## 5. Conclusion

Depuis une décennie, les pays de la Caraïbe, et particulièrement l'AL, subissent les impacts induits par l'arrivée massive et régulière de Sargasses sur leurs zones côtières. Cette étude présente l'application d'une approche de clustering pour déterminer les types de circulations de courants de surface intégrant la dérive supplémentaire du vent et leurs liens éventuels avec les échouages de Sargasses observés sur les côtes de Los Angeles. L'archipel guadeloupéen a été choisi comme site d'observation des échouages pour la période 2019-2020. Cette analyse a été réalisée à l'aide des versions les plus récentes des modèles 3D de courants océaniques, Mercator et HYCOM. Les données de vitesse du vent de surface du modèle ERA-5 ont également été utilisées. Le regroupement des champs de courants de surface spatio-temporels, y compris la dérive, a été produit à l'aide de l'algorithme k-mean combiné à la métrique de distance experte.

Pour cette région (8-17°N, 66-55°W) divisée en trois sous-régions, nous identifions quatre modèles cohérents à partir d'ensembles de données. Ils contiennent les structures actuelles liées aux courants de Guyane, les branches du gyre atlantique subtropical, le front et les anneaux de réflexion liés au NBC.

La résolution plus fine de l'analyse HYCOM a fourni des informations plus détaillées sur les vitesses des courants de surface près des îles que les champs Mercator (c'est-à-dire une différence de vitesse locale moyenne de 0,3 ms<sup>-1</sup>). Au large, ces différences restent très faibles.

Les liens entre les clusters et les échouages observés en Guadeloupe ont été étudiés en tenant compte des vents, des distributions de vitesse parangon et des cartes d'abondance mensuelles. Les circulations de courants de surface caractérisant les paires d'amas (MC3 ; HC3) et (MC4 ; HC2) semblaient les plus favorables au transport et à l'échouage des Sargasses sur les côtes des Petites Antilles.

L'analyse de clustering sur les rétroséquences du courant d'échouage basée sur les champs HYCOM a confirmé que la récurrence de HC3 (Seq2, entre décembre et juin) et HC2 (Seq4, entre juillet et novembre) induirait des échouages importants sur les côtes guadeloupéennes durant ces périodes respectives. Alors que le régime de courant HC2 est caractérisé par la prédominance du gyre nord-atlantique avec de faibles vitesses zonales, le régime de courant HC3 est marqué par l'influence du NBC, les anneaux de réflexion induits et le fort courant de Guyane conduisant à des vitesses méridiennes plus élevées dans la région LA3 .

Des algorithmes d'apprentissage automatique (KMS, ED, classificateur d'arbre de décision) ont été appliqués pour estimer la probabilité d'échouages de sargasses en Guadeloupe, sur la base : des prévisions de courant de surface, des séquences rétrogrades du régime actuel et de plusieurs combinaisons de probabilités. Le score de performance de ce modèle prédictif a montré que la résolution plus fine de HYCOM (c'est-à-dire inférieure à l'échelle de 5 km) semble plus adaptée pour reproduire les schémas de courants à petite échelle induisant ou non des échouages dans les Petites Antilles. La précision de l'arbre de décision atteint respectivement 50,8% et 73,3% pour Mercator et HYCOM. Cette précision pourrait être améliorée en pondérant et hiérarchisant les différents modules. De nouveaux modules seraient également ajoutés comme les observations de télédétection des Sargasses.

En raison de la disponibilité très récente de la version de nouvelle génération HYCOM sélectionnée, la présente étude n'a été menée que sur deux ans (c'est-à-dire 2019-2020). La période étudiée pourrait être étendue à plusieurs années pour intégrer la variabilité interannuelle des courants de surface.

Néanmoins, les résultats obtenus sont très encourageants et ouvrent de nouvelles possibilités pour la prévision de ce type de risque naturel. Les méthodes d'apprentissage automatique développées dans cette analyse se sont révélées utiles dans la prévention d'un risque naturel en fonction de combinaisons physiques multifactorielles.



400 Le présent système prédictif d'analyse de regroupement pourrait être appliqué à d'autres Petites Antilles en modifiant le site  
d'échouage d'observation. L'association de méthodes de clustering et d'arbres de décision nécessitant de faibles coûts de calcul  
peut améliorer les systèmes opérationnels existants pour aider les décideurs dans la gestion du risque sargasses. Maréchal et al.  
(2017) ont limité le point de départ de leur système opérationnel de prévision à court terme à 50-100 km des côtes de Los Angeles  
afin de réduire les erreurs de prévision. Cette limite géographique correspondrait à une période de prévision de 1 à 2 jours avant  
405 l'échouage. Les informations régionales actuelles sur la dynamique des courants conduisant à l'arrivée des Sargasses à proximité  
des îles seraient utiles pour étendre cette limite. De cette façon,

**Disponibilité des données.** Les données de cette recherche ne sont pas accessibles au public. Les chercheurs intéressés peuvent contacter l'auteur correspondant de cet  
article.

410 **Les contributions de l'auteur.** L'étude a été principalement conceptualisée et rédigée par DB et EB. RC1, RC2, NS ont commenté les résultats et  
examiné le manuscrit. RC2 et NS ont contribué au traitement des données d'observation des échouages.

**Intérêts concurrents.** Les auteurs déclarent n'avoir aucun conflit d'intérêts.

**Remerciements.** Cette étude a été soutenue par le projet FEDER/C3AF (numéro de subvention : CR/16-115). Les auteurs tiennent à  
remercier Danièle Frison qui a aidé à la traduction.

#### 415 **Références**

Anses : Expositions aux émanations d'algues sargasses en décomposition aux Antilles et en Guyane, Rapport technique, ANSES, Maisons-  
Alfort, France, 162 pp, disponible sur : <https://www.anses.fr/en/system/files/AIR2015SA0225Ra.pdf>, 2017.

Allshouse, MR, Ivey, GN, Lowe, RJ, Jones, NL, Beegle-Krause, CJ, Xu, J. et Peacock, T. : Impact du vent sur les structures cohérentes  
lagrangiennes de la surface de l'océan, Environ. Fluid Mech. 17, 473-483, <https://doi.org/10.1007/s10652-016-9499-3>, 2017. Arnault, S.,

420 Thiria, S., Crépon, M. et Kaly, F. : A tropical Atlantic analyse dynamique en combinant l'apprentissage automatique et les données  
satellitaires, Advances in Space Research, 68, 2, 467-486, <https://doi.org/10.1016/j.asr.2020.09.044>. 2021.

Baklouti, M., Devenon, J.-L., Bourret, A., Froidefond, J.-M., Ternon, J.-F. et Fuda, J.-L. : Nouvelles perspectives sur le continent  
guyanais circulation du plateau et sa relation avec la rétroflexion du courant du nord du Brésil, J. Geophys. Res., 112, C02023,  
doi:10.1029/2006JC003520., 2007.

425 Berline, L., Ody, A., Jouanno, J., Chevalier, C., André, J.-M., Thibaut, T. et Ménard, F. : rétrospective de la dispersion 2017 des algues  
sargasses dans le Nord tropical Atlantique, Marine Pollution Bulletin, 158, <https://doi.org/10.1016/j.marpolbul.2020.111431>, 2020.

Bernard, D., Biabiany, E., Sekkat, N., Chery, R., et Cécé, R. : Echouage massif d'algues sargasses pélagiques sur les côtes des Antilles  
françaises : Analyse des situations observées avec l'analyse et la prévision océaniques globales Opérationnelles Mercator système. 24e

430 Congrès Français de Mécanique, Brest, <https://cfm2019.sciencesconf.org/258628/document>, 2019.

Bertola, LD, Boehm, JT, Putman, NF, Xue, AT, Robinson, JD, Harris, S., Baldwin, CC, Overcast, I. et Hickerson, MJ : Flux génétique  
asymétrique dans cinq syngnathidés co-distribués expliqués par courants océaniques et propension au rafting, Proc. R. Soc. B.,  
287, <https://doi.org/10.1098/rspb.2020.0657>, 2020.



