

Capacités d’imitation de la loi Gamma Généralisée pour la modélisation statistique d’images radar haute résolution.

Hélène SPORTOUCHE, Jean-Marie NICOLAS, Florence TUPIN

Institut Mines-Télécom ; Télécom ParisTech ; CNRS LTCI ;
46 rue Barrault, 75013 Paris, France.

helene.sportouche@telecom-paristech.org, jean-marie.nicolas@telecom-paristech.fr,
florence.tupin@telecom-paristech.fr

Résumé – Dans le cadre de la modélisation en intensité des données radar haute résolution, nous analysons dans ce papier les capacités de la distribution Gamma Généralisée, décrite par 3 paramètres, à imiter d’autres distributions, grâce à sa flexibilité. Nous comparons d’abord son aptitude et celle de la distribution de Fisher à mimer une loi \mathcal{K} de référence, grâce à la méthode des log-cumulants et nous montrons quantitativement la légère prédominance de Gamma Généralisée selon des critères fondés sur les divergences de Kullback-Leibler. Nous exposons ensuite les limites du potentiel d’imitation de la loi Gamma Généralisée et nous démontrons que cette loi peut approximativement mimer une loi Log-Normale, suivant une méthode originale qui identifie la Log-Normale comme la limite asymptotique d’une Gamma Généralisée.

Abstract – In the framework of intensity modeling for high resolution SAR data, we analyze in this paper the competences of the Generalized Gamma distribution, described by 3 parameters, to imitate (or “mimic”) other distributions, thanks to its flexibility. We first compare its ability and the one of the Fisher distribution to mimic a reference \mathcal{K} law, thanks to the method of log-cumulants and we quantitatively show the slight predominance of Generalized Gamma according to criteria based on Kullback-Leibler divergences. We then expose the limits of the imitating potential of Generalized Gamma and we demonstrate that this law is able to approximately mimic a Log-Normal law, through an original method that identifies the Log-Normal as the asymptotic limit of a Generalized Gamma.

1 Introduction

Le choix d’une loi de probabilité appropriée pour la modélisation statistique des images radar est essentiel pour l’application de plusieurs traitements comme le débruitage, la segmentation ou la classification. Plusieurs lois sophistiquées ont été utilisées dans la littérature pour décrire les données radar haute résolution en intensité (module au carré de l’onde électromagnétique rétrodiffusée vers le capteur) et s’affranchir des limites des lois traditionnelles. En particulier, la loi Gamma Généralisée [1] (définie par trois paramètres) a fait l’objet de nombreux travaux (sa flexibilité est démontrée dans [2] sur des extraits homogènes et hétérogènes). Toutefois, celle-ci est-elle suffisamment générique pour s’approprier différents types de comportements et pour pouvoir être recommandée à la place d’un dictionnaire de distributions ?

Dans ces travaux, nous analysons d’abord la capacité de la loi Gamma Généralisée à imiter une loi de référence, choisie ici comme la loi \mathcal{K} [3] (définie aussi par trois paramètres et résultant d’une modélisation en Gamma du speckle et d’une modélisation en Gamma de la texture de la scène), en comparant son potentiel de mime à celui d’une autre loi très connue pour sa souplesse, la loi de Fisher [4] (définie aussi par trois paramètres). Cette étude met en œuvre une méthode d’estimation des paramètres des lois imitatrices fondée sur les log-cumulants et utilise des critères d’évaluation développés à par-

tir des divergences de Kullback-Leibler. Nous exposons ensuite les limites théoriques et pratiques du potentiel de mime de la loi Gamma Généralisée. Nous proposons alors une méthode alternative originale pour l’imitation approximative par celle-ci d’une loi Log-Normale [5] (définie par deux paramètres et dont l’usage s’avère pratique pour convertir un problème de bruit multiplicatif en un problème de bruit additif Gaussien), en identifiant cette dernière comme la limite asymptotique d’une loi Gamma Généralisée, grâce à des propriétés de convergence des fonctions Polygamma.

Ce papier est organisé comme suit : En section 2, les expressions analytiques des lois sont rappelées ; en section 3, l’étude comparative pour le mime d’une loi \mathcal{K} est réalisée ; en section 4, le cas du mime de la Log-Normale est traité ; enfin, en section 5, les principales conclusions sont résumées.

