

Apport des copules pour améliorer la détection de cibles à l'aide d'un radar polarimétrique

Roxane LEDUC^{1,3} Audrey GIREMUS¹ Eric GRIVEL² Adrien GILLIOT³ Yoan VEYRAC³ Vincent CORRETJA³

¹Université de Bordeaux, UMR CNRS 5218, IMS, 351 cours de la libération, Talence, France

²Bordeaux INP, UMR CNRS 5218, IMS, 351 cours de la libération, Talence, France

³Thales Defense Mission System France SAS, Campus Thales, 75-77 Av. Marcel Dassault, Merignac, France

Résumé – Les radars polarimétriques sont des dispositifs émergents qui ouvrent des perspectives prometteuses, notamment en matière de détection de cibles maritimes. Pour concevoir une solution tirant parti de ces nouvelles capacités, il est essentiel de développer un cadre théorique mettant à profit la diversité des mesures polarimétriques. Comme ces dernières offrent plusieurs perspectives d'une même scène, leur corrélation doit être modélisée de manière appropriée. Dans cette communication, nous proposons l'utilisation de la théorie des copules pour capturer et quantifier ces dépendances, permettant ainsi d'améliorer les performances de détection de cibles.

Abstract – Polarimetric radars are emerging devices that open up promising prospects, particularly for maritime target detection. To design a solution that takes advantage of these new capabilities, it is essential to develop a theoretical framework that integrates the diversity of polarimetric measurements. The latter offer several perspectives of the same scene, their correlation needs to be appropriately modeled. In this paper, we propose the use of copula theory to capture and quantify these dependencies, enabling improved target detection performance.

1 Introduction

Le traitement statistique du signal radar en environnement maritime aborde plusieurs problématiques, notamment la caractérisation et la modélisation des perturbations telles que le fouillis de mer, la mise en œuvre de tests de détection, ainsi que l'estimation des paramètres de localisation des objets détectés. Selon les cas, exploiter des informations variées pour une même situation est soit nécessaire pour la détection et l'estimation, soit un atout permettant d'en améliorer les performances. Traditionnellement, cette diversité est obtenue par le biais de différentes dimensions, notamment spatiales et fréquentielles. Dans cette communication, nous explorons un autre type de modalité lié à la polarisation : le choix de la polarisation des ondes électromagnétiques influence significativement la réflexion du signal radar par les objets. En effet, le fouillis de mer ne renvoie pas nécessairement des échos de même puissance selon l'orientation du champ électromagnétique, particulièrement autour de l'angle de Brewster [9].

Pour les générations passées et actuelles de radars de détection aéroportés, la polarisation, généralement horizontale (HH)¹ ou verticale (VV), est déterminée lors de la phase de conception. De ce fait, elle n'est pas nécessairement la plus adaptée au scénario d'emploi du radar. À l'avenir, des antennes et des chaînes d'émission/réception permettant de changer le plan de polarisation dynamiquement pourront être utilisées.

Dans cette communication, nous nous plaçons dans ces nouvelles conditions opérationnelles. La problématique réside alors dans la manière de traiter les données reçues afin de détecter la présence d'objets. Peu de recherches ont été réalisées sur ce sujet, mis à part en imagerie, où la polarimétrie

facilite la segmentation des images SAR [5]. Divers modèles stochastiques et physiques ont été développés pour représenter les échos radar en mer, couvrant de larges plages de géométrie d'observation et d'état de mer pour un type de polarisation d'émission-réception donné [7]. Cependant, à notre connaissance, leur utilisation conjointe a été peu explorée principalement en raison du manque de données de fouillis de mer acquises par des radars polarimétriques.

En modifiant la polarisation du radar au cours du temps, une approche pourrait consister à traiter les données issues d'une polarisation HH et d'une polarisation VV indépendamment. La problématique serait alors de combiner les décisions des tests de détection. Une alternative est de traiter ces données conjointement. Dans ce cas, il apparaît nécessaire de considérer la dépendance entre les mesures des plans de polarisation horizontal et vertical². Pour ce faire, nous proposons d'évaluer l'apport d'un outil statistique, les copules, qui établissent une loi conjointe entre plusieurs variables aléatoires. Elles structurent le type et le niveau de dépendance entre ces variables et permettent notamment de séparer les lois marginales de la structure de dépendance. À ce stade, se pose la question de la sélection de la copule la plus adaptée pour représenter la dépendance entre les mesures disponibles. Dans cette communication, nous proposons de nous appuyer sur le critère d'information d'Akaike (AIC). Par la suite, nous évaluons l'impact de la prise en compte de la dépendance dans un test de détection fondé sur le rapport de vraisemblance généralisé (GLRT en anglais).