- 435 Biabiany, E., Bernard, D., Page, V., et Paugam-Moisy, H. : Design of an expert distance metric for climate clustering : Le cas des pluies aux Petites Antilles, *Informatique & Géosciences*, 145, <https://doi.org/10.1016/j.cageo.2020.104612>, 2020.
- Bo, X., Chunfen, Z., Xinning, D. et Jiayue, W. : L'application d'un arbre de décision et d'un modèle forestier stochastique dans la prévision des précipitations estivales à Chongqing. *Atmosphère*, 11(5), 508, <https://doi.org/10.3390/atmos11050508>, 2020.
- Boé, J. et Terray, L. : Régimes météorologiques et downscaling, *La Houille Blanche*, 94, 45-51, <https://doi.org/10.1051/lhb:2008016>, 2008.
- 440 Brooks, MT, Coles, VJ, Hood, RR et Gower, JF : Facteurs contrôlant la distribution saisonnière des sargasses pélagiques, *Mar. Ecol. Programme. Ser.*, 599, 1-18, <https://doi.org/10.3354/meps12646>, 2018.
- Cassou, C., Terray L., Hurrell, JW et Deser, C. : Régimes climatiques hivernaux de l'Atlantique Nord : asymétrie spatiale, stationnarité dans le temps et forçage océanique. *J. Climat.* 17, 1055-1068, [https://doi.org/10.1175/1520-0442\(2004\)017<1055:NAWCRS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<1055:NAWCRS>2.0.CO;2), 2004.
- 445 Copernicus PSY4V3R1 Mercator 1/12 degré 3D : <https://datastore.cls.fr/catalogues/mercator-model-psy4v3-velocity-112>, dernier accès : 15 juin 2021.
- Cuevas, E., Uribe-Martínez, A. et Liceaga-Correa, M. : Une approche multi-indices de télédétection par satellite pour discriminer les sargasses pélagiques dans les eaux de la péninsule du Yucatan, Mexique, *Int. J. Remote Sens.*, 39, 3608-3627, <https://doi.org/10.1080/01431161.2018.1447162>, 2018.
- 450 Elzinga, CH et Studer, M. : Séquences d'orthographe, proximités d'état et mesures de distance, méthodes et recherches sociologiques, 44, 3-47, <https://doi.org/10.1177/0049124114540707>, 2015.
- Fereday, DR, Knight, JR, Scaife, AA, Folland, CK et Philipp, A. : Analyse par grappes des types de circulation nord-atlantique-européen et des liens avec les températures de surface de la mer du Pacifique tropical, *Journal of Climate*, 21(15), 3687 -3703, <https://doi.org/10.1175/2007JCLI1875.1>, 2008.
- 455 Fossette, S., Putman, NF, Lohmann, KJ, Marsh, R. et Hays, GC : Un guide du biologiste pour évaluer les courants océaniques : une revue, *Mar. Ecol. Programme. Ser.*, 457, 285-301, <http://www.jstor.org/stable/24876354>, 2012.
- Florenne, T., Guerber, F., et Colas-Belcour, F. : Le phénomène d'échouages des sargasses dans les Antilles et en Guyane, ministère de l'Outre-mer, ministère de l'Environnement, de l'Énergie et de la Mer, ministère de l'Agriculture, Agroalimentaire et forêt, Paris, France, 406 pp, [https://agriculture.gouv.fr/sites/minagri/files/cgaaer\\_15113\\_2016\\_rapport.pdf](https://agriculture.gouv.fr/sites/minagri/files/cgaaer_15113_2016_rapport.pdf), (dernier accès : 10 avril 2021), 2016.
- 460 Franks, JS, Johnson, DR, Ko, DS, Sanchez-Rubio, G., Hendon, JR et Lay, M. : Afflux sans précédent de sargasses pélagiques le long des côtes des îles des Caraïbes pendant l'été 2011. Dans : Actes du Golfe et des Caraïbes Institut des pêches, 64, 6-8, <http://aquaticcommons.org/21307/>, 2012.
- Franks, JS, Johnson DR et Ko DS : Pelagic Sargassum in the Tropical North Atlantic. *Gulf and Caribbean Research*, 27, <https://doi.org/10.18785/gcr.2701.08>, 2016.
- 465 Fratantoni, DM et Glickson, DA : Génération et évolution de l'anneau de courant du nord du Brésil observées avec SeaWiFS, *Journal of Physical Oceanography*, 32(3), 1058-1074. [https://doi.org/10.1175/1520-0485\(2002\)032<1058:NBCRGA>2.0.CO;2](https://doi.org/10.1175/1520-0485(2002)032<1058:NBCRGA>2.0.CO;2), 2002.
- Fratantoni, DM, et Richardson, PL: L'évolution et la disparition des anneaux de courant du nord du Brésil, *Journal of Physical*



Océanographie, 36, 1241-1264, <https://doi.org/10.1175/JPO2907.1>, 2006.

470 Fučkar, NS, Guemas, V., Johnson, NC, Massonet, F., and Doblas-Reyes, FJ : Grappes de variabilité interannuelle de la glace de mer dans l'hémisphère nord. *Clim Dyn*, 47, 1527-1543, <https://doi.org/10.1007/s00382-015-2917-2>, 2016.

Gabadinho, A., Ritschard, G., Müller, N., Studer, M. : Analyser et visualiser des séquences d'états dans R avec TraMineR, *Journal of Statistical Software*, 40(4), 1-37, <https://doi.org/10.18637/jss.v040.i04>, 2011.

475 Gasparin, F., Guinehut, S., Mao, C., Mirouze, I., Rémy, E., King, RR, Hamon, M., Reid, R., Storto, A., Le Traon, PY., Martin, MJ et Masina, S. : Exigences pour un système intégré d'observation in situ de l'océan Atlantique à partir d'expériences de simulation de systèmes d'observation coordonnés, *Frontiers in Marine Science*, 83(6), <https://doi.org/10.3389/fmars.2019.00083>, 2019.

Gower, J., Hu, C., Borstad, G. et King, S. : Ocean Color Satellites Show Extensive Lines of Floating Sargassum in the Gulf of Mexico, dans *IEEE Transactions on Geoscience and Remote Sensing*, 44, 3619-3625, <https://doi.org/10.1109/TGRS.2006.882258>, 2006.

480 Gower, J. et King, S. : Répartition des sargasses flottantes dans le golfe du Mexique et l'océan Atlantique cartographié à l'aide de MERIS, *International Journal of Remote Sensing*, 32, 1917-1929, <https://doi.org/10.1080/01431161003639660>, 2011.

Gower, J., Young, E. et King, S. : Les images satellite suggèrent une nouvelle région source de Sargassum en 2011, *Remote Sensing Letters*, 4:8, 764-773, <https://doi.org/10.1080/2150704X.2013.796433>, 2013.

485 Haza, AC, Paldor, N., Ozgokmen, TM, Curcic, M., Chen, SS et Jacobs, GA : Estimations basées sur le vent des courants de surface océaniques à partir d'amas massifs de bouées dérivantes dans le golfe du Mexique. *Journal of Geophysical Research : Océans*, 124, 5844-5869. <https://doi.org/10.1029/2018JC014813>, 2019.

Hu, C. : Un nouvel indice de couleur des océans pour détecter les algues flottantes dans les océans mondiaux, *Remote Sens. Environ.*, 113, 2118-2129, <https://doi.org/10.1016/j.rse.2009.05.012>, 2009.

Hu, C., Feng, L., Hardy, R. F et Hochberg, EJ : Exigences spectrales et spatiales des mesures à distance des macroalgues pélagiques Sargassum, *Remote Sens. Environ.*, 167, 229-246, <https://doi.org/10.1016/j.rse.2015.05.022>, 2015.

490 Hu, C., Murch, B., Barnes, B., Wang, M., Maréchal, J., Franks, J., Johnson, D., Lapointe, B., Goodwin, D., Schell J. et Siuda, A. : La surveillance des sargasses avertit des algues entrantes, *Eos*, 97, <https://doi.org/10.1029/2016EO058355>, 2016.

Huang, M., Liang, X., Zhu, Y., Liu, Y. et Weisberg, RH : les tourbillons relient l'océan Atlantique tropical et le golfe du Mexique. *Lettres de recherche géophysique*, 48, e2020GL091277. <https://doi.org/10.1029/2020GL091277>, 2021.

495 Johns, WE, Townsend, TL, Fratantoni, DM, Douglas Wilson, DW : Sur l'afflux de l'Atlantique vers la mer des Caraïbes, *Deep Sea Res. Première partie Oceanogr. Rés. Pap.*, 49, 211-243, [https://doi.org/10.1016/S0967-0637\(01\)00041-3](https://doi.org/10.1016/S0967-0637(01)00041-3), 2002.

Johns, EM, Lumpkin, R., Putman, NF, Smith, RH, Muller-Karger, FE, Rueda, D., Hu, C., Wang, M., Brooks, MT, Gramer, LJ, Werner FE : Établissement d'une population pélagique de sargasses dans l'Atlantique tropical : conséquences biologiques d'un événement de dispersion à longue distance à l'échelle du bassin, *Progrès en océanographie* (2020), 182, 25 pp, <https://doi.org/10.1016/j.pocean.2020.102269>, 2020.

500 Johnson, D., Ko, DS, Franks, JS, Moreno, P., Sanchez-Rubio, G. : L'invasion des Sargasses dans les Caraïbes orientales et sa dynamique des actes de l'Institut des pêches de l'Atlantique Nord équatorial, du Golfe et des Caraïbes, 65, 102-103, 2014.