2 Préliminaires

2.1 Présentation des lois

Nous rappelons ici la densité de probabilité (ddp) de la loi Gamma Généralisée et les log-cumulants d’ordre 1 à 3 des lois Gamma Généralisée, Fisher, \mathcal{K} et Log-Normale [6] (par manque de place, nous ne redonnons pas ces trois dernières distributions).

Loi Gamma Généralisée La ddp $\mathcal{G}\mathcal{G}_{pdf}(\mu_{GG}, L_{GG}, \eta_{GG})$ est définie par :

$$\mathcal{G}\mathcal{G}_{pdf}(x) = \frac{|\eta_{GG}| L_{GG}^{\left(\frac{1}{\eta_{GG}}\right)} \Gamma(L_{GG}) \left(\frac{L_{GG}}{\mu_{GG}} x\right)^{\eta_{GG} L_{GG} - 1}}{\mu_{GG} \Gamma(L_{GG})} - \left(\frac{L_{GG}}{\mu_{GG}} x\right)^{\eta_{GG}} \quad (1)$$

où $\Gamma(x)$ est la fonction Gamma, μ_{GG} un paramètre d'échelle, L_{GG} un paramètre de forme et η_{GG} un paramètre de puissance. Notons que pour $\eta = 1$, le premier moment est égal à μ_{GG} .

Ses trois premiers log-cumulants sont définis par :

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{G}\mathcal{G}_{pdf} &= \ln(\mu_{GG}) + \frac{\Psi(L_{GG}) - \ln(L_{GG})}{\eta_{GG}} \\ \widetilde{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf} &= \frac{\Psi(1, L_{GG})}{\eta_{GG}^2} \\ \widetilde{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf} &= \frac{\Psi(2, L_{GG})}{\eta_{GG}^3} \end{aligned} \quad (2)$$

où $\Psi(x)$ et $\Psi(n, x)$ représentent respectivement la fonction Digamma et la fonction Polygamma d'ordre n.

Loi de Fisher Ses trois premiers log-cumulants sont définis par :

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{F}_{pdf} &= \ln(\mu_{\mathcal{F}}) + \Psi(L_{\mathcal{F}}) - \ln(L_{\mathcal{F}}) - \Psi(M_{\mathcal{F}}) + \ln(M_{\mathcal{F}}) \\ \widetilde{\kappa}_2 \mathcal{F}_{pdf} &= \Psi(1, L_{\mathcal{F}}) + \Psi(1, M_{\mathcal{F}}) \\ \widetilde{\kappa}_3 \mathcal{F}_{pdf} &= \Psi(2, L_{\mathcal{F}}) - \Psi(2, M_{\mathcal{F}}) \end{aligned} \quad (3)$$

où $\mu_{\mathcal{F}}$ est un paramètre d'échelle et $L_{\mathcal{F}}$ et $M_{\mathcal{F}}$ sont deux paramètres de forme.

Loi K Ses trois premiers log-cumulants sont définis par :

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{K}_{pdf} &= \ln(\mu_{\mathcal{K}}) + \Psi(L_{\mathcal{K}}) - \ln(L_{\mathcal{K}}) + \Psi(M_{\mathcal{K}}) - \ln(M_{\mathcal{K}}) \\ \widetilde{\kappa}_2 \mathcal{K}_{pdf} &= \Psi(1, L_{\mathcal{K}}) + \Psi(1, M_{\mathcal{K}}) \\ \widetilde{\kappa}_3 \mathcal{K}_{pdf} &= \Psi(2, L_{\mathcal{K}}) + \Psi(2, M_{\mathcal{K}}) \end{aligned} \quad (4)$$

où $\mu_{\mathcal{K}}$ est un paramètre d'échelle et $L_{\mathcal{K}}$ et $M_{\mathcal{K}}$ sont deux paramètres de forme. Notons que le premier moment de la loi K est égal à $\mu_{\mathcal{K}}$. Les cas spéciaux de distributions \mathcal{K} où l'on a l'égalité $L_{\mathcal{K}} = M_{\mathcal{K}}$ correspondent à des distributions appelées « caustiques \mathcal{K}_c » dans ce papier.