Cette communication est organisée comme suit : la section 2 traite de la théorie des copules. Les sections 3 et 4 sont respectivement dédiées à l'AIC et au test de détection. Conclusions et perspectives sont présentées en section 5.

¹Dans ce cas l'onde électromagnétique émise par le radar et l'onde réfléchie reçue par l'antenne sont toutes deux polarisées horizontalement.

²Le radar est ici supposé non-cohérent.

2 Théorie des copules

Après avoir défini les copules bivariées et rappelé quelques familles bien connues, nous nous intéressons aux méthodes d'estimation des paramètres qui les caractérisent.

2.1 Introduction sur les copules bivariées

De manière pratique, on appelle copule bivariée, toute fonction de répartition bivariée C ayant pour marginales des lois uniformes sur $[0, 1]$.

Soit $Y = (Y_1, Y_2)$ un vecteur aléatoire prenant ses valeurs dans $\mathbb{R}^2$. On note f sa densité de probabilité. Considérons F_1, F_2 et f_1, f_2 , les fonctions de répartition et densités de probabilité de Y_1 et Y_2 respectivement. Si les marginales F_1 et F_2 sont continues, le théorème de Sklar [6] stipule qu'il existe une unique copule $C : [0, 1]^2 \rightarrow [0, 1]$, telle que,

$$\forall (y_1, y_2) \in \mathbb{R}^2 :$$

$$f(y_1, y_2) = f_1(y_1)f_2(y_2)c(F_1(y_1), F_2(y_2)) \quad (1)$$

avec $c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}$ la densité de la copule. Dans le cas où les variables aléatoires Y_1 et Y_2 sont indépendantes, la densité de la copule dans (1) est égale à un.

2.2 Familles de copules paramétriques

Un large éventail de copules peut être envisagé pour définir des relations de dépendance [6]. Certaines d'entre elles sont construites à partir de lois multivariées existantes. C'est le cas de la famille des copules elliptiques, qui inclut :

- la copule gaussienne, qui ne présente pas de dépendance de queue. Elle est définie à partir de la distribution normale bivariée avec une matrice de corrélation associée au coefficient de corrélation de Pearson ρ .
- la copule de Student, qui présente au contraire des queues de dépendance lourdes. Construite de manière similaire à la copule gaussienne, elle est définie à partir de la distribution de Student bivariée, avec une matrice de corrélation associée à ρ et un degré de liberté ν . Lorsque ν tend vers l'infini, elle tend vers la copule gaussienne.

Contrairement aux copules elliptiques, la famille des copules dites « archimédiennes » a l'avantage de décrire des structures de dépendance très diverses, notamment asymétriques, pour lesquelles les coefficients de queues inférieures et supérieures diffèrent. Une copule est dite archimédienne de fonction génératrice $\phi : [0, 1] \rightarrow [0, +\infty]$, de classe $\mathcal{C}^2$, si elle s'exprime sous la forme suivante :

$$C(u_1, u_2) = \begin{cases} \phi^{-1}(\phi(u_1) + \phi(u_2)) & \text{si } \phi(u_1) + \phi(u_2) \leq \phi(0) \\ 0 & \text{sinon.} \end{cases} \quad (2)$$

avec $\phi(1) = 0$, $\phi'(u) < 0$ et $\phi''(u) > 0$, $\forall u \in [0, 1]$.

Cette famille de copules inclut les copules de Gumbel, Clayton et Frank, modélisant respectivement une dépendance de queue supérieure, inférieure et symétrique. Chacune possède sa propre fonction génératrice ϕ et dépend d'un paramètre θ . Cf. Table 1. Un paramètre θ plus élevé (en valeur absolue) correspond généralement à une dépendance plus forte, tandis qu'un paramètre plus proche de sa borne inférieure indique une faible dépendance. Il existe une relation explicite entre ces paramètres θ et le tau de Kendall τ_K , qui est une mesure d'association permettant de quantifier des relations non nécessairement linéaires entre des variables et ce, indépendamment du modèle de copule choisi [6].