- Jouanno, J., Benschila, R., Berline, L., Soulié, A., Radenac, M.-H., Morvan, G., Diaz, F., Sheinbaum, J., Chevalier, C., Thibaut, T., Changeux, T., Menard, F., Berthet, S., Aumont, O., Ethé, C., Nabat, P. et Mallet, M. : Un modèle basé sur NEMO de la distribution des Sargasses dans la zone tropicale Atlantique : description du modèle et analyse de sensibilité (NEMO-Sarg1.0), *Geosci. Modèle Dev.*, 14, 4069–4086, <https://doi.org/10.5194/gmd-14-4069-2021>, 2021.
- Jouini, M., Béranger, K., Arsouze, T., Beuvier, J., Thiria, S., Crépon, M. et Taupier-Letage, I. : La circulation de surface du canal de Sicile revisitée à l'aide d'une analyse de une simulation haute résolution, *J. Geophys. Rés. Océans*, 121, 4545– 4567, doi : 10.1002/2015JC011472, 2016.
- Kullback, S., et Leibler, R. : Sur l'information et la suffisance, *Annals of Mathematical Statistics*, 22, 79-86, <https://doi.org/10.1214/aoms/1177729694>, 1951.
- Larmarange, J., Mossong, J., Bärnighausen T., Newell ML : Dynamique de la participation dans la surveillance longitudinale du VIH basée sur la population en Afrique du Sud rurale, *PLOS ONE*, 10(4), 16 pp, <https://doi.org/10.1371/journal.pone.0123345>, 2015.
- Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., et Le Traon, P.-Y. : Mises à jour récentes du Copernicus Marine Service global ocean monitoring and système de prévision en temps réel haute résolution 1/12°, *Ocean Science*, 14, 1093-1126, <https://doi.org/10.5194/os-14-1093-2018>, 2018.
- Lumpkin, R. et Garzoli SL : Circulation près de la surface dans l'océan Atlantique tropical, *Deep Sea Res. Première partie Oceanogr. Rés. Pap.*, 52, 495-518, <https://doi.org/10.1016/j.dsr.2004.09.001>, 2005.
- Maréchal, J.-P., Hellio, C., Hu, C. : Une méthode simple, rapide et fiable pour prédire l'échouement des sargasses dans les Petites Antilles, *Remote Sens. Appl. : Soc. Environ.*, 5, 54-63, <https://doi.org/10.1016/j.rsase.2017.01.001>, 2017.
- Michelangeli, P., Vautard, R., Legras, B. : Occurrence et quasi stationnarité des régimes météorologiques., *J. Atmos. Sci.*, 52, 1237-1256, [https://doi.org/10.1175/1520-0469\(1995\)052](https://doi.org/10.1175/1520-0469(1995)052), 1995.
- Miron, P., Olascoaga, MJ, Beron-Vera, FJ, Putman, NF, Triñanes, J., Lumpkin, R. et Goni, GJ : Regroupement de dériveurs de type débris marins et sargasses expliqué par la dynamique des particules inertielles, *Geophys. Rés. Lett.*, 47, <https://doi.org/10.1029/2020GL089874>, 2020.
- Pham, BT, Phong, TV, Nguyen-Thoi, T., Parial, K., Singh SK, Ly, H.-B., Nguyen, KT, Ho, LS, Van Le, H. et Prakash, I. , Modélisation d'ensemble de la susceptibilité aux glissements de terrain à l'aide d'un apprenant de sous-espace aléatoire et de différents classificateurs d'arbres de décision, *Geocarto International*, <https://doi.org/10.1080/10106049.2020.1737972>, 2020.
- Putman, NF, Goni, GJ, Gramer, LJ, Hu, C., Johns, EM, Trinanes, J. et Wang, M. : Simulation des voies de transport des sargasses pélagiques de l'Atlantique équatorial à la mer des Caraïbes, *Prog. Oceanogr.*, 165, 205-214, <https://doi.org/10.1016/j.pocean.2018.06.009>, 2018.
- Putman, NF, Lumpkin, R., Olascoaga, MJ, Trinanes, J. et Goni, GJ : Améliorer les prévisions de transport des sargasses pélagiques, *Journal of Experimental Marine Biology and Ecology*, 529, <https://doi.org/10.1016/j.jembe.2020.151398>, 2020.
- Swain, PH et Hauska, H. : Le classificateur d'arbre de décision : conception et potentiel, *IEEE Transactions on Geoscience Electronics*, 15, <https://doi.org/10.1109/TGE.1977.6498972>, 1977.



Resiere, D., Valentino, R., Nevière R., Banydeen, R., Gueye, P., Florentin, J., Cabié, A., Lebrun, T., Mégarbane B., Guerrier G. et Mehdaoui, H. : Algues sargasses des îles des Caraïbes : un problème international de santé publique, *The Lancet*, 392, 2691, [https://doi.org/10.1016/S0140-6736\(18\)32777-6](https://doi.org/10.1016/S0140-6736(18)32777-6), 2018.

540 Rousseeuw, PJ : Silhouettes : une aide graphique à l'interprétation et à la validation de l'analyse par grappes, *Journal of Computational and Applied Mathematics*, 20, 53-65, [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7) , 1987.

Sargassum Watch System (SAWS), Université de Floride du Sud, <https://optics.marine.usf.edu/projects/SaWS.html>, dernier accès : 15 juin 2021.

Schell, JM, Goodwin, DS et Siuda, ANS : Événements récents d'inondation de sargasses dans les Caraïbes : les observations à bord révèlent la dominance d'une forme auparavant rare, *Oceanography*, 28, 8–11, <https://www.jstor.org/stable/24861895>, 2015.

545 Széchy, MTM, Guedes, PM, Baeta-Neves, MH et Oliveira, EN : Vérification de *Sargassum natans* (Linnaeus) Gaillon (Heterokontophyta : Phaeophyceae) de la mer des Sargasses au large des côtes du Brésil, océan Atlantique occidental, *Liste de contrôle*, 8 , 638-641, <http://dx.doi.org/10.15560/8.4.638>, 2012.

Studer, M. et Ritschard, G. : Ce qui compte dans les différences entre les trajectoires de vie : un examen comparatif des mesures de dissemblance de séquence, *Journal de la Royal Statistical Society, série A*, 179(2), 481-511, <https://doi.org/10.1111/rssa.12125>, 2016.

550 Van Tussenbroek, BI, Hernández Arana, HA, Rodríguez-Martínez, RE, Espinoza-Avalos, J., Canizales-Flores, HM, González-Godoy, CE, Barba-Santos, MG, Vega-Zepeda, A. et Collado -Vides, L. : Impacts sévères des marées brunes causées par *Sargassum* spp. sur les communautés d'herbiers côtiers des Caraïbes, *Mar. Pollut. Bull.*, 122, 272-281. <https://doi:10.1016/j.marpolbul.2017.06.057>, 2017.

555 Wang, M. et Hu, C. : Cartographier et quantifier la distribution et la couverture des sargasses dans le centre-ouest de l'Atlantique à l'aide des observations MODIS, *Remote Sens. Environ.*, 183, 350-367, <https://doi.org/10.1016/j.rse.2016.04.019>, 2016.

Wang, M., et Hu, C. : Prédire les efflorescences de sargasses dans la mer des Caraïbes à partir des observations MODIS, *Geophys. Rés. Lett.*, 44, 3265–3273, <https://doi.org/10.1002/2017GL072932>, 2017.

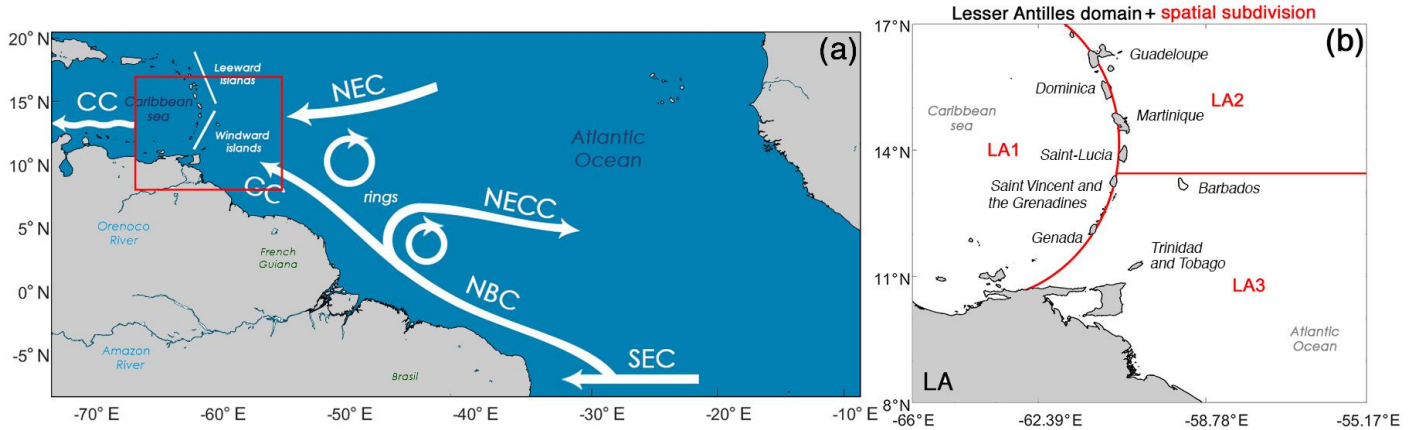
Wang, M., Hu, C., Cannizzaro, J., English, D., Han, X., Naar, D., B., Lapointe, R., Brewton, F., Hernandez. : Télédétection des sargasses biomasse, nutriments et pigments. *Géophys. Rés. Lett.*, 45, 12 359– 12 367. <https://doi.org/10.1029/2018GL078858>, 2018.

560 Wang, M., Hu, C., Barnes, BB, Mitchum, G., Lapointe, B., Montoya, JP : La grande ceinture atlantique de sargasses, *Science*, 365, 83-87, <https://doi.org/10.1126/science.aaw7912>, 2019.

Webster, RK et Linton T. : Développement et mise en œuvre du *Sargassum* Early Advisory System (SEAS), *Shore & Beach*, 81, 1–6, [http://www.sargassoseacommission.org/storage/Webster\\_et\\_linon\\_2013\\_1.pdf](http://www.sargassoseacommission.org/storage/Webster_et_linon_2013_1.pdf), 2013.