Loi Log-Normale Ses trois premiers log-cumulants sont définis par :

$$\begin{aligned} \widetilde{\kappa}_1 \mathcal{LN}_{pdf} &= \mu_{\mathcal{LN}} \\ \widetilde{\kappa}_2 \mathcal{LN}_{pdf} &= \sigma_{\mathcal{LN}}^2 \\ \widetilde{\kappa}_3 \mathcal{LN}_{pdf} &= 0 \end{aligned} \quad (5)$$

où $\mu_{\mathcal{LN}}$ est un paramètre d'échelle et $\sigma_{\mathcal{LN}}$ un paramètre d'écart-type. Les log-cumulants de la loi Log-Normale sont nuls pour tout ordre strictement supérieur à 2.

2.2 Représentation dans le diagramme des log-cumulants d'ordre 2 et 3

On présente en figure 1 les positions de ces distributions (régions pour celles à trois paramètres et courbes pour celles à

deux paramètres). Les larges régions recouvertes par les lois Gamma Généralisées et les lois de Fisher démontrent l'aspect générique de ces lois (à noter toutefois que l'axe des ordonnées est exclu de la région la plus étendue correspondant aux Gamma Généralisées). En particulier, on observe que les lois \mathcal{K} peuvent être théoriquement imitées par une loi Gamma Généralisée comme par une loi de Fisher.

Pour appliquer une méthode d'estimation de paramètres fondée sur les log-cumulants (méthode dite « MLC », offrant des performances très compétitives), il est nécessaire, pour la loi Gamma Généralisée et pour la loi de Fisher, que les log-cumulants empiriques de ces distributions vérifient les équations [7] :

$$\frac{\widehat{\kappa}_2 \mathcal{G}\mathcal{G}_{pdf}}{\widehat{\kappa}_3 \mathcal{G}\mathcal{G}_{pdf}} > \frac{1}{4} \quad (6)$$

$$\widehat{\kappa}_2 \mathcal{F}_{pdf} \geq \Psi(1, \Phi(2, -|\widehat{\kappa}_3 \mathcal{F}_{pdf}|)) \quad (7)$$

où $\Phi(n, x)$ représente l'inverse de la Polygamma d'ordre n.

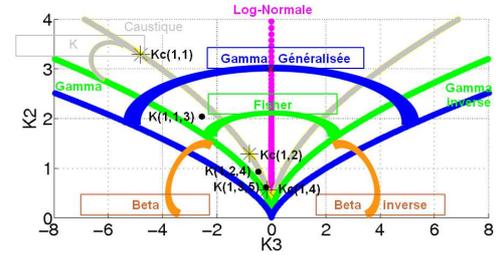


FIGURE 1 – Représentation des lois Gamma Généralisées, Fisher, \mathcal{K} et Log-Normales dans le diagramme $\widehat{\kappa}_2 - \widehat{\kappa}_3$.

Notons qu'en raison d'une propriété essentielle des fonctions Polygamma ($\Psi(r, x) \sim \frac{\Gamma(r)}{x^r}$), il est tout à fait justifié de se restreindre aux trois premiers log-cumulants pour caractériser et comparer les lois étudiées ici.

3 Mime d'une loi \mathcal{K} par Gamma Généralisée : comparaison avec Fisher

3.1 Méthodologie

3.1.1 Estimation des paramètres des lois imitatrices à partir des log-cumulants

On suppose que les paramètres de la loi \mathcal{K} de référence mimée sont inconnus, comme c'est le cas en pratique, et que seuls des échantillons distribués selon cette loi sont disponibles (il s'agit de N_s valeurs en intensités de pixels appartenant à une région, dont l'histogramme peut être bien modélisé par \mathcal{K}).

L'estimation des paramètres de la loi imitatrice (ici, Gamma Généralisée ou Fisher) est réalisée à partir des trois premiers log-cumulants empiriques de la loi \mathcal{K} mimée selon :

1. Calcul des log-cumulants empiriques $\widehat{\kappa}_2$, $\widehat{\kappa}_3$ puis $\widehat{\kappa}_1$ à partir des échantillons (voir [6]).

2. Identification de ces valeurs avec les expressions analytiques de $\widehat{\kappa}_{2GG_{pdf}}$, $\widehat{\kappa}_{3GG_{pdf}}$ puis $\widehat{\kappa}_{1GG_{pdf}}$ ou $\widehat{\kappa}_{2F_{pdf}}$, $\widehat{\kappa}_{3F_{pdf}}$ puis $\widehat{\kappa}_{1F_{pdf}}$, pour estimer les paramètres L_{GG} , η_{GG} puis μ_{GG} ou $L_{\mathcal{F}}$, $M_{\mathcal{F}}$ puis $\mu_{\mathcal{F}}$ des lois imitatrices.