TABLE 1 : Copules archimédiennes

Copule	$\phi(u)$	$C(u_1, u_2)$
Clayton ($\theta > 0$)	$u^{-\theta} - 1$	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$
Gumbel ($\theta \geq 1$)	$(-\ln(u))^\theta$	$e^{-((-\ln(u_1))^\theta + (-\ln(u_2))^\theta)^{\frac{1}{\theta}}}$
Frank ($\theta \neq 0$)	$-\ln(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1})$	$-\frac{1}{\theta} \ln(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1})$

2.3 Estimation des paramètres

Même si le type de dépendance entre deux variables et donc l'expression de la copule est connue, la valeur des paramètres de leur distribution doit être ajustée en pratique. Dans cette section, nous rappelons plusieurs méthodes d'estimation appliquées à notre cas d'étude, en distinguant les approches paramétriques des approches semi-paramétriques [4].

Par la suite, pour une impulsion radar, nous noterons $\{(y_{1i}, y_{2i})\}_{i=1, \dots, n}$, les n couples de mesures de puissance collectées le long des cases distances, provenant respectivement des plans de polarisation HH et VV. Leurs fonctions de répartition et densités de probabilités seront notées F_1 et f_1 (respectivement F_2 et f_2). Dans cette étude, nous travaillons en basse résolution distance. Les couples sont donc supposés mutuellement indépendants.

- L'approche paramétrique impose un modèle spécifique pour la copule ainsi que pour les distributions marginales. La méthode FML (*Full Maximum Likelihood*) estime conjointement les paramètres des distributions marginales, notées β_i avec $i = 1, 2$, et ceux de la copule, notées Θ . Pour cela, on maximise la fonction de log-vraisemblance donnée par :

$$l(\Theta, \beta_1, \beta_2) = \sum_{i=1}^n \ln(c(F_1(y_{1i}, \beta_1), F_2(y_{2i}, \beta_2), \Theta)) + \ln(f_1(y_{1i}, \beta_1)) + \ln(f_2(y_{2i}, \beta_2)) \quad (3)$$

Cette méthode, asymptotiquement efficace au sens de la borne de Cramér-Rao, nécessite cependant des calculs intensifs, rendant parfois la résolution du problème d'optimisation difficile. Une procédure alternative en deux étapes, sous-optimale mais moins calculatoire, appelée IFM (*Inference Functions for Margins*), a alors été proposée. Dans un premier temps, les paramètres des distributions marginales sont estimés :

$$\hat{\beta}_j = \max_{\beta_j} \sum_{i=1}^n \ln(f_j(y_{ji}, \beta_j)) \quad (4)$$

Puis, on déduit le vecteur des paramètres de la copule, compte tenu des estimations de l'étape précédente :

$$\hat{\Theta} = \max_{\Theta} \sum_{i=1}^n \ln(c(F_1(y_{1i}, \hat{\beta}_1), F_2(y_{2i}, \hat{\beta}_2), \Theta)) \quad (5)$$

Tout comme la FML, l'inconvénient de cette méthode est qu'elle suppose de connaître l'expression de la distribution marginale.

- L'approche semi-paramétrique suppose quant à elle un modèle paramétrique pour la copule et non-paramétrique pour les

distributions marginales. La méthode CML (*Canonical Maximum Likelihood*) est semi-paramétrique avec les distributions marginales estimées par la distribution empirique suivante :

$$\hat{F}_j(y_{ji}) = \frac{1}{n} \sum_{l=1}^n \mathbb{1}_{y_{jl} \leq y_{ji}} \quad (6)$$

avec $\mathbb{1}$, la fonction indicatrice. La méthode du maximum de vraisemblance est alors utilisée pour estimer uniquement Θ .

3 Sélection de la copule

Notre objectif est à présent de proposer une méthode permettant de sélectionner une copule paramétrique caractérisant au mieux la dépendance entre les plans de polarisation horizontal et vertical. À cet effet, nous proposons d'appliquer l'AIC. Étant donné que la structure de dépendance peut ne correspondre exactement à aucune des copules testées, nous appliquons également le test de Cramér-Von-Mises pour comparer la copule choisie à la copule empirique issue des données observées.