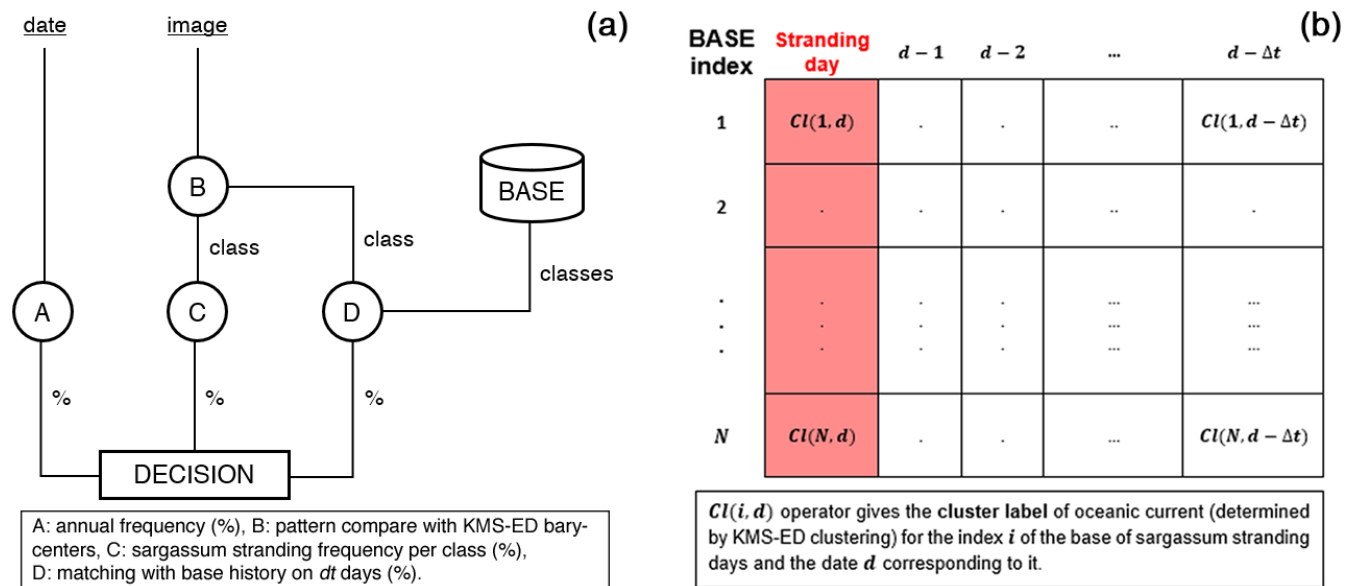
565 Witherington, B., Hiram S. et Hardy, R.: Jeunes tortues marines de la communauté de dérive pélagique dominée par les Sargasses: utilisation de l'habitat, densité de population et menaces, *Mar. Ecol. Programme. Ser.*, 463, 1-22, <https://doi.org/10.3354/meps09970>, 2012.

Zhao, Q., Basher, Z. et Costello, MJ : Cartographie des écosystèmes marins mondiaux près de la surface grâce à l'analyse par grappes de données environnementales. *Recherche écologique.*; 35 : 327–342. <https://doi.org/10.1111/1440-1703.12060>, 2020.



570

Figure 1 : (a) Principaux courants océaniques se produisant et interagissant dans les régions de l'Atlantique central et des Petites Antilles ; Courant des Caraïbes (CC), courant nord-équatorial (NEC), courant nord-brésilien (NBC), contre-courant nord-équatorial (NECC), courant sud-équatorial (SEC). Domaine des Petites Antilles (LA) : le rectangle rouge correspond à la zone d'étude (55-66° W, 8-17° N) ; (b) Subdivision spatiale de la zone d'étude en trois sous-zones : LA1 (c'est-à-dire la mer des Caraïbes), LA2 (c'est-à-dire l'Atlantique nord-tropical au-dessus de la Barbade (13,2° N)) et LA3 (c'est-à-dire l'Atlantique nord tropical en dessous de 13,2° N).



580

Figure 2 : (a) Schéma du classificateur d'arbre de décision pour prédire la probabilité d'échouement de Sargassum. (b) Base de combinaison d'étiquettes de clusters de courants océaniques obtenues par KMS-ED de chaque jour d'échouage à  $\Delta t$  jours avant.



585

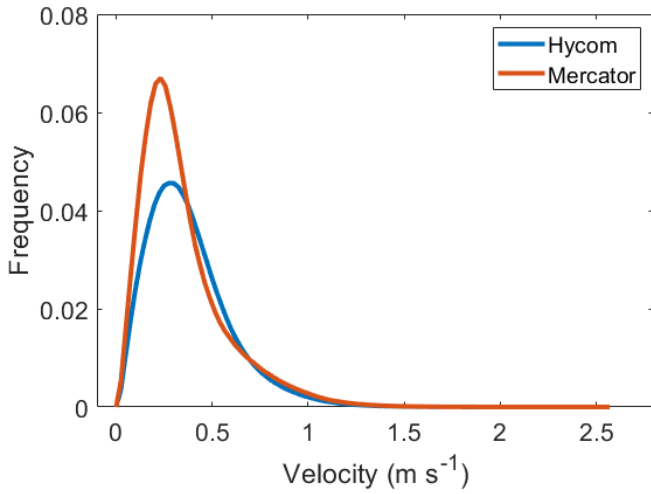
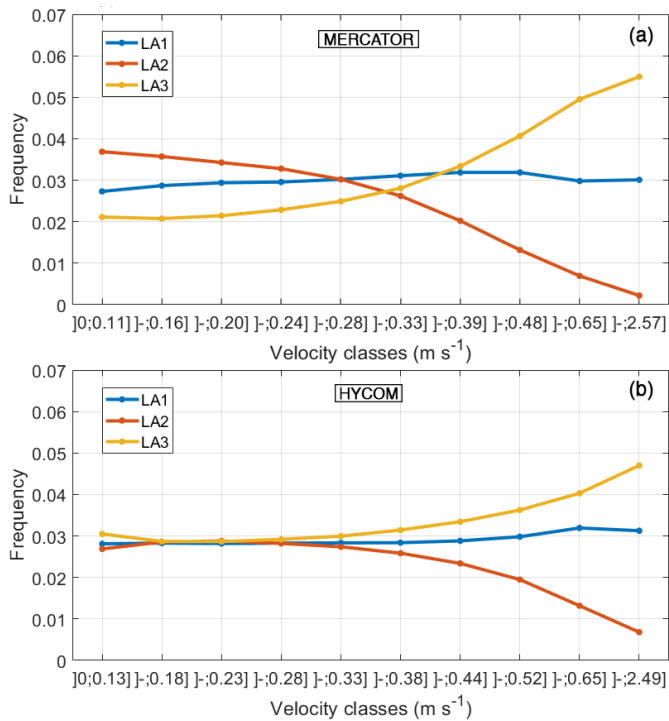


Figure 3 : Distributions des courants de surface océaniques, y compris la dérive pour les deux modèles, ensembles de données HYCOM (bleu) et Mercator (rouge).



590 Figure 4 : Distribution relative des fréquences des vitesses des courants pour les trois sous-régions hauturières autour des Petites Antilles (2019-2020), LA1 (bleu), LA2 (rouge), LA3 (jaune). (a) Mercator avec dérive ERA-5 et (b) HYCOM avec dérive ERA-5.

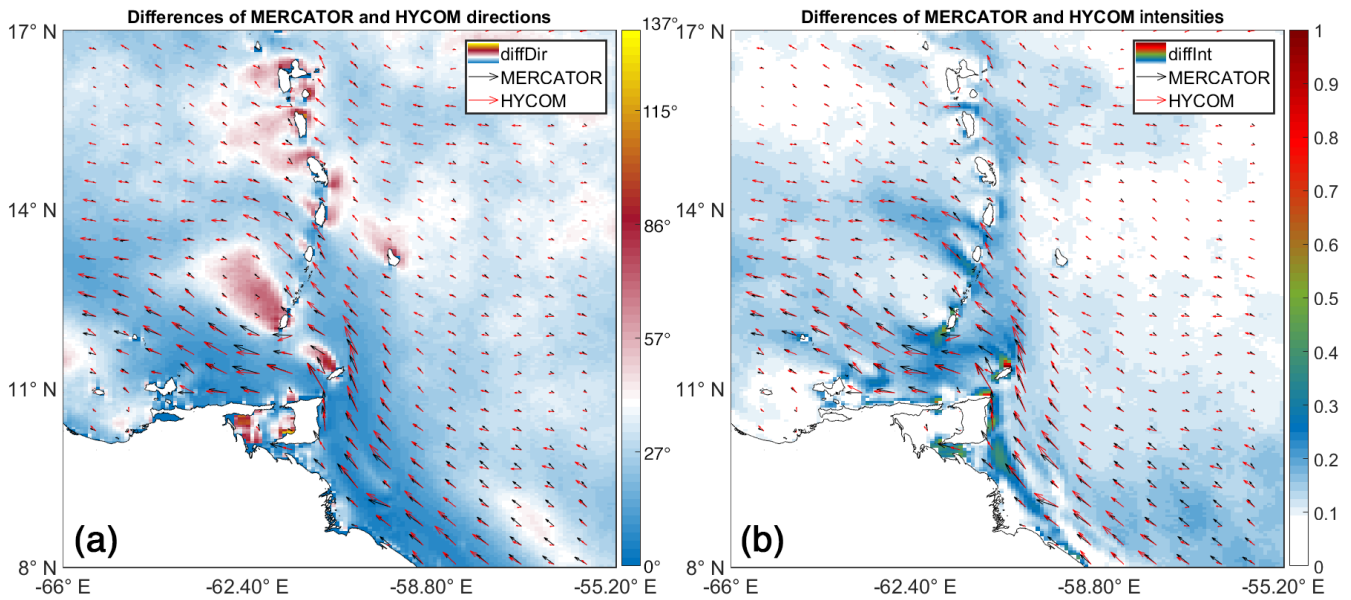


Figure 5 : Comparaison entre les courants de surface Mercator et HYCOM de 2019 à 2020 : différences médianes de direction du courant en degré (a), différences médianes d'intensité du courant en ms<sup>-1</sup>(b).