3.1.2 Critères de comparaison fondés sur les divergences de Kullback-Leibler (KLD)

Nous rappelons que la KLD, qui mesure l'approximation de la distribution \mathcal{P} (observation) par \mathcal{Q} (modèle) est définie par :

$$KLD(\mathcal{P}_{pdf}(x), \mathcal{Q}_{pdf}(x)) = \int_0^{\infty} \mathcal{P}_{pdf}(x) \ln \left(\frac{\mathcal{P}_{pdf}(x)}{\mathcal{Q}_{pdf}(x)} \right) dx \quad (8)$$

Nous proposons d'estimer le paramètre d'échelle et la variance de plusieurs KLD calculées entre \mathcal{K} et Gamma Généralisée ou entre \mathcal{K} et Fisher selon :

1. Définition (i) d'un jeu de paramètres initiaux décrivant une loi \mathcal{K} mimée et (ii) d'une taille d'échantillon
2. Génération de plusieurs réalisations d'échantillons distribués selon cette loi \mathcal{K}
3. Pour chaque réalisation, (i) estimation des paramètres de la loi imitatrice (Gamma Généralisée ou Fisher) selon 3.1.1 et (ii) calcul de la KLD entre la loi mimée \mathcal{K} (dont les paramètres ont été initialement fixés) et la loi estimée
4. Calcul de la moyenne (pour une analyse dite « en précision ») et de la variance (pour une analyse dite « en robustesse ») de ces KLD sur les différentes réalisations.

Dans ces travaux, une petite moyenne des KLD sera caractéristique d'une bonne précision, tandis qu'une petite variance des KLD sera caractéristique d'une bonne robustesse.

Il est possible de donner une forme simplifiée des expressions analytiques des KLD entre les lois \mathcal{K} /Gamma-Généralisée et les lois \mathcal{K} /Fisher (voir [8]). Pour certaines valeurs de paramètres, celles-ci peuvent être évaluées numériquement en utilisant un logiciel de calcul formel tel que Maple.

3.2 Résultats et observations

Plusieurs cas d'études ont été analysés pour comparer le potentiel d'imitation de Gamma Généralisée et de Fisher pour mimer une loi \mathcal{K} (voir [8] et figure 1, où les étoiles et les disques indiquent la localisation des points associés aux six jeux traités) et ont permis d'aboutir à des conclusions similaires. Lors du calcul des moyennes et variances de KLD, 50 réalisations ont été considérées pour chaque taille d'échantillon N_s entre 250 et 2500. Seuls les résultats du jeu $\mathcal{K}(1, 1, 3)$ sont présentés en figure 2, où l'on peut voir respectivement l'évolution des moyennes et des variances des KLD en fonction de N_s .

Nous observons que la loi Gamma Généralisée semble légèrement meilleure que la loi de Fisher en précision (moyennes des KLD inférieures) quelle que soit la taille d'échantillon et en robustesse (variances des KLD inférieures) à partir d'une certaine taille (ici, au-delà de 500 pixels).

Par ailleurs, si on suppose que les paramètres de la loi \mathcal{K} de référence mimée sont connus, il est possible de calculer les valeurs constantes des KLD entre la loi \mathcal{K} mimée et la loi imitatrice, estimée cette fois non pas à partir des log-cumulants empiriques de la loi \mathcal{K} mais à partir de ses log-cumulants vrais. Ces distances correspondent aux valeurs vers lesquelles convergent les moyennes des KLD, lorsque N_s tend vers l'infini (indiquées en trait plein sur la figure 2). Il est aussi possible alors de superposer sur un même graphe la ddp de la loi \mathcal{K} mimée (ses paramètres étant connus) et celle de la loi imitatrice, afin d'apprécier qualitativement le potentiel de mime de cette dernière. Il est ainsi montré dans [8] que, visuellement, la loi Gamma Généralisée mime la loi \mathcal{K} légèrement mieux que la loi de Fisher, bien que les deux lois fournissent une très bonne adéquation.