3.1 Sélection de la copule par critère d'Akaike

Le critère AIC permet de choisir le modèle de probabilité le plus adapté pour décrire des données parmi un ensemble de modèles candidats. Il favorise la parcimonie en trouvant un compromis entre la qualité de l'ajustement et la simplicité du modèle. Pour le mettre en œuvre, il est nécessaire d'estimer les paramètres au sens du maximum de vraisemblance. Nous avons opté pour l'approche semi-paramétrique CML (Cf. Section 2.3) afin de rendre notre démarche la plus générique possible. Dans ce cas, le critère associé à chaque modèle peut s'exprimer comme suit :

$$AIC = -2 \sum_{i=1}^n \ln(c(\hat{F}_1(y_{1i}), \hat{F}_2(y_{2i}), \hat{\Theta})) + 2k \quad (7)$$

avec k le nombre de paramètres de la copule étudiée. Les marginales étant identiques, on remarque que seule la copule intervient dans l'équation (7). Le modèle retenu minimise (7). Afin de nous assurer que la copule choisie représente fidèlement la structure de dépendance observée, nous procédons à la validation ou au rejet de l'hypothèse H_0 , qui stipule que la structure de dépendance d'une distribution bivariable est correctement représentée par une famille paramétrique spécifique de copules, en utilisant le test de Cramér-Von Mises [1]. Cette approche est fondée sur une estimation empirique de la copule. Nous proposons donc de comparer cette estimation à la copule théorique déterminée par le critère AIC, en évaluant la distance quadratique suivante :

$$\begin{aligned} T &= n \int_{[0,1]^2} (\hat{C}(u_1, u_2) - C_{\hat{\theta}}(u_1, u_2))^2 d\hat{C}(u_1, u_2) \\ &= \sum_{i=1}^n (\hat{C}(u_{1i}, u_{2i}) - C_{\hat{\theta}}(u_{1i}, u_{2i}))^2 \end{aligned} \quad (8)$$

avec $\hat{C}(u_{1i}, u_{2i}) = \frac{1}{n} \sum_{l=1}^n \mathbb{1}_{u_{1l} \leq u_{1i}, u_{2l} \leq u_{2i}}$, la copule empirique, $C_{\hat{\theta}}$ la copule théorique déterminée par AIC et $u_{jl} = \hat{F}_j(y_{jl})$, $\forall j = 1, 2$, en accord avec l'équation (6).

3.2 Étude expérimentale

N'ayant pas de données réelles à notre disposition pour le moment, nous commençons par générer des données synthétiques

en utilisant la méthode d'inversion [6]. La simulation d'un vecteur (Y_1, Y_2) , de marginales F_1 et F_2 ayant une structure de dépendance bivariable fondée sur une copule C entre les plans de polarisation, se fait en deux étapes :

1/ les variables U_1 et U_2 , uniformément réparties sur $[0, 1]$, sont générées en utilisant la copule C .

2/ les variables $Y_1 = F_1^{-1}(U_1)$ et $Y_2 = F_2^{-1}(U_2)$, qui possèdent les marginales souhaitées tout en étant corrélées selon la copule choisie, sont obtenues.

Il existe différentes méthodes permettant de générer cette paire de variables uniformes [6]. Par exemple, pour des copules de Clayton ou de Frank, on peut utiliser l'inverse de la fonction de répartition conditionnelle. Pour les copules elliptiques, on peut directement utiliser leurs distributions bivariées. La méthode de Marshall-Olkin [3] est couramment employée dans le cas de la copule de Gumbel.

Dans le cadre de la détection des cibles maritimes, le fouillis de mer est souvent modélisé par une loi K, une distribution composée qui présente l'intérêt d'être applicable aux deux plans de polarisation [7]. Elle dépend d'un paramètre de forme noté ν_1 (respectivement ν_2 , selon la polarisation) et d'une puissance moyenne notée μ_1 (respectivement μ_2).

Nous avons mené des simulations pour différentes valeurs de n (entre 100 et 1000). On peut noter que pour cette plage de valeurs, le test de Cramér-Von-Mises a systématiquement validé les choix effectués par l'AIC. Des matrices de confusion sont présentées en Figure 1 pour $n = 100$ et 500 couples d'observations. Les paramètres des marginales ont été fixés à $\nu_1 = 1.4$, $\nu_2 = 9.7$, $\mu_1 = 7.4 \times 10^{-12}W$ et $\mu_2 = 1,1 \times 10^{-11}W$. On peut constater qu'en dessous d'une centaine d'échantillons, les performances de l'algorithme de sélection deviennent moins fiables, en particulier pour les copules elliptiques.