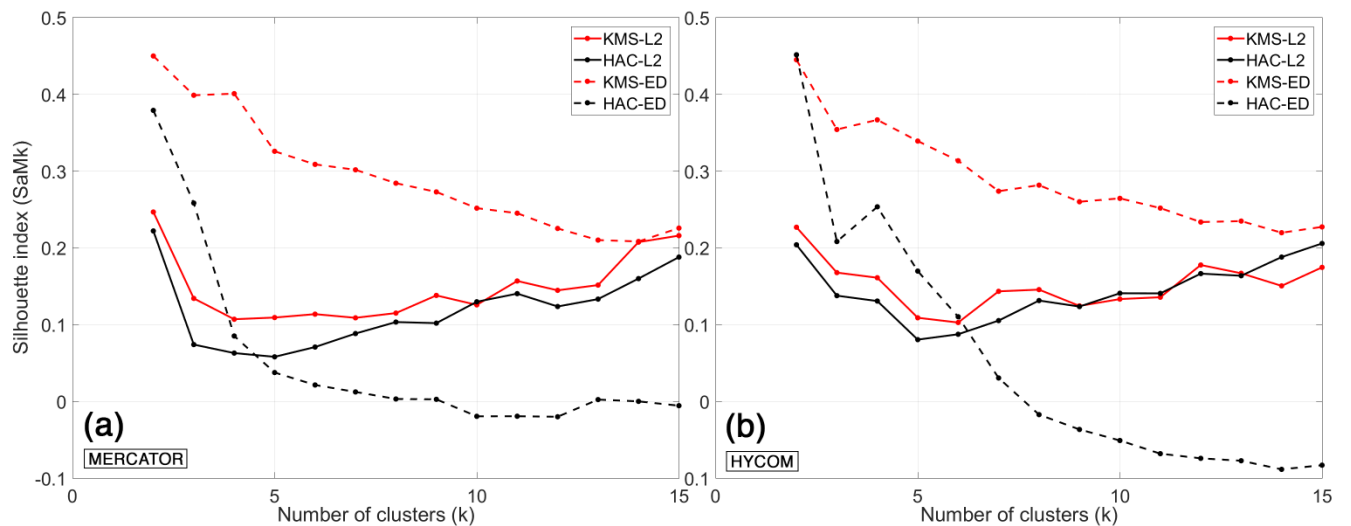


Figure 6 : Évolution de l'indice de silhouette SaMik (par méthode) en fonction du nombre de clusters k, Mercator (a) et HYCOM (b) : méthode HAC (noir), méthode KMS (rouge), avec métrique L2 (plein ligne) et la métrique ED (ligne pointillée).



605

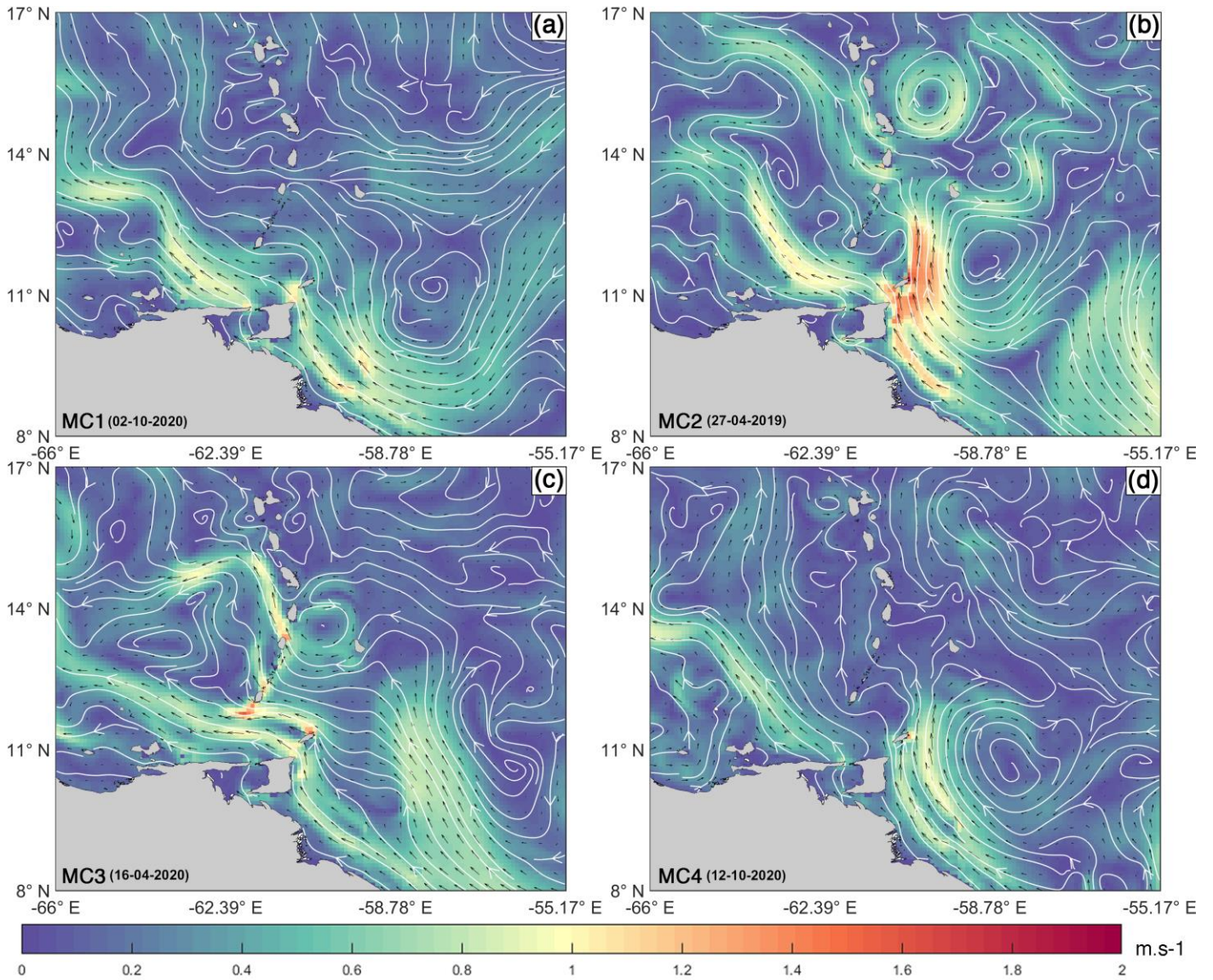
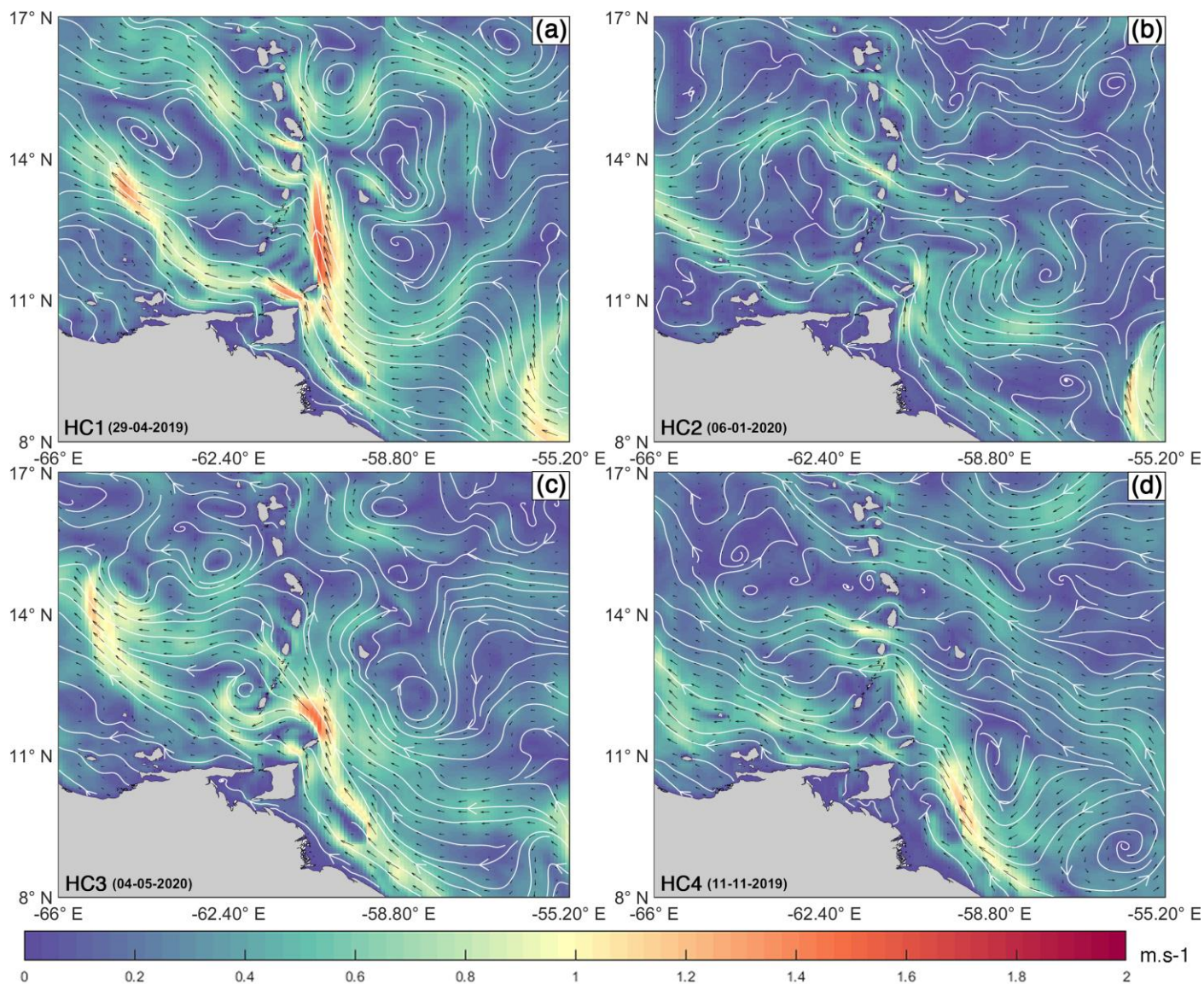
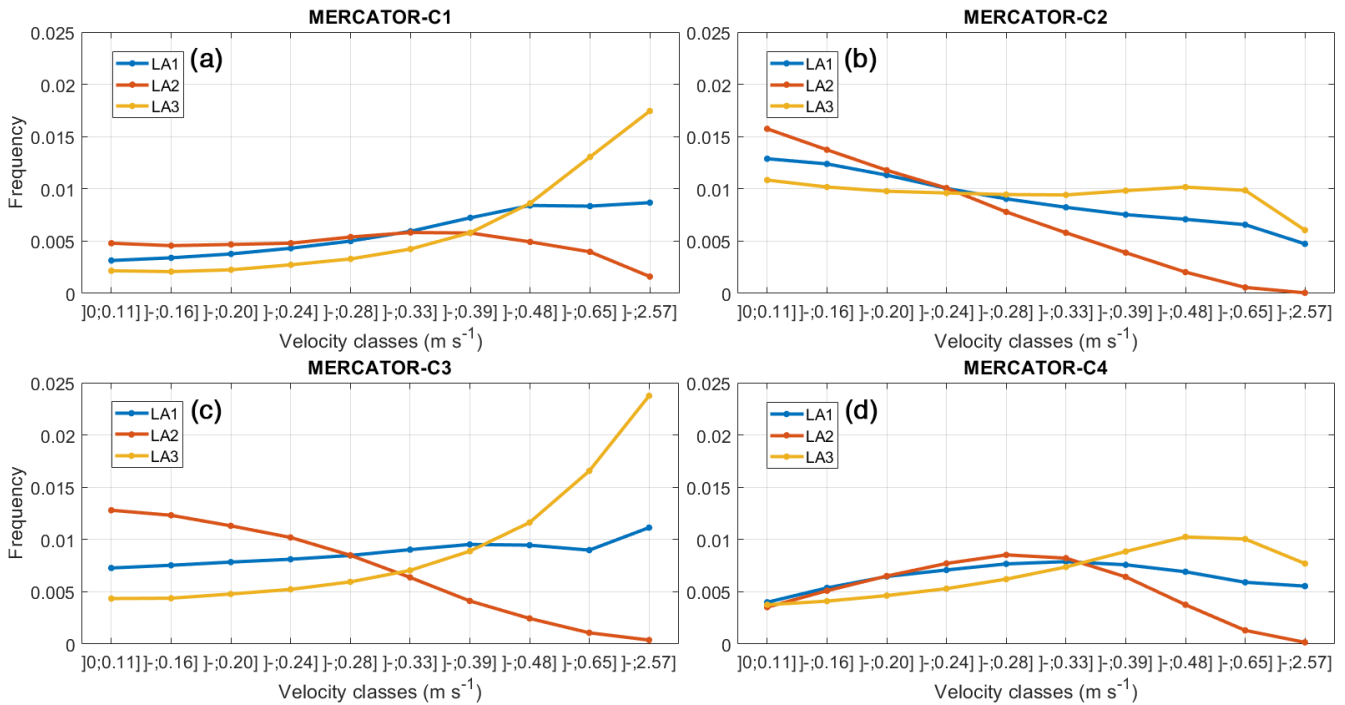


Figure 7 : Éléments représentatifs des clusters issus des données de courant Mercator combinées au vent ERA-5 (méthode KMS-ED avec  $k = 4$ ) : MC1 (jour 02-10-2020) (a), MC2 (jour 27-04-2019) (b), MC3 (jour 16-04-2020) (c), MC4 (jour 12-10-2020) (d).

610



615 Figure 8 : Éléments représentatifs des clusters issus des données de courant HYCOM combinées au vent ERA-5 (méthode KMS-ED avec  $k = 4$ ) : HC1 (jour 29-04-2019) (a), HC2 (jour 06-01-2020) (b), HC3 (jour 04-05-2020) (c), HC4 (jour 11-11-2019) (d).



**Figure 9 :** Distribution relative des fréquences des vitesses des courants pour les trois sous-régions hauturières : MC1 (a), MC2 (b), MC3 (c) et MC4 (d). Les éléments représentatifs ont été obtenus après regroupement KMS-ED pour Mercator.



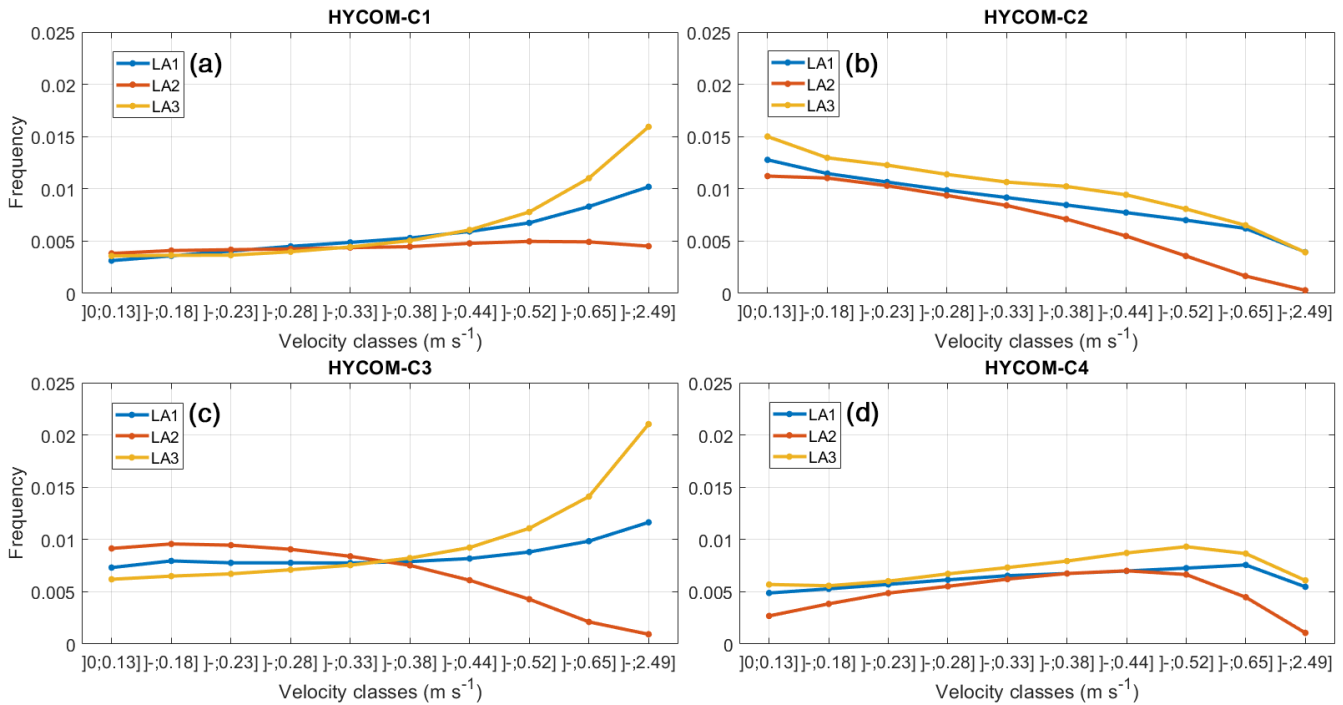
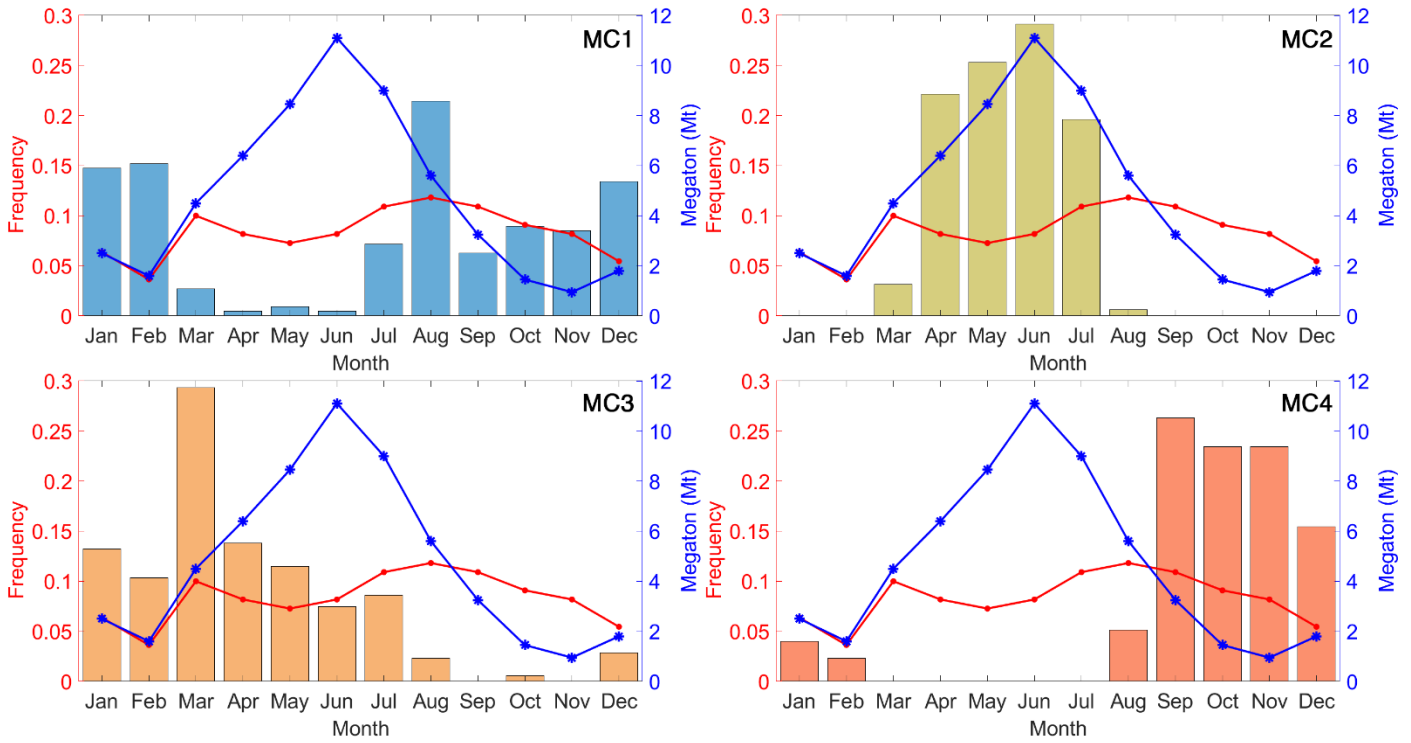


Figure 10 : Distribution relative des fréquences des vitesses des courants pour les trois sous-régions offshore : HYCOM-C1 (a), HYCOM-C2 (b), HYCOM-C3 (c) et HYCOM-C4 (d). Les éléments représentatifs ont été obtenus après regroupement KMS-ED pour HYCOM.



625

Figure 11 : Répartition mensuelle de l'occurrence des clusters à partir des sorties Mercator, de 2019 à 2020, aux Petites Antilles (55-66°W, 8-17°N) : MC1 (a), MC2 (b), MC3 (c) et MC4 (d). La ligne rouge montre la répartition mensuelle des échouages de Sargasses sur les côtes de la Guadeloupe durant la même période. La ligne bleue indique la distribution mensuelle de l'abondance des sargasses dans l'océan Atlantique et la mer des Caraïbes (en mégatonne).

630

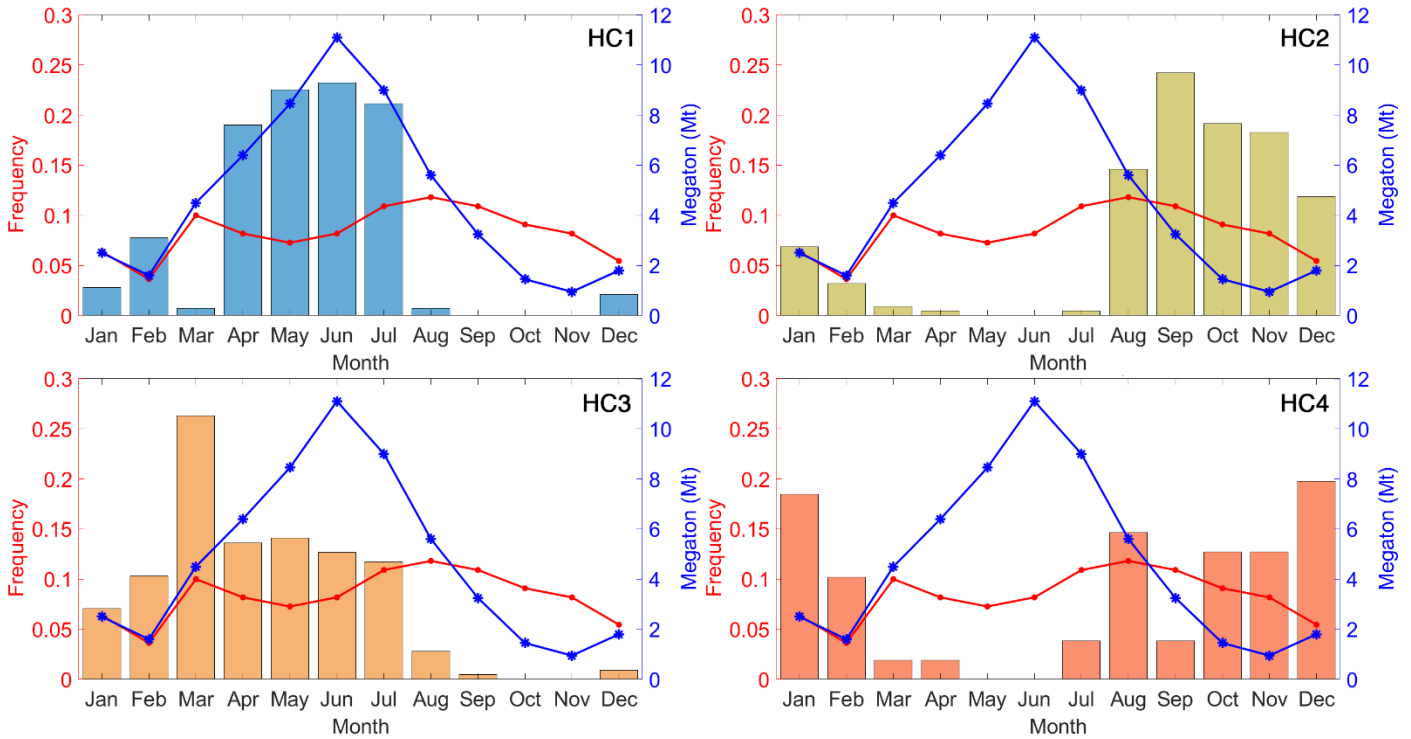
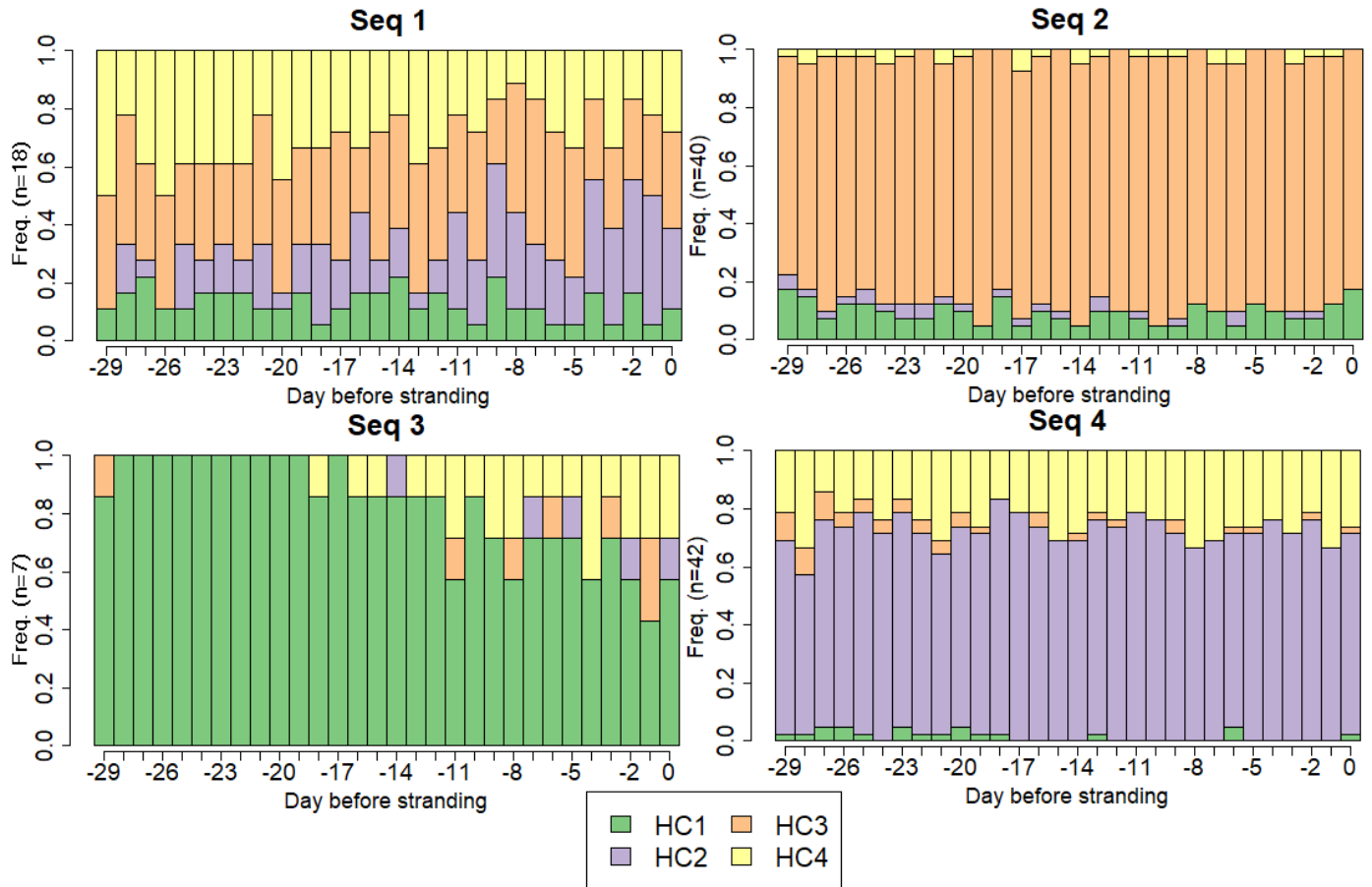
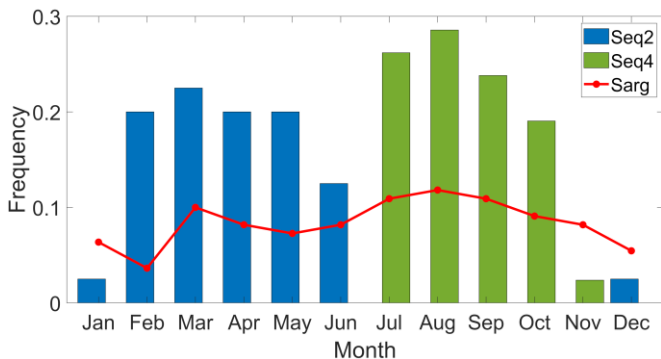