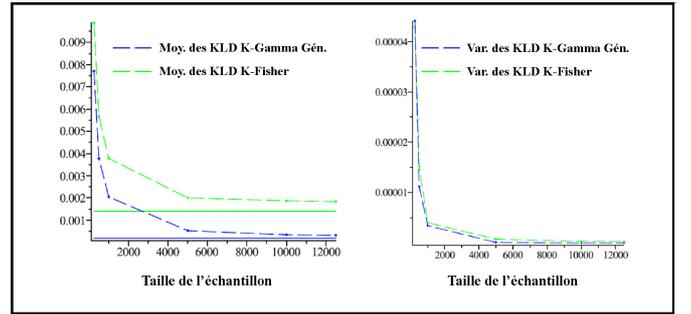


FIGURE 2 – Évolution en fonction de la taille d'échantillon des moyennes et des variances des KLD, calculées entre la loi \mathcal{K} mimée et la loi imitatrice estimée (Gamma Généralisée ou Fisher), pour le jeu $\mathcal{K}(1, 1, 3)$.

4 Limites du potentiel de mime de la loi Gamma Généralisée et imitation d'une loi Log-Normale

4.1 Considérations théoriques et pratiques

Indépendamment des conclusions dérivant de critères d'évaluation de mime, il est important de prendre en compte les points suivants pour la recommandation d'une loi de modélisation sur données réelles. (1) L'applicabilité de la méthode MLC doit être vérifiée (voir équation 6) pour l'estimation des paramètres de la loi Gamma Généralisée (voir [2] et [7] pour les cas pratiques). (2) La loi Gamma Généralisée n'étant théoriquement pas définie analytiquement sur l'axe des ordonnées, il convient aussi de vérifier dans le diagramme que, parmi les couples $(\widehat{\kappa}_{2GG_{pdf}}, \widehat{\kappa}_{3GG_{pdf}})$ du nuage estimé empiriquement sur des fenêtres, aucun ne se situe sur l'axe de la Log-Normale. (3) Le risque d'instabilités numériques lors de l'inversion de la fonction Polygamma doit également être souligné dans les zones « critiques » : ici, autour de l'axe de la Log-Normale, le paramètre η_{GG} tend vers zéro. Ceci est illustré dans [8], où l'on observe sur un extrait TerraSAR-X un nombre important

d'itérations avant convergence. Dans ces deux derniers types de situation, l'usage de la loi Gamma Généralisée avec la méthode MLC usuelle ne semble plus être adapté et des solutions alternatives doivent être préconisées (voir [2] et ci-dessous).

4.2 La loi Log-Normale, limite asymptotique de la loi Gamma Généralisée

En réalité, il est possible, par passage à la limite de la loi Gamma Généralisée, d'approximer une loi Log-Normale, suivant la méthode originale d'estimation des paramètres de la loi imitatrice proposée ci-dessous. En effet, en posant L_{GG} sous la forme particulière $L_{GG} = \frac{S_{GG}}{\eta_{GG}^2}$, on obtient (voir [8]) :

$$\lim_{\eta_{GG} \rightarrow 0} \widetilde{\kappa}_2 \mathcal{G}_{pdf} = \lim_{\eta_{GG} \rightarrow 0} \frac{\Psi(1, \frac{S_{GG}}{\eta_{GG}^2})}{\eta_{GG}^2} = \frac{1}{S_{GG}} \quad (9)$$

et on montre que tous les log-cumulants $\widetilde{\kappa}_r \mathcal{G}_{pdf}$ d'ordre r strictement supérieur à 2 sont nuls, lorsque η_{GG} tend vers zéro.

En identifiant alors les log-cumulants d'ordre 2 et d'ordre 1 de la loi Gamma Généralisée avec ceux de la loi Log-Normale, on obtient (voir [8]), lorsque η_{GG} tend vers zéro, que $\frac{1}{S_{GG}} = \sigma_{LN}^2$ et que $\ln(\mu_{GG}) = \mu_{LN}$, grâce à l'approximation de la fonction Digamma $\Psi(x) \sim \log(x)$, pour x tendant vers l'infini.