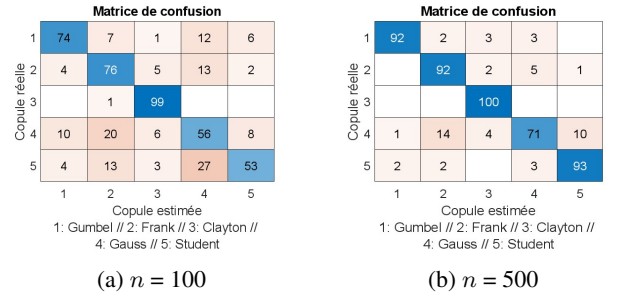


FIGURE 1 : Analyse des performances pour une modélisation par loi K, effectuée à partir de 100 simulations par copule, avec des paramètres sélectionnés aléatoirement dans leurs plages de données respectives.

4 Test de détection

Après avoir proposé une approche permettant d'estimer le type et le niveau de dépendance entre les données polarimétriques (Cf. Section 3), nous déterminons dans cette section le gain apporté par la prise en compte de cette dépendance dans un test de détection.

4.1 Approche par GLRT

Un test classiquement utilisé en détection radar est le GLRT [2]. Il s'agit d'un test composite, qui permet de décider entre

deux hypothèses de distribution de probabilité dépendant de paramètres inconnus. Appliqué à une case distance, il vise à décider entre les deux hypothèses :

- H_0 , présence seule de fouillis,
- H_1 , présence de cible en plus du fouillis.

Il procède en comparant une statistique T dépendant des données à un seuil γ , obtenu à partir d'une probabilité de fausse alarme $P_{fa} = \mathbb{P}(T > \gamma | H_0)$ souhaitée. La mesure de performance d'une statistique se fait enfin par calcul de la probabilité de détection $P_d = \mathbb{P}(T > \gamma | H_1)$.

Nous proposons de comparer le GLRT classique à une variante prenant en compte la dépendance entre les plans de polarisation au moyen d'une copule. La statistique de test peut être exprimée, pour chaque couple de données $\{(y_{1i}, y_{2i})\}_{i=1, \dots, n}$ de la manière suivante :

$$T(y_{1i}, y_{2i}) = \frac{f(y_{1i}|\hat{\beta}_{11})f(y_{2i}|\hat{\beta}_{21})}{f(y_{1i}|\hat{\beta}_{10})f(y_{2i}|\hat{\beta}_{20})} \times \frac{c(F_1(y_{1i}|\hat{\beta}_{11}), F_2(y_{2i}|\hat{\beta}_{21})|\hat{\Theta}_1)}{c(F_1(y_{1i}|\hat{\beta}_{10}), F_2(y_{2i}|\hat{\beta}_{20})|\hat{\Theta}_0)} \quad (9)$$

L'estimation des paramètres (Cf. Section 2.3) se fait par la méthode IFM lorsque la copule est prise en compte, et par la méthode du maximum de vraisemblance appliqué aux marginales dans le cas contraire. On souhaite ici favoriser la précision afin de chiffrer au mieux le gain possible.

4.2 Étude expérimentale

Dans cette section, les données, générées selon une loi K pour les deux plans de polarisation sous H_0 et selon une loi K+Rayleigh [8] sous H_1 , ont été corrélées à l'aide d'une copule de Gumbel. Pour améliorer les performances de détection, 7 observations de la même case distance sont utilisées, ce qui équivaut sous hypothèse d'indépendance à multiplier leurs variables de test (9). La Figure 2 illustre la comparaison entre les résultats obtenus en intégrant la copule dans le calcul des rapports de vraisemblance et ceux obtenus en ne tenant pas compte de cette intégration (copule indépendante). Cette comparaison est réalisée pour plusieurs valeurs de taux de Kendall τ_K . Les valeurs des paramètres des marginales sont les mêmes qu'en Section 3.2. La puissance moyenne de la cible est de $9.5 \times 10^{-12}W$ pour les deux plans de polarisation (RSB de 3.6dB en HH et 2.7dB en VV). Pour différents seuils, les probabilités de fausse alarme et de détection ont été estimées par méthode de Monte Carlo à partir de 1000 simulations du test.