Figure 12 : Répartition mensuelle de l'occurrence des clusters à partir des sorties HYCOM, de 2019 à 2020, dans les Petites Antilles (55-66°W, 8-17°N) : HC1 (a), HC2 (b), HC3 (c) et HC4 (d). La ligne rouge montre la répartition mensuelle des échouages de Sargasses sur les côtes de la Guadeloupe durant la même période. La ligne bleue indique la distribution mensuelle de l'abondance des sargasses dans l'océan Atlantique et la mer des Caraïbes (en mégatonne).

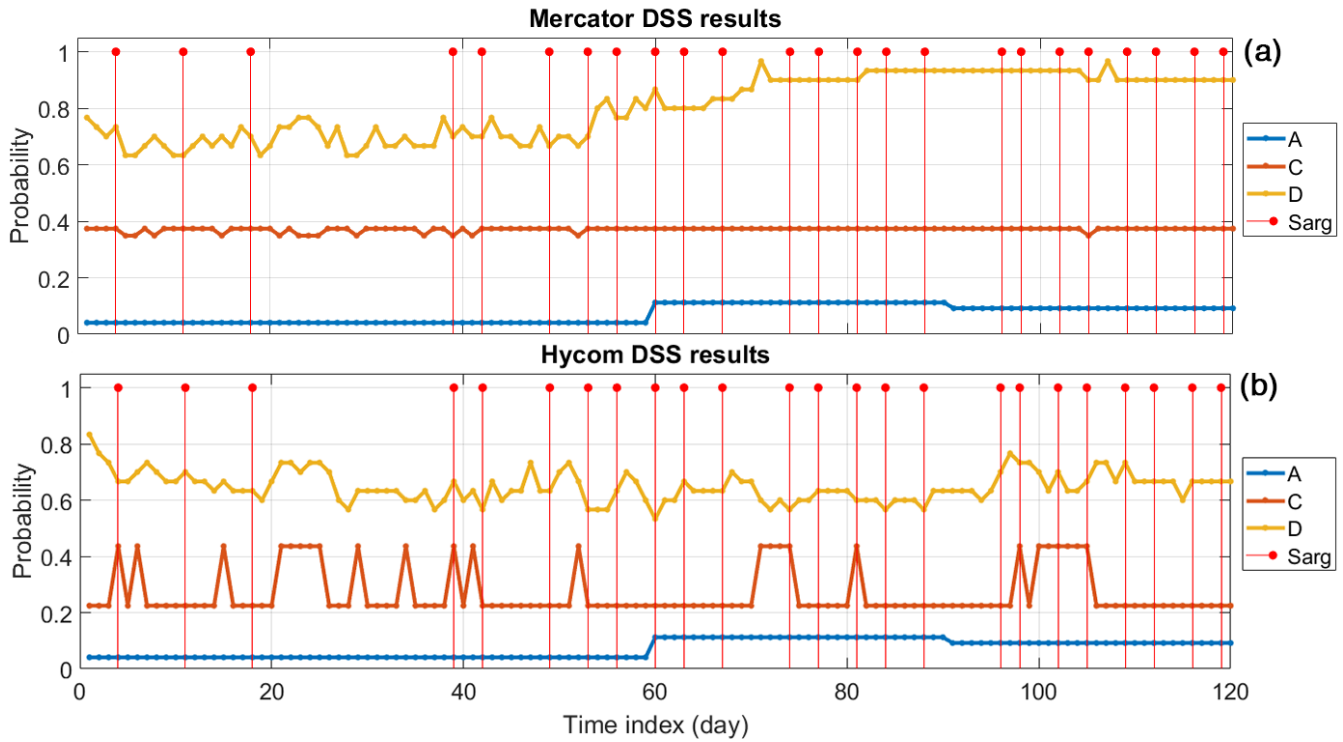
635



640 **Figure 13 :** Répartition des régimes de courant sur les séquences rétrogrades d'échouage de 30 jours : HC1 en vert, HC2 en violet, HC3 en orange et HC4 en jaune.



645 **Figure 14 :** Répartition mensuelle des principales séquences rétrogrades d'échouage observées : Seq2 (bleu) et Seq4 (vert). La ligne rouge représente la distribution des jours d'échouage observés.



650 Figure 15 : Résultats du système d'aide à la décision (DSS) : probabilité d'échouage obtenue par module. Fréquence d'échouage mensuelle obtenue pour le module A (ligne bleue), fréquence d'échouage par cluster pour le module C (ligne rouge), pourcentage de correspondance pour le module D (ligne jaune). Jour d'échouage observé sur les côtes guadeloupéennes (points rouges) : Mercator (a) et HYCOM (b).

650

655



Déciles (D <sub>je</sub> )	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	Max	moyenne	Sigma
Mercator (ms <sup>-1</sup> )	0,11	0,16	0,20	0,24	0,28	0,32	0,39	0,48	0,65	2,57	0,33	0,22
HYCOM (ms <sup>-1</sup> )	0,13	0,18	0,23	0,28	0,32	0,38	0,44	0,52	0,65	2,49	0,36	0,21

660

Tableau 1 : Limites des classes d'histogramme utilisées pour quantifier les données de vitesse des courants de surface avec Sigma comme écart-type.

Jeux de données	C1	C2	C3	C4
<b>MERCATOR</b>	<b>224</b> (30,7%)	158 (21,6%)	174 (23,8%)	175 (23,9%)
<b>HYCOM</b>	142 (19,4%)	<b>219</b> (29,9%)	213 (29,1%)	157 (21,5%)

665Tableau 2 : Nombre de jours correspondant à chaque cluster pour les jeux de données MERCATOR et HYCOM.

		HYCOM			
		C1	C2	C3	C4
MERCATOR	C1	8,3 %	9,6 %	7,1 %	<b>50,6 %</b>
	C2	<b>60,4 %</b>	(-)	12,4 %	1,3 %
	C3	0,3 %	4,8 %	<b>56,7 %</b>	4,7 %
	C4	(-)	<b>69,8 %</b>	0,8 %	3,1 %

670Tableau 3 : Tableau de correspondance entre les 4 clusters générés avec les jeux de données MERCATOR et HYCOM, le pourcentage exprime la proportion de jours communs entre deux clusters ((-) pour 0%).

Jeux de données	C1	C2	C3	C4
<b>MERCATOR</b>	33	15	34	28
<b>HYCOM</b>	14	35	40	21

Tableau 4 : Répartition des jours d'échouage de Sargasses observés (côtes de la Guadeloupe) dans les clusters MERCATOR et HYCOM.

675



Afficher les statistiques de publication

En arrière séquence (HYCOM)	Seq1	Seq2	Seq3	Seq4
<small>Des classes</small>				
<b>n</b>	18	40	7	42
<b>%</b>	16.8	37,4	6.5	39.3

Tableau 5 : Répartition des classes de séquence vers l'arrière.

680

Jeux de données	TP	FN	TN	PF	Précision
Mercator (SOMME)	9	15	52	44	61
<b>Mercator (%)</b>	37,5	62,5	54,2	45,8	<b>50,8</b>
HYCOM (SOMME)	dix	14	78	18	88
<b>HYCOM (%)</b>	41,7	58,3	81,3	18,8	<b>73,3</b>

Tableau 6 : Précision de l'arbre de décision : vrai positif (TP), faux négatif (FN), vrai négatif (TN) et faux positif (FP).