On a ainsi montré qu'il est possible d'identifier la loi Log-Normale comme la limite d'une loi Gamma Généralisée, en définissant, à travers la forme spécifique de L_{GG} , une nouvelle loi à deux paramètres (le troisième paramètre de la loi Gamma Généralisée permet d'affiner la précision de l'approximation). Ce passage à la limite (η_{GG} tendant vers zéro) se justifie, car on peut facilement montrer avec le logiciel Maple que la fonction caractéristique de deuxième espèce de la loi Gamma Généralisée $\phi_{GG_{pdf}}(s)$, décrite par la relation $L_{GG} = \frac{S_{GG}}{\eta_{GG}^2}$, est définie pour un voisinage de $s = 1$ (voir [8]).

En supposant connus les paramètres de lois Log-Normales mimées et en superposant leurs ddps et celles des lois Gamma Généralisées imitatrices, on peut apprécier dans [8] la qualité des imitations obtenues. Plusieurs exemples de Gamma Généralisées mimant des Log-Normales sont présentés en figure 3 dans le diagramme $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$. On note au niveau de l'axe vertical l'allure plate des lois imitatrices (estimées selon la méthode ci-dessus), qui pourrait assurer une certaine stabilité des résultats.

5 Conclusions et perspectives

Dans une première contribution, nous avons montré la légère prédominance de la loi Gamma Généralisée sur la loi de Fisher pour mimer une loi \mathcal{K} de référence. Dans une seconde contribution, nous avons prouvé que les lois Gamma Généralisées peuvent assurer une continuité sur l'axe vertical du diagramme $\widetilde{\kappa}_2 - \widetilde{\kappa}_3$. Dans nos futures recherches, la comparaison du potentiel de mimes des lois Gamma Généralisée et Fisher pourrait être menée avec d'autres lois de référence et les aptitudes de ces lois génériques à modéliser différents comportements pourraient être analysées sur des extraits radar réels variés (voir [8]).

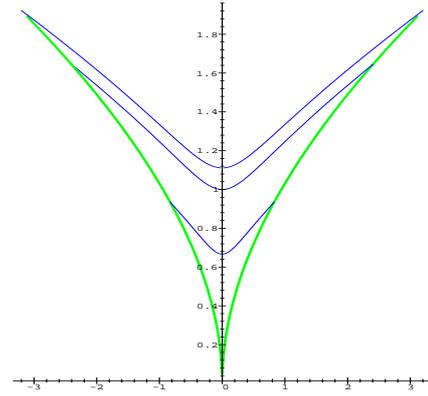


FIGURE 3 – Imitations de la loi Log-normale par la loi Gamma Généralisée pour des valeurs de η_{GG} entre -1 et +1 et pour $S_{GG} \in \{0.9; 1; 1.5\}$. La loi Gamma et la loi Gamma Inverse sont tracées (en vert) de part et d'autre de ces imitations (en bleu) pour des valeurs de $L_G \in [0.9, \infty[$ et de $L_{GI} \in [0.9, \infty[$.

Références

- [1] E. W. Stacy, "A Generalization of the Gamma distribution," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1187–1192, 1962.
- [2] H.-C. Li, W. Hong, Y.-R. Wu, and P.-Z. Fan, "On the empirical-statistical modeling of SAR images with generalized Gamma distribution," *IEEE J-STSP*, vol. 5, no. 3, pp. 386–397, 2011.
- [3] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE TAP*, vol. 24, no. 6, pp. 806–814, 1976.
- [4] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maître, "A new statistical model for markovian classification of urban areas in high-resolution SAR images," *IEEE TGRS*, vol. 42, no. 10, pp. 2046–2057, 2004.
- [5] G. V. Trunk and S. F. George, "Detection of targets in non-Gaussian sea clutter," *IEEE TAES*, no. 5, pp. 620–628, 1970.
- [6] J.-M. Nicolas, "Application de la transformée de Mellin : étude des lois statistiques de l'imagerie cohérente," Telecom ParisTech, Paris, France, Tech. Rep. 2006D010, 2006.
- [7] V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "On the method of logarithmic cumulants for parametric probability density function estimation," *IEEE TIP*, vol. 22, no. 10, pp. 3791–3806, 2013.
- [8] H. Sportouche, M. M. Horta, F. Tupin, J.-M. Nicolas, and N. D. A. Mascarenhas, "Statistical modeling of High Resolution SAR data : Fisher Versus Generalized Gamma distributions ?" *currently under writting for coming submission to Pattern Recognition*.