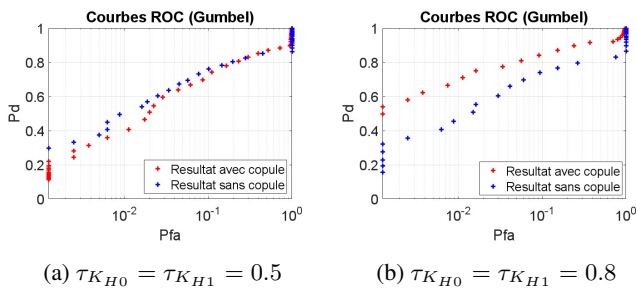


FIGURE 2 : Mesure de performance des statistiques de test avec (croix rouges) et sans (croix bleues) prise en compte de la dépendance entre les plans de polarisation pour différents τ_K .

Ces simulations sur données synthétiques indiquent que l'utilisation de la copule de Gumbel est avantageuse lorsque les taux

de corrélation, mesurés par le coefficient de Kendall, sont élevés. Lorsque τ_K est trop faible (inférieur à 0.6), l'intégration de cette copule dans les calculs tend à dégrader sensiblement les performances du fait de l'ajout de paramètres inconnus. De manière générale, plus la dépendance entre les variables est forte, plus les copules améliorent les tests de détection. Toutefois, pour certaines des copules présentées dans ces travaux, une probabilité de fausse alarme trop faible réduit considérablement l'intérêt de leur utilisation.

5 Conclusions et perspectives

Ces travaux mettent en évidence l'intérêt d'utiliser les copules afin d'améliorer les tests de détection lors de la réception de données polarimétriques, à condition que ces données présentent une corrélation suffisante. Plus largement, on pourrait envisager d'appliquer cette méthodologie à davantage de polarisations (HH, VV, HV, VH) ou encore dans d'autres contextes, par exemple à des mesures provenant de capteurs différents mais observant une même scène (scénario de combat collaboratif). Par ailleurs, une prochaine étape pourrait consister à aborder la problématique des modèles mal spécifiés [10], en évaluant les conséquences de la négligence de la dépendance entre les mesures polarimétriques ou en faisant une estimation incorrecte de cette dépendance en choisissant une copule inadéquate (cas de données réelles).

Références

- [1] C. GENEST, B. RÉMILLARD et D. BEAUDOIN : Goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and economics*, 44(2):199–213, 2009.
- [2] S. HERMAN et J. ASH : A generalized gamma copula model for high resolution polarimetric sar change detection. In *Algorithms for Synthetic Aperture Radar Imagery XXXI*, volume 13032, pages 113–129. SPIE, 2024.
- [3] M. HOFERT : Efficiently sampling nested archimedean copulas. *Computational Statistics & Data Analysis*, 55(1):57–70, 2011.
- [4] G. KIM, M. J. SILVAPULLE et P. SILVAPULLE : Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, 51(6):2836–2850, 2007.
- [5] Z. MA, R. ZHANG et W. YANG : Sea-land segmentation in polarimetric sar images. In *2021 2nd China International SAR Symposium (CISS)*, pages 1–5. IEEE, 2021.
- [6] Roger B NELSEN : *An Introduction to Copulas*. Springer, New York, NY, USA, second édition, 2006.
- [7] L. ROSENBERG et S. WATTS : Continuous sea clutter models for the mean backscatter and k-distribution shape. In *International Conference on Radar Systems (Radar 2017)*, pages 1–6. IET, 2017.
- [8] L. ROSENBERG, S. WATTS et M. S. GRECO : Modeling the statistics of microwave radar sea clutter. *IEEE Aerospace and Electronic Systems Magazine*, 34(10):44–75, 2019.
- [9] D. B. TRIZNA : A model for brewster angle damping and multipath effects on the microwave radar sea echo at low grazing angles. *IEEE Transactions on Geoscience and Remote Sensing*, 35(5):1232–1244, 1997.
- [10] H. WHITE : Maximum likelihood estimation of misspecified models. *Econometrica : Journal of the econometric society*, pages 1–25, 1